Analysis and Interpretation of Freshwater Fisheries Data

is a special project of the
Education Section of the American Fisheries Society

Analysis and Interpretation of Freshwater Fisheries Data

Edited by

CHRISTOPHER S. GUY
U.S. Geological Survey
Montana Cooperative Fishery Research Unit
Montana State University

and

MICHAEL L. BROWN
Department of Wildlife and Fisheries Sciences
South Dakota State University

American Fisheries Society
Bethesda, Maryland, USA
2007

Suggested citation formats are

Entire book

Guy, C. S., and M. L. Brown, editors. 2007. Analysis and interpretation of freshwater fisheries data. American Fisheries Society, Bethesda, Maryland.

Chapter in book

Hayes, D. B., J. R. Bence, T. J. Kwak, and B. E. Thompson. 2007. Abundance, biomass, and production. Pages 327–374 *in* C. S. Guy and M. L. Brown, editors. Analysis and interpretation of freshwater fisheries data. American Fisheries Society, Bethesda, Maryland.

Printed in the United States of America on acid-free paper.

Library of Congress Control Number 2007925455

ISBN 978-1-888569-77-3

American Fisheries Society Web site: *www.fisheries.org*

American Fisheries Society
5410 Grosvenor Lane, Suite 110
Bethesda, Maryland 20814-2199
USA

In Memoriam

A special dedication to the memory of statistical consultant, colleague, and friend Jeffrey Pontius (1954–2006), who passed away during the development of this book.

Contents

Contributors

Micheal S. Allen (Chapter 9): The University of Florida, Department of Fisheries and Aquatic Sciences, 7922 Northwest 71st Street, Gainesville, Florida 32653, USA; msal@ufl.edu

Douglas J. Austen (Chapter 2): Pennsylvania Fish and Boat Commission, 1601 Elmerton Avenue, Harrisburg, Pennsylvania 17106, USA; Dausten@state.pa.us

T. Douglas Beard Jr. (Chapter 3): U.S. Geological Survey, Aquatic and Endangered Resources, MS-301, 12201 Sunrise Valley Drive, Reston, Virginia 20192, USA; dbeard@usgs.gov

David A. Beauchamp (Chapter 16): U.S. Geological Survey, University of Washington, Washington Cooperative Fish and Wildlife Research Unit, School of Aquatic and Fisheries Sciences, Box 355020, Seattle, Washington 98195, USA; davebea@u.washington.edu

James R. Bence (Chapter 8): Michigan State University, Department of Fisheries and Wildlife, 13 Natural Resources Building, East Lansing, Michigan 48824, USA; bence@msu.edu

Phillip W. Bettoli (Chapter 6): U.S. Geological Survey, Tennessee Cooperative Fishery Research Unit, Tennessee Technological University, Box 5114, Cookeville, Tennessee 38505, USA; pbettoli@tntech.edu

Michael L. Brown (Coeditor, Chapter 1): South Dakota State University, Department of Wildlife and Fisheries Sciences, Box 2140B, Brookings, South Dakota 57007, USA; michael.brown@sdstate.edu

Steven R. Chipps (Chapter 11): U.S. Geological Survey, South Dakota Cooperative Fish and Wildlife Research Unit, South Dakota State University, Department of Wildlife and Fisheries Sciences, Box 2140B, Brookings, South Dakota 57007, USA; steven.chipps@sdstate.edu

Mary C. Fabrizio (Chapter 7): Virginia Institute of Marine Science, Department of Fisheries Science, Post Office Box 1346, Gloucester Point, VA 23062, USA; mfabrizio@vims.edu

James E. Garvey (Chapter 11): Southern Illinois University Carbondale, Fisheries and Illinois Aquaculture Center and Department of Zoology, Carbondale, Illinois 62901, USA; jgarvey@siu.edu

Timothy B. Grabowski (Chapter 5): The University of Georgia, Warnell School of Forestry and Natural Resources, Athens, Georgia, 30602, USA; tgrabow@uga.edu

Christopher S. Guy (Coeditor, Chapter 1): U.S. Geological Survey, Montana Cooperative Fishery Research Unit, Montana State University, Fish and Wildlife Management Program, 301 Lewis Hall, Bozeman, Montana 59717, USA; cguy@montana.edu

Michael J. Hansen (Chapter 3): University of Wisconsin-Stevens Point, College of Natural Resources, 800 Reserve Street, Stevens Point, Wisconsin 54481, USA; mhansen@uwsp.edu

Kyle J. Hartman (Chapter 12): West Virginia University, Wildlife and Fisheries Resources Program, Division of Forestry, 320 Percival Hall, Morgantown, West Virginia 26506, USA; hartman@wvu.edu

Daniel B. Hayes (Chapters 3 and 8): Michigan State University, Department of Fisheries and Wildlife, 13 Natural Resources Building, East Lansing, Michigan 48824, USA; hayesdan@msu.edu

Robert S. Hayward (Chapter 12): University of Missouri, Department of Fisheries and Wildlife Sciences, 302 Anheuser-Busch Natural Resources Building, Columbia, Missouri 65211, USA; haywardr@missouri.edu

Wayne A. Hubert (Chapter 7): U.S. Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, University of Wyoming, Department 3166, 1000 East University Avenue, Laramie, Wyoming 82071, USA; WHubert@uwyo.edu

J. Jeffery Isely (Chapter 5): U.S. Geological Survey, South Carolina Cooperative Fish and Wildlife Research Unit, Clemson University, G-20 Lehotsky Hall, Clemson, South Carolina 29634, USA; jisely@clemson.edu

Donald A. Jackson (Chapter 18): University of Toronto, Department of Ecology and Evolutionary Biology, 25 Harbord Street, Toronto, Ontario M5S 3G5, Canada; jackson@zoo.utoronto.ca

Brett M. Johnson (Chapter 16): Colorado State University, Department of Fish, Wildlife, and Conservation Biology, 1474 Campus Delivery, Fort Collins, Colorado 80523, USA; brett@cnr.colostate.edu

Carter G. Kruse (Chapter 10): Turner Enterprises, Inc., 1123 Research Drive, Bozeman, Montana 59718, USA; carter.kruse@retranches.com

Thomas J. Kwak (Chapters 8 and 15): U.S. Geological Survey, North Carolina Cooperative Fish and Wildlife Research Unit, North Carolina State University, Box 7617, Raleigh, North Carolina 27695, USA; tkwak@ncsu.edu

Michael J. Maceina (Chapter 4): Auburn University, Department of Fisheries, 203 Swingle Hall, Alabama 36849, USA; mmaceina@acesag.auburn.edu

Leandro E. Miranda (Chapter 6): U.S. Geological Survey, Mississippi Cooperative Fish and Wildlife Research Unit, Mississippi State University, Post Office Box 9691, Mississippi 39762, USA; smiranda@cfr.msstate.edu

Robert M. Neumann (Chapter 9): In-Fisherman, 7819 Highland Scenic Road, Baxter, Minnesota 56425, USA; rneumann@in-fisherman.com

Tammy J. Newcomb (Chapter 17): Michigan Department of Natural Resources, Post Office Box 30446, East Lansing, Michigan 48909, USA; newcombt@michigan.gov

Richard L. Noble (Chapter 2): North Carolina State University, Department of Zoology and Forestry (Emeritus), Raleigh, North Carolina 27695, USA; pnoble@mindspring.com

Donald J. Orth (Chapter 17): Virginia Polytechnic Institute and State University, Department of Fisheries and Wildlife Sciences, 100 Cheatham Hall, Blacksburg, Virginia 24061, USA; dorth@vt.edu

Mark A. Pegg (Chapter 2): University of Nebraska-Lincoln, School of Natural Resources, 402 Hardin Hall, Lincoln, Nebraska 68583, USA; mpegg2@unl.edu

Donald L. Pereira (Chapter 4): Minnesota Department of Natural Resources, Division of Fish and Wildlife, 500 Lafayette Road, Box 20, St. Paul, Minnesota 55155, USA; don.pereira@dnr.state.mn.us

James T. Peterson (Chapter 15): U.S. Geological Survey, Georgia Cooperative Fish and Wildlife Research Unit, University of Georgia, Warnell School of Forest Resources, Athens, Georgia 30602, USA; peterson@smokey.forestry.uga.edu

Kevin L. Pope (Chapter 10): U.S. Geological Survey, Nebraska Cooperative Fish and Wildlife Research Unit, University of Nebraska-Lincoln, 103 Miller Hall, Lincoln, Nebraska 68583, USA; kpope2@unl.edu

Michael Power (Chapter 13): University of Waterloo, Department of Biology, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada; m3power@scimail.uwaterloo.ca

Frank J. Rahel (Chapter 18): University of Wyoming, Department of Zoology and Physiology, Department 3166, 1000 E. University Ave., Laramie, Wyoming 82071, USA; FRahel@uwyo.edu

Kevin B. Rogers (Chapter 14): Colorado Division of Wildlife, Post Office Box 775777, Steamboat Springs, Colorado, 80477, USA; Kevin.Rogers@state.co.us

Dean Stauffer (Chapter 17): Virginia Polytechnic Institute and State University, Department of Fisheries and Wildlife Sciences, 100 Cheatham Hall, Blacksburg, Virginia 24061, USA; dstauffe@vt.edu

Bradley E. Thompson (Chapter 8): Michigan State University, Department of Fisheries and Wildlife, 13 Natural Resources Building, East Lansing, Michigan 48824, USA; thomp149@msu.edu

David H. Wahl (Chapter 16): Illinois Natural History Survey, Kaskaskia Biological Station, RR1 Box 157, Sullivan, Illinois 61951, USA; d_wahl@uiuc.edu

Gary C. White (Chapter 14): Colorado State University, Department of Fish, Wildlife, and Conservation Biology, 1474 Campus Drive, Fort Collins, Colorado 80523, USA; gwhite@cnr.colostate.edu

Preface

Our rationale for the development of this book was simple but twofold. First, there are numerous statistics reference texts available for biostatistical analyses. Our intent was not to repeat those efforts but to take a more functional approach. Rather than supply another test-oriented book (e.g., analysis of variance or multiple regression) with fisheries examples, the major fisheries data types dictated the development and presentation of statistical approaches. Secondarily, this book provides the profession with a frame of reference to encourage appropriate sample design, analysis, and interpretation of freshwater fisheries data. We narrowed the scope of the book to freshwater data types because we believed by including marine analytical methods the book would become excessively cumbersome. Although there is much overlap in freshwater and marine analytical methods, there are inherent differences with regard to research problems and data collection.

This book was also developed to complement *Fisheries Techniques* (Murphy and Willis 1996.) and *Inland Fisheries Management in North America* (Kohler and Hubert 1999). Several years ago we observed a lack of textbooks regarding data analysis with user-friendly examples. Thus, we proposed this book to fill this void and help practicing fisheries professionals and students move beyond sampling and into analysis of data. Accordingly, this book was designed as a reference for practicing fisheries professionals, and we intend the book to be used in advanced undergraduate and graduate courses. We intentionally asked authors to minimize text regarding the theory, derivation, and development of formulas related to statistical testing or mathematical modeling. Our intent for this book was to minimize complexity and make it user friendly. There are several excellent books that explain statistical theory and model development which are referenced in this book—our intent was not to duplicate those efforts. If we have succeeded in our efforts, the reader should be able to follow the methods and box examples and obtain the same results as the authors by using the data found on the accompanying compact disc.

Following the successful approach taken in the production of *Fisheries Techniques* and *Inland Fisheries Management of North America,* we attempted to select a minimum of two authors per chapter. We also had several reviewers per chapter.

At least one fisheries professional who was knowledgeable of the data type was recruited for review. Aside from our technical reviews, Drs. Jeffrey Pontius (Kansas State University) and Kenneth Gerow (University of Wyoming) were contracted as statistical consultants to ensure that appropriate statistical approaches were followed. Lastly, several chapters were reviewed by fisheries graduate students to gauge reading level and class utility. We are greatly indebted to all of these reviewers for their conscientious efforts.

Fisheries data analysis is not easy and can often be conducted in several ways. For example, two fisheries scientists may analyze the same data set differently, given they are trying to answer the same question. Similarly, statistical analyses are complex and several statistical approaches may be used on a single data set. Two statistical consultants may suggest different methods for analyzing the same data set. Therefore, this book is not a "how to guide" but rather a reference to provide a better understanding of data analysis techniques and increase awareness by using current, appropriate analytical methods. Chapter 1 includes some text on the more complex and novel approaches to analyzing data. However, many of the techniques outlined in the remaining chapters focus on standard techniques that have been used and proven useful for many years. It would be impossible to include every data analysis technique, but we attempted to include the more common methods and encourage readers to go beyond the text and develop additional methods that can better address specific fisheries management problems.

There are numerous software packages that are available to analyze data, and we were forced to choose a single software package for consistency among chapters. We selected SAS (SAS Institute, Cary, North Carolina) given that many of the authors were familiar with this software and most statisticians are familiar with the software. We recognize the shortcomings associated with SAS but also understand that there would be shortcomings with any software package we would have selected. Additionally, different SAS versions will yield slightly different output and results may differ from output as depicted in the box examples derived from the code on the accompanying compact disc. We encourage readers to develop program language using other software packages (e.g., R; available http://www.r-project.org; January 2007) with the data used in the box examples. We are exploring the potential for an affiliated webpage for the book where programs written using software other than SAS could be included along with additional developed examples.

We thank the following for their volunteer efforts as reviewers: P. L. Angermeier, J. R. Bence, M. S. Bevelhimer, B. G. Blackwell, J. C. Boxrucker, M. Bozek, P. J. Braaten, J. E. Breck, S. R. Chipps, M. A. Colvin, S. J. Cooke, J. S. Diana, W. G. Duffy, T. E. Essington, C. P. Ferreri, W. L. Fisher, L. S. Fore, B. D. S. Graeb, R. Gresswell, D. B. Hayes, C. W. Hoagstrom, E. R. Irwin, D. A. Isermann, T. E. McMahon, M. H. Meeuwig, P. H. Michaletz, C. L. Milewski, B. R. Murphy, J. J. Ney, D. L. Parrish, C. P. Paukert, M. C. Quist (two chapters), C. F. Rabeni, B. A. Rich, S. M. Sammons, J. Schreer, R. A. Stein, T. M. Sutton, J. S. Tillma, M. J. Van Den Avyle (two chapters), D. W. Willis, and A. V. Zale. We thank I. Davis for the cover art. We

thank *In-Fisherman* (J. Simpson and R. M. Neumann) for the blue catfish photo, S. L. Denson for the river photo, and M. H. Meeuwig for the photo of fisheries scientists.

The editors thank Brian Murphy, Chuck Scalet, and Dave Willis for their ideas regarding the development of this book. The editors also thank the Presidents of Education Section during the development of the book, Chris Kohler, Al Zale, Tom Coon, and Rob Neumann (in chronological order of presidency). Mike Maceina and Don Pereira (Chapter 4) thank Terrance Quinn II and Sandy Weisberg for assistance with statistical models. Jeff Isley and Tim Grabowski (Chapter 5) thank William Bridges, Jr. Wayne Hubert and Mary Fabrizio (Chapter 7) thank Ken Gerow, Jeff Pontius, Mike Maceina, and Ann Zimmerman for assistance and guidance throughout the development of their chapter. Rob Neumann and Mike Allen (Chapter 9) thank Ramon Littell, Rich Cailteux, and Marty Hale for helpful suggestions. Steve Chipps and Jim Garvey (Chapter 11) thank Roy Stein for his advice and comments. Michael Power (Chapter 13) thanks Marshall Adams, Martin Attrill, Lynn McCarty, James Reist, and Brian Dempson for discussions on the topic of population bioassessment. Particular thanks go to Geoff Power who provided exacting comments on the initial draft of the Chapter 13. Kevin Rogers and Gary White (Chapter 14) thank Eric Bergersen and Greg Langer for the use of their research in the examples. Dave Beauchamp, Dave Wahl, and Brett Johnson (Chapter 16) thank Alison Cross, Lisa Einfalt, Tracy Galarowicz, Marci Koski, Mike Mazur, Jim Matilla, Jenifer McIntyre, Jamal Moss, Nathanael Overman, and Ruth Wagner for their helpful suggestions and contributions. Frank Rahel and Don Jackson (Chapter 18) thank Ken Gerow, Wayne Hubert, Nathan Nibbelink, Amy Schrank, and Dan Isaack for helpful comments and statistical insight. Research support for Don Jackson was provided by the Natural Sciences and Engineering Research Council of Canada.

The editors thank Eva Silverfine for her outstanding work as technical editor. Finally, the editors thank Debby Lehman and Aaron Lerner with the American Fisheries Society for their patience and professionalism throughout the development of this book.

Mention of trade names or product vendors does not imply endorsement by the American Fisheries Society, authors, editors, or the employers of the book's contributors.

Christopher S. Guy
Michael L. Brown

References

Kohler, C. C, and W. A. Hubert, editors. 1999. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Murphy, B. R., and D. W. Willis, editors. 1996. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

List of Species

Fish

alewife	*Alosa pseudoharengus*
American eel	*Anguilla rostrata*
American shad	*Alosa sapidissima*
Arkansas River shiner	*Notropis girardi*
Atlantic cod	*Gadus morhua*
Atlantic menhaden	*Brevoortia tyrannus*
Atlantic salmon	*Salmo salar*
Atlantic sturgeon	*Acipenser oxyrinchus*
banded darter	*Etheostoma zonale*
bay anchovy	*Anchoa mitchilli*
bigeye tuna	*Thunnus obesus*
bighead carp	*Hypophthalmichthys nobilis*
black basses	*Micropterus* spp.
black bullhead	*Ameiurus melas*
black crappie	*Pomoxis nigromaculatus*
black redhorse	*Moxostoma duquesnei*
blackside darter	*Percina maculata*
bloater	*Coregonus hoyi*
bluefish	*Pomatomus saltatrix*
bluegill	*Lepomis macrochirus*
bluntnose minnow	*Pimephales notatus*
brook silverside	*Labidesthes sicculus*
brook trout	*Salvelinus fontinalis*
brown trout	*Salmo trutta*
bull trout	*Salvelinus confluentus*
bullhead	*Cottus gobio*
bullhead minnow	*Pimephales vigilax*
burbot	*Lota lota*

carps	Family Cyprinidae
central mudminnow	*Umbra limi*
chain pickerel	*Esox niger*
channel catfish	*Ictalurus punctatus*
Chinook salmon	*Oncorhynchus tshawytscha*
cisco	*Coregonus artedi*
ciscoes	*Coregonus* spp. and *Prosopium* spp.
coho salmon	*Oncorhynchus kisutch*
common carp	*Cyprinus carpio*
common snook	*Centropomus undecimalis*
crappies	*Pomoxis* spp.
cunner	*Tautogolabrus adspersus*
cutthroat trout	*Oncorhynchus clarkii*
darters	*Etheostoma* spp. and Family Percidae
desert pupfish	*Cyprinodon macularius*
English sole	*Parophrys vetulus*
Eurasian perch	*Perca fluviatilis*
fathead minnow	*Pimephales promelas*
flathead catfish	*Pylodictis olivaris*
Florida largemouth bass	*Micropterus salmoides floridanus*
gizzard shad	*Dorosoma cepedianum*
golden redhorse	*Moxostoma erythrurum*
golden shiner	*Notemigonus crysoleucas*
goldfish	*Carassius auratus*
grass carp	*Ctenopharyngodon idella*
green sunfish	*Lepomis cyanellus*
greenback cutthroat trout	*Oncorhynchus clarkii stomias*
haddock	*Melanogrammus aeglefinus*
herrings	Family Clupeidae
hornyhead chub	*Nocomis biguttatus*
hybrid striped bass	*M. chrysops* ♂ × *Morone saxatilis* ♀
jaguar guapote	*Cichlasoma managuense*
johnny darter	*Etheostoma nigrum*
kokanee	*Oncorhynchus nerka*
Lahontan cutthroat trout	*Oncorhynchus clarkii henshawi*
lake trout	*Salvelinus namaycush*
lake whitefish	*Coregonus clupeaformis*
largemouth bass	*Micropterus salmoides*
leopard darter	*Percina pantherina*
logperch	*Percina caprodes*
longear sunfish	*Lepomis megalotis*
longnose dace	*Rhinichthys cataractae*
longnose gar	*Lepisosteus osseus*
mimic shiner	*Notropis volucellus*
minnows	Family Cyprinidae

molly miller	*Scartella cristata*
mottled sculpin	*Cottus bairdii*
muskellunge	*Esox masquinongy*
North American catfishes	Family Ictaluridae
northern hog sucker	*Hypentelium nigricans*
northern largemouth bass	*Micropterus salmoides salmoides*
northern pike	*Esox lucius*
northern pikeminnow	*Ptychocheilus oregonensis*
northern redbelly dace	*Phoxinus eos*
orangespotted sunfish	*Lepomis humilis*
Pacific cod	*Gadus macrocephalus*
Pacific salmon	*Oncorhynchus* spp.
paddlefish	*Polyodon spathula*
perches	Family Percidae
pike	Family Esocidae
pumpkinseed	*Lepomis gibbosus*
quillback	*Carpiodes cyprinus*
rainbow smelt	*Osmerus mordax*
rainbow trout	*Oncorhynchus mykiss*
red snapper	*Lutjanus campechanus*
redfin shiner	*Lythrurus umbratilis*
redside shiner	*Richardsonius balteatus*
river redhorse	*Moxostoma carinatum*
rock bass	*Ambloplites rupestris*
rosyface shiner	*Notropis rubellus*
round goby	*Neogobius melanostomus*
sablefish	*Anoplopoma fimbria*
salmons	Family Salmonidae
sand shiner	*Notropis stramineus*
sauger	*Sander canadensis*
sculpins	Family Cottidae
sea lamprey	*Petromyzon marinus*
shads	Family Clupeidae
shorthead redhorse	*Moxostoma macrolepidotum*
shovelnose sturgeon	*Scaphirhynchus platorynchus*
silver redhorse	*Moxostoma anisurum*
slenderhead darter	*Percina phoxocephala*
smallmouth bass	*Micropterus dolomieu*
smallmouth buffalo	*Ictiobus bubalus*
smelts	Family Osmeridae
sockeye salmon	*Oncorhynchus nerka*
southern flounder	*Paralichthys lethostigma*
spot	*Leiostomus xanthurus*
spotfin shiner	*Cyprinella spiloptera*
spotted bass	*Micropterus punctulatus*

spotted sucker	*Minytrema melanops*
steelhead	*Oncorhynchus mykiss*
striped bass	*Morone saxatilis*
striped shiner	*Luxilus chrysocephalus*
sturgeons	Family Acipenseridae
suckermouth minnow	*Phenacobius mirabilis*
suckers	Family Catostomidae
sunfishes	Family Centrarchidae
temperate basses	*Morone* spp.
threadfin shad	*Dorosoma petenense*
Topeka shiner	*Notropis topeka*
trouts	Family Salmonidae
walleye pollock	*Theragra chalcogramma*
walleye	*Sander vitreus*
weakfish	*Cynoscion regalis*
westslope cutthroat trout	*Oncorhynchus clarkii lewisi*
white bass	*Morone chrysops*
white crappie	*Pomoxis annularis*
white perch	*Morone americana*
white sturgeon	*Acipenser transmontanus*
white sucker	*Catostomus commersonii*
whitefishes	*Coregonus* spp. and *Prosopium* spp.
winter flounder	*Pseudopleuronectes americanus*
yellow perch	*Perca flavescens*
yellowfin sole	*Limanda aspera*
Yellowstone cutthroat trout	*Oncorhynchus clarkii bouvieri*
zander	*Stizostedion lucioperca*

Other Animals

American beaver	*Castor canadensis*
American lobster	*Homarus americanus*
anchor worm	*Lernaea cyprinacea*
brine shrimp	*Artemia* spp.
cladoceran	*Daphnia pulicaria*
exotic cladoceran	*Daphnia lumholtzi*
loggerhead sea turtles	*Caretta caretta*
mysid shrimp	Family Mysidae
northern clearwater crayfish	*Orconectes propinquus*
opossum shrimp	*Mysis relicta*
virile crayfish	*Orconectes virilis*
zebra mussels	*Dreissena polymorpha*

1 Science and Statistics in Fisheries Research

Michael L. Brown and Christopher S. Guy

1.1 INTRODUCTION

Fisheries science is considered to be a relatively young profession in North America, with its origins in the late 1800s (Nielsen 1999). Considerably younger are many of the analytical tools currently used by fisheries scientists to develop interpretations of data during the decision-making process. Accordingly, the use of statistics in fisheries science has paralleled the development of statistical theory, approaches, and computing tools that facilitate both simple and complex analyses. Efron (1998) noted that 1925 was "the year that statistical theory became of age, the year statistics went from an ad hoc collection of ingenious techniques to a coherent discipline." Thus, in a way, fisheries and statistical sciences are of a similar vintage.

In the recent past, most freshwater fisheries management activities have been centered on controlling processes and population dynamics associated with single species, with the goal of maximizing numbers and sizes available to anglers. Mathematical treatments of inland fisheries data have followed that trend. However, the need to accommodate research and management at the ecosystem level has recently promoted the collection of diverse community and habitat data to address questions about processes and interconnectedness (Krueger and Decker 1999). The growing body of study designs and statistical analyses often causes some consternation within the profession; however, we are responsible for utilizing the most appropriate data management and statistical tools to formulate and assess research and management activities. Thus, by necessity, fisheries scientists must become more knowledgeable about the increasingly diverse array of statistical and data management tools and potential applications. This book, as well as several other texts, provides acceptable approaches to the analysis of common fish and fisheries data (Table 1.1).

A recent occurrence has been the use of alternative approaches to null hypothesis testing, such as methods based in decision, information, and Bayesian theories. All of these methods contribute to a growing body of analytical literature. Regardless of method, the intention of alternative methods is to minimize uncertainty in the decision-making process. Thoughtful review of the literature and consultation with a statistician on complex analytical approaches will greatly aid

Table 1.1 Guide to analytical approaches for specific fish and fishery data. Topics are addressed in the listed chapters, which are contained in the sources *Fisheries Techniques* (FT, Murphy and Willis 1996), *Inland Fisheries Management in North America* (IFM, Kohler and Hubert 1999), *Methods for Fish Biology* (MFB, Schreck and Moyle 1990), and this book (AIFFD). Collection and statistical analysis of angler data are addressed in *Angler Survey Methods and Their Applications in Fisheries Management* (Pollock et al. 1994).

	Source			
Topic	FT	IFM	MFB	AIFFD
Age and growth		6	11	5
Behavior			17	
Bioenergetics			12	12, 16
Community assessment		7	19	15
Condition	15			10
Food habits				11
Genetics			2	
Habitat				17
Mortality rate		6		6
Population bioassessment				13
Population size		6		
Predator–prey interactions				16
Production and yield		6		8
Recruitment		6		4
Relative abundance	21	7		7
Size structure	15	7		9
Telemetry	19			14
Toxicology			15	
Watershed				18

study design and data analysis. This chapter does not center on debates among statistical philosophies or methods but describes the science and the basic conditions for each approach. We also discuss general issues confronted in inference but refer readers to Chapter 2 and Chapter 3 for specific information concerning study design and sampling issues.

1.2 FISHERIES SCIENCE

To begin, what do we mean by fisheries science? To help answer this question we first define fisheries and science separately. Fisheries (plural for fishery) include (1) a population or assemblage of fishes used for commercial or recreational purposes, (2) habitats, and (3) associated humans. For example, the rainbow trout fishery in the Madison River, Montana, is a highly regarded recreational fishery by salmonid anglers throughout the world. Another example is the commercial fishery for sockeye salmon, which is arguably the most important salmonid fishery in North America (Behnke 2002). Strahler (1992) stated, "Science is the acquisition of reliable but not infallible knowledge of the real world, including explanations of the phenomena." The preceding quote is good because it includes the words

"infallible knowledge." Science is not without error, but science and the scientific method (see below) allow scientists to learn from past misconceptions. Science is really a way of obtaining reliable knowledge about the universe. Thus, fisheries science is the process of obtaining reliable knowledge about fisheries through scientific inquiry.

Scientific inquiry is associated with measurable metrics and is based on empirical evidence, not value judgments (Lee 1999). For example, to determine the effects of angling on a fishery we would establish hypotheses and predictions, design an appropriate experiment (Chapter 3), and measure population metrics such as total annual mortality (Chapter 6), exploitation rate (Chapter 6), and growth rate (Chapter 5). We would use these empirical data to test our hypotheses. It would be inappropriate and not scientific to assess the impact that angling might have on the beauty of a fishery because this is a value judgment. Similarly, we need to be aware of "belief fields," for which knowledge is based on belief. Belief is something that cannot be observed to exist physically, thus it is not science (Strahler 1992).

In addition to being careful about incorporating value judgments in science, scientists need to control subjectivity in science. Nevertheless, scientists are not machines and are quite capable of adding subjectivity into science (Lee 1999). Subjectivity can unknowingly enter science whereby the scientist has preconceived hypotheses about the way the world operates and unintentionally designs experiments that support those hypotheses. One control for subjectivity is peer review. Peer review is important at all levels of science, particularly in the study design and publication stages. Peer review helps maintain integrity in research. Issues related to fisheries science are becoming ever more entangled with the social, political, and economic fabric of society, thus it is important that fisheries scientists guide their research with the utmost integrity. Maintaining high ethical standards in research will help ensure the public's trust and support of research in fisheries science. The National Research Council outlined eight practices that a scientist should follow (IOM 2002): (1) intellectual honesty in proposing and reporting research; (2) accuracy in representing contributions to research proposals and reports; (3) fairness in peer review; (4) collegiality in scientific interactions, including communications and sharing resources; (5) transparency in conflicts of interest or potential conflicts of interest; (6) protection of human subjects in the conduct of research; (7) humane care of animals in the conduct of research; and (8) adherence to the mutual responsibilities between investigators and their research teams. The most important trait in a scientist is integrity; this is above intelligence, creativity, or determination (Lee 1999).

1.3 SCIENTIFIC METHOD AND RESEARCH

The fisheries profession was founded on animal husbandry and natural history observations, much like our counterpart, the wildlife profession (see Garton et al. 2005). Thus, many decisions regarding inland fisheries management are based on observational associations rather than experimental studies and the scientific method. The failure of the wildlife profession to follow the scientific method was

eloquently described by Romesburg (1981), and his arguments apply to the fisheries profession. Romesburg (1981) stated that the wildlife profession provides "unreliable knowledge" because researchers often do not follow accepted approaches to sound scientific inquiry, such as the hypothetico–deductive method. Popper (1959, 1968) popularized the hypothetico–deductive method after Chamberlin's (1965) work (first printed in 1890) emphasizing the need to examine multiple working hypotheses. Garton et al. (2005) describe the hypothetico–deductive method as "a circular process where previous information is synthesized into a theory, predictions are deduced from the theory, the predictions are stated explicitly in the form of hypotheses, hypotheses are tested through an investigation involving experimentation, observation, or quantitative models, the theory is supported, modified, or expanded on the basis of the results of these tests, and the process starts again." Some other early classic papers regarding the scientific method were written by Dewey (1938) and Platt (1964).

In defense of our profession, many of the studies we conduct are in highly variable environments, and we do not have the ability to control variables associated with the study such as physicists or chemists can do in their studies. Subsequently, much of the research in the fisheries profession is descriptive and has broad objectives. Scientists in other disciplines can regularly conduct their research under the most stringent rigors of the scientific method, repeat experiments under highly controlled conditions, determine cause and effect, and obtain repeatable results. It is often difficult for fisheries scientists to conduct science in a similar manner. For example, conducting a study to determine unequivocally the factors influencing recruitment of yellow perch in a large reservoir is impractical because it would be difficult to determine cause and effect given the constantly changing factors, such as density, mortality, growth, habitat, or weather. We might measure abundance of age-0 and age-1 yellow perch, water levels, water temperature, and spawning habitat. From these data, we might find that year-class strength of yellow perch was related to high-water levels in the early spring. Thus, we would develop a water level management plan for the reservoir to enhance year-class strength of yellow perch. However, many alternative hypotheses may exist and should be investigated. Does adult density relate to year-class strength? Is this pattern consistent among years and water bodies? Does predator density vary with water levels? Alternative hypotheses are rarely investigated in field studies often because of monetary and logistic reasons. Laboratory studies can determine cause and effect, but the applicability of these studies to the field is often questionable. It would be unrealistic to think that fisheries scientists will commonly conduct science similar to the physics or chemistry professions. Nevertheless, fisheries scientists need to be cognizant of the scientific method and attempt to use the hypothetico–deductive method and experimental studies to construct sound conclusions based on scientific inquiry.

Fisheries studies are often conducted using an inductive approach by which the scientist collects data, analyzes the data, and then develops a conclusion based on those observations. Inductive reasoning takes specific information and makes generalizations. For example, from sampling 10 black crappies with seven dorsal

spines each you would conclude that all black crappie have seven dorsal spines. Rarely are fisheries studies conducted at a large enough scale or over a sufficient time frame to develop theories from the inductive approach. A deductive approach is that by which the fisheries scientist starts with a theory, collects data, and then analyzes the data to address the theory. This approach is often more useful and allows the researcher to target the study design and data collection directly at testing the theory. Deductive reasoning sequentially proceeds from general to specific. Thus, from the previous black crappie example, we would deduce that the next black crappie sampled will have seven dorsal spines. Lee (1999) stated: "Deduction is believed to be more efficient than induction and less likely to lead scientists astray. When little is known about the subject, however, scientists may find it advantageous to use induction in order to gain an initial understanding of the phenomenon."

Some fisheries research can be categorized as reactive research. For example, a fisheries scientist notices a decline in the number of bull trout redds over time and subsequently develops a study to determine the cause of the decline. Reactive research can be considered deductive because the fisheries scientist often has a theory, or theories, regarding the observed phenomenon. For example, one theory regarding the decline in bull trout redds may be related to the presence and increased abundance of nonnative lake trout in the system. That theory could yield several testable hypotheses.

1.3.1 The Research Framework—Definition of the Problem

The first step, and likely the easiest step, in scientific inquiry is defining the problem (Figure 1.1). Problems can be classified as basic or applied. Applied research projects are often a function of political or sociological issues (e.g., evaluating the impact recreational angling has on white sturgeon in the Columbia River) and often have management implications. Rarely are fisheries studies conducted for the sake of gaining more knowledge (i.e., basic research). However, the line between basic and applied is frequently blurred in fisheries studies. That is, information about the life history of a species is often required (i.e., basic research) to answer applied problems. For example, to enhance our understanding of the impacts of recreational angling on white sturgeon in the Columbia River, basic information on reproductive physiology and behavior, as well as other aspects of the species biology, is required. A problem stimulates a question, or question set, which can be formally stated as research questions (Figure 1.1). Research questions are typically general questions that may be too broad to be addressed in one study. The foundation for research questions is typically obtained through a review of the scientific literature. However, the foundation can also be obtained through a descriptive study or a review of historical data. Some research questions are amenable to formal experimentation, and others are predisposed to descriptive study. Natural phenomena are frequently observed and described in descriptive field studies. In general, descriptive studies include broad objectives that do not lend themselves to the scientific method. However, information based on the

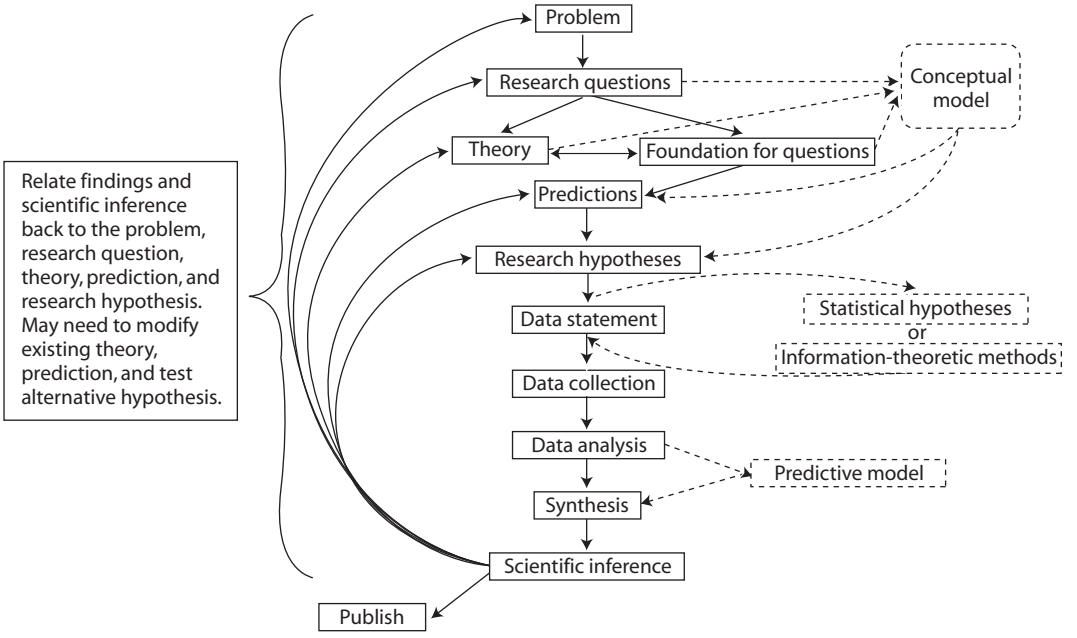


Figure 1.1 The scientific method process (modified from Ford 2000 and Garton et al. 2005).

outcome of a descriptive study may form the basis for theories that can be formally tested through experimental studies. There is a necessity to have both types of research activities in the fisheries profession.

1.3.2 Identification of the Theory

Through the process of the literature review and historical data analysis, the fisheries scientist may be able to identify a theory (Figure 1.1) that could be used to answer the question being asked (Ford 2000). A theory is a logical statement regarding the explanation of a phenomenon that directs the general objective of the research (Ford 2000). Kerlinger and Lee (2000) state that a theory is "a set of interrelated constructs (concepts), definitions, and propositions that present a systematic view of phenomena by specifying general relations among variables, with the purpose of explaining and predicting the phenomena." There are several popular theories in ecology, such as the optimal foraging theory. Stating a theory for your research is important because it allows you to think about your research in a larger context. Also, a theory for a given problem or research question does not necessarily have to subscribe to theories found in pedagogical textbooks.

1.3.3 Development of Predictions and Research Hypotheses

The next step in the scientific method involves developing predictions (Figure 1.1). Predictions are tentative propositions about the relationship among variables (Garton

et al. 2005). Stating predictions allows you to think about what may be observed and establish the foundation for research hypotheses. Research hypotheses are a rewording of the predictions in a testable format (Garton et al. 2005). Research hypotheses can be highly variable with respect to complexity, but they should be broad enough to be appealing to a wide range of fisheries scientists and specific enough to be answered comprehensively. Research hypotheses should be based on deduction, the fisheries scientist having conducted a preliminary descriptive study or analyzed historical data and conducted a literature review to develop well-establish scientific knowledge (i.e., foundation for the research question). However, fisheries scientists do not always have the luxury of preliminary data and may have to base research hypotheses on intuition. Research hypotheses differ from statistical hypotheses in that they do not specify a null hypothesis and the statistical test and assumptions associated with the test. It is important to keep these two types of hypotheses separate because you always need research hypotheses, but statistics and statistical hypotheses are not always needed in creating new knowledge. The literature review, preliminary data, and historical data can be used to develop a conceptual model to understand better the complexities among the variables of interest and place the problem in a larger context (Figures 1.1 and 1.2). The conceptual model can offer explanations and possible solutions to the problem and is useful in developing predictions (Garton et al. 2005).

1.3.4 Development of the Data Statement

A data statement (Figure 1.1) defines the assessment procedure for deciding the logical outcome of the research hypothesis (Ford 2000). The data statement (modified from Ford 2000) includes (1) the scientific procedure to investigate the research hypothesis, (2) measurements to be made for each component of the research hypothesis, and (3) the statistical hypotheses and specific requirements for any statistical tests. Statistical null hypotheses are fundamentally different from research hypotheses because the null hypothesis (H_0) frames the research hypothesis in terms of representing the case of no difference between population parameters. Conversely, the alternative hypothesis (H_a) represents a unidirectional or bidirectional difference. Statistical hypotheses are not always necessary and do not determine whether the fisheries scientist followed the scientific method. Methods are available to assess research hypotheses without using frequentist statistical approaches based on null hypothesis testing (see section 1.4).

1.3.5 Data Interpretation and Synthesis

After data collection (see *Fisheries Techniques* [Murphy and Willis 1996] for methods on sampling aquatic biota) and data analysis (see Chapters 2–18 in this book), the difficult task of synthesis, making scientific inference, and developing alternative hypotheses begins. Making the connection from the synthesis of the data and conclusion back to the problem, research question, theory, prediction, and research hypothesis is important (Figure 1.1). Remember, the credibility of a research

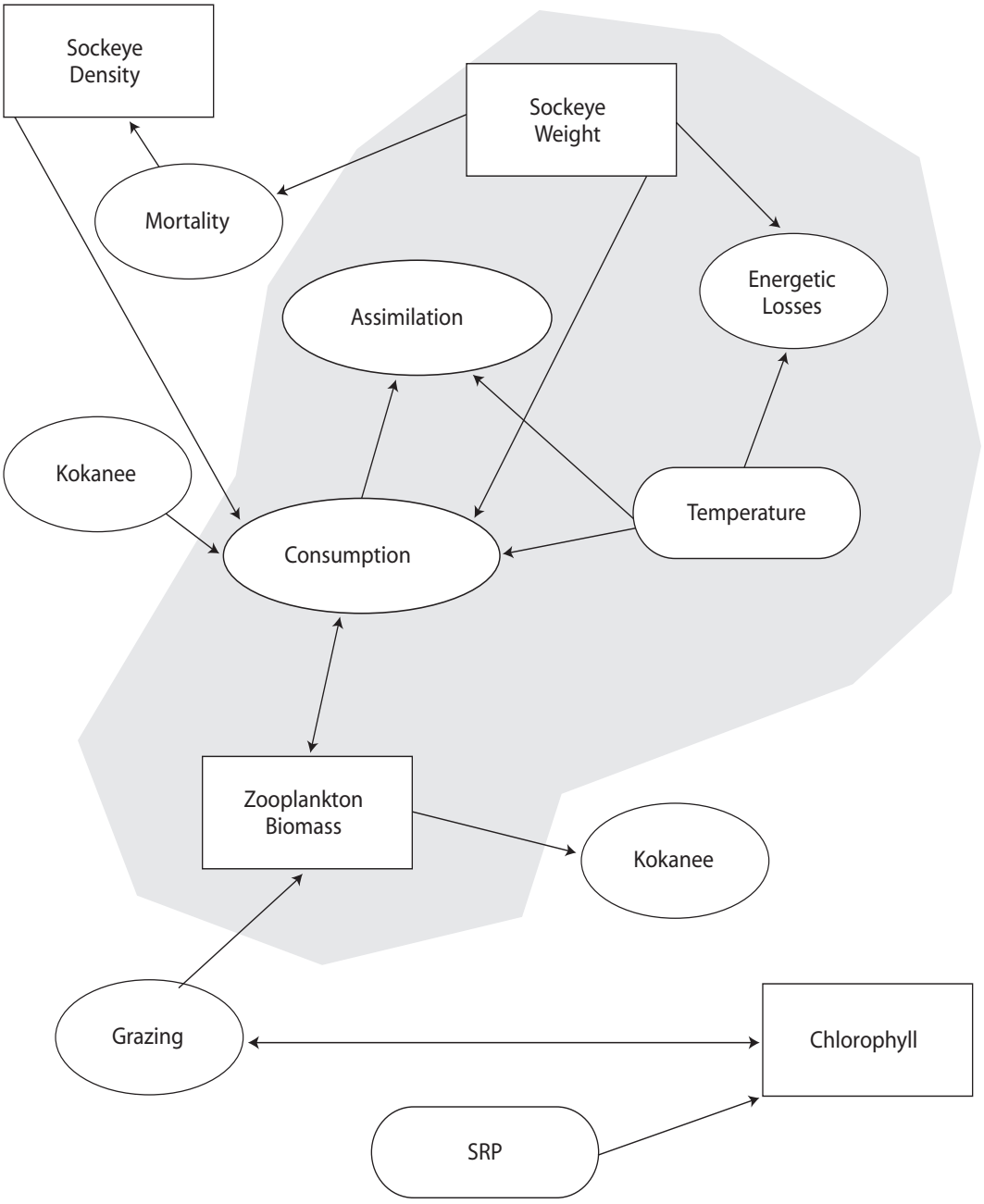


Figure 1.2 Example of a conceptual model in a paper published by Luecke et al. (1996). The conceptual model is used to predict sockeye salmon growth and production using bioenergetics and lake carrying capacity. Bioenergetics components are contained within the shaded area. Components outside the shaded area are carrying capacity inputs (e.g., soluble reactive phosphorus, SRP). State variables (i.e., points of accumulation) are represented as rectangles and functions are represented as ovals.

hypothesis is increased by more of its predictions being supported and those of alternative hypotheses being rejected. The feedback to the original theory allows the fisheries scientist to develop alternative hypotheses and modify the study. Modifying the hypotheses and study design may allow for increasing the scope of inference because the study design and sampling method highly influence the ability to make inferences (Figure 1.3). Fisheries scientists must determine what they expect in terms of inference space and certainty of conclusions prior to developing a study. There are many options for study designs, and no single method is perfect. For example, natural experiments such as floods, disease outbreaks, and hurricanes provide a large inference space, but the certainty of conclusion remains limited because they lack replication (Figure 1.3). Conversely, a laboratory study often provides clear and certain results but may have little applicability to conditions in the field (Figure 1.3). A combined approach in which laboratory studies and field studies are integrated provides large inference space and more certain conclusions, but these types of studies are usually costly and take many years to complete.

Fisheries scientists must be careful not to make conclusions beyond the scope of the data. Data interpretation must stay within the scale of the study. For example, if we found that recruitment of black crappies in South Dakota natural lakes was related to spring water levels and wave action, we cannot imply that recruitment of black crappies is influenced by these factors in all natural lakes containing black crappie. Replication of studies across time, space, and life histories

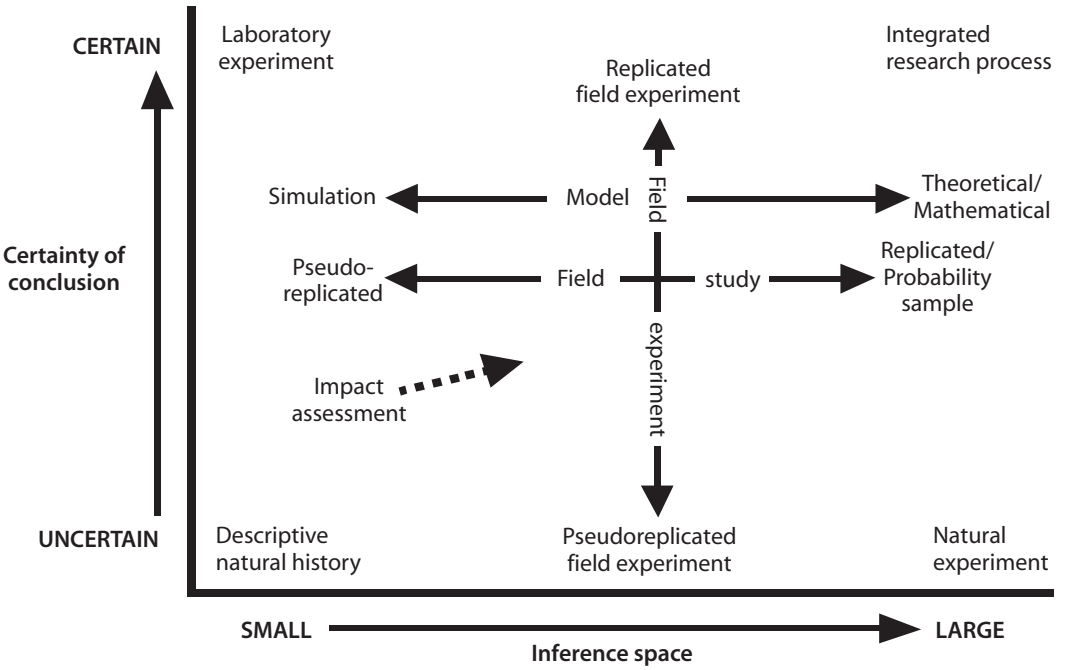


Figure 1.3 The relationship between inference space and certainty of conclusion for several study types (from Garton et al. 2005; reproduced with permission from The Wildlife Society).

provides great confidence in our findings and increases the inference space. Research often generates more questions than answers, and rarely does a single research project provide an unequivocal answer. Designing simple testable hypotheses and conducting multiple studies at this scale is useful because we cannot describe every component of a system in one study. The ability to speculate with respect to findings is a critical aspect of science (Garton et al. 2005). However, speculation must be identified as such and not be confused with conclusions based on the data (Garton et al. 2005). Speculation often takes the form in alternative hypotheses that should be tested: the foundation for the hypothetico–deductive method. For example, if we found sonic-tagged adult lake trout occupying depths below the thermocline in the summer months we cannot conclude that adult lake trout are using the area as thermal refugia without additional data. There could be several alternative hypotheses such as lake trout use the area below the thermocline because it is occupied by their prey (lake whitefish) or it has the highest concentration of dissolved oxygen. Testing these hypotheses we might find that lake trout are feeding on lake whitefish, and lake whitefish are occupying the area directly below the thermocline because that is where zooplankton density (prey for lake whitefish) is the highest. We may then hypothesize that the phytoplankton must be highest directly below the thermocline and we are observing a "bottom-up" response in habitat use by lake trout.

Developing theories, predictions, and research hypotheses is complex and requires critical and imaginative thinking skills. Thinking critically means that you carefully monitor your ideas to determine if they make sense (Moore and Parker 2007). Critical thinkers do not accept the beliefs of others without carefully analyzing the subject themselves, and they strive to incorporate all relevant knowledge into their thoughts (Paul and Binker 1990). The human mind is comfortable with biases, falsehoods, and half-truths, and it takes a special effort to evaluate our own creations critically (Paul 1990). Thus, critical thinking is acquired and must be developed throughout the career of the fisheries scientist. Paul and Binker (1990) developed 35 strategies of critical thinking, and Lee (1999) restated them with application to science. One of our favorites is under the heading "Developing intellectual courage." Lee (1999) states: "Scientists should deal with ideas directly and honestly and that includes confronting unreasonable aspects of popular ideas and acknowledging reasonable parts of unpopular ideas. Nonconformity, when justified, is a sign of courageous critical thought, even though the consequences of nonconformity sometimes are severe." Under the heading "Developing intellectual perseverance," Lee states: "Critical thought is not easy. Before making a judgment on an idea, the scientist should take the time and make the effort to be sure that the idea has been analyzed fully and carefully." For more information on the 35 strategies of critical thinking see Paul and Binker (1990) and Lee (1999).

1.4 STATISTICAL APPROACHES IN SCIENTIFIC STUDIES

The dominant statistical paradigms that have guided data analysis in fisheries science are Neyman–Wald frequentist and Fisherian. Other methods such as Bayesian

analysis and model selection have been used more recently by researchers. For example, 38 and 66 articles appearing from 1996 to 2006 in *Transactions of the American Fisheries Society* contain results from model selection and Bayesian methods, respectively. Yet, a myriad of results from null hypothesis testing have been reported since the early 1900s in that journal. Thus, the fisheries publication record indicates that frequentist (e.g., analysis of variance [ANOVA] and *t*-tests; Sheskin 2000) and Fisherian (e.g., general linear models; McCullagh and Nelder 1989) methods have been the statistical mainstay, but nonfrequentist approaches have gained in popularity. Development of computer-intensive methods such as resampling (Manly 1997) and likelihood analysis have facilitated use of nonfrequentist inference. As a testament to the impact that computing has had on statistical science, Efron (1998) commented that "a year's combined computational effort by all statisticians of 1925 wouldn't equal a minute of modern computer time."

The frequentist approach uses tests of significance that suppose that the null hypothesis is true. That is, the statistical test determines the probability of the data given the null hypothesis. Conversely, Bayesian statistics determine the probability of a model given the data. This inference process entails fitting a probability model to data. Hilborn and Mangel (1997) popularized the application of Bayesian inference for ecological data. They demonstrated how likelihood and Bayesian methods provided more meaningful conclusions for field studies (ecological reality) than did null hypothesis testing. Similarly, Burnham and Anderson (2002) used likelihood methods and model selection to focus on which model(s) best fit (i.e., weighted by evidence) the data. The latter approach deemphasizes the use of probability values, as used in null hypothesis testing, and emphasizes model utility.

The generalization, or tendency, to be made here is that frequentist methods may be more often applicable to experimental studies where true replication of treatments is possible; nonfrequentist approaches may be more often applicable to observational studies. However, this is likely an oversimplified description of these paradigms because frequently null hypothesis testing of observational data are appropriate, particularly when assessing univariate causality (Stephens et al. 2005). Chapters 4 through 18 of this book rely heavily on Fisherian and frequentist methods, the statistical mainstream of fisheries data analysis. In the following sections we provide a general discussion of null hypothesis testing, Bayesian inference, and model selection.

1.4.1 Null Hypothesis Testing

1.4.1.1 Statistical Hypothesis Formulation and Testing

The classical decision approach of statistical inference involves hypothesis testing. Simply put, the outcomes of experiments generally are not clear-cut, and a decision has to be made between competing hypotheses. Hypothesis testing provides an objective, uniform framework for making decisions rather than an individual, subjective approach of decision making based on simply looking at the

data. The research questions and theory crafted during study development provide research hypotheses (Figure 1.1). To evaluate a research hypothesis statistically, the hypothesis is restated in the form of two or more (competing) statistical hypotheses, the null (H_0) and one or more alternatives (H_a). The ultimate goal of hypothesis testing is to determine the probability of the null statement being true given the data.

The statistical null hypothesis is simply a statistical statement alternative to what the researcher believes, and from the data analysis the researcher fully expects to contradict the statistical null. If the investigator finds that the data contradicts the prediction of the null hypothesis, then either the null hypothesis is actually false or a low probability event has occurred. The development and testing of null hypotheses with statistical methods is consistent with the scientific method (falsification) as described by Popper (1959, 1968).

Formal statistical hypothesis testing involves a stepwise approach. The first step in statistical hypothesis testing is to specify the H_0 and the H_a. Generally, the H_0 is a statement that the population parameter has a specified value or that parameters (e.g., the mean, μ) from two or more populations are similar, such as $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$. The alternative is that at least two samples were derived from populations with different means (H_a: $\mu_i \neq \mu_j$). The next step is to establish the a priori significance level (α; see fourth step). The most commonly selected significance levels used are 0.05 and 0.01. More recently, scientists conducting field studies are using 0.10 as a significance level. The third step is to select a statistical test and calculate a statistic analogous to the parameter specified by H_0. If H_0 were defined by the parameters as $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$, then the statistic for $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \ldots = \bar{y}_k$ would be computed (e.g., *F*-statistic of an ANOVA). The fourth step is to calculate the probability value (often called the *P*-value), which is the probability of obtaining a statistic as or more different ($\leq$) from the parameter specified in H_0 as the statistic computed from the data. (These calculations are made under the assumption that H_0 is true.) Thus, the decision to reject or fail to reject H_0 is based on that test statistic and whether the calculated value falls in the rejection (or critical) region of the statistic distribution, as defined by the a priori significance level.

If H_0 is rejected, the outcome is said to be "statistically significant"; if one fails to reject H_0 the outcome is said to be "not statistically significant." If the outcome is statistically significant, then H_0 is rejected in favor of the alternative hypothesis, H_a. We never reject H_a or accept H_0; if we do not reject H_0 it suggests that there is not enough evidence to support H_a over H_0. Conversely, if H_0 is rejected then it suggests H_a may be true. Thus, the conclusion is based on the concept of proof by contradiction of H_0. If we fail to reject H_0, what are the odds that we are wrong? Statistical tests allow us to calculate these odds, which are expressed as probability values. The lower the probability value, the more likely it is that H_0 is not true.

1.4.1.2 Statistical Errors in Hypothesis Testing

There are two common types of statistical errors, known as type I and type II errors (see Table 1.2), that may be committed during hypothesis testing. The

Table 1.2 The four possible outcomes for null hypothesis (H_0) testing.

	True state	
Decision	H_0 is true	H_0 is false
Reject H_0	Type I error (α)	Correct ($1 - \beta$)
Fail to reject H_0	Correct ($1 - \alpha$)	Type II error (β)

probabilities associated with these errors (α = the probability of type I error; β = the probability of type II error) provide a measure of the goodness of the statistical test. If H_0 is actually true and we fail to reject H_0, or if H_a is actually true and we reject H_0, the correct decision has been made. However, if H_0 is true and we reject H_0 or if H_a is true and we fail to reject H_0, then we have committed a statistical error. When a decision has to be made, it would be convenient if we could always arrive at a correct conclusion. Unfortunately, this is statistically impossible, because these decisions are based on probability, sample size, and variation. As shown in Table 1.2, there are four common outcomes of a statistical test. Consequently, we control the probability, or risk, with which an error occurs. To control these errors we assign a small type I error probability to them. As previously mentioned, the most frequently used significance levels are an α 0.01 or 0.05; the former is more conservative because stronger evidence is required to reject H_0 at the 0.01 than at the 0.05 α-level. The probability assigned to each error will depend on the seriousness of the error. It should be noted that type I and type II errors are inversely related; that is, α increases when β decreases, and vice versa. An α-value of 0.05 is the most commonly used α-level; however, the α-level should be established with respect to the willingness to accept a type I error. Although hypothesis testing is considered objective, the selection of α is subjective.

1.4.1.3 Power Analysis

Power analysis provides a mathematical means to determine the probability of obtaining a statistically significant result given a true effect actually occurs in a population. Thus, power may be broadly defined as the ability of a statistical test to detect an effect, given that the effect actually exists. The power of a statistical test is technically defined as 1 minus the probability of a type II error, or $1 - \beta$. In that vein, power analysis tells us how likely we are to find a significant difference given that H_a is true. If the power is too low, then we have little chance of detecting a significant difference (i.e., an analysis would yield nonsignificant statistical results), even though there may be real differences. It would be desirable to use statistical tests that minimize α and β; however, this would require a compromise because making α small involves rejecting the H_0 less often, whereas making β small involves failing to reject H_0 less often. These, of course, are contradictory actions. As an alternative we fix α at a specified significance level and then apply the statistical test that maximizes the power.

There are several important pieces of information required to conduct a statistical power analysis. First, the significance level (α = probability of a type I error) must be established. Again the conventional and common choices are 0.01 and 0.05. Second, the power ($1 - \beta$) required to detect an effect is established. The probability of $1 - \beta = 0.80$ is a common choice. Third, the effect size (e.g., the biological change or significance) that needs to be detected is determined; effect size is based on the actual units of the response. Effect size and the ability to detect it are indirectly related; therefore, as effect size becomes smaller, its detection becomes more difficult. Fourth, the extent of variation (i.e., standard deviation, SD) associated with the response variable in the population(s) is determined. Typically, the SD used in a power analysis can be determined from a similar study previously conducted or a pilot study.

The components of a power analysis are interdependent; any combination of the four components dictates the outcome of the fifth component (e.g., sample size). The usual objectives of a power analysis are to calculate the sample size based on the level of significance (α), power, effect size, and variability (SD). A larger sample size generally leads to parameter estimates with smaller variances. Small variance then provides a greater ability to detect a significant difference. For studies for which the maximum sample size obtainable is known (often related to the budget), power analysis is a useful tool to determine if sufficient power exists for specified values of α, effect size, and SD. At this point the investigator must decide whether the study should be conducted based on the power to detect an effect.

If inadequate sample planning takes place, a lack of statistical power can present a problem in the final evaluation of management or research results. All study plans should include some initial calculations of the power of statistical tests that will be obtained with the sample sizes that are planned (Hoenig and Heisey 2001). Although power analysis can be done by hand, these calculations are extremely tedious, especially for complex designs. Fortunately, there are numerous standalone and statistical software programs that provide power computations for a variety of statistical tests. For a listing and review of software programs see Thomas and Krebs (1997).

1.4.1.4 Parametric and Nonparametric Statistical Tests

Parametric statistical tests (e.g., *z*-test, *t*-test, and *F*-test) assume that sampled data are from populations that follow a certain distribution (i.e., normal), have interval or ratio scales (parametric tests are inappropriate for nominal or ordinal data), and have similar variability (if multiple samples are compared). However, measured biological data do not always follow a normal distribution precisely, often because of low sample size. Fortunately, many kinds of quantitative data follow a bell-shaped distribution that is approximately normal. Because parametric statistical tests work well even if the distribution is only approximately normal, parametric statistical tests are commonly applied.

An alternative approach, nonparametric testing (Table 1.3), does not assume that (interval or ratio) data follow a normal distribution. In nonparametric testing,

Table 1.3 Typical nonparametric tests and proximate parametric counterparts (derived from Conover 1980 and Sheskin 2000).

Nonparametric	Parametric or alternative test
Mann–Whitney *U*-test	Two-sample *t*-test; median test
Kruskal–Wallis test	One-way *F*-test; median test
Squared rank's test (equal variance)	Two-sample *F*-test
Spearman's rho	Regression
Kendall's tau	Regression
Wilcoxon matched-pairs signed-ranks test	Paired *t*-test
Mood test	*F*-test
Moses test	*F*-test
Friedman test (two sample)	Two-sample *t*-test
Friedman test (*k* samples)	*F*-test

observations are categorized (nominal) or values are ranked (ordinal) from low to high, and the analyses are based on the distribution of ranks. Consequently, nonparametric tests make fewer assumptions about the distribution of the data and are often called assumption- or distribution-free tests, which is not entirely true (Marascuilo and McSweeney 1977). The primary drawback with nonparametric tests is that they are sometimes less robust than are parametric tests. In practice this means that probability values (associated with errors) tend to be higher, making it harder to detect real differences as being significant. If sample size is large the difference in relative efficiency is minor. Nonparametric tests have reduced power to detect differences with small sample sizes, and that artifact must be considered in the interpretation.

Certain situations cause investigators to defer to a nonparametric test automatically. If the variable is a rank or score, or if the sample contains fewer than a dozen or so observations, a nonparametric test should be used. If the sample distribution is clearly nonnormal there may be two options. First, if sample size is sufficient, a transformation (section 1.4.1.5) may cause the data to approximate a normal distribution; however, if that effort fails, then an appropriate nonparametric test may be used. Second, in some situations the sample distribution(s) is normal but the data contains extreme outliers; therefore, it may be inappropriate to analyze these data with a parametric test. In such cases a nonparametric test provides a robust approach for analysis of data because the ranks of the values are used.

Frequently, it is difficult to decide whether to select a nonparametric test. The normality assumption regards the underlying distribution of the population. Examine the scatter of data from previous experiments that measured the same variable. Also consider the source of the scatter. When variability is contributed by numerous independent sources, the underlying distribution may be assumed to be normal. Although testing to determine whether data were sampled from a normal distribution is helpful, normality testing (i.e., residual analysis) is often less useful than we would like. It is difficult to determine whether the data came

from a normal distribution or not when examining the distribution of a small sample. Furthermore, a normality test may not simplify the decision. The tests simply have little power to discriminate (detect deviations) between normal and nonnormal distributions with small sample sizes. Thus, the interpretation of a normality test should hinge on the probability value and the sample size. Ultimately, the decision to choose a parametric or nonparametric test for interval or ratio data is most critical when sample size is low.

1.4.1.5 Data Transformation

Data may be transformed to provide a distribution that is conducive to a particular statistical analysis. Common transformations are logarithmic (base 10 or natural), power, square-root, and arcsine transformations (Table 1.4). A transformation is simply the process of converting or changing the numerical scale. For example, if a variable is not normally distributed, transforming the values may produce a normal distribution. If the distribution of the population is known, transforming the values to approximate a normal distribution may be appropriate, as it allows the use of parametric statistical tests. It is important to point out that inferences are based on the scale at which data are analyzed and care should be taken when making statements about the original scale if data have been transformed.

Many statistical software programs provide tests for deviations from common distributions. For example, the normality test is used to determine the closeness of a data distribution to the normal distribution. Software programs will test for normality using the Shapiro–Wilk (SW) test or the Kolmogorov–Smirnov (KS) test. The SW statistic (represented in some programs as W) and KS statistic (represented in some programs as D) quantifies the difference between the data distribution and an ideal normal distribution; a larger value denotes a larger discrepancy. The statistics are not informative themselves but are used by the software to

Table 1.4 Common data types and distribution characteristics with associated transformations to normalize data distributions.

Data and distribution	Normalizing transformation
Count data	Square root of x or square root of $(x + c)$ if there are 0 values
Positive (right) skew	
Poisson distribution	
Percentages or proportion data	Arcsine of square root of x
Platykurtotic	
Binomial distribution	
Measurement data	$\text{Log}_{10}(x)$ or $\log_{10}(x + c)$ if there are 0 values
Positive (right) skew	
Lognormal distribution	
Time or duration data	Reciprocal: $1/x$
Positive (right) skew	
Other distributions	
Negative (left) skew	x^c where $c = 2, 3$; or use e^x
Positive (right) skew	x^c where $c = -0.5, -1$, or greater negative value

compute a probability value that may be used to interpret whether the data distribution follows a normal distribution.

The traditional KS method cannot be used to calculate the probability of normality unless the true mean and SD of the population are known. When analyzing data the population mean and SD are rarely known. Thus, parameter estimates derived from the sample are used in the testing of normality by an approximation method. Software programs often use an approximation method such as the Lilliefors' test (Dallal and Wilkinson 1986). Approximation methods are most accurate with small *P*-values; some software programs may simply report "$P > 0.10$" for large probabilities.

1.4.2 Bayesian Inference

Bayesian methods provide an alternative to hypothesis testing but are not yet commonly applied by fisheries scientists. However, Bayesian analysis of certain areas (e.g., age- and size-based stock assessment) of fish population dynamics is becoming more popular (Box 1.1). Bayesian analysis is also likely to gain popularity because the process easily allows the integration of new information, a necessary feature in planning activities in adaptive management. Bayesian inference draws heavily from Bayes' theorem (more below), hence its name, and provides methods to account for uncertainty in model selection. Based on conditional probabilities, the objective of the Bayesian approach is to incorporate prior knowledge in combination with new data or information to make statistical inferences. Existing or prior information could be results derived from previous studies or comparable experiments. Bayesian analysis can also be useful when there is a lack of prior information for a given problem but there exists a strong understanding of the mechanisms that may affect the problem. Using prior information about the mechanisms (parameters), a posterior distribution for the mechanisms is determined and inferences about the model parameters can be interpreted (see Gelman et al. 1995 for thorough coverage of Bayesian statistics or Press 1989 and Sivia 1996 for overviews).

Bayesian inference is based upon Bayes' theorem, a result of probability theory, which allows different event probabilities to be related. That is, for two events the probability of event 1 conditional on event 2 will differ from the probability of event 2 conditional on event 1. The relationship that exists between these probabilities characterizes Bayes' theorem. Bayesian probability differs from frequentist probability; in the pure Bayes' form, probabilities for unknown information also can be assigned.

Bayes' theorem is founded on conditional probabilities of stochastic events. The basic model derived from this theorem (Gelman et al. 1995), which relates conditional and marginal probabilities for events *A* and *B,* is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{L(A|B) \cdot P(A)}{P(B)}, \qquad (1.1)$$

Box 1.1 A Bayesian Application to Fisheries Management

A whole-lake fertilization project was conducted on Chilko Lake, British Columbia, during the late 1980s and early 1990s for the purpose increasing abundance of sockeye salmon for commercial harvest (Bradford et al. 2000). It was concluded that fertilization had a positive effect on recruitment (number of recruits per spawner) but that estimates were highly imprecise due to the short duration of the fertilization program and high natural variability in recruitment before and during fertilization (Bradford et al. 2000; Maxwell et al. 2006). Thus, there was considerable uncertainty about the success of the fertilization experiment and the resulting economic benefit to the fishery (i.e., cost of fertilization relative to the additional number of fish available for harvest). Consequently, Maxwell et al. (2006) used a Bayesian approach to "describe uncertainties in the stock–recruitment relationship for the periods prior to and during lake fertilization and propagating those uncertainties through to the economic calculations."

Several competing models were used to estimate the effect of fertilization. Four candidate models were developed, each using the Ricker stock–recruit model (Chapter 4) as a core, to reflect different hypotheses. In addition to stock–recruit parameters, the candidate model set included a density-independent model, a density-dependent model, and two other models that contained the Fraser index (FI) as an additional parameter to expand the density-independent and density-dependent models. The FI parameter, based on the dynamics of other sockeye salmon populations, accounted for annual variability not due to fertilization. The best-supported models were the density-independent model (DIFI, $\Delta_i = 0.00$) followed by the density-dependent model (DDFI, $\Delta_i = 0.92$), both containing the FI parameter (see table below). The DIFI and DDFI models estimated 5.4 and 4.4 million recruits, respectively.

Table Least-squares best-fit parameter estimates (*a* – *d* and *g*) for four models considered in an analysis of the effect of Chilko Lake fertilization on sockeye salmon abundance. Not all parameters are applicable to all models (NA). Models were density-independent (DI), density independent + Fraser index (DIFI), density-dependent (DD), and density-dependent + Fraser index (DDFI). The Δ_i values are the relative differences between the small-sample-corrected Akaike's Information Criterion (AIC_c) of a given model and the DIFI model AIC_c. (Analysis and interpretation adapted from Maxwell et al. 2006.)

	Model			
Parameter	DI	DIFI	DD	DDFI
a	2.71	2.58	2.55	2.50
b	3.3×10^{-6}	2.4×10^{-6}	2.1×10^{-6}	1.7×10^{-6}
c	0.63	0.57	1.65	1.19
d	NA	NA	3.7×10^{-6}	2.3×10^{-6}
g	NA	0.63	NA	0.55
Δ_i	7.48	0.00	5.67	0.92

The range of prior distributions (uniform) was first based on the best fit of each model parameter (±1 SE) derived from regression analysis. Then a single, combined prior distribution was developed for each parameter to encompass the entire range of possibilities for a parameter across the four models. To estimate uncertainty (i.e., acknowledging the range of possible values for each parameter in each model), posterior probabilities were calculated such that each set of parameter values described the stock–recruit relation given the observed data. Posterior probability estimates were then used to determine the number of additional recruits attributable to fertilization and also the benefit–cost ratio.

The authors found that fertilization provided biological and economic benefit. Bayesian analysis indicated an increase of 0.5 million additional recruits above levels estimated from best-fit DIFI and DDFI models. Posterior probabilities of at least 81% supported that there was an increase in sockeye salmon abundance due to fertilization. Similarly, posterior probabilities of at least 84% supported that the benefit–cost ratio exceeded 1.

where the likelihood (L) of event A given event B for some fixed value of B is

$$L(A|B) = P(B|A). \tag{1.2}$$

The terms in the Bayesian model are defined as the conditional, or posterior, probabilities ($P[A|B]$, the conditional probability of A given a specified B value, or $P[B|A]$, the conditional probability of B given a specified A value) and the prior probabilities ($P[A]$, the prior probability of A that does not contain any information about B, or $P[B]$, the prior probability of B that does not contain any information about A).

In other words, the prior probability is the probability of the model being true before any data are observed. The posterior probability is the probability that a model is true following the incorporation of observed data or information. The prior probability, $P(B)$, functions as a normalizing constant so that the posterior probability is proportional to the likelihood (L) times the prior probability. The likelihood describes the conditional probability of the data given the model. Although we presented the simple model above, conditional probabilities and Bayes' theorem can be applied to multivariate data and multiple hypotheses (Gelman et al. 1995).

There are several model alternatives and versions that subscribe to Bayes' theorem, but perhaps the model most beneficial to fisheries scientists is the empirical form that allows the evaluation of a set of hypotheses (Gelman et al. 1995; Haddon 2001). For example, if we interpret A to be the observed data and B as our set of different hypotheses, then the model becomes

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i), \tag{1.3}$$

where $P(A)$ is the combined probability for the data and all hypotheses under consideration. If the parameters considered by hypotheses are discrete, then the model is stated as

$$P(H_i|\text{data}) = \frac{L(\text{data}|H_i) \cdot P(H_i)}{\sum_{i=1}^{n} [L(\text{data}|H_i) \cdot P(H_i)]}. \tag{1.4}$$

If the parameters are continuous, the model becomes

$$P(H_i|\text{data}) = \frac{L(\text{data}|H_i) \cdot P(H_i)}{\int L(\text{data}|H_i) \cdot P(H_i) \, \mathrm{d}H_i}. \tag{1.5}$$

The individual hypotheses (H_i) associated with these models would be individual models, each with a unique set of parameter values. The data are the posterior observations being considered given the set of hypotheses.

Gelman et al. (1995) generally define the process of Bayesian analysis in three steps. Step one consists of developing the prior distributions. Priors are the distributions of parameters (or hypotheses) derived from probability models. The prior,

or marginal, distribution is so named because it is not conditioned on previous aspects of the process. There are two types of priors, informative and noninformative, and either can play an influential role in Bayesian analysis. Informative priors make use of the best available information (e.g., previous data) to estimate model parameters (prior probability distributions). If no data are available then noninformative priors are usually applied for which equal probabilities are assigned for competing hypotheses. Noninformative priors are distributions having no basis (e.g., no prior information or vague) and have less influence on the posterior distribution; the uniform distribution is commonly used as this prior. If no data or strong opinion preexists for an informative prior, an investigator would choose a noninformative prior.

Step two consists of defining the posterior distribution. Following data collection, the observed data are used to condition the model; that is, the prior distribution is now combined with sample information to provide an updated estimate. This Bayesian estimate is functionally a weighted average estimate, based on the prior and posterior probabilities. Haddon (2001) stated that there are three elements that are required to produce the posterior distribution when comparing hypotheses. These are (1) the individual hypotheses to be considered; (2) the likelihood required to determine the probability of the observed data given each hypothesis (H_i); and (3) the prior probability for each hypothesis.

At this point, the full model is the joint probability distribution that contains all observable and unobservable quantities or information associated with the specific problem or question.

In step three, the model fit is evaluated in conjunction with an interpretation of the reasonability of the posterior distribution. In essence, does the model fit the data and provide a logical conclusion.

The development of priors is likely the most problematic aspect of Bayesian analysis. Berger (1994) identified several characteristics to consider in choosing a class of priors. Priors should be easy to derive and interpret, computationally simple, large enough to reflect prior uncertainty, and extendable to higher orders or dimensions. Kass and Wasserman (1996) describe formal procedures for selecting noninformative priors. Also, there are parametric and nonparametric classes of priors, a topic too lengthy to expand on here, but there is considerable literature that specifically discusses these prior classes (e.g., Walley 1991; Wasserman 1992; Dey et al. 1998; Geweke 1998). Lastly, careful consideration should be given to the presentation of Bayesian analyses, particularly choices made regarding priors. Summary results should be reported but accompanied by solid explanations. Results from model checking (e.g., posterior predictive results) should be reported as well (Rubin 1984; Gelman et al. 1995).

1.4.3 Model Selection

Model selection provides another inference approach that is based on information theory (Kullback and Leibler 1951; Burnham and Anderson 2002). This approach is particularly appealing to researchers conducting field studies for which

experimental manipulations (treatment and replication logistics) are cost prohibitive. Rooted in a philosophy similar to Bayesian analysis, model selection focuses on the existence of a knowledge base from which a suite of realistic competing models can be derived prior to data collection and analysis. In general, the candidate model best supported by the data is interpreted to be the best model. Further, the best model should be objective and repeatable (Burnham and Anderson 2002). In comparison, Bayesian analysis uses the prior distribution, model(s), and observed data to make inferences about a posterior distribution, whereas information theory compares performance of a priori selected models in how well they describe the observed data.

Conceptually, the mechanics of the selection approach are fairly straightforward. Typically, a global model containing variables thought to be biologically relevant to the question is developed and combined with reductions of that model to compose the model set. The goal in variable selection is to develop the simplest (parsimonious) model that encompasses cause and effect relations. Too few variables (underfitting) and a selected model may be very precise but will contain high bias. Too many variables (overfitting) will result in low precision but a model with low bias. Thus, a balance needs to be struck somewhere in the range of model parameters. Although selection of parameters generally should be similar among investigators for a particular question, different parameters would yield different results. A likelihood criterion is used to compare among the competing models.

The basis for the evolution of model selection procedures is Kullback–Leibler (K–L) information, or distance (Kullback and Leibler 1951). Conceptually, if the full truth (reality) is known, then the distance from the full truth could be determined for a model set being used to approximate the full truth. The model that deviates least (smallest K–L distance) from full truth has the least information loss. Obviously, full truth is not likely to be known in fisheries studies.

Model selection gained substantial utility when Akaike (1973) introduced a model comparison procedure. Akaike (1973, 1974) demonstrated that relative K–L distance could be estimated by the asymptotic result of the empirical log-likelihood function. The final, applicable expression of Akaike's original finding is called Akaike's Information Criterion (AIC):

$$\mathrm{AIC} = -2\log_e(L) + 2K, \tag{1.6}$$

where K is the number of estimable parameters (for bias correction) and L is the maximum likelihood function. If normally distributed errors are assumed, then AIC is calculated as

$$\mathrm{AIC} = n \cdot \log_e\left(\frac{\mathrm{RSS}}{n}\right) + 2K, \tag{1.7}$$

where n is sample size and RSS is the residual sum of squares.

The AIC provides a measure of the best model by quantifying the goodness or lack of fit of a set of models, given the observed data. The preferred model has

the lowest criterion value (minimum information loss). The AIC is sensitive to sample size; models containing numerous parameters may be found to have good fit but are overfitted and suffer from low precision. Because AIC is inadequate if the number of parameters in relation to sample size is too high, a second-order criterion (AIC_c) is recommended as a small sample bias adjustment (Hurvich and Tsai 1989). The AIC_c correction is calculated as

$$AIC_c = \frac{AIC + 2K(K+1)}{(n-K-1)}. \tag{1.8}$$

This criterion functionally penalizes the likelihood of the model given the number of parameters required to estimate from the observed data. As a general rule, Burnham and Anderson (2002, 2004) suggest the use of AIC_c over AIC when the ratio of n:K is less than 40.

Another commonly used criterion in ecological studies is quasi-AIC (QAIC), which is an adjusted AIC applicable to overdispersed (i.e., sample variance exceeds theoretical variance) binomial data (Lebreton et al. 1992). Generally, QAIC should be considered for count data when the variance inflation factor (c) for the global model exceeds 1. The underlying reason for accounting for variance inflation is that parameter estimates may be unbiased under overdispersive conditions, but the model-based variances likely would overestimate precision (Burnham and Anderson 2002). The QAIC is similar to AIC but includes the estimate of the variance inflation factor ($\hat{c}$, an overdispersion parameter estimate), as

$$QAIC = -[2\log_e(L)/\hat{c}] + 2k. \tag{1.9}$$

The variance inflation factor ($\hat{c}$) can be estimated as $c = \chi^2/df$ of the global model (Cox and Snell 1989). Note that if $\hat{c} = 1$, then QAIC reduces to AIC. The small sample size adjustment ($QAIC_c$) for QAIC is similar to AIC_c and is computed as

$$QAIC_c = QAIC + 2K(2K+1)/(n-K-1). \tag{1.10}$$

Other, less commonly used criteria include Takeuchi's Information Criterion (TIC) and Bayesian Information Criterion (BIC) (Burnham and Anderson 2002).

The basic objective in model selection is to determine which model is most appropriate. Because individual criterion values are not interpretable, differences between information criterion values of candidate models are used to rank the models. For AIC (used here as an example) these differences are determined as

$$\Delta_i = AIC_i - AIC_{minimum}. \tag{1.11}$$

These Δ_i values provide an interpretation of how well each model explains variation in the observed data. The model having the smallest Δ_i is determined to be the best-fit model among the candidate models. Generally, models with Δ_i less than 2 have good support, whereas model Δ_i values exceeding 10 have little or no support.

Further evidence of model support is obtained by calculating Akaike's weights (Burnham and Anderson 2002). Akaike's weights (w_i) are used to determine the likelihood of each model (model probability) in the set (R) and are calculated as

$$w_i = \frac{\exp^{(-\Delta i/2)}}{\sum_{j=1}^{R} \exp^{(-\Delta j/2)}}, \tag{1.12}$$

where exp is the base on the natural logarithm.

These normalized likelihoods conveniently sum to 1.0 ($\Sigma w_i = 1$), providing a further means of comparing model strength.

Garton et al. (2005) point out two weaknesses associated with information theory and AIC. First, information theoretic approaches currently do not assist with critical issues during study design, such as sample size formulation. However, an investigator could use resampling techniques to estimate sample sizes providing that proximate data distributions were known. Second, although AIC results allow objective selection of the best model the approach does not actually specify model performance. Investigators are referred to Burnham and Anderson (2002) who provided a detailed overview of information theory development, criterion development and comparisons, and application mechanics. See Box 1.2 for an example application of the model selection approach.

1.5 PUBLISHING

Publishing research results is an important step in the scientific method (Figure 1.1). We are hesitant to say that publishing is the final step in the scientific method because it is important to publish research findings while testing alternative hypotheses. Fisheries scientists must publish when they believe they have sufficient evidence to address a research hypothesis because no single research project is the last word on any issue. Thus, nothing would be published if all scientists waited for the final word.

Enormous amounts of research dollars are wasted because the knowledge gained from a research project is never published (Garton et al. 2005). Research conducted and stored in file cabinets or computer files is useless to the profession and causes redundancy in research efforts. Publishing research results is difficult and time consuming. Further, the publishing process can be humbling, especially after the peer-review process. Nevertheless, the peer-review process is vital to maintaining our profession's standards of publishing reliable knowledge. Fisheries scientists should not be discouraged to defend their research; this is part of the review process. Some research may not be suitable for publication, and this is where careful planning and developing a sound study design at the beginning are especially important. It is rare that a well-designed research study is not publishable. Some well-designed studies have a difficult time in the peer-review process because they are not well written. Thus, clear and concise writing is paramount in technical writing. We suggest the primer by Hunter (1990) for tips on technical writing.

Box 1.2 A Model Selection Application to Fisheries Management

Bunnell et al. (2006) used model selection with AIC_c to investigate crappie (combined black and white crappies) recruitment relations in 11 Ohio reservoirs. The objectives of the study were to determine whether stock–recruit models improved with inclusion of environmental parameters and to determine which life stage best inferred recruitment. The observed data consisted of environmental variables (seasonal reservoir elevation and chlorophyll *a*) and larval (density), age-2 (recruit catch per unit effort [*C/f*]), and adult (spawning stock *C/f*) crappies. The approach consisted of using Ricker or Beverton–Holt stock–recruit models (Chapter 4) that either included or did not include environmental parameters. Sixteen candidate models were considered for each of three stock–recruit relations.

The following provides a summary of model selection results modified from Bunnell et al. (2006). Model selection was used to explain the variation in white and black crappie larval density or *C/f* of recruited (age-2) crappies. Of the candidate models included in the analysis, only the five highest-ranking models in each group are given. Rank was determined by AIC_c. All models included the parameters *a, b,* and *e* from either the Ricker or Beverton–Holt models. Some models also included one or more of the following environmental variables: chlorophyll *a,* winter water elevation, spring water elevation, and summer water elevation. Data include the number of parameters estimated, the residual sum of squares divided by sample size, the difference between each model and the model with the minimum AIC_c, and Akaike weights.

Table Summary of model selection results modified from Bunnell et al. (2006) to explain the variation in crappie larval density or *C/f* of recruited (age-2) crappies. Models may include chlorophyll *a* (CHL), winter water elevation (WI), spring water elevation (SP), and summer water elevation (SU). The stock–recruitment model is specified as Ricker (R) or Beverton–Holt (BH). The measures represented are *K* (the number of parameters estimated), σ^2 (the residual sum of squares divided by *N*), AIC_c, Δ_i (the difference between each model and the model with the minimum AIC_c), and w_i (Akaike weights).

Environmental variables	Stock–recruit model	*K*	σ^2	AIC_c	Δ_i	w_i
Models to explain variation in larval density						
1) Adult *C/f*, CHL	BH	4	0.502	–0.378	0.00	0.46
2) Adult *C/f*, CHL	R	4	0.518	0.152	0.53	0.35
3) Adult *C/f*, SU, CHL	BH	5	0.497	3.554	3.93	0.06
4) Adult *C/f*, SP, CHL	BH	5	0.499	3.626	4.00	0.06
5) Adult *C/f*, WI, CHL	BH	5	0.502	3.735	4.11	0.06
Models to explain variation in *C/f*						
6) Larval density	BH	3	0.602	–2.029	0.00	0.31
7) Larval density	R	3	0.604	–1.969	0.06	0.30
8) Larval density, CHL	BH	4	0.546	–0.649	1.38	0.16
9) Larval density, CHL	R	4	0.552	–0.445	1.58	0.14
10) Larval density, SP	R	4	0.583	0.620	2.65	0.08
Models to explain variation in recruit *C/f*						
11) Adult *C/f*, CHL	R	4	0.219	–2.039	0.00	0.51
12) Adult *C/f*	R	3	0.402	–0.595	1.44	0.25
13) Adult *C/f*	BH	3	0.454	0.733	2.77	0.13
14) Adult *C/f*, CHL	BH	4	0.314	1.937	3.98	0.07
15) Adult *C/f*, WI	R	4	0.354	3.236	5.28	0.04

As shown in the above table, chlorophyll *a* was the most common environmental parameter found to contribute to candidate models. The AIC_c, Δ_i, and w_i values for models 1 and 2 indicate that Beverton–Holt and Ricker stock–recruit models that included chlorophyll *a* provided the best-supported models for variation in larval densities. Best supported stock–recruit models (6 and 7) for variation in *C/f* based on larval density did not include environmental parameters. The Ricker model containing adult *C/f* and chlorophyll *a* (model 11) was best supported in explaining variation in recruit *C/f*.

Correctly reporting results in publications, particularly statistical results, can help the reader determine, among other relevant measures, the statistical test(s) used, effect size, biological versus statistical significance, and sample size. By incorrectly reporting results the author can mislead the reader and make the manuscript confusing, which often leads to the manuscript being rejected or needing major revision.

Authors need to define clearly the sample population (in the statistical sense) and experimental unit from which the statistical inference will be drawn. When reporting measures of central tendency, using the mean is appropriate, but when nonparametric statistics are used medians and modes should be considered for summarizing data. All measures of central tendency and variability should be accompanied by a sample size (e.g., $N = 121$). There are many measures of variability (e.g., SD [a descriptive statistic], standard error [SE, an inferential statistic], coefficient of variation [CV], and confidence interval [CI]), and each has its specific use. However, because we are often interested in inference about the population mean we recommend SE or CIs. Confidence intervals are extremely useful because they infer with a given level of confidence for the interval within which the true parameter lies. Several authors that criticize hypothesis testing do support the use of CIs (Yoccoz 1991; Johnson 1999; Anderson et al. 2001). One of the most common ways to report these data are to give the mean ± 95% CI (for symmetric confidence intervals; report the upper and lower confidence intervals for asymmetric intervals) along with the sample size (e.g., 15 ± 4, $N = 95$).

Care should be taken when reporting information from null hypothesis testing. For example, making statements such as "there were no significant differences in length among treatments" says nothing about the statistical technique used, sample size, level of significance, or effect size. We recommend that authors should report the test statistic value and probability value to two decimal places (e.g., $F = 2.31$, $P = 0.03$, $N = 85$), unless there is statistical justification and need for reporting more significant digits. Also, use of two decimal places corresponds with most published probability levels. Actual probability levels are more useful than broader values (e.g., $P < 0.05$). Presenting the actual probability level allows the reader to consider the statistical or biological significance of the result. The predefined significance level may not be the absolute limit between nonsignificant and significant findings under practical conditions. For example, what might we conclude if the mean length of a species did not differ significantly between two populations and the *P*-value was 0.06 ($\alpha = 0.05$). In many cases authors would conclude no difference in mean length. We suggest that authors mention the effect size and state the difference was significant at $P = 0.06$. It is possible for results to be biologically significant but not statistically significant. One of the shortcomings associated with hypothesis testing is that the *P*-value is closely associated with sample size. Thus, increasing the sample size can increase the likelihood of observing significant results. Determining what is biologically significant (i.e., effect size) should be determined prior to the study. For example, is a mean length difference of 1 mm biologically important? What about 10 cm? Identifying what is biologically significant at the beginning of a study is as important as establishing α and β values.

Regression and correlation analyses are important tools commonly used in the fisheries profession. Minimally, the regression line (i.e., in figures), parameter estimates, and the coefficient of determination (r^2 or R^2) should be included when reporting regression results. When linear regression is used for prediction, the prediction interval instead of the CI should be used. The prediction interval and CI of the regression line are hyperbolic upper and lower boundaries to the regression line, and the prediction interval is farther from the regression line than is the CI. The prediction interval is the interval within which we are 95% ($\alpha = 0.05$) confident that a single future observation will fall. The CI provides the range within which the mean of additional observations will fall. When using nonlinear models authors need to state clearly why a nonlinear model was used (e.g., how did it improve the fit over a linear model), and all parameters in the model should be described. Multiple regression models are usually presented in tables because they are difficult to depict graphically. A table reporting the multiple regression results should contain at a minimum the parameter estimates, SEs, R^2, and *P*-values. Generally, results that report only the correlation coefficient (r, strength of the linear association) and significance level have less value in the scientific literature.

It seems that many of the problems authors have in the publication and peer-review process center around experimental design and statistical tests. This is rather unfortunate given both of these issues can be easily avoided. A good experimental design can be developed for most studies that are truly experiments. Complex statistical tools, frequentist methods, or information theory are not a prerequisite to publishing research. Problems arise when observational studies are treated like experimental studies in the design and analysis stage. Further, it is inappropriate to choose a statistical analysis and then attempt to fit the data to that analytical method. Data analysis and statistical tools must be appropriate for the question. Clearly defining a study at its beginning and using the appropriate data analysis tools will make publishing a more satisfying, contributory process. Rarely are papers rejected because the question was not important or relevant to fisheries science.

1.6 SUMMARY

Conducting high-quality fisheries science is challenging. To conduct quality science, fisheries scientists need to apply the scientific method, use the correct experimental design, and use the most appropriate analytical tools. Being judicious throughout all aspects of the scientific process will help ensure that the fisheries profession gains reliable knowledge. Not all research fits the model of experimental design, hypothesis testing, or information theory; nevertheless, these studies (i.e., observational) can provide useful information regarding the natural history of a species or case history of a management application, to name a few. With that said the fisheries profession should strive to increase efforts in experimental research to provide a more solid foundation for interpretation and conclusions. Further, the fisheries profession should implement research based on deductive reasoning as opposed to inductive reasoning. Fisheries scientists have a tremendous amount of

responsibility with regard to how they influence aquatic ecosystems. Thus, understanding and managing these systems correctly can only be achieved through reliable knowledge.

1.7 REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petrox and F. Caski, editors. Proceedings of the second international symposium on information theory. Akademiai Kiado, Budapest, Hungary.

Akaike, H. 1974. A new look at the statistical model identification. IEEE (Institute of Electrical and Electronics Engineers) Transactions on Automatic Control 19:716–723.

Anderson, D. R., W. A. Link, D. H. Johnson, and K. P. Burnham. 2001. Suggestions for presenting the results of data analyses. Journal of Wildlife Management 65:373–378.

Behnke, R. J. 2002. Trout and salmon of North America. Simon and Schuster, New York.

Berger, J. O. 1994. An overview of robust Bayesian analysis. Test 3:5–124.

Bradford, M. J., B. Pyper, and K. S. Shortreed. 2000. Biological responses of sockeye salmon to the fertilization of Chilko Lake, a large lake in the interior of British Columbia. North American Journal of Fisheries Management 20:661–671.

Bunnell, D. B., R. S. Hale, M. J. Vanni, and R. A. Stein. 2006. Predicting crappie recruitment in Ohio reservoirs with spawning stock size, larval density, and chlorophyll concentrations. North American Journal of Fisheries Management 26:1–12.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd edition. Springer Science and Business Media, New York.

Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods and Research 33:261–304.

Chamberlin, T. C. 1965. The method of multiple working hypotheses. 1890. Reprint, Science 148:754–759.

Conover, W. J. 1980. Practical nonparametric statistics, 2nd edition. Wiley, New York.

Cox, D. R., and E. J. Snell. 1989. Analysis of binary data, 2nd edition. Chapman and Hall, New York.

Dallal, G. E., and L. Wilkinson. 1986. An analytic approximation to the distribution of Lilliefors' test statistic for normality. The American Statistician 40:294–296.

Dewey, J. 1938. Scientific method: induction and deduction. Pages 419–441 *in* J. Dewey, editor. Logic—the theory of inquiry. Henry Holt, New York.

Dey, D., P. Muller, and D. Sinha. 1998. Practical nonparametric and semiparametric Bayesian statistics. Springer-Verlag, New York.

Efron, B. 1998. R. A. Fisher in the 21st century. Statistical Science 13:95–122.

Ford, E. D. 2000. Scientific method for ecological research. Cambridge University Press, Cambridge, UK.

Garton, E. O., J. T. Ratti, and J. H. Giudice. 2005. Research and experimental design. Pages 43–71 *in* C. E. Braun, editor. Techniques for wildlife investigations and management, 6th edition. The Wildlife Society, Bethesda, Maryland.

Gelman, A., J. B. Carline, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall/CRC Press, Boca Raton, Florida.

Geweke, J. 1998. Simulation methods for model criticism and robustness analysis. Pages 275–299 *in* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors. Bayesian statistics, volume 6. Oxford University Press, Oxford, UK.

Haddon, M. 2001. Modeling and quantitative methods in fisheries. Chapman and Hall/CRC Press, Boca Raton, Florida.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey.

Hoenig, J. M., and D. M. Heisey. 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. The American Statistician 55:19–24.

Hunter, J., editor. 1990. Writing for fishery journals. American Fisheries Society, Bethesda, Maryland.

Hurvich, C. M., and C. L. Tsai. 1989. Regression and time series model selection in small samples. Biometrika 76:297–307.

IOM (Institute of Medicine). 2002. Integrity in scientific research. National Academy of Sciences, Washington, D.C.

Johnson, D. H. 1999. The insignificance of statistical significance testing. Journal of Wildlife Management 63:763–772.

Kass, R. E., and L. Wasserman. 1996. The selection of prior distributions by formal rules (with comments). Journal of the American Statistical Association 91:1343–1370.

Kerlinger, F. N., and H. B. Lee. 2000. Foundations of behavioral research, 4th edition. Harcourt College Publishers, New York.

Kohler, C. C, and W. A. Hubert, editors. 1999. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Krueger, C. C., and D. J. Decker. 1999. The process of fisheries management. Pages 31–59 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Kullback, S., and R. A. Leibler. 1951. On information and sufficiency. Annals of Mathematical Statistics 22:79–86.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monograph 62:67–118.

Lee, J. A. 1999. The scientific method: a primer on scientific principles and practice. Addison Wesley Longman, San Francisco.

Luecke, C., W. A. Wurtsbaugh, P. Budy, H. P. Gross, and G. Steinhart. 1996. Simulated growth and production of endangered Snake River sockeye salmon. Fisheries (Bethesda) 21(6):18–25.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology, 2nd edition. Chapman and Hall, London.

Marascuilo, L. A., and M. McSweeney. 1977. Nonparametric and distribution free methods for the social sciences. Brooks/Cole Publishing Company, Monterey, California.

Maxwell, M. R., R. M. Peterman, M. J. Bradford, and E. A. MacIsaac. 2006. A Bayesian analysis of biological uncertainty for a whole-lake fertilization experiment for sockeye salmon in Chilko Lake, British Columbia, and implications for the benefit–cost ratio. North American Journal of Fisheries Management 26:418–430.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall/CRC Press, Boca Raton, Florida.

Moore, B. N., and R. Parker. 2007. Critical thinking, 8th edition. McGraw-Hill, New York.

Murphy, B. R., and D. W. Willis, editors. 1996. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Nielsen, L. A. 1999. History of inland fisheries management in North America. Pages 3–30 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Paul, R. W. 1990. Critical thinking in North America. Pages 18–43 *in* A. J. A. Binker, editor. Critical thinking: what every person needs to survive in a rapidly changing world. Center for Critical Thinking and Moral Critique, Rohnert Park, California.

Paul, R. W., and A. J. A. Binker. 1990. Strategies: thirty-five dimensions of critical thinking. Pages 305–349 *in* A. J. A. Binker, editor. Critical thinking: what every person needs to survive in a rapidly changing world. Center for Critical Thinking and Moral Critique, Rohnert Park, California.

Platt, J. R. 1964. Strong inference. Science 146:347–353.

Pollock, K. H., C. M. Jones, and T. L. Brown, editors. 1994. Angler survey methods and their applications in fisheries management. American Fisheries Society, Special Publication 25, Bethesda, Maryland.

Popper, K. R. 1959. The logic of scientific discovery. Hutchinson, London.

Popper, K. R. 1968. Conjectures and refutations: the growth of scientific knowledge, 2nd edition. Harper and Row, New York.

Press, S. J. 1989. Bayesian statistics: principles, models, and applications. Wiley, New York.

Romesburg, H. C. 1981. Wildlife science: gaining reliable knowledge. Journal of Wildlife Management 45:293–313.

Rubin, D. B. 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. Annals of Statistics 12:1151–1172.

Schreck, C. B., and P. B. Moyle, editors. 1990. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Sheskin, D. J. 2000. Handbook of parametric and nonparametric statistical procedures, 2nd edition. Chapman and Hall, Boca Raton, Florida.

Sivia, D. S. 1996. Data analysis: a Bayesian tutorial. Oxford University Press, Oxford, UK.

Stephens, P. A., S. W. Buskirk, G. D. Hayward, and C. Martinez Del Rio. 2005. Information theory and hypothesis testing: a call for pluralism. Journal of Applied Ecology 42:4–12.

Strahler, A. N. 1992. Understanding science: an introduction to concepts and issues. Prometheus Books, Amherst, New York.

Thomas, L., and C. J. Krebs. 1997. A review of statistical power analysis software. Bulletin of the Ecological Society of America 78:126–139.

Walley, P. 1991. Statistical reasoning with imprecise probabilities. Chapman and Hall, London.

Wasserman, L. A. 1992. Recent methodological advances in robust Bayesian inference. Pages 483–502 *in* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors. Bayesian statistics, volume 4. Oxford University Press, Oxford, UK.

Yoccoz, N. G. 1991. Use, overuse, and misuse of significance tests in evolutionary biology and ecology. Bulletin of the Ecological Society of America 72:106–111.

2 Fisheries Management Study Design Considerations

Richard L. Noble, Douglas J. Austen, and Mark A. Pegg

2.1 INTRODUCTION

Fisheries management is broadly defined as the art and science of providing sustained aquatic resource productivity. Specifically, management is usually the process of maintaining fish populations at some target level, commensurate with the capacity of the environment and in accordance with established management objectives that are set with consideration of user or constituent needs (e.g., Bennett 1971). In short, fisheries management is the integration of the fish, habitat, and user dimensions of the resource to yield a given product or set of products. Usually this is done through manipulation of one or more of these dimensions (the "action"; Krueger and Decker 1999); however, a stable, viable resource may need only protection or, in some cases, no manipulation or management of any kind. Hence, the adage "no management is management" pertains as well. In either case, resource studies by fisheries scientists, both researchers and managers, are an integral part of understanding the aquatic resources and how best to manage them.

It is a common misperception that fisheries scientists must be data rich to conduct effective management. Perhaps that is where the art of management enters and the conventional-wisdom approach (Johnson 1999) is employed. For example, consider that any of us fortunate enough to have a home with a lawn is a resource manager. Much like a fishery manager who assumes responsibility for a resource in his or her jurisdiction, we buy our first, probably previously owned, home, and we inherit a more or less managed lawn resource. We want something quantitatively or qualitatively different, so we visualize what we want and begin to think of management strategies to meet our goals. Do we gather data on the exact size of the lawn, its precise mix of species, the amount of unvegetated space between clumps, or the rate of growth before and following each cutting? Do we develop computer models that integrate functional relations among the grass, the environment (e.g., soil, water, and fertility), and the lawnmower? Probably we do not. Instead, we begin a management program, and the lawn likely improves, sometimes even to the point of our satisfaction. Our goal has been achieved. At most we may have had a soil test, checked with retail or extension consultants for recommendations on grass varieties, fertilized according to general prescriptions, applied pesticides

remedially, watered during droughts, and periodically adjusted the height at which we mowed. A level of success in management was achieved without an assessment-rich or data-rich approach.

On the other hand, a lawn is a quite simple system, and much is known about the relations of grass variety, environmental conditions, and mowing (harvest) strategies. Furthermore, with vaguely defined objectives, the resulting condition, though deemed acceptable, may have been far from optimal—as may have been the management itself. Further investigation might be required to determine whether the appropriate varieties had been planted; complex measures would be essential to optimize use of fertilizers, pesticides, and water; and refinement of mowing gear and technique might also be considered. The system would require a more complete conceptualization and more data to make effective management decisions.

With adequate experience, fisheries scientists may be able to manage for general objectives largely on the basis of observation analogous to the lawn scenario presented above. However, it is more informative and meaningful to integrate a broad information base (Krueger and Decker 1999) to address a rather specific question or set of questions. The most applicable and defensible component of that information base probably is a set of data collected in studies directly conducted on the specific resource of interest following a proper sample design. Strategies used to obtain such data could include a determination of status at a point in time–usually in anticipation of the need for management (the "best current data" approach; Johnson 1999), evaluation of change over time in response to a specific manipulation or perturbation, or simply long-term monitoring of attributes of interest without specific intentions to implement management. All of these strategies have important ramifications in how the data are collected and used that must be considered at the onset of any new sampling effort.

2.1.1 Study Objectives versus Management Objectives

There are subtle, yet substantial differences in the perception and implementation of sample designs when comparing the objectives of a fisheries management plan versus that of a research study (Figure 2.1). The differences will be discussed in section 2.1.2 in greater detail, but the crux of the differences between the two types of objectives is that the management objective approach is aimed at evaluating a response to achieve a goal with some understanding of the tools that might be useful to achieve the objective whereas research objectives are less concerned with predetermined, desired results. Both approaches do follow parallel paths in that they should be statistically rigorous and follow the philosophy of the scientific method (see Chapter 1). Their development should follow a logical progression from a conceptual model of the problem to formulation into an analytical framework specifying the variables and their probable interrelations (Box 2.1; Waters and Erman 1990). Success in achievement of these objectives is usually measured in terms of the statistical confidence in the results through data analyses, as well as by the manager's confidence of their applicability to management decisions.

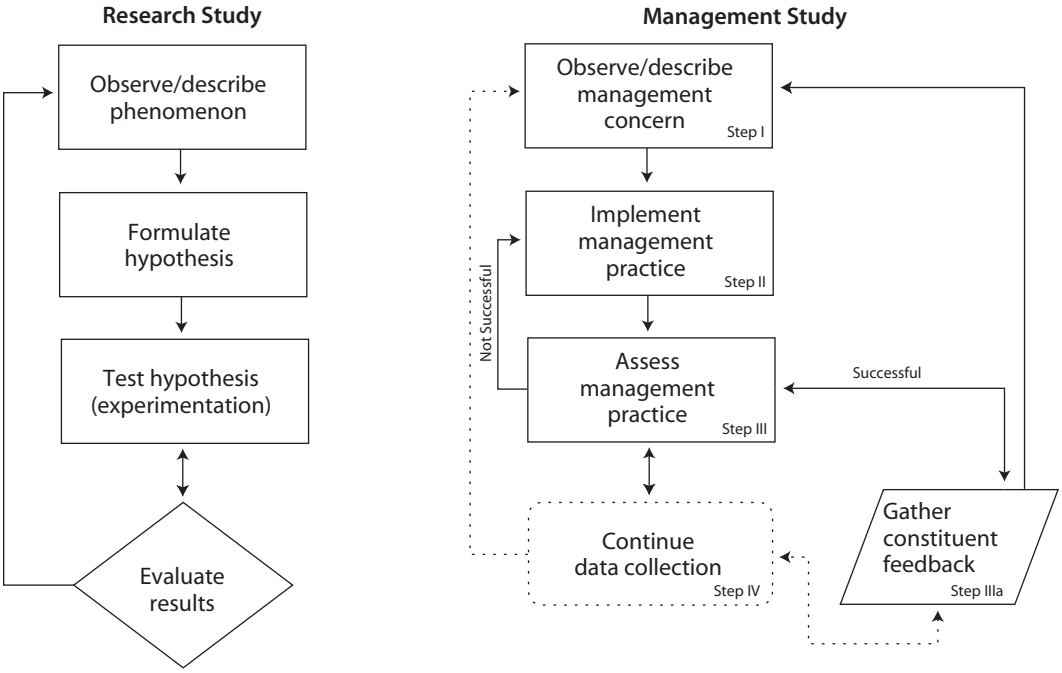


Figure 2.1 Simplified, conceptual diagram of a research study (left) and management study (right) outlining the similarities and differences between the two study types. Dashed lines represent long-term data collection that can provide future, iterative feedback to the management study as the management goals are refined. See text and Box 2.1 for detailed examples of the management study process.

Sample designs and efforts that assess management objectives can be more limited because the manager needs very explicit information to answer questions at hand. This is primarily because fisheries scientists manage resources according to goals and objectives that target specific outcomes. Ultimately, meeting the needs of the resource users—historically anglers interested in one or more key species but increasingly the general public concerned with the integrity of the system—is the goal. This goal is commonly embraced somehow under the umbrella term "optimum yield" (Anderson 1975; Roedel 1975; Malvestuto and Hudgins 1996). However, the manager may set specific objectives at various system levels, assuming that benefits accrue sequentially toward goals established for higher system levels (Noble 1986). For example, objectives might be set to achieve a prescribed minimum catch rate by anglers, a minimum proportion of satisfied anglers for a specific fishery, or a level of faunal diversity acceptable to the general public. Management of one or more components of the fishery is anticipated to achieve the objective, but there is a need for assessment to measure other potential outcomes of any implemented management practice. A comprehensive study of all aspects of the fishery—fish community dynamics, habitat, and users–would be

Box 2.1 Implementation Process of a Largemouth Bass Management and Monitoring Program

Implementation of management studies can be somewhat different than studies conducted purely for research. Although there are distinct similarities in the thought process and evolution of management and research studies, there are also distinct differences. The largest dichotomy of the management versus research process is that a management study is limited in statistical comparability but still has some means of direct feedback to answer the question, Did the management tool have the desired outcome in the system being managed? The answer to this question can be revealed through the same mechanisms used in research studies, but often there is an additional resource constituent input that influences the process too (Figure 2.1). There is also often a need for longer-term assessments due to this iterative process (Figure 2.1) that only continued data collection can provide. The management of the B. E. Jordan Reservoir (Jordan Lake) largemouth bass population provides an example of the management study process.

Jordan Lake is a 5,720-ha impoundment in the Piedmont Region of North Carolina. Jordan Lake was impounded and filled between 1981 and 1983. The lake's proximity to population centers in the state suggested it would become a popular fisheries resource. As new reservoirs are prone to do, Jordan Lake experienced a period of dynamic fish community response to impoundment followed by eventual stabilization. The adult largemouth bass population quickly became characterized by large fish in excellent condition, typical of a eutrophic system with low recruitment and ample prey. The management goal for Jordan Lake was to maintain a quality largemouth bass fishery (step I, Figure 2.1), so the management agency implemented a 406-mm-minimum-size limit in 1987—only 4 years after full impoundment (step II, Figure 2.1).

Largemouth bass stocks in lakes and reservoirs are typically assessed by shoreline electrofishing of adult fish during the spring. That sampling is designed to monitor abundance (catch per unit effort, *C/f*), size and age composition, and condition (relative weight) simultaneously. After the initial sampling to assess immediate response of the largemouth bass population (step III, Figure 2.1), the agency established an assessment objective of "a favorable size structure at moderate to high densities (electrofishing *C/f*)." As the management goal was refined and constituent satisfaction established (step IIIa, Figure 2.1), the assessment objective was further modified to maintain electrofishing *C/f* between 0.02 and 0.04 largemouth bass per minute for the 380–509-mm size-group; to maintain mean relative weight at or above 100 for that size-group; and to maintain a fish total length of at least 406 mm (the minimum size limit) at age 5. The agency implemented an annual sampling regime that started in 1983 to evaluate the management regulations (steps III and IV, Figure 2.1). Through continued monitoring, fishery managers responsible for Jordan Lake know that the regulations set in 1987 have fulfilled the management goal to date (McRae and Oakley 2005).

The added value of long-term monitoring can also relate to improved efficiencies in the future management of a given system (step IV, Figure 2.1). Initially 16–17 sites were sampled by electrofishing annually at Jordan Lake to provide the necessary statistics for all components of the assessment objective. However, stability in the system and consistent annual catch rates have allowed the number to be reduced to five to eight sites without major loss of information.

The goal of a management study is ultimately to assess tools used to sustain the resource. While there have been some changes to the sampling regime as more information is learned about the system, this standardized approach conducted over nearly 20 years has satisfactorily indexed the status of the Jordan Lake largemouth bass population to judge the need for and response to management.

desirable, but the fishery manager is likely to find it more feasible to assess key attributes of the system that correlate with the desired management goal rather than to measure a full complement of constituent responses directly. The important thing here is that the sample design and data collected must allow the manager to evaluate effectively the problem being addressed (Brown and Austen 1996). Common methods for each component of the fishery discussed above have been recently summarized and provide good insight on many of the issues centered around sample design and data collection. Murphy and Willis (1996) emphasize biotic sampling, Bain and Stevenson (1999) stress habitat assessment, and Pollock et al. (1994) discuss elaborate techniques for angler surveys.

2.1.2 Management Studies Differ from Research

Knowledge of conceptual relations is critical for fisheries scientists, whether working in research or management roles. Fisheries scientists manage according to conceptual relations in their systems but usually focus their assessment efforts on key attributes believed to be important indicators of the status of their resource. Therefore, management studies are likely to differ substantially from research studies. Rather than elucidating mechanisms based on hypotheses or concepts, the manager is likely to assume that a fundamental model pertains to the system and focus on monitoring status of key components of the stock, environment, or users. In doing so, the manager is apt to depend on indices and trends rather than getting absolute estimates (e.g., use of catch per unit effort to index density; see Chapter 7). Rarely will the manager rely upon measures associated with a single resource attribute, so studies commonly assess several variables concurrently and frequently during the same sampling endeavor. Internal consistencies among data sets (e.g., increases in condition indices corresponding to decreases in catch rates) are then examined to reinforce confidence in the validity of individual study results.

Also in contrast to research, assessment by the manager is likely to be site- or resource-specific, with results aimed at being of local rather than broad applicability. Unlike original research, these studies may be repetitive of other studies due to their limited inference capabilities. In fact, agencies commonly establish standardized protocols for resource investigations to optimize comparisons across time and space (Noble 2002).

Although alternative approaches are available, management studies typically lack replication and controls. This then results in concerns over the strength of the overall sampling design and scientific rigor compared with research investigations. Therefore, many management studies, while potentially highly valuable, are poorly communicated to the profession because of their potentially narrow focus and perceived limited contribution to the science of fisheries management. That does not have to be the case, as well-designed management studies that are properly executed, rigorously analyzed, and written with regard to broader scientific literature commonly have been published in the primary journals as articles, management briefs, and case histories.

2.2 TYPES OF MANAGEMENT INVESTIGATIONS

Fisheries management assessment takes many forms based on the objectives and amount of knowledge available for the system being assessed. Generally, these assessments can be classified as feasibility studies, pilot studies, treatment (management) evaluations, or resource monitoring. These study classifications represent an evolution of knowledge about the system being studied from relatively simplistic feasibility studies to more complex evaluation and monitoring studies. In many cases, sample designs implemented in one type of assessment can readily progress to another classification when properly planned.

2.2.1 Feasibility Studies

Feasibility studies are usually precursory to more advanced studies, exploratory in nature, and entirely dependent on existing information and experiences to determine if a problem can be resolved. The feasibility study will help identify the working aspects of future study by providing information on the operations and technical issues, such as the appropriate sampling gear, the appropriate temporal and spatial scales, the behavior of the system to be sampled, and sociological concerns. Other issues, such as the economics of implementing a management plan, may need to be addressed. Experience by investigators in other systems is likely to have identified the appropriate gear to use, as well as potential biases associated with the gear, and should provide some insight into temporal considerations. For example, it may already have been established that nets need to be set overnight or water chemistry needs to be measured at a certain time of day to be comparable both among samples within a given water body and among other studies. Likewise, some indications of the frequency of sampling (e.g., sporadically in response to environmental events or at monthly, seasonal, annual, or even multiyear intervals) probably can be obtained. However, future sampling will need to be modified to meet specific study objectives and the behavior of the specific system to be assessed.

Spatial variations in the attributes to be studied, as well as the applicability and efficiency of sampling gear, are cause for much concern in establishing sampling programs. Feasibility studies will help define areas that are conducive to sampling (e.g., snag-free areas to seine), those that provide the most information (e.g., riffle areas as aquatic insect production habitat), and those of appropriate sampling depths (dependent on the variables of interest and the limitations of the sampling gear). In evaluating these aspects, it will be necessary to incorporate knowledge of anticipated weather effects, varying from changes in water level to limits on accessibility. In many cases, information on statistical characteristics can also be obtained through the feasibility study. Literature may be available that provides good information on sampling needs across spatial and temporal scales, expected sample variance, approximate sample sizes needed, and analytical power (e.g., Lubinski et al. 2001).

The fisheries scientist should explore existing knowledge from data already collected by the management entity to existing literature and also incorporate

personal experiences when study objectives have been established. The gray literature (e.g., reports and correspondence) may prove as fruitful as the published literature. Particularly helpful may be reports of fisheries management studies in the USA supported by Federal Aid in Fish Restoration Act. Fisheries scientists have communicated management study progress and results over the past several decades through media such as the newsletter of the Fisheries Management Section of the American Fisheries Society and regionally through working groups and committees of this society. Networking may be the best approach for obtaining investigators' personal experiences pertinent to the study. Professional society list servers provide many valuable Web sites and a quick way to obtain a variety of perspectives on feasibility aspects.

At the completion of the feasibility study, it should be possible to estimate reasonably the time and personnel required to conduct a given amount of sampling and to weigh those costs against the likely benefits of doing the study. Other dimensions of the project, such as public relations aspects of the fieldwork, can also be explored during the feasibility study.

2.2.2 Pilot Studies

Pilot studies are typically smaller-scale studies that offer an opportunity to evaluate assessment alternatives identified under the feasibility study. Essentially, pilot studies are used to determine what sampling is likely to be effective in achieving the specific objectives under the conditions of the investigation to be conducted before carrying out larger studies. They may also be of value in situations in which the concept of the study as a whole is not well accepted, and the pilot study can give assurance that a larger commitment to a full study will be a reasonable investment. In addition to testing the overall methodology anticipated for the investigation, pilot studies are directed at establishing statistical reliability (i.e., precision and accuracy; see Chapter 3). The pilot study should estimate the sample variability and allow calculation of sample sizes needed to attain the desired level of precision. For sampling programs that use fixed time or distance sampling, choices can be made on the relative precision of data from large numbers of small samples versus small numbers of large samples as well as the effects of travel time between samples (e.g., Miranda et al. 1996). The investigator also has the opportunity to evaluate the effects of deviations in technique during the pilot study (e.g., whether it matters that a sample is taken an hour late because of a flat tire on the way to the site). Management studies commonly establish more liberal levels of accepted variability than do research projects (e.g., confidence limits of ± 25% for management studies versus ± 10% for research; Robson and Regier 1964). Therefore, sample sizes for management studies are often much smaller than they are for research investigations.

Opportunities for, and limitations to, random sampling can become apparent over the course of feasibility and pilot studies due to limitations of gear and availability of personnel to conduct a truly random sampling effort. Therefore, management studies characteristically entail sampling at subjectively selected fixed sites. Likewise, for many stream sampling programs, samples are taken near bridges

due to access and logistical limitations. These are constraints that must be realized and dealt with appropriately as such sampling poses the possibility of introducing estimation biases when performing statistical analyses (e.g., Maceina et al. 1994; Manly 2001). Pilot studies can help elucidate the extent of bias associated with fixed-site sampling.

Because the fisheries scientist may be collecting more than one kind of management data simultaneously (e.g., fish density and population structural indices), there is a temptation to select sites that have high densities or sites where fish have high vulnerability to sampling. High catch rates result in greater precision in estimation of structural indices and, conversely, reduce the number of samples required for a given precision. However, changes in catch rates at these sites may not necessarily correspond directly to actual changes in overall population density. Furthermore, structural indices based on these high-density sample sites may also be biased (Hubbard and Miranda 1988). A pilot study offers the opportunity to establish whether the sampling produces representative data. It is not unusual for pilot studies to lead to changes in protocols suggested by the feasibility study.

2.2.3 Measurement of Treatment Effects

Evaluating the impact of a management action is typically the most intensive study done by a manager. In its simplest form, such an evaluation is a before and after study of a single system (without replication). Fisheries based on long-lived species may take several years to respond to a management action, so such studies may need to be of long duration. Before-and-after studies (termed BACI for before–after comparison of impact; Green 1979) suffer from their inability to elucidate causal mechanisms, so sampling during the intervening years, though not essential to meeting the principal objectives, is commonly used to provide supplemental information on the rate and process of response to the treatment. Because most of these studies are site specific and without replication, they serve as case histories that document what happened in response to a management action. A sample design may provide tests of change with designated statistical probability, but such studies cannot provide information about the probability of similar responses in another system, thus limiting inference capabilities.

Broad inferences are facilitated by replicated treatments. If systems are similar enough that treatments can be applied to them as replicates, comparison of treatment effects under varied conditions across the landscape can provide insight into mechanisms of response. Such approaches improve confidence in more general applications of inferences but require large commitments of resources and high levels of coordination.

Development of a management strategy originates from changes detected during routine monitoring, new opportunities for optimization recognized by the manager, and public demands for changes in management. In the interest of responding to the need promptly, and recognizing that it may take years for the effects of management to be manifested in the fishery, the manager tends to

move quickly to implement the management action, thereby minimizing the "before treatment" component of the study. It is common for the baseline data to consist only of the data available from monitoring studies (see section 2.2.4), thereby limiting rigor for detecting specific changes in response to the management action. When possible, the pretreatment study should be intensified to ensure statistical rigor and broadened to include investigation of additional factors and potential responses. Consistency of protocols for the pre- and post-treatment studies is essential for interpreting the response to the change in management. There are always concerns that temporal trends, unrelated to the treatment, can occur. Therefore, BACI investigations are most reliable if experimental controls are employed. Limitations of this approach have been examined in substantial detail (e.g., Underwood 1994; Stewart-Oaten and Bence 2001).

Resource management is continually evolving as more data are obtained resulting in an approach termed adaptive resource management (Walters 1986). This approach typically entails replicated treatments and controls to address uncertainty in the hypothesized responses and observed responses to management through stepwise changes in management and related assessment (Lancia et al. 1996; Johnson 1999). However, the adaptive management approach goes significantly beyond the commonly used monitor and modify approach (Johnson 1999). In general, adaptive resource management has been difficult for natural resource agencies to implement because of insufficient commitments to carry out the relatively large amount of sampling required for adaptive resource management compared with that typically needed for most local management studies. Nevertheless, the approach has been applied effectively on large- scale fisheries in some situations (e.g., McConnaha and Paquet 1996).

2.2.4 Monitoring

Monitoring is crucial to managing fish populations, and the basic need for long-term data is widely acknowledged (Likens 1992; Thomas 1999). This is particularly true given the need for understanding how management practices change fish communities, stocks, populations, or other units of interest to achieve specific goals and objectives. However, management agencies are often criticized for their vast files of data collected for monitoring purposes and archived for future use. While monitoring is not necessarily active management in the sense that fisheries scientists are assessing a specific management practice (although that can certainly be the case), ongoing monitoring is essential to establish the status of the resource and detect the need for active management. Studies conducted for short periods of time represent one relatively quick snap-shot that may not be an accurate representation of the fishery (see Box 2.1). In fact, fisheries managers accept that much of their work is driven by questions and decisions waiting to happen that require data on historical trends and current status. Consequently, long-term monitoring data can be an invaluable source of information that should be given high priority.

Long-term monitoring studies are designed to provide a periodic update on the status of the resource, detect major changes, and establish trends. Such monitoring

studies commonly require a dedicated effort. Periodicity of sampling frequently depends on the dynamics of the system where monitoring of a relatively stable resource (e.g., a well-established fish stock comprising many age groups) may be conducted at multi-year intervals or a more variable resource may be sampled at annual, seasonal, or other more frequent intervals. Monitoring can demand a large commitment of personnel and staging time, so it is not unusual for such assessments to monitor multiple attributes of the resource at one time (U.S. Geological Survey 1999). Different sample sizes are likely to be required for the various objectives identified. Fisheries scientists usually define a time of year when sampling gear is most effective, thereby ensuring high catches or high probability of obtaining desired samples. This approach further minimizes the amount of sampling effort required for certain objectives.

When establishing long-term monitoring programs, great care should be exercised to anticipate change from current conditions. If fixed-site sampling is used (see section 2.3.3), one must recognize that long-term changes in the landscape may occur (e.g., a homeowner may bulkhead the shoreline at a seine site or new water level management regimes for a reservoir may inundate sample sites during the sampling season) that should be dealt with by contingency planning. Likewise, changes in gear technologies are inevitable, leading ultimately to using a more efficient gear or lower number of gears (Ickes and Burkhardt 2002) than initially employed. In either of these cases, sampling concurrently with the previously used and newly adopted gears or sites under a range of conditions can provide a basis for adjusting results needed for long-term comparisons.

2.3 DESIGN OF THE STUDY

Monitoring studies and management evaluations provide major challenges when determining minimum sample sizes (number of samples) and the distribution of samples through time and space while still maintaining a high level of statistical reliability. In trying to develop a sampling scheme, the fisheries scientist is likely to experience bewilderment and frustration with the conflicts of science and practice. For example, the methods used to calculate sample size are straightforward (Chapter 3), but the commonly used approach of sampling simultaneously for multiple objectives complicates sample size determination. Furthermore, effectiveness in the distribution of samples through time and space (e.g., stratification) affects minimum sample sizes. Ultimately, management studies are a compromise that allows inferences from a study based on (1) effective sampling to produce representative data and (2) adjustment of the sample design to accommodate real field conditions. Philosophies and methods on dealing with such design complexities in fisheries studies are discussed in Chapter 3 and also reviewed by Johnson and Nielsen (1983), Brown and Austen (1996), Willis and Murphy (1996), and Ney (1999). In general, simplicity in study design makes communication of results more straightforward and the interpretation to constituents more effective.

2.3.1 Data Collection to Achieve Multiple Objectives

A review of Box 2.1 illustrates the challenge of establishing sample sizes when multiple objectives are simultaneously pursued by the manager. Variation in *C/f* (calculated from a pilot study or data from previous years) and the number of samples required to detect a given amount of change can be calculated (see Chapter 3). However, these samples also need to provide data on length–age distributions and condition (weight–length relations). Information requirements differ among these objectives as the investigator only needs to differentiate those fish below 380 mm and over 509 mm long to achieve the *C/f* objective, whereas specific lengths, weights, and age data are required to measure the other responses. This issue is further compounded by the fact that the electrofishing transect provides the needed data for *C/f*, yet information for the remaining variables is obtained from individual fish. The number of fish required varies depending on the specific estimates being sought. For example, as many as 100 fish may be required to get a reliable index of size structure to achieve the desired precision for a given objective (Weithman et al. 1980). If *C/f* is low, insufficient numbers of fish may be available to estimate size structure or mean relative weight with desired precision. Conversely, if *C/f* is high, subsampling may be necessary for efficient use of labor. A contingency plan should be developed in advance for addressing these situations (Willis and Murphy 1996).

Managers have long sought response variables or metrics that concurrently incorporate multiple attributes of their systems, in part to alleviate problems of simultaneous sampling for multiple objectives. Classic examples of these metrics include those that attempt to reflect predator–prey relations (Swingle 1950; Anderson 1976; Jenkins and Morais 1978; Chapter 16). Statistical properties of these metrics are generally unclear, both in terms of appropriate sample sizes and analytical models. Community structure likewise can be indexed by diversity measures (Shannon and Weaver 1949; Chapter 15) and multiple fish attributes (Karr 1981). Statistical properties of some of these indices have been examined in detail (Karr and Chu 1999), and knowledge of those properties should be used in designing sampling programs for assessment of changes in fish assemblages and community structure.

2.3.2 Statistical Efficiency in Sampling

Fisheries scientists must first and foremost have confidence their study design will produce data that will allow the pertinent question(s) to be answered. Key components of that confidence are the representativeness of the data and precision of the estimate. Decisions on how to allocate samples through time and space are paramount to achieving adequate accuracy and precision. Representativeness is most likely achieved through random sampling, either simple (totally) random, stratified random, or cluster sampling (Chapter 3). However, systematic sampling can also provide representative results under most circumstances but will probably have inflated variability (Manly 2001).

Refinements of sample design beyond simple random sampling can be made to reduce sampling effort, thereby reducing time, personnel, and fiscal requirements when major trends in resource attributes are known. For example, sampling to monitor a fish population or community is seldom conducted at random times throughout the year but rather during a specific period when organisms are highly susceptible to collection, environmental conditions are stable, and sampling gear most efficient. Furthermore, the attribute of interest (e.g., density of fish or water chemistry) may follow a general trend or have similarity among subunits of space (e.g., upstream to downstream) or time (e.g., seasonally) such that stratified random sampling may be employed to increase precision or, more likely, decrease the number of samples required to attain the desired precision. Stratification has become a standardized technique in fisheries management and research studies. For instance, angler surveys are commonly stratified over time (i.e., season, day of week, and time of day) in relation to temporal variation in participation (Pollock et al. 1994). The concept of stratification is based on sample data being more similar within subunits than for the resource component as a whole. The risk of reducing precision through stratification is relatively small, whereas the gains can be high if variation among strata is high and within strata is low (Chapter 3).

In the strictest sense, statistical variability among samples should be the natural variability associated with the parameter of interest in relation to the sample technique. For example, the number of fish passing through each of a set of parallel weirs varies at any given time due to the tendency of many organisms to aggregate or school. In fact, the underlying distribution of sample data commonly is such that the variance increases disproportionately with the mean, such that a large number of small samples produce superior precision to a small number of large samples (e.g., Anscombe 1949). However, under field conditions, time and labor considerations commonly favor large samples in smaller numbers as the relation between sample duration and travel time may influence overall catch rates (Miranda et al. 1996). Likewise, when allocating samples that estimate density (e.g., C/f) under a stratified random sampling design, it is likely that precision will be increased if high-density strata are sampled more intensively than low-density strata.

Sample variability is further inflated by any inconsistency in sampling. Consequently, it is extremely important that sampling gear be in a consistent state of repair each time it is used, that sampling conditions and durations be as uniform as possible, that operator error (including that of data recording) be minimized, and so on. Protocols should be established in advance to establish acceptable deviations and alleviate many of the inconsistency issues that can arise without a formal design.

2.3.3 Fixed Sites: Compromise between Statistical and Practical Considerations

Study design is heavily dependent upon understanding the decision-making process, the quality of the information the study demands, and the possible consequences

of differing conclusions from the study (Box 2.2). For the fishery manager, indices, rather than estimates, are frequently used to meet decision-making requirements. In such cases, sampling at fixed sites is commonly accepted. Randomly or systematically selected fixed sites can offer some statistical advantages over subjectively selected sites, but sites are commonly selected subjectively for convenience, consistency, and efficiency of field time and labor.

Reasons for subjective selection of sample sites are many. Access is a common factor. For example, water quality conditions in a stream might be monitored at a highway bridge, where the stream is easily accessible and private property issues can be avoided. Another important factor is interaction of environmental conditions and gear efficacy. For example, use of fixed seine or gill-net sites ensures that sampling variability is not inflated by incomplete samples due to unpredictable snags. Any sampling to assess density by measures of C/f depends on uniform catchability of fish over time and space. Therefore, use of fixed stations can help ensure that the data are not influenced by variations in catchability due to unpredictable sampling conditions. Subjective selection of sampling times is also commonly employed. Weekend sampling may be avoided to minimize potential conflicts with the public or water sampling may be restricted to conditions below flood stage for personnel safety, among many other reasons.

The investigator is faced with decisions about whether to establish sites that are judged to represent the environment, to represent extremes in the environment, to be similar to one another, or to produce highest catches. Each will have its individual effect on variance among samples for the index. For example, fixed site samples chosen for similarity should produce data that have a smaller variance than do samples chosen to represent extremes, so significant differences in the index values could be more easily detected. Nevertheless, the chances of the conclusions being wrong due to nonrepresentative data are higher with this approach. Therefore, random sampling conducted in conjunction with fixed sites can provide the manager with confidence in the degree of representativeness of the data (Manly 2001).

Fisheries scientists commonly select sites subjectively to represent the types of habitats available (e.g., shoreline gill-net sites on a variety of slopes and substrates) or to encompass a variety of habitats at given sites (e.g., shoreline electrofishing transects that traverse a combination of cove and non-cove areas in reservoirs or riffles, pools, and runs in streams). Even if sample sites are subjectively selected as being representative, the data must be recognized as indices reflecting those sites rather than being population parameter estimates. Even though data collected at fixed sites must be considered indices, statistical precision still should be maximized to allow detection of changes in index means. Consequently, attention should be paid to opportunities for stratification (see Chapter 3). Likewise, statistical analyses can be more complex due to issues such as increases in the likelihood of nonindependence of sequential samples and other analytical issues, such as pseudoreplication, if analyses are not performed appropriately (Hurlbert 1984; Chapters 14 and 18). Failure to take these issues into account can lead to misinterpreted results, thus specific statistical approaches must be used (e.g., repeated

Box 2.2 The Importance of Long-Term Monitoring of Fish Populations

The Illinois River was known as one of the fishing and hunting destinations for presidents and as the second most productive inland commercial fishery in the USA until the early 1900s. However, this same river has since suffered ecological decimation from Chicago, other sources of urban and industrial sewage, levees that isolated the river from its floodplain, and water level manipulation. By the mid- to late-1900s, though, efforts to remediate at least the most chronic of these conditions, in this case, water quality due to human health concerns, began in earnest. Subsequent restoration of some of the water quality aspects of the Illinois River through implementation of the Clean Water Act of 1972 markedly improved water conditions in the later quarter of the twentieth century.

Concurrent to these water improvement measures, fisheries managers began to document the current status of the fishery in the Illinois River to establish information that would assist with documenting future changes. That sampling effort, now termed the Illinois River Fish Population Monitoring Program, has resulted in a study conducted by the Illinois Natural History Survey that has provided nearly continuous information on the fish populations in the Illinois River since 1957 through the present (Pegg and McClelland 2004).

The program currently samples 27 fixed sites over nearly 450 river kilometers by means of standardized protocols that detail the specific gear used (e.g., AC boat-mounted electrofishing) and under what hydrological and thermal conditions the sampling should be conducted during an early fall sampling window. With minor exceptions, the sites have been consistent since the inception of sampling. The result is an over four-decade data set that has documented a positive fish response in the upper half of the river where water quality was most improved (Pegg and McClelland 2004). Continued sampling through this program will provide future insight into systemic changes as additional restoration efforts are implemented by numerous federal and state management programs.

This study does illustrate the issue raised with fixed site (section 2.3.3) compared with random site selection. The Illinois River Fish Population Study has continued a fixed site tradition that has allowed data comparison since 1957. However, the use of fixed sites can restrict inferences. In contrast, a random selection, or one stratified based on habitat type, would have allowed the investigators to make more robust inferences to the entire river. Nevertheless, this long-term monitoring study has documented the gradual increase in exotic species such as the bighead carp, the rebound of sport fish populations, and some responses to changes in water quality, such as the incidence of external abnormalities like lesions or tumors (McClelland and Pegg 2005). The data have been critically useful in understanding the effects of anthropogenic influences on an expansive and valuable resource.

measures for sites sampled through time). Wilde and Fisher (1996) discuss some of the risks associated with making inferences from sample designs for which data are obtained by subjectively chosen sites in fisheries studies.

2.3.4 Other Practical Considerations

Beyond planning for statistical precision and accuracy, the fisheries scientist must also be realistic in the efficient use of financial and personnel resources. Personnel time and operating funds for assessment activities are always limited. Field

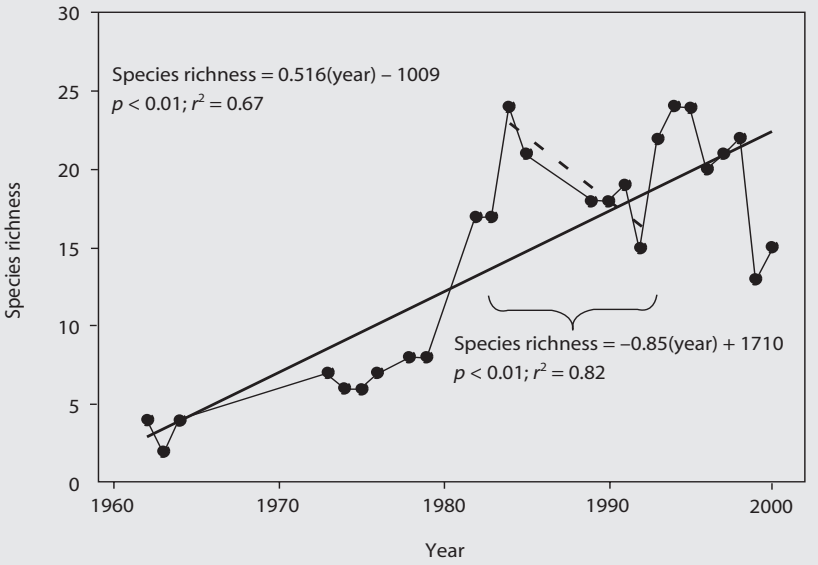


Figure Number of fish species caught by the Illinois River Fish Population Monitoring Program at sites in the Dresden Reach of the Illinois River from 1961 through 2000. The overall trend has been an increase in species richness during the period of record as indicated by the regression line and statistical summary in the upper left corner. The segment of the data represented by the regression equation in the lower right corner of the figure illustrates a different outcome if the study had been conducted over a shorter time frame and highlights the need for fisheries scientists to consider the temporal aspects of their fishery. Figure modified from Pegg and McClelland 2004.

The Illinois River Fish Population Monitoring Program also highlights the need for long- term studies. For example, consider the implications of species richness on the Illinois River. The number of species caught at sites in the Dresden Reach of the Illinois River has exhibited a significant increase since the early 1960s (see figure). However, if a study had been conducted solely during the 1980s to early 1990s, the conclusions would have been considerably different and may have led to implementation of management practices that were not needed or may have been ineffective. There are myriad reasons for different responses at different time scales that likely meld life history, biotic, climatic, and numerous other interactions to elicit responses. Those responses also likely operate at varying time scales. Therefore, it is critical that fisheries scientists consider the temporal time frame and responses of organisms when designing studies.

operations need to be carefully planned, including coordination with other program activities, and all personnel need to be fully oriented to individual responsibilities. Review of operational protocols should be used not only as a means to assure data quality but also as a continual reinforcement of safety policies for field sampling (Berry 1996) and proper procedures for animal care and use (ASIH et al. 1988). Contingency plans should also be in place so that it is quite clear how to respond to deviations from the expected routine when encountered in the field.

Safety is a major consideration when making decisions of whether to conduct a planned assessment as scheduled. Extreme events such as storms, floods, heat,

and cold are becoming increasingly predictable, such that it is reasonable to anticipate interference or interruption of sampling. In many cases it is important to have all samples taken during a single sample period, so it may be preferable to postpone rather than have to complete sampling at another time with differing field conditions. Nevertheless, many statistical analyses readily deal with missing data, so sampling expeditions may be terminated before completion without complete loss of data. Impacts of such deviations from schedules should be evaluated during planning stages of field studies.

2.4 CONCLUSIONS

To a large extent, the fishery manager assesses what is reasonably measurable with the resources available, integrating a number of measures to address the pertinent management question(s). In doing so, experience with sampling gear and data provide an appreciation of the nature of biases and their potential impact on conclusions. Furthermore, familiarity with variability in the data allows some insight on the precision of estimates and confidence in the conclusions. Knowledge of the inference space provided by the study design must be clear so the data can be properly interpreted in terms of whether the data are directly applicable to the parameters of interest or only to the samples themselves. Likewise, the sample design can have significant effects on both the accuracy and precision of estimates, therefore on the defensibility of conclusions drawn from the study. In a sense, the art and science of management is paralleled by an art and science of sample design that reflects an ability to integrate effectively statistical and practical considerations in sampling.

Duration of the study must be sufficient, sample sizes must be adequate, and allocation of samples must be appropriate for results to be conclusive. These considerations make it highly advisable to consult a statistical advisor during the design or re-design stage of any study regardless of the experience level of the investigator. Otherwise, considerable effort, especially when exerted in a monitoring or evaluation study, will have been wasted. Like the researcher, the manager should adopt the rule of thumb that statistical rigor should be sufficient to pass the test of eventual scientific publication of results, even though the study may be site specific, unreplicated, and focused on amount of change more so than on mechanisms of change.

Consistency in implementation of sample design is extremely important. Deviations will occur, and they will tend to increase variability, possibly reduce accuracy, and inevitably reduce overall conclusiveness of the study. Nevertheless, the manager needs to remain flexible, especially for long-term studies.

2.5 REFERENCES

Anderson, R. O. 1975. Optimum sustainable yield in inland recreational fisheries management. Pages 29–38 *in* P. M. Roedel, editor. Optimum sustainable yield as a concept in fisheries management. American Fisheries Society, Special Publication 9, Bethesda, Maryland.

Anderson, R. O. 1976. Management of small warmwater impoundments. Fisheries (Bethesda) 1(6):5–7, 26–28.

Anscombe, F. J. 1949. Sampling theory of the negative binomial and logarithmic series distributions. Biometrika 37:358–382.

ASIH, AFS, and AIFRB (American Society of Ichthyologists and Herpetologists, American Fisheries Society, and American Institute of Fishery Research Biologists). 1988. Guidelines for use of fishes in field research. Fisheries (Bethesda) 13(2):16–23.

Bain, M. B., and N. J. Stevenson, editors. 1999. Aquatic habitat assessment: common methods. American Fisheries Society, Bethesda, Maryland.

Bennett, G. W. 1971. Management of lakes and ponds, 2nd edition. Van Nostrand Reinhold, New York.

Berry, C. R., Jr. 1996. Safety in fisheries work. Pages 63–81 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Brown, M. L., and D. J. Austen. 1996. Data management and statistical techniques. Pages 17–62 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. Wiley-Interscience, New York.

Hubbard, W. D., and L. E. Miranda. 1988. Competence of non-random electrofishing sampling in assessment of structural indices. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 40(1986):79–84.

Hurlbert, S. H. 1984. Psuedoreplication and the design of ecological experiments. Ecological Monographs 54:187–211.

Ickes, B. S., and R. W. Burkhardt. 2002. Evaluation and proposed refinement of the sampling design for the Long Term Resource Monitoring Program's fish component. U.S. Geological Survey, Upper Midwest Environmental Sciences Center, LTRMP 2002-T001, La Crosse, Wisconsin.

Jenkins, R. M., and D. I. Morais. 1978. Predator–prey relations in the predator-stocking-evaluation reservoirs. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 30(1976):141–157.

Johnson, B. L. 1999. The role of adaptive management as an operational approach for resource management agencies. Conservation Biology 3(2):8. (Available at http://www.consecol.org/vol3/iss2/art8.)

Johnson, D. L., and L. A. Nielsen. 1983. Sampling considerations. Pages 1–22 *in* L. A. Nielsen and D. L. Johnson, editors. Fisheries techniques. American Fisheries Society, Bethesda, Maryland.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries (Bethesda) 6(6):21–27.

Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters: better biological monitoring. Island Press, Washington, D.C.

Krueger, C. C., and D. J. Decker. 1999. The process of fisheries management. Pages 31–59 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Lancia, R. A., and nine coauthors. 1996. ARM! for the future: adaptive resource management in the wildlife profession. Wildlife Society Bulletin 24:436–442.

Likens, G. E. 1992. The ecosystem approach: its use and abuse. Ecology Institute, Oldendorf/Luhe, Germany.

Lubinski, K., R. Burkhardt, J. Sauer, D. Soballe, and Y. Yin. 2001. Initial analyses of change detection capabilities and data redundancies in the Long Term Resource Monitoring Program. U.S. Geological Survey, Upper Midwest Environmental Sciences Center, LTRMP 2001-T001, La Crosse, Wisconsin.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot analysis of variance design for repeated measures fishery data. Fisheries (Bethesda) 19(3):14–20.

Malvestuto, S. P., and M. D. Hudgins. 1996. Optimum yield for recreational fisheries management. Fisheries (Bethesda) 21(6):6–17.

Manly, B. F. J. 2001. Statistics for environmental science and management. Chapman and Hall/CRC, Boca Raton, Florida.

McClelland, M. A., and M. A. Pegg. 2005. The long-term Illinois River fish population monitoring program. Annual report to the Illinois Department of Natural Resources. Illinois Natural History Survey Technical Report 05/08.

McConnaha, W. E., and P. J. Paquet. 1996. Adaptive strategies for the management of ecosystems: the Columbia River experience. Pages 410–421 *in* L. E. Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society, Symposium 16, Bethesda, Maryland.

McRae, B. J., and N. C. Oakley. 2005. Characteristics of largemouth bass, crappie, and striped bass in B. E. Jordan Reservoir, 2004. North Carolina Wildlife Resources Commission, Summary Report Federal Aid in Fish Restoration F-23, Raleigh.

Miranda, L. E., W. D. Hubbard, S. Sangare, and T. Holman. 1996. Optimizing electrofishing sample duration for estimating relative abundance of largemouth bass in reservoirs. North American Journal of Fisheries Management 16:324–331.

Murphy, B. R., and D. W. Willis, editors. 1996. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Ney, J. J. 1999. Practical use of biological statistics. Pages 167–191 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Noble, R. L. 1986. Stocking criteria and goals for restoration and enhancement of warmwater and cool-water fisheries. Pages 139–146 *in* R. H. Stroud, editor. Fish culture in fisheries management. American Fisheries Society, Bethesda, Maryland.

Noble, R. L. 2002. Reflections on 25 years of progress in black bass management. Pages 419–431 *in* D. P. Philipp and M. S. Ridgway, editors. Black bass: ecology, conservation, and management. American Fisheries Society, Symposium 31, Bethesda, Maryland.

Pegg, M. A., and M. A. McClelland. 2004. Assessment of spatial and temporal fish community patterns in the Illinois River. Ecology of Freshwater Fish 13:125–135.

Pollock, K. H., C. M. Jones, and T. L. Brown. 1994. Angler survey methods and their applications in fisheries management. American Fisheries Society, Special Publication 25, Bethesda, Maryland.

Robson, D. S., and H. A. Regier. 1964. Sample size in Petersen mark–recapture experiments. Transactions of the American Fisheries Society 93:215–226.

Roedel, P. M., editor. 1975. Optimal sustainable yield as a concept in fisheries management. American Fisheries Society, Special Publication 9, Bethesda, Maryland.

Shannon, C. E., and W. Weaver. 1949. The mathematical theory of communication. University of Illinois Press, Urbana.

Stewart-Oaten, A., and J. R. Bence. 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs 71:305–339.

Swingle, H. S. 1950. Relationships and dynamics of balanced and unbalanced fish populations. Agricultural Experiment Station Bulletin 274, Auburn University, Auburn, Alabama.

Thomas, L. 1999. Monitoring long-term population change: why are there so many analysis methods? Ecology 77:49–58.

Underwood, A. J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications 41:3–15.

U.S. Geological Survey. 1999. Ecological status and trends of the Upper Mississippi River System 1998: a report of the Long Term Resource Monitoring Program. U.S. Geological Survey, Upper Midwest Environmental Sciences Center, LTRMP 99-T001, La Crosse, Wisconsin.

Walters, C. J. 1986. Adaptive management of renewable resources. Macmillan, New York.

Waters, W. E., and D. C. Erman. 1990. Research methods: concept and design. Pages 1–34 *in* C. B. Schreck and P. B. Moyle, editors. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Weithman, A. S., J. B. Reynolds, and D. E. Simpson. 1980. Assessment of structure of largemouth bass stocks by sequential sampling. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 33(1979):415–424.

Wilde, G. R., and W. L. Fisher. 1996. Reservoir fisheries sampling and experimental design. Pages 397–409 *in* L. E. Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society, Symposium 16, Bethesda, Maryland.

Willis, D. W., and B. R. Murphy. 1996. Planning for sampling. Pages 1–15 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

3 Sampling and Experimental Design

Michael J. Hansen, T. Douglas Beard Jr., and Daniel B. Hayes

3.1 INTRODUCTION

Even though many different methods are used to sample fish populations, their habitats, and anglers, sampling plans often share common traits. The appropriate gear and protocol to collect data are important, but the sampling design and the characteristics of the population determine the statistical properties of the estimates obtained. Similarly, when experiments are conducted to determine the response of fishes or their habitats to treatment, the overall experimental design and underlying variability determine the power of the experiment and can limit the questions or hypotheses that can be addressed. The goal of this chapter is to describe some of the most common sampling and experimental designs used in fisheries science. Our principal intent is not to teach the theory underlying these topics but rather to illustrate common data analysis approaches based on that statistical theory.

3.1.1 Populations and Samples

Fisheries scientists take samples from populations because data or information from all individuals in the population typically cannot be obtained. Fundamental to the idea of sampling is that a population of sampling units exists from which samples are taken. Ideally, all sampling units in the population can be sampled, but in many field sampling programs the sample frame, or the set of sample units that are actually available to be sampled, may be only a subset of the entire target population. In general, whenever the sample units in the sample frame differ from the units in the target population, the design may provide results that reflect the sample frame but not the target population (termed bias; see section 3.1.2). The degree of bias due to this mismatch is generally case specific and is virtually impossible to determine. Throughout this chapter, we assume that the sample units in the sampling frame match the units in the target population and that all units are sampled with equal efficiency.

Definition of the sampling unit is not always straightforward and often depends on the objectives of the study. For example, individual fish are sampling units in a

telemetry study of fish home range if the investigator wishes to know how individual fish in a single population use available habitat. If the telemetry study is conducted in several lakes, each lake may be viewed as a sampling unit, with individual fish as secondary sampling units. In both of these cases, the sampling units are naturally defined units. In contrast, consider a situation where sampling units are defined as possible seining sites (Figure 3.1). Seining-site boundaries are defined by the investigator, not by natural boundaries. The critical concept underlying this example is that after the size of a seine site is defined, and an arbitrary starting point is determined, a finite population of nonoverlapping sampling units is defined. This example also illustrates a case in which some sample units are not part of the sample frame because they cannot be sampled with the gear used. Whenever some sites that are part of the target population are not part of the

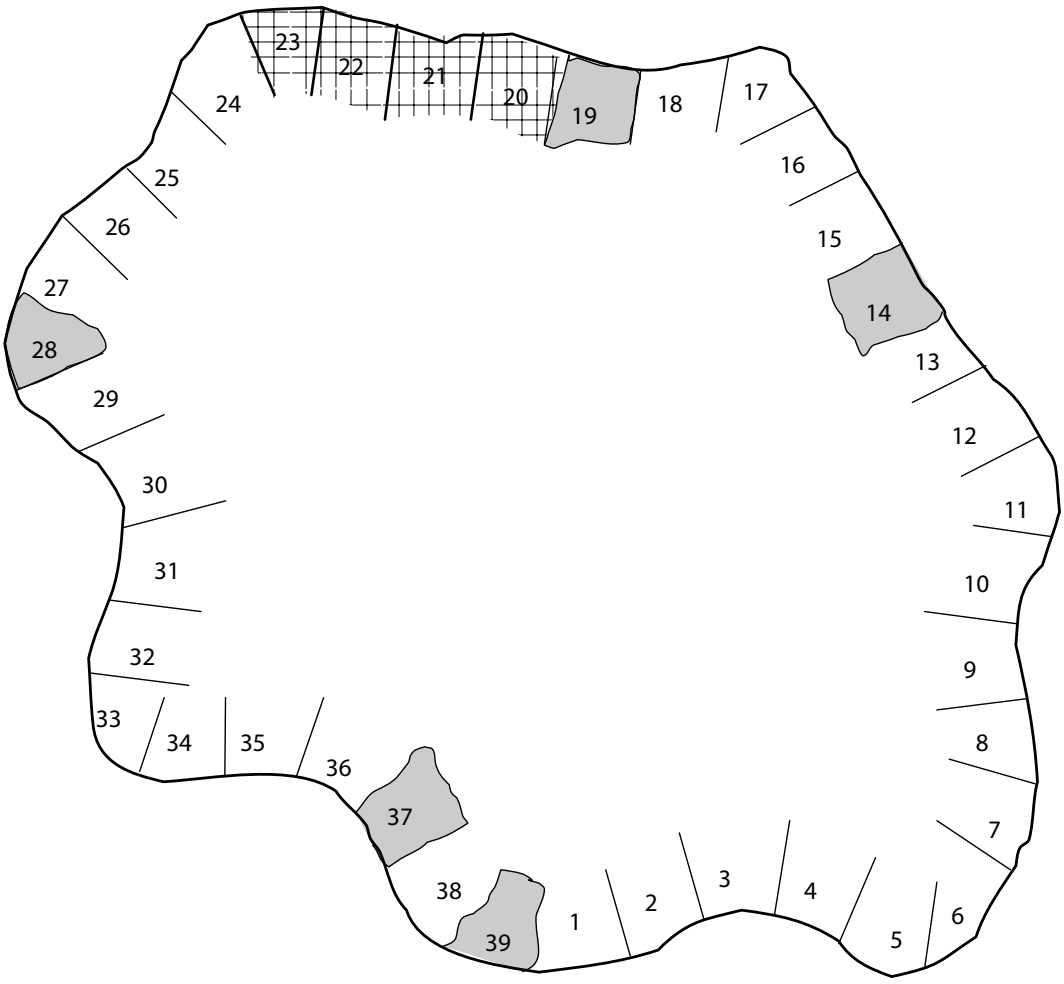


Figure 3.1 Example of a population of seining sites in a lake. Sites selected for sampling are shaded in gray, whereas sites that could not be sampled because of obstructions or soft bottom are crosshatched and were not considered part of the sample frame.

sampling frame, the attribute being estimated will reflect the sample frame but not the target population, so results of sampling will be biased in relation to the true value for the target population.

3.1.2 Bias

The goal of a sampling program is to provide estimates about the characteristics of a population. When used in the context of statistical sampling, bias is generally defined as the difference between the true value of the population attribute and the expected value (i.e., the mean across all possible samples) of the estimator (Cochran 1977). As indicated earlier, a significant source of bias can occur when units in the sample frame differ from units in the entire population or when sampling units within the population are not all sampled with equal efficiency. Therefore, the investigator must define the target population in a way that reflects this mismatch or consider using gears and protocols that produce samples that more accurately characterize the target population. Appropriate sampling gear and protocols are prerequisite for applying the methods covered in this chapter.

In addition to the potential biases described above, estimators such as ratio estimators (section 3.2.2.3) may also result in biased estimates. Although a biased estimator sounds like something to be avoided, for estimators such as ratio estimators some degree of bias is unavoidable (see section 3.2.2.3). However, these biases differ from biases due to sampling frame problems in that the amount of bias can be estimated (remembering that this is based on an average), and a decision can be made whether the bias is acceptable. In some cases, the amount of bias introduced is negligible and is more than offset by gains in precision. Bias is generally evaluated in combination with precision (described below), and their combination is expressed as mean square error (MSE, in squared units of measure), which is calculated as

$$\mathrm{MSE}(\hat{y}) = \mathrm{bias}^2 + \mathrm{var}(\hat{y}) . \tag{3.1}$$

The word accurate is often used as a synonym for unbiased. In common use, however, the word accurate is often used to convey more than simply being unbiased but is used to mean correct. Therefore, we discourage the use of the term accurate in discussions of the statistical properties of sampling programs.

3.1.3 Precision and Confidence Intervals

In addition to obtaining a point estimate of some characteristic of a population, the degree of confidence we have in that estimate is also important to determine. Unless we sample the entire population, our point estimate is unlikely to match the true population value exactly. Thus, a critical concept is that the precision of an estimate is a measure of how likely it is that our estimate is close to the unknown true value. The precision of estimates is often expressed as the standard error (SE) of the estimate or as a confidence interval (CI) around the estimate.

When estimates are viewed as one possible outcome of many possibilities, such that repeating the same procedure would likely result in a different sample being taken with a different point estimate, the estimates can be treated as coming from a statistical distribution, and the SE is simply the square root of the variance of that distribution. For a normal distribution, 68% of the estimates would fall between one SE on either side of the true mean. Similarly, a CI can be thought of as a range within which most, commonly 95%, of the estimates would be expected to lie.

Assuming an unbiased estimator, the precision of estimates is affected by the inherent variability of the attribute being measured and the number of observations of the attribute that are obtained during sampling, in addition to the sampling design (covered later). In general, to obtain estimates of a given level of precision, more samples are required for population attributes that are highly variable than for population attributes that are relatively invariable. Before sampling, the investigator should decide what level of precision is acceptable and then determine how many samples are needed based on prior knowledge of the level of variability that is expected for the attribute of interest, such as would be obtained from a preliminary survey. The level of precision that is acceptable is often determined from the practical needs of the investigator or agency. For example, a fishery scientist may wish to know the mean length of walleyes in a particular lake following a new regulation and would like to have 95% CIs around the mean that were less than ± 50 mm. Here, the acceptable level of precision (±50 mm) could be set to exceed the level of interannual variation that would occur in the absence of a regulation change. Alternatively, the acceptable level of precision could be set at some arbitrary level that is deemed by the agency or investigator to be acceptable.

3.1.4 Random versus Nonrandom Sampling

Generally, samples must be drawn randomly from the population of interest to ensure the sample is representative of the entire population. When samples are drawn nonrandomly or using subjective criteria, measured attributes will usually be biased, though the degree of bias cannot be determined from the samples. For example, sampling is sometimes focused in areas where fish are known to aggregate to avoid spending sampling effort in areas where fish are known to be scarce. Unfortunately, attributes (e.g., length) of fish in areas where they aggregate may differ from attributes of fish in areas where fish are scarce. Thus, such a sample may not be representative of the entire population. The only way to avoid such potential bias is to sample units randomly from the population.

3.2 SAMPLING DESIGNS

3.2.1 Overview of Sampling Designs

In this section, we discuss four commonly used sampling designs, for which both a naturally defined sampling unit (individual fish) and an artificially defined sampling unit (sampling grid locations) can be used (Figure 3.2). Although other

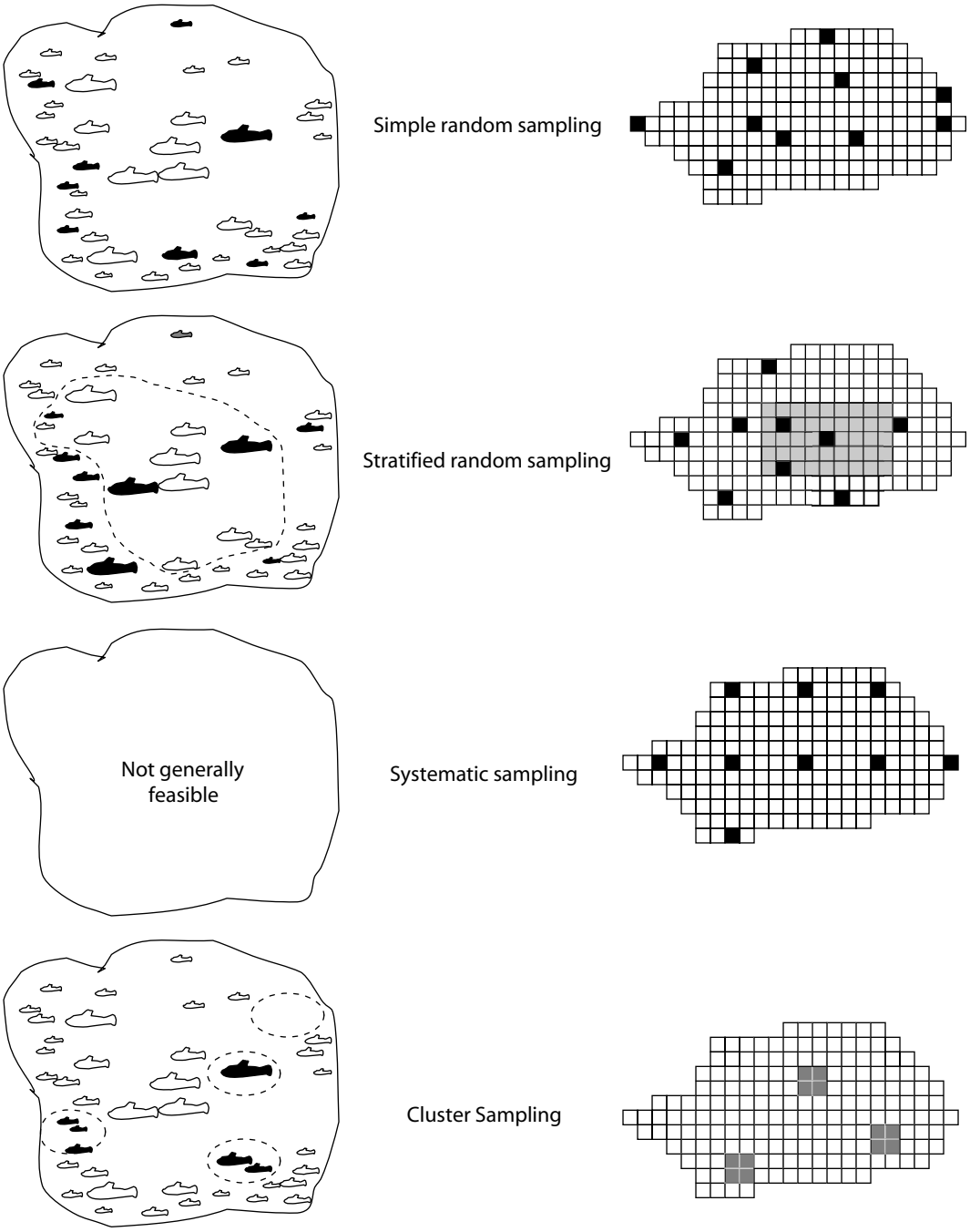


Figure 3.2 Sample frames and sampling unit selection for basic sampling designs. The column on the left illustrates the case in which the sampling unit is individual fish or groups of fish for cluster sampling. The column on the right illustrates the case in which the sampling unit is a grid location in a lake or a group of grids for cluster sampling. Units shaded in black are selected for sampling. The dotted line or shaded region in stratified random sampling indicates stratum boundaries.

sampling designs are available (e.g., Cochran 1977; Thompson 1992; Lohr 1999), we will cover the basic designs that are applicable in most situations. The critical consideration distinguishing each sampling design is how sampling units in the population are defined and how they are selected. In simple random sampling, each sampling unit in the population has an equal probability of being included in the sample, and each sampling unit is selected independently of other units (section 3.2.2). In stratified random sampling, each sampling unit in the population is first assigned to a stratum, and then a simple random sample is independently drawn from each stratum (section 3.2.3). In cluster sampling, the population is divided into primary sampling units, termed clusters, and secondary sampling units, termed elements (section 3.2.4). In systematic sampling, all the sampling units in the population are arranged in a sequence, and then from a random starting point, every kth sampling unit is included in the sample (section 3.2.5). In systematic sampling, selection of the first sampling unit determines all other units in the sample, so sampling units are not independently selected. In large populations, lack of independence does not generally lead to biased estimates of the mean but typically results in a biased estimate of the variance if dependency among sample units is not taken into account. Each of these sampling designs is described in more detail below. To present computational methods concisely and facilitate comparisons among sampling designs, we present the basic formulae for estimating the mean in Table 3.1, which are summarized from Cochran (1977), Thompson (1992), and Lohr (1999).

3.2.2 Simple Random Sampling

In simple random sampling, a sample of size n is randomly selected from a population with N sampling units. Implementing simple random sampling is easiest when all of the sampling units can be enumerated before sampling begins, as in the example of seining sites within a lake (Figure 3.1) or in the example of sampling grids within a lake (Figure 3.2). Sample units are often selected for sampling without replacement (selecting each sample unit no more than once) by using a random number table or generator (Wilde and Fisher 1996). When the sampling frame is unknown before sampling begins, such as in the example of fish within a lake (Figure 3.2), implementing true simple random sampling may be impossible, thereby leading to the use of an alternative design (e.g., cluster sampling, section 3.2.4). Simple random sampling is often less efficient and less precise than are other designs but illustrates concepts and estimators that are inherent in other designs and also provides a basis for understanding the efficiency of other designs, as we will illustrate in the ensuing parts of section 3.2.

3.2.2.1 Estimation of Mean Values

An important property of simple random sampling is that the sample mean and variance provide unbiased estimates of the population mean and variance regardless of the shape of the distribution in the population being sampled (Cochran 1977; Lohr 1999). The sample mean is simply calculated as the sum of

the observations (y_i) divided by the sample size (n; Table 3.1; Box 3.1). Several equivalent formulae are available for calculating the sample variance, but one approach is to sum the squared deviations (differences) between observations of the sampling units (y_i) and the sample mean, and divide by $n - 1$ (Table 3.1; Box 3.1). One can view the sample mean as a simple prediction for each observation, and the sample variance is the average squared deviation between observations and their predicted value. In most fisheries applications, the proportion of the population that is sampled (also known as the sampling fraction, n/N) is small, and the SE of the sample mean is estimated by taking the square root of the sample variance, divided by the sample size (Table 3.1). When the sampling fraction is large (e.g., $n/N > 0.5$), the confidence in estimates of the sample mean is increased, and consequently the SE is reduced:

$$\mathrm{SE}(\bar{y}) = \sqrt{\left(1 - \frac{n}{N}\right)\frac{s^2}{n}}\,. \tag{3.2}$$

The term $(1 - n/N)$ is called the finite population correction. As the sampling fraction approaches 1 (the entire population is sampled), the finite population correction approaches 0, and the SE also approaches 0. In a census of the entire population, the SE of the sample mean would be 0 because all possible sampling units in the population would be included in the sample. Many books on sampling theory include formulae and derivations that include the finite population correction, but it is typically negligible in practice. Therefore, our summary of formulae (Table 3.1) excludes the finite population correction factor.

The SE is often used as a measure of precision of estimates and describes the variability that would be expected if the sampling process could be repeated a large number of times. For a normal distribution, approximately 68% of the distribution is found between 1 SE above and below the mean. An alternative method of conveying the precision is to estimate confidence limits (CLs) on point estimates. Estimating CLs on the mean requires knowledge of the distribution of the mean or assumptions about the shape of the distribution. For large samples (e.g., $n \geq 50$; Zar 1999), the distribution of the mean approaches a Student's t-distribution with $n - 1$ df, so CLs can be estimated as

$$\begin{aligned} &\text{Lower CL} = \bar{y} - t_{\alpha,\, n-1}\mathrm{SE}(\bar{y}), \text{ and} \\ &\text{Upper CL} = \bar{y} + t_{\alpha,\, n-1}\mathrm{SE}(\bar{y}), \end{aligned} \tag{3.3}$$

where $t_{\alpha,\, n-1}$ is the value of the t-distribution (commonly available from a table in a statistics book or equation in a spreadsheet) for an α equal to the probability of making a type I error (often 0.05) and an n of a given sample size. The $\mathrm{SE}(\bar{y})$ is the SE of the mean (see Box 3.1 for an example). The t-distribution is often used to estimate approximate confidence limits for small samples, but the bootstrap method can also be used to estimate CIs and is often recommended for small sample sizes (see Efron and Tibshirami [1998] for details on this approach).

Table 3.1 Summary of formulae for computing the mean and associated measures of precision for common sampling designs. Note that these formulae do not include the finite population correction factor ($1 - n/N$; where n/N is the proportion of the population sampled). Variable definitions are as follows. For simple random sampling, y_i = measurement of item i and n = sample size. For stratified random sampling, y_{ih} = measurement of item i in stratum h; h = stratum index; N_h = number of elements in stratum h; n_h = units sampled from within each stratum; L = total number of strata; and N = number of elements in population. For cluster sampling, y_{ij} = measurement of element j in cluster i; n = number of primary units or clusters sampled; m_i = number of secondary elements i sampled; and M_i = number of secondary elements in cluster i. For regression, or double sampling, $\bar{x}$ = the mean of estimates for the subsample, $\bar{y}$ = the mean of measured values for the subsample, and $\bar{X}$ = the mean of estimates for the entire sample.

	Mean	Variance	SE (mean)
Simple random sampling	$\bar{y} = \frac{\sum y_i}{n}$	$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1}$	$\mathrm{SE}(\bar{y}) = \sqrt{\frac{s^2}{n}}$
Stratified random sampling	$\bar{y} = \sum_{h=1}^{L} \frac{N_h}{N} \bar{y}_h$ $= \sum_{h=1}^{L} W_h \bar{y}_h$	$s_h^2 = \frac{\sum (y_{ih} - \bar{y}_h)^2}{n_h - 1}$	$\mathrm{SE}(\bar{y}) = \sqrt{\sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h}}$
Single-stage cluster sampling	cluster total $= y_i = \sum y_{ij}$	Can treat (clusters) as simple random sample	Treat clusters as simple random sample of size n
	mean cluster total $= \bar{y} = \frac{\sum y_i}{n}$		
	mean per secondary unit $= \frac{\sum y_i}{\sum M_i} = \hat{R}$	See ratio estimators	$\mathrm{SE}(\hat{R}) = \frac{1}{\sqrt{n}\,\bar{M}} \sqrt{\frac{\sum (y_i - \hat{R} M_i)^2}{n-1}}$

Two-stage cluster sampling	Estimated cluster total $\hat{y}_i = \frac{M_i}{m_i}\sum_{j=1}^{m} y_{ij}$ Mean cluster total $\bar{y} = \sum_{i=1}^{n} \frac{\hat{y}_i}{n}$ Mean per secondary unit within cluster $\bar{y}_i = \sum_{i=1}^{n} \frac{y_{ji}}{m_i}$ Estimated population mean per secondary unit $\bar{\bar{y}} = \frac{\sum \hat{y}_i}{\sum M_i}$	$s_1^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$ $s_2^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij} - \bar{y}_i)^2}{n(m-1)}$	$\mathrm{SE}(\bar{y}) = \sqrt{\frac{s_1^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n(n-1)}}$ $\mathrm{SE}(\bar{\bar{y}}) = \sqrt{\frac{\sum_{i=1}^{n} M_i^2(\bar{y}_i - \bar{\bar{y}})^2}{n\bar{M}^2(n-1)}}$
Systematic sampling (one start point)	$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$	Cannot be directly estimated	Cannot be directly estimated
Systematic sampling (two start points)	$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$	Can treat data associated with each start point as a cluster	See single-stage cluster sampling
Regression or double sampling	$\bar{y}_{\mathrm{reg}} = \bar{y} + b(\bar{X} - \bar{x})$		$\mathrm{SE}(\bar{y}_{\mathrm{reg}}) = \sqrt{\frac{1}{n(n-2)}\left\{\sum (y_i - \bar{y})^2 - \frac{[\sum (y_i - \bar{y})(x_i - \bar{x})]^2}{\sum (x_i - \bar{x})^2}\right\}}$

Box 3.1 Example of Estimating the Mean Based on Simple Random Sampling

Fifteen sites were randomly selected from an *X* – *Y* grid superimposed on a shallow lake. At each site, the catch of central mudminnow in a throw trap (assumed to be equally efficient at all sites in the lake) was recorded. The goal of the sampling was to determine the mean density of central mudminnows in the lake.

Table Catch of central mudminnow from 15 lake sites randomly selected on an *X* – *Y* grid.

Coordinate and total			
X	*Y*	Catch	(Catch – mean)^2
5	18	3	1.96
15	16	0	2.56
4	9	4	5.76
14	3	1	0.36
11	8	4	5.76
12	5	1	0.36
2	4	1	0.36
3	20	0	2.56
11	7	1	0.36
1	8	0	2.56
2	15	2	0.16
11	2	2	0.16
3	17	3	1.96
3	12	2	0.16
1	10	0	2.56
Total		24	27.60

3.2.2.2 Estimation of Proportions

Many characteristics of fishes or their habitats cannot be expressed quantitatively as a continuous variable but can be expressed qualitatively as a categorical variable. For example, the sex of a fish is a qualitative (categorical) trait. Sampling is often undertaken to estimate the proportion of the population (p) that possess some quality or attribute. An attribute that takes on one of only two values (e.g., male or female or mature or immature) is a single or binary classification system, whereas an attribute that falls into one of several categories or classes (e.g., species of fish or length intervals) is a multiple classification system.

For single classification variables, the observation (y_i) is coded as 1 if the individual possesses one attribute or trait and 0 if it possesses the other attribute or trait. The proportion of individuals that possess the trait in the population (p) can

The mean catch and associated measures of precision are calculated as follows, using formulae from Table 3.1.

$$\bar{y} = \frac{\sum y_i}{n} = \frac{24}{15} = 1.6.$$

$$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{27.6}{15-1} = 1.97.$$

$$\mathrm{SE}(\bar{y}) = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{1.97}{15}} = 0.36.$$

Although the sample size is not large (<30), a normal approximation can be used to estimate approximate 95% confidence limits (CLs).

$$\text{Lower CL} = \bar{y} - (t_{\alpha, n-1})(\mathrm{SE}) = 1.60 - 2.145 \cdot 0.36 = 0.83.$$
$$\text{Upper CL} = \bar{y} + (t_{\alpha, n-1})(\mathrm{SE}) = 1.60 + 2.145 \cdot 0.36 = 2.37.$$

This example can also be used to illustrate how to compute estimates of target sample sizes (equation [3.9]). For example, if we wanted to compute the mean catch with a SE of 0.10, we would start by guessing a sample size of 60 might be adequate. Using this preliminary guess (which is needed to get an initial estimate of the *t*-statistic used in the formula for sample size), we would estimate that the necessary sample size, to the nearest integer, would be

$$n = \frac{(1.97)(2.0)^2}{0.10} \approx 79.$$

Even though this is different than our initial guess of 60, the actual *t*-statistic for 79 is 1.990, which would change our integer estimate of the necessary sample size to only 78.

be estimated by summing the y_i and dividing by n (Table 3.2; Box 3.2; Cochran 1977). The proportion of individuals lacking the trait (q) is termed the complement of p, and is computed as $q = 1 - p$. As with estimates of the mean for quantitative measurements, the estimate of p is unbiased in simple random sampling, as long as the attribute is identified correctly for each individual examined. Although the estimator for the SE of p (Table 3.2; Box 3.2; Cochran 1977) is unbiased, when p is close to 0 or 1 the distribution around p is skewed because p cannot be less than 0 or greater than 1. Therefore, we recommend that CLs for p be estimated from the *F*-distribution. The lower CL (L_1) for p is (Zar 1999)

$$L_1 = \frac{a}{a + (\mathrm{n} - a + 1)F_{\alpha, v_1, v_2}}, \tag{3.4}$$

Table 3.2 Summary of formulae for proportions, and their associated measures of precision, for common sampling designs. The proportion of individuals that possess the trait in the population is given by p and the proportion of individuals lacking the trait, q, is computed as $q = 1 - p$. Note that these formulae do not include the finite population correction factor. Variable definitions for simple random sampling are y_i = observation on unit coded 0 if not in class C and 1 if in class C; a = number of units in class C in sample; n = sample size; and m_i = number of units in cluster i. See Table 3.1 for additional variable definitions.

Sampling design	Proportion in class C	SE(p)
Simple random sampling	$\hat{p} = \frac{\sum y_i}{n} = \frac{a}{n}$	$\mathrm{SE}(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n-1}}$
Stratified random sampling	$\hat{p}_h = \sum_{h=1}^{H} \frac{y_{ih}}{n_h}$ $\hat{p} = \sum_{h=1}^{H} \frac{N_h \hat{p}_h}{N}$ $= W_h \hat{p}_h$	$\mathrm{SE}(p_{st}) = \sqrt{\sum_{h=1}^{L} \frac{W_h^2 \hat{p}_h \hat{q}_h}{(n_h - 1)}}$
Single-stage cluster sampling	$\hat{p}_i = \frac{\sum y_i}{n} = \frac{a}{n} = p$ $\hat{p}_i = \frac{\sum a_i}{\sum m_i}$	$\mathrm{SE}(\hat{p}) = \sqrt{\frac{1}{n\bar{m}^2} \frac{\sum a_i^2 - 2\hat{p}\sum a_i m_i - \hat{p}^2 \sum m_i}{n-1}}$

Box 3.2 Example of Estimating a Proportion in Simple Random Sampling

One hundred sixteen brown trout were collected at random from a population in a stream with the goal of estimating the proportion in each age-group. The age of each fish was estimated from scales to produce the following data.

Table Age distribution of a random sample of 116 brown trout from a stream.

Age	*n*
0	55
1	22
2	10
3	18
4	6
5	3
6	1
7	1

The proportion in each age-class was estimated as follows.

$$\hat{p}_0 = \frac{\sum y_i}{n} = \frac{55}{116};\ \hat{p}_1 = \frac{22}{116};\ \hat{p}_2 = \frac{10}{116};\ \hat{p}_3 = \frac{18}{116};\ \hat{p}_4 = \frac{6}{116};\ \hat{p}_5 = \frac{3}{116};\ \hat{p}_6 = \frac{1}{116};\ \text{and}\ \hat{p}_7 = \frac{1}{116}.$$

For example, for age 0 the SE (Table 3.2) and CLs (equations [3.4] and [3.5]) were calculated as follows.

$$\text{SE}(\hat{p}_0) = \sqrt{\frac{\hat{p}_0\hat{q}_0}{n-1}} = \sqrt{\frac{0.47 \cdot (1-0.47)}{116-1}} = 0.047.$$

$$\text{Lower CL} = \frac{a}{a + (n - a + 1)F_{\alpha, v_1, v_2}} = \frac{55}{55 + (116 - 55 + 1)F_{0.05,\, 2\cdot(116-55+1),\, 2\cdot 55}} = 0.3948.$$

$$\text{Upper CL} = \frac{(a+1)F_{\alpha, v_1, v_2}}{n - a + (a+1)F_{\alpha, v_1, v_2}} = \frac{(55+1)F_{0.05,\,(2\cdot 55)+2,\,2\cdot(116-55+1)-2}}{116 - 55 + (55+1)F_{0.05,\,(2\cdot 55)+2,\,2\cdot(116-55+1)-2}} = 0.5545.$$

Estimates of the proportion in each age-class, and appropriate measures of precision, are given in the table below.

Table Estimates of the proportion of brown trout in each age-class (p_{age}) and measures of precision.

Age	p_{age}	SE	Lower CL	Upper CL
0	0.47	0.047	0.3948	0.5545
1	0.19	0.037	0.1320	0.2596
2	0.09	0.026	0.0475	0.1418
3	0.16	0.034	0.1028	0.2214
4	0.05	0.021	0.0228	0.0995
5	0.03	0.015	0.0071	0.0655
6	0.01	0.009	0.0004	0.0402
7	0.01	0.009	0.0004	0.0402

where $a = \Sigma y_i$; n = sample size; and the F-statistic is evaluated for the two-tailed level of α, the numerator df $v_1 = 2(n - a + 1)$, and the denominator df $v_2 = 2a$. Similarly, the upper CL (L_2) for p is (Zar 1999)

$$L_2 = \frac{(a+1)F_{\alpha,\, v_1',\, v_2'}}{n - a + (a+1)F_{\alpha,\, v_1',\, v_2'}}, \tag{3.5}$$

where the F-statistic is evaluated for the two-tailed level of α, the numerator df $v_1' = v_2 + 2$, and the denominator df $v_2' = v_1 - 2$ (Box 3.2). The upper and lower CLs for q are obtained by subtracting the upper and lower CLs for p from 1.

For attributes in a multiple classification system, the problem can be simplified by focusing on one class at a time and treating the attribute as a single classification variable where the individual either has the attribute or not. The proportion within any single class and the associated SE is then estimated exactly as for the single classification situation. Box 3.2 illustrates how to calculate proportions, SEs, and CLs for a multiple classification system, where any one class of the multiple classification system can be used to illustrate a single classification variable.

3.2.2.3 *Estimation of Ratios*

Attributes of fishes or habitats are often expressed as ratios of variables that both vary among units, which contrasts with proportions that describe the fraction of a sample that possess a certain attribute, as in section 3.2.2.2. A familiar example is angler catch per effort where both catch and effort vary among individual anglers. Unfortunately, situations in which ratios are estimated are often confused with situations in which a proportion is being estimated. For example, in diet studies, the amount of food consumed, by weight, among various prey taxa is commonly referred to as a proportion but is more appropriate to view as a ratio of the weight consumed of each prey taxon to the total weight consumed. These types of data are best treated as a ratio because the weight consumed of each prey taxon varies among sampling units and the total weight consumed varies among sampling units.

In simple random sampling, the population ratio (R) is estimated from the ratio of the sums of the sampled quantities (Box 3.3; Cochran 1977; Lohr 1999):

$$\hat{R} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i}. \tag{3.6}$$

For catch per effort data, the numerator of equation (3.6) is the sum of catch and the denominator is the sum of effort. For diet data, the numerator is the total weight of one prey taxon and the denominator is the total weight of all prey taxa (as in Box 3.3). This estimator is biased, but the bias tends toward 0 as the sample size increases (Cochran 1977). Cochran (1977) showed that the degree of bias relative to the SE of the estimated ratio can be approximated as

Box 3.3 Example of Estimating a Ratio in Simple Random Sampling

Twenty yellow perch were randomly sampled from a lake, and weights of zooplankton, benthos, and fish in each yellow perch stomach were measured. The goal was to determine the ratio of each prey category to total weight of prey in the diet of the yellow perch population.

Table Ratio of three prey categories to total weight of prey in the diets of 20 yellow perch. Squared deviations of the observed (y) minus the predicted ($\hat{R}x$, where x is the total weight for each fish and $\hat{R}$ is estimated as shown below the table) allows for estimation of population ratios for each prey category (sums of sampled quantities $[y - Rx]^2$).

	Weight in stomach				$(y - Rx)^2$		
Fish and total	Zoo-plankton	Benthos	Fish	Total Weight	Zoo-plankton	Benthos	Fish
1	0.000	0.000	16.217	16.217	4.257	36.537	65.735
2	0.200	2.501	0.000	2.701	0.021	2.233	1.824
3	0.593	0.054	0.000	0.647	0.261	0.035	0.105
4	0.356	0.741	0.000	1.097	0.047	0.110	0.301
5	0.070	1.112	0.000	1.182	0.006	0.450	0.349
6	0.191	1.734	0.000	1.925	0.003	1.033	0.926
7	0.012	0.022	0.000	0.034	0.000	0.000	0.000
8	0.017	2.822	0.000	2.839	0.119	3.111	2.015
9	0.400	2.796	0.000	3.196	0.000	2.575	2.554
10	0.202	2.154	0.000	2.356	0.010	1.627	1.388
11	0.591	0.559	0.000	1.150	0.198	0.017	0.331
12	0.737	0.902	0.000	1.639	0.280	0.085	0.672
13	0.095	0.098	0.000	0.193	0.005	0.001	0.009
14	0.000	0.000	12.090	12.090	2.366	20.306	36.534
15	0.747	1.913	0.000	2.660	0.167	0.849	1.769
16	0.663	0.600	0.000	1.263	0.253	0.017	0.399
17	0.937	0.354	0.000	1.291	0.598	0.016	0.417
18	0.664	0.213	0.000	0.876	0.305	0.013	0.192
19	0.623	0.448	0.000	1.072	0.237	0.002	0.287
20	0.103	2.077	0.000	2.181	0.030	1.600	1.189
Total	7.202	21.099	28.307	56.608	9.160	70.617	116.997

The ratio (equation [3.6]) and SE (equation [3.8]) of each prey category in the diet is estimated as follows.

$$\hat{R}_{\text{zooplankton}} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{7.202}{56.608} = 0.127.$$

$$\text{SE}(\hat{R}_{\text{zooplankton}}) = \frac{1}{\sqrt{n}\,\bar{x}} \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{R}x_i)^2}{n-1}} = \frac{1}{\sqrt{20} \cdot 2.830} \sqrt{\frac{9.160}{20-1}} = 0.055.$$

(Box continues)

Box 3.3 *(continued)*

$$\hat{R}_{\text{benthos}} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{21.099}{56.608} = 0.373.$$

$$\text{SE}(\hat{R}_{\text{benthos}}) = \frac{1}{\sqrt{n}\,\bar{x}} \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{R}x_i)^2}{n-1}} = \frac{1}{\sqrt{20} \cdot 2.830} \sqrt{\frac{70.617}{20-1}} = 0.152.$$

$$\hat{R}_{\text{fish}} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{28.307}{56.608} = 0.500.$$

$$\text{SE}(\hat{R}_{\text{fish}}) = \frac{1}{\sqrt{n}\,\bar{x}} \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{R}x_i)^2}{n-1}} = \frac{1}{\sqrt{20} \cdot 2.830} \sqrt{\frac{116.997}{20-1}} = 0.196.$$

$$\frac{\text{bias}(\hat{R})}{\text{SE}(\hat{R})} = \frac{\text{SE}(\bar{x})}{\bar{X}}, \tag{3.7}$$

where $\bar{X}$, the population mean, generally can be estimated without bias using the sample mean, $\bar{x}$ (see Cochran 1977). Because the SE of x decreases to 0 as sample size increases, the degree of bias also decreases as sample size increases. This expression can be easily computed from sample data to determine if bias is large enough to be problematic. Importantly, a ratio should not be estimated by averaging the ratios for individuals (often termed a mean of ratios) but rather as a ratio of totals (often termed a ratio of means), because a mean of ratios has a larger degree of bias than does a ratio of means, and this bias does not diminish as sample size increases (Cochran 1977). The SE of a ratio is derived from deviations between the numerator of the ratio (y_is) and the product of the denominator of the ratio (x_is) and the ratio ($\hat{R}$) (Box 3.3; Cochran 1977; Lohr 1999):

$$\text{SE}(\hat{R}) = \frac{1}{\sqrt{n}\,\bar{x}} \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{R}x_i)^2}{n-1}}. \tag{3.8}$$

Confidence limits around the estimated ratio are typically approximated using the *t*-distribution, which is reasonable for large sample sizes but may not represent skewness in the distribution of the estimate for small sample sizes.

3.2.2.4 *Estimation of Sample Size*

The sample size required to estimate the population mean (for example) can be derived for simple random sampling from knowledge of the variance of the y_i in the population when the desired degree of precision is specified. Various strategies have been developed to simplify the process of estimating sample size requirements (Wilde and Fisher 1996), but ultimately the investigator must specify expectations for the outcome of sampling and acceptable levels of precision in estimates. Precision can be expressed on an absolute scale (e.g., ± 10 mm) or on a relative scale (e.g., ± 8% of the mean). When desired precision is expressed in absolute terms (δ), the sample size needed (n; ignoring the finite population correction factor) can be estimated as (Cochran 1977):

$$n = \frac{s^2 t^2_{n-1}}{\delta^2}, \tag{3.9}$$

where s^2 is estimated from a pilot study or prior experience with similar situations and the t-statistic is defined for a given α level from a statistical table or spreadsheet function. Because t_{n-1} depends on the sample size, the estimate of n must be solved by trial and error. In practice, the first guess does not need to be close to the true value, and only two or three iterations are necessary to obtain the solution because the t-statistic does not vary greatly with n.

When the desired precision is expressed in relative terms (r), an estimate of the mean must also be included. As above, a preliminary estimate of the mean can be obtained by a pilot study or prior experience, and sample size can be estimated as

$$n = \left(\frac{t_{n-1} s}{r\bar{y}}\right)^2. \tag{3.10}$$

The sample size that is necessary for specified precision for proportions is analogous to that for a mean value (Cochran 1977):

$$n = \frac{s^2 t^2_{n-1}}{\delta^2}, \tag{3.11}$$

where the variance is estimated as the product of p and q and the investigator specifies the absolute error (δ).

3.2.3 Stratified Random Sampling

Stratified random sampling performs as well or better than simple random sampling in nearly all cases and results in substantial improvement in precision when variation within the strata is less than variation among the strata. In stratified random sampling, the total sample frame containing N sample units is divided

into L subpopulations or strata, each containing N_h sample units. Within each stratum, a simple random sample of n_h sample units is drawn independently. For example, in Figure 3.2 a lake with N sample grids is divided into two depth strata (L = deep and shallow) with N_h sample grids in each stratum, and n_h sample grids are then sampled from within each stratum.

An estimate for the whole population is obtained by weighting estimates from each stratum $\bar{y}_h$ by the fraction of the whole population contained in each stratum ($W_h = N_h/N$). Stratified random sampling is advantageous over simple random sampling because sampling can be allocated disproportionately among strata to ensure adequate precision can be obtained for subpopulations represented by strata. Stratified random sampling requires that the entire sampling frame be divided into strata before sampling begins, so it should not be applied to situations where the strata are defined a posteriori.

3.2.3.1 *Construction of Strata*

To be most efficient, the strata means should differ widely from one another, so that variability between strata is large and variability within each stratum is small. However, the data necessary to specify strata that best partition the variability in the population would require the investigator to complete the survey. Consequently, other features that are readily obtainable and are correlated to the attribute of interest are often used to construct strata. For example, when sampling fish, we often assume that fish associate themselves with habitat conditions (such as water depth), and we construct strata that coincide with habitat boundaries (as in Figure 3.2). Therefore, prior information about the attribute of interest can be used to construct strata whenever available.

How many strata to develop is also a difficult question to answer. The number of strata that can be sampled is obviously limited by the sample size ($L < n$), but a minimum of at least two sample units must be sampled per stratum to allow calculation of the within-stratum variance. In our experience, the number of strata should depend on the quality and amount of available information, so you should use few strata when prior information is not available and more strata when better prior information is available. The sample size within each stratum should be large enough (e.g., at least 10) to provide reasonable estimates for each subpopulation.

3.2.3.2 *Estimation of Mean Values*

The mean value for a stratified random sample is estimated from the mean values within the strata, weighted by the fraction of the entire population of sample units in each stratum (Table 3.1; Box 3.4; Cochran 1977). Stratum means are estimated as described for a simple random sample (section 3.2.2.1), and each stratum mean is weighted (W_h) by the number of units in the stratum sample (N_h) divided by the total number of units in the population (N). The SE of the mean value for a stratified random sample is a weighted sum of the variances of the mean values for the individual strata (Table 3.1; Box 3.4; Cochran 1977). As in simple random sampling, estimates of the stratified mean and its SE are unbiased. Likewise, estimating CLs on the mean requires an assumption regarding the

Box 3.4 Example of Stratified Random Sampling

A grid was superimposed on the map of a shallow lake, and all grid cells were classified as being in one of three depth strata (0–2 m, 2–4 m, >4 m). Ten grid cells were sampled in each depth stratum, and at each site the catch of age-0 yellow perch in a throw trap (assumed to be equally efficient at all sites in the lake) was recorded. The goal of the sampling program was to estimate the mean density of age-0 yellow perch.

Table Catch of age-0 yellow perch at three depth strata within a shallow lake. Variance in parentheses below mean.

0–2-m stratum		2–4-m stratum		>4-m stratum	
Catch and mean	(Catch – mean)^2	Catch and mean	(Catch – mean)^2	Catch and mean	(Catch – mean)^2
0	2.89	4	1.21	7	1.69
2	0.09	2	0.81	5	0.49
2	0.09	3	0.01	7	1.69
2	0.09	5	4.41	7	1.69
3	1.69	2	0.81	5	0.49
1	0.49	4	1.21	5	0.49
3	1.69	1	3.61	7	1.69
2	0.09	3	0.01	6	0.09
2	0.09	2	0.81	3	7.29
0	2.89	3	0.01	5	0.49
1.7 (1.122)		2.9 (1.433)		5.7 (1.789)	

Within each stratum, the mean catch and variance were computed using formulae for a simple random sample (Table 1; Box 3.1 example). The lake contained 320 grid cells, which included 172 in the 0–2-m stratum, 80 in the 2–4-m stratum, and 68 in the > 4-m stratum, so the weight for each stratum (W_h) was

$$W_h = \frac{N_h}{N}.$$

$$W_{0-2} = \frac{N_{0-2}}{N} = \frac{172}{320} = 0.5375.$$

$$W_{2-4} = \frac{N_{2-4}}{N} = \frac{80}{320} = 0.250.$$

$$W_{>4} = \frac{N_{>4}}{N} = \frac{68}{320} = 0.2125.$$

The stratified mean catch was

$$\bar{y} = \sum_{h=1}^{L} W_h \bar{y}_h = (0.5375 \cdot 1.7) + (0.2500 \cdot 2.9) + (0.2125 \cdot 5.7) = 2.85.$$

The SE of the stratified mean catch per effort was

$$\mathrm{SE}(\bar{y}) = \sqrt{\sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h}} = \sqrt{\left(\frac{0.5375^2 \cdot 1.122}{10}\right) + \left(\frac{0.2500^2 \cdot 1.433}{10}\right) + \left(\frac{0.2125^2 \cdot 1.789}{10}\right)} = 0.222.$$

Approximate 95% confidence intervals can be computed (assuming normality) using the same approach as for simple random sampling (Box 3.1).

sampling distribution of the stratified mean. If the stratum means are normally distributed, the *t*-distribution can be used to estimate CLs. In most situations, the degrees of freedom are calculated as the total sample size minus the number of strata. However, Satterthwaite (1946) showed that the effective number of degrees of freedom should be reduced when allocation of sampling effort is not proportional to the weight for each stratum (when n_h/n is not equal to N_h/N). If the finite population correction term is ignored, the effective number of degrees of freedom can be estimated by (derived from Cochran 1977)

$$df = \frac{\left(\sum \frac{s_h^2}{W_h}\right)^2}{\sum \frac{s_h^4}{W_h(n_h - 1)}}. \tag{3.12}$$

3.2.3.3 *Allocation of Samples within Strata*

An important feature of stratified random sampling is that estimates of the mean are unbiased, regardless of the distribution of the target population and regardless of the sampling effort allocated to each stratum (assuming that at least one sample is taken per stratum). Because of this property, estimates of the stratified mean in different periods are directly comparable if the sampling allocation is altered, or even if the strata boundaries are altered (assuming that the sampling frame remains the same). Sampling effort is often allocated to each stratum proportionally to the weight for each stratum ($n_h = W_h \times n$). Although this generally results in higher precision than simple random sampling, the sampling effort can be allocated to minimize the variance of the resulting estimate. Three general rules have been developed to guide the allocation of sampling effort to minimize the SE of the stratified mean. Using these rules, greater sampling effort should be allocated to strata where (1) the stratum is larger, (2) the stratum has a larger variance, or (3) sampling cost per unit is less expensive in the stratum. If the cost per sample c_h varies among strata, then the optimal allocation of sampling effort is (Cochran 1977; Lohr 1999)

$$n_h = n\left(\frac{\frac{N_h s_h}{\sqrt{c_h}}}{\sum_{h=1}^{H} \frac{N_h s_h}{\sqrt{c_h}}}\right), \tag{3.13}$$

where n is the total sample size, N_h is the total number of units in stratum h, s_h is the standard deviation (SD) in stratum h, and c_h is the cost per sample in stratum h. If the cost per sample is the same in all strata, the optimal allocation, termed the Neyman allocation, is (Cochran 1977)

$$n_h = n\left(\frac{N_h s_h}{\sum_{h=1}^{H} N_h s_h}\right). \tag{3.14}$$

If variances are specified correctly, the Neyman allocation will always give estimates with smaller SEs than will proportional allocation because larger samples will be drawn from strata with larger variance (s_h in equations [3.13] and [3.14]), thereby reducing the SEs of the stratum means ($SE\bar{y}_h$ in Table 3.1), which are inversely related to sample size.

3.2.3.4 Estimation of Proportions

Estimates of the proportion of sampling units in a population that fall into a defined class are computed much like the stratified mean. Essentially, the proportion is estimated for each stratum using the formula for a simple random sample and the stratum-specific proportions are combined using the stratum weights (Table 3.2; Cochran 1977). Similarly, the SE of the estimate for the proportion in the entire population is a weighted sum of the individual stratum variances (Table 3.2).

3.2.4 Cluster Sampling

In cluster sampling, the population is divided into primary sampling units (clusters) and secondary sampling units (elements). In an example of sampling fish, the secondary sampling units are individual fish and the primary sampling units are groups of fish as might be caught together in a net (Figure 3.2). In an example of sampling grid locations, the secondary sampling units are the individual grid locations and the primary sampling units are blocks of four grid locations (Figure 3.2). In cluster sampling, the primary units are selected independently at random, which in the fish example may be thought of as having randomly selected netting locations. Cluster sampling is single stage if each element (e.g., individual fish) in each cluster (e.g., net) is included in the sample and two stage if only a subsample of each element from each cluster is included in the sample (e.g., individual fish are subsampled from each net). Cluster sampling is distinguished from other designs in that the primary units are sampled independently but the secondary units are potentially correlated. Put another way, fish caught in a net may not be independent because they may be more similar to each other than to randomly selected fish from the entire population.

Cluster sampling is commonly used when the sampling frame is difficult or impossible to construct or the sampling process naturally results in clusters of secondary units. For example, when fish are collected with nets set at random points on a grid (as described in Box 3.1), the net is the primary sampling unit, and the individual fish collected are secondary units. Although catch per net (in numbers or weight) is appropriately treated as coming from a simple random sampling design, the mean weight of fish estimated from this sampling design should be treated as a cluster sample because individual fish within a net may not be sampled independently.

In practice, cluster sampling often results in a situation in which individual elements within each cluster are similar and differences in the means are larger among clusters than within clusters. In the above example, a truly random sample of individual fish in a lake would be very difficult to obtain because trying to

collect one fish at a time would be very inefficient and would likely lead to a much smaller sample size than would using nets that can capture multiple fish. In this case, redundant information is provided by each fish measured (because of their similarity or correlation), and the precision of the overall mean is reduced relative to a simple random sample with the same sample size. Although this seems like a poor sampling strategy relative to simple random sampling, the advantage of cluster sampling design is that sampling is often less expensive and a substantially greater sample size can be obtained.

3.2.4.1 *Single-Stage Cluster Sampling*

In cluster sampling, estimates of several different quantities can be obtained (Box 3.5). For example, when individual fish caught in randomly placed nets are counted and weighed, estimates of the mean number of fish caught per net can be obtained by the usual estimator for simple random sampling. In addition, the mean total weight of fish caught per net can be estimated, but this quantity is an example of a cluster total. Cluster totals are sometimes interpretable statistics (such as presented here) but in other cases are hard to visualize or interpret. If, for example, the lengths of individual fish were measured instead of weight, the cluster total would represent the total length of all the fish caught, a statistic that is of little use. When the statistic of interest focuses on cluster totals, single-stage cluster sampling reduces to a simple random sample for which each cluster total is treated as a single observation.

The mean weight (or length) of individual fish is an example of the mean per secondary unit, which is another statistic that can be computed in cluster sampling. A further complication of single-stage cluster sampling is that simpler formulae may be used when the clusters are of equal size (i.e., the number of secondary units is equal in all clusters) than when the clusters are of unequal size. In most fisheries applications, clusters are of unequal size, so we will emphasize the formulae relevant to such situations.

3.2.4.2 *Estimation of the Mean per Secondary Unit*

As indicated above, the process for estimating mean cluster totals follows simple random sampling (Box 3.5). Estimating the mean per secondary unit, such as the mean weight of individual fish, is conceptually related to stratified random sampling because the mean for each cluster is weighted by the number of secondary units in each cluster. The principal difference, however, is that all strata are sampled in stratified random sampling, whereas only a sample of all clusters is selected in cluster sampling. Estimating the mean per secondary unit is also related to ratio estimation because cluster totals and numbers of elements (secondary units) in each cluster are both random variables, and estimation of the mean per secondary unit naturally uses the number of elements (secondary units) as a divisor (Table 3.1). The mean per secondary unit is simply the sum of cluster totals divided by the total number of secondary units (Table 3.1). Estimating the SE is similar to the procedure for estimating the SE of a ratio, where each cluster total is the y variable and the number of secondary elements in each

Box 3.5 Example of Cluster Sampling

Five throw nets were deployed at random locations along the shoreline of a lake to collect age-0 bluegill. Greater sampling effort would usually be required, but data from these five nets are used to illustrate the procedure. The weight (g) of each of the age-0 bluegill was measured. The goal of sampling was to estimate the mean biomass of age-0 bluegill per net, the mean catch per net, and the mean weight of individual age-0 bluegill.

Table Catch per net and weight per individual of age-0 bluegill caught in five throw nets. Mean fish weight is given by for which computations are shown below table.

Measure and summary statistic	Net 1	Net 2	Net 3	Net 4	Net 5
Catch (M_i)	10	5	7	0	3
Weight (g)	0.495	0.319	0.514		0.610
	0.391	0.419	0.497		0.572
	0.274	0.503	0.374		0.681
	0.470	0.451	0.457		
	0.309	0.491	0.388		
	0.369		0.521		
	0.381		0.539		
	0.308				
	0.420				
	0.326				
Cluster total (y_i)	3.743	2.183	3.290	0	1.863
$(y_i - \hat{R}M_i)^2$	0.47197	0.00102	0.03572	0	0.28516
Mean weight per fish per net	0.374	0.437	0.470		0.621

The mean catch per net is

$$\overline{M} = (10 + 5 + 7 + 0 + 3)/5 = 5.0 \text{ fish.}$$

The mean cluster total (mean biomass per net) is

$$\overline{y}_i = (3.743 + 2.183 + 3.290 + 0 + 1.863)/5 = 2.216 \text{ g.}$$

The mean weight per fish, the SE of the mean, and the CLs are

$$\text{Mean fish weight} = (\hat{R}) = \frac{\sum y_i}{\sum M_i} = \frac{3.743 + 2.183 + 3.290 + 0 + 1.863}{10 + 5 + 7 + 0 + 3} = 0.443;$$

$$\text{SE}(\hat{R}) = \frac{1}{\sqrt{n}\,\overline{M}} \sqrt{\frac{\sum (y_i - \hat{R}M_i)^2}{n-1}} = \frac{1}{\sqrt{5} \cdot 5} \sqrt{\frac{0.47197 + 0.00102 + 0.03572 + 0 + 0.28516}{5-1}} = 0.040;$$

$$\text{Lower CL} = \hat{R} - t_{\alpha, n-1}\text{SE}(\hat{R}) = 0.443 - 2.776 \cdot 0.040 = 0.332; \text{ and}$$

$$\text{Upper CL} = \hat{R} + t_{\alpha, n-1}\text{SE}(\hat{R}) = 0.443 + 2.776 \cdot 0.040 = 0.554.$$

(Box continues)

Box 3.5 *(continued)*

Now, instead of treating the data as a single-stage cluster sample, consider the situation where the same number of fish per net are weighed but 48 fish are caught in Net 1 and 20 fish are caught in Net 3, thereby leading to a two-stage cluster sample with different numbers of fish caught in each net.

Table Catch per net and weight per individual of age-0 bluegill in five throw nets. Computation of the mean weight per secondary unit ($\bar{\bar{y}}$) is given below table.

Measure and summary statistic	Net 1	Net 2	Net 3	Net 4	Net 5
Catch (M_i)	48	5	20	0	3
Weight (g)	0.495	0.319	0.514		0.610
	0.391	0.419	0.497		0.572
	0.274	0.503	0.374		0.681
	0.470	0.451	0.457		
	0.309	0.491	0.388		
	0.369		0.521		
	0.381		0.539		
	0.308				
	0.420				
	0.326				
Estimated cluster total ($\hat{y}_i$)	17.966	2.183	9.400	0	1.863
Mean weight per fish per net ($\bar{y}_i$)	0.374	0.437	0.470		0.621
$M_i^2 (\bar{y}_i - \bar{\bar{y}})^2$	3.5044	0.0144	1.2996	0.0000	0.3894

The mean per secondary unit (mean fish weight) is

$$\bar{\bar{y}} = \frac{\sum \hat{y}_i}{\sum M_i} = \frac{17.966 + 2.183 + 9.400 + 0 + 1.863}{48 + 5 + 20 + 0 + 3} = 0.413.$$

The SE of the mean per secondary unit is approximated by

$$\mathrm{SE}(\bar{\bar{y}}) = \sqrt{\frac{\sum_{i=1}^{n} M_i^2 (\bar{y}_i - \bar{\bar{y}})^2}{n\bar{M}^2(n-1)}} = \sqrt{\frac{3.5044 + 0.0144 + 1.2996 + 0 + 0.3894}{5 \cdot 15^2 \cdot (5-1)}} = 0.034.$$

cluster is the x variable (Table 3.1). Confidence intervals are obtained as in estimating ratios using the t-distribution.

3.2.4.3 *Two-Stage Cluster Sampling*

In two-stage cluster sampling, a simple random sample of n clusters is selected and then a simple random sample of the elements (secondary units) is subsampled

from within each sampled cluster. In our example of fish in nets (Figure 3.2), the secondary sampling units (fish) are subsampled from the primary sampling units (nets). This differs from single-state cluster sampling, where secondary sampling units are all sampled completely rather than subsampled. Two levels of sampling are employed, so means and variances at two levels are defined. First, the mean per secondary unit within the ith cluster is estimated as in Table 3.1 (Box 3.5; Cochran 1977), where y_{ij} is the measured value for the jth element in the ith cluster and m is the number of elements sampled within each cluster (which has M_i secondary units). From this, the total for each cluster and the mean cluster total are estimated as in Table 3.1 (Box 3.5; Cochran 1977). From the estimated cluster totals, the overall mean per secondary unit is estimated as in Table 3.1 (Box 3.5; Cochran 1977). The variance of the overall mean includes two components that represent the variation between clusters and the variation due to subsampling within clusters. When the number of clusters sampled is small relative to the number of clusters in the population, the SE of the mean per secondary unit can be approximated as in Table 3.1 (Cochran 1977).

3.2.5 Systematic Sampling

In systematic sampling, all sampling units in the population are arranged in a sequence, and then from a random starting point every kth sampling unit is included in the sample. Systematic sampling is often used for ease of execution and convenience. Also, systematic samples are usually spread more evenly over the population, so population attributes may be estimated more precisely than with simple random sampling. However, a major difficulty with systematic sampling based on a single starting point is that the variance and SE of the estimates cannot be directly determined. This occurs because systematic sampling with a single starting point is equivalent to cluster sampling with just one cluster being sampled (the samples are not independent). One way of alleviating this problem is to take a systematic sample with two or more randomly selected starting points. For example, a lake could be divided into grids and several rows of grids could be randomly selected as starting points of evenly spaced grids (Figure 3.2). When systematic sampling is implemented with multiple starting points, the formulae for single-stage cluster sampling apply, with each group of observations associated with each start point treated as a cluster.

Systematic sampling with a single starting point should be avoided whenever sample units are ordered in a linear or nonlinear pattern. Under such circumstances, stratified random sampling produces more precise estimates because strata can be constructed to account for the pattern in the sample units and thereby reduce within-stratum variance that would not be accounted for by systematic sampling (Cochran 1977). Therefore, the choice of systematic sampling versus simple random sampling or stratified random sampling needs to be judged on a case-by-case basis.

The mean value for a single-starting-point systematic sample is estimated in the same way as for a simple random sample (Table 3.1), that is, as the sum of the

observations divided by the number of observations. Assuming that the sampling fraction is relatively small, estimates of the mean using systematic sampling are unbiased, for the same reason that estimates of the mean using simple random sampling are unbiased (section 3.2.2.1). A small amount of bias may occur when the sampling fraction is large, if the number of sampling units in the population (N) is not evenly divisible by the sample size (n) because some units would have a lower probability of being included in a sample than would others. When multiple starting points are used, the mean is estimated as it would be for a single-stage cluster sample (Table 3.1) and is also unbiased if the sampling fraction is relatively small.

Several methods have been developed for approximating the SE for a systematic sample with a single starting point, but we do not recommend their use because they can lead to strongly biased estimates of the SE (Cochran 1977). With two or more starting points, valid estimates of the SE of the mean can be obtained using a single-stage cluster sampling approach (Box 3.6).

3.2.6 Model-Based Estimators

All of the designs we have discussed to this point have focused on sampling a single variable or attribute of interest. Further, all of the designs, when properly implemented, are designed to be unbiased for estimating the mean and proportions. However, in many fisheries investigations, several variables are of interest. Further, relationships among the variables measured provide an opportunity to extract more information than is provided by each variable alone. A familiar example is sampling fish when length and weight are both measured on a subset of fish collected and only length is collected on the remaining fish. Because the two are related, we can infer the weight of fish where only length data are collected. This situation is an example of a model-based design.

Model-based designs are limitless in their variations, given the number of variables that can be measured and the number of relationships among variables that might be considered. This being the case, we will describe a model-based design that illustrates a commonly used approach. The key advantage of model-based designs is that the additional information contained in auxiliary variables can substantially improve precision of estimates. Improvements in precision, however, come at the cost of losing the property of being design unbiased. When choosing models (such as a linear regression) to represent a relationship among variables, the right model is often uncertain. Thus, if the wrong model is chosen, estimates of the mean or proportion can be biased. This is not to say that such an approach is necessarily worse than using a sampling design that is design unbiased. In some situations, the gains in precision may more than offset the bias introduced by having the wrong model. As indicated earlier, precision is appropriately expressed as mean square error (MSE) when bias is present. From equation (3.1), an estimator that reduces the variance component faster than the bias^2 term produces estimates with a smaller MSE. In addition, the concept of MSE applies to situations that are unbiased, but the bias term drops out (being equal to zero).

Box 3.6 Example of Systematic Sampling with Two Starting Points

The width of a stream was measured at sampling locations arranged every 20 m from two random starting points, with 15 points sampled for each random starting point.

Table Stream width measurements based on systematic sampling with two starting points.

Starting point 1		Starting point 2	
Distance upstream (m)	Width (m)	Distance upstream (m)	Width (m)
3	6.1	4	10.3
23	11.4	24	6.5
43	13.7	44	9.0
63	11.3	64	7.6
83	11.7	84	6.3
103	13.3	104	12.0
123	12.1	124	6.2
143	11.5	144	13.1
163	6.4	164	10.2
183	34.8	184	26.5
203	31.7	204	32.1
223	27.2	224	28.6
243	26.7	244	29.2
263	24.1	264	33.5
283	10.1	284	6.0

The estimated mean width and its associated SE are

$$\text{Mean width} = (\hat{R}) = \frac{\sum y_i}{\sum M_i} = \frac{6.1 + 11.4 + 13.7 + \ldots + 33.5 + 6.0}{15 + 15} = \frac{489.2}{30} = 16.31, \text{ and}$$

$$\text{SE}(\hat{R}) = \frac{1}{\sqrt{n}\,\bar{M}} \sqrt{\frac{\sum (y_i - \hat{R}M_i)^2}{n-1}} = \frac{1}{\sqrt{2} \cdot 15} \sqrt{\frac{56.347 + 56.347}{2-1}} = 0.50.$$

Regression or double sampling. A commonly used model-based approach is regression sampling, sometimes referred to as double sampling. In this method, a sample is collected where the auxiliary variable (x_i = independent variable) is measured on all units. A subsample is then selected where the variable of interest (y_i = dependent variable) is also collected. A linear regression between the two variables is formed, and the regression is used to incorporate the information contained in the auxiliary variable into the estimate of the mean for the variable of interest. Typically, this approach is implemented when the auxiliary variable is much less expensive to measure than is the variable of interest, thereby allowing for a greater sample size. An example of this situation is that visual estimates of stream width can be collected much more rapidly (and hence, less expensively) than can actual

measurements of stream width. Thus, the precision of estimates of the mean width of a stream may be improved by taking many visual estimates of stream width while measuring only a subsample of sites to provide a calibration, via linear regression, between measured and visually estimated stream width (Box 3.7). Another example is that the percent water content of fish tissue is easily determined by weighing, drying, and reweighing the tissue and provides a reasonable predictor of the fat content of the tissue (Hartman and Brandt 1995). Fat content is often determined on dry tissues and requires time-consuming extraction by use of solvents in sophisticated equipment. Thus, the mean fat content may be estimated by collecting many measurements on percent water content and only a few concurrent measurements of actual fat content.

An important assumption of this method is that a linear relationship exists between the two variables. If this assumption does not hold, estimates of the mean can be biased, thereby offsetting any gain in precision. Another important consideration is that the value of the y_i for sampling units (e.g., stream width at a particular location or individual fish for fat content), where only the auxiliary (x_i = independent) variable is measured, can be estimated with the regression equation. Therefore, the precision of individual sample units may be relatively poor because the prediction does not match the value that would be obtained by direct measurement. However, the precision of the estimated population mean will usually be increased because of increased sample size. This general principal of sampling reflects the fact that the SE of the mean is inversely related to the sample size; that is, an increased sample size reduces the SE of the mean, which increases precision of the estimated mean.

The formula for estimating the mean using a double-sampling approach is given in Table 3.1, and an example of the application of this method is given in Box 3.7. This example only brushes the surface of the diversity of applications of model-based designs. For a more in-depth treatment see Draper and Smith (1981) for linear regression models and Seber and Wild (1989) and Bates and Watts (1988) for nonlinear regression models.

3.2.7 Advanced Designs

The designs we describe and illustrate above are intended to provide a basis for the appropriate design and analysis of sampling programs. These relatively simple designs provide useful approaches in many situations, and are sufficient for many of the questions posed by fisheries scientists. For more complex situations, these designs can be combined and adapted to suit the needs of the investigator. Many other designs have been developed for specialized situations (e.g., hydroacoustic surveys). Thompson (1992) covers additional designs. One extension to normal sampling designs we would like to highlight are adaptive designs. In adaptive designs, additional sampling is concentrated near sampling points where something interesting happens. For example, in surveys trying to estimate the density of rare species, additional sampling can be concentrated near sampling locations where the rare species is found. Adaptive designs can provide improved precision of density estimates and also have the advantage that more specimens can be collected for length,

Box 3.7 Example of Regression or Double Sampling

The percent of coverage by woody material was visually (Table, column 2) estimated at 25 randomly selected points along a stream (Table, column 1), and the actual amount of woody material coverage was measured at 10 of these points (Table, column 3), with the goal of estimating mean woody material coverage for this reach. The regression between the visually estimated coverage and the measured coverage gave the following equation:

$$\text{Measured coverage} = 7.2937 + 1.0357(\text{estimated coverage}).$$

Table Measurement and visual estimation of percent of woody material coverage along a stream.

	Woody cover (%)				
Stream location and mean	Visually estimated (x_i)	Measured (y_i)	$(y-\bar{y})^2$	$(y-\bar{y})(x-\bar{x})$	$(x-\bar{x})^2$
1	30	34	29.16	5.4	1
2	40	52	158.76	113.4	81
3	60	66	707.56	771.4	841
4	20	26	179.56	147.4	121
5	20	21	338.56	202.4	121
6	20	25	207.36	158.4	121
7	30	38	1.96	1.4	1
8	0	14	645.16	787.4	961
9	40	54	213.16	131.4	81
10	50	64	605.16	467.4	361
11	40				
12	70				
13	80				
14	20				
15	50				
16	80				
17	50				
18	20				
19	80				
20	80				
21	40				
22	90				
23	30				
24	40				
25	40				

The estimated mean coverage using double sampling is

$$\bar{y}_{\text{reg}} = \bar{y} + b(\bar{X} - \bar{x}) = 39.4 + 1.0357 \cdot (44.8 - 31.0) = 53.69,$$

where $\bar{x}$ = the mean of visual estimates for the 10 subsampled stream locations, $\bar{y}$ = the mean of measured values for the 10 subsampled stream locations, and $\bar{X}$ = the mean of visual estimates for all 25 stream locations.

The SE of this estimate is

$$\text{SE}(\bar{y}_{\text{reg}}) = \sqrt{\frac{1}{n(n-2)}\left\{\sum(y_i-\bar{y})^2 - \frac{[\sum(y_i-\bar{y})(x_i-\bar{x})^2]}{\sum(x_i-\bar{x})^2}\right\}} = \sqrt{\frac{1}{10(10-2)}\left\{3086.4 - \frac{[2786]}{2690}\right\}} = 1.58.$$

age, or other biological variables. Because sampling is concentrated near hot spots, the additional samples are not independent of the original sampling locations, and specialized formulae must be used to reduce or remove biases that would occur if the data were treated as coming from a sample of independent observations.

3.3 EXPERIMENTAL DESIGNS

Developing an adequate design to an experiment is perhaps the trickiest and most difficult task that a fisheries scientist faces. Fisheries scientists must balance the need to control the experiment to understand better the results with the need to assure that the design is relevant to natural systems (Yandell 1997). Many experimental designs used by fisheries scientists come from disciplines such as agriculture, where experiments are easier to develop and factors are easier to manipulate. By necessity, fisheries scientists often rely on experimental units, such as lakes or fish, over which they have little control. Lack of control over experimental units is an important reason why developing a sound experimental design and analysis is critical to the success of any fisheries experiment.

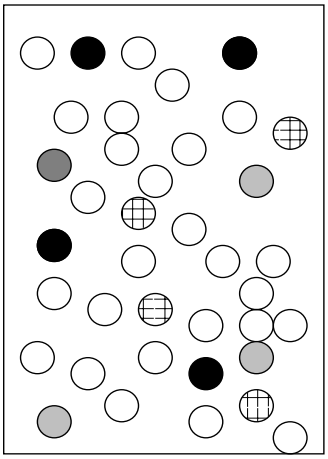
The first step in designing an experiment is to develop a clear statement of objectives for the experiment (Cochran and Cox 1957; Yandell 1997). This step should include the questions that are being asked in the experiment or the hypotheses being tested. In the first section of this chapter, we focused on sampling designs for which the goal is generally to describe the attributes of a population. When an experiment is conducted, the goal is often to answer questions focusing on the response to a treatment or to determine the influence of natural or anthropogenic factors. Questions should be clearly focused and reasonably answered. For example, the question, Are more large largemouth bass present after an increase in the minimum length limit? is too vague to be answered through an experiment. A clearer way to phrase the question is, Did the population density of largemouth bass longer than 35 cm increase in lakes where the minimum length limit was increased from 25 cm to 35 cm? The second question is more specific than the first and helps to determine how the experiment should be designed. The second question can now be turned into a testable statistical hypothesis. Hypothesis testing is the formal approach that is used to assess whether evidence supports your question. Hypotheses are set up in two competing claims, the null hypothesis (H_0) and the alternative hypothesis (H_a). The statistical test is set up either to support or not to support the null hypothesis (see Chapter 1). In our example, we could formulate the null hypothesis as the number of largemouth bass longer than 35 cm is the same in lakes with the increased size limit and in lakes with no change in the size limit. The alternative hypothesis is then the number of largemouth bass longer than 35 cm is not the same in lakes with the increased size limit and in lakes with no change in the size limit. Remember the statistical test is set up to support or not support the null hypothesis, so we can either fail to reject or reject that the number of largemouth bass greater than 35 cm differs in lakes with and without the size limit. However, we cannot conclude that the alternative hypothesis is true.

After specifying the research question, the experiment should be described in terms of experimental units, units being sampled (subsamples being taken from each experimental unit), number and type of treatments, number of replicates per treatment, and target population (Cochran and Cox 1957; Brown and Austen 1996). For the largemouth bass question in the previous paragraph, the experimental units are lakes and the treatment is a regulation change. To complete the study design, the fishery scientist would need to specify a population of lakes that would be subjected to the regulation change, to select randomly a sample of lakes for study (lakes = replicates), and to choose a sampling plan with an appropriate gear for capturing largemouth bass longer than 35 cm. To differentiate effects of the regulation change from background variation, sampling would also need to begin before implementing the rule change (temporal controls) and need to include a sample of lakes on which the regulation change was not implemented (spatial controls). Finally, a well-designed experiment should also have an outline of the method of analysis to be applied to the data after the experiment is completed (Cochran and Cox 1957). Methods of analysis for experimental designs are provided in box examples below.

3.3.1 Completely Randomized Design

A completely randomized design is a design in which treatments are applied to the experimental units completely at random, so that each experimental unit has an equal probability of being selected for each treatment (Figure 3.3; Cochran and Cox 1957). For example, we may want to determine how different creel limits affect angler catch rates for walleye in Wisconsin lakes (Box 3.8). In this example, the statistical population being considered includes lakes in Wisconsin that contain walleye, and the experimental unit is an individual lake. Although angler catch rates are determined through a creel survey program (thus constituting a subsample of all anglers), we will treat catch rates as a single observation per lake. In this example, we decided a priori that we were interested only in one factor (i.e., bag limits) and we were interested in three levels of this factor: bag limits of one, two, or five walleye per day. The experiment was designed to have six replicates, seven replicates, and nine replicates for the one, two, and five walleye bag limit treatment, respectively, for a total of 22 lakes in the experiment. Each lake was assigned a treatment level at random, completing the experimental design.

As with simple random sampling designs described earlier, completely randomized designs provide a basic standard against which to compare other designs. Completely randomized designs have the advantage of allowing complete flexibility in the number of treatments and replicates allowed for the experiment (Cochran and Cox 1957). Further, statistical analysis is relatively easy regardless of the number of replicates and treatments (Cochran and Cox 1957). This holds true even when treatments or data are missing in the experiment (Cochran and Cox 1957), a common problem with large-scale field experiments. Finally, completely randomized designs have the advantage of maximizing the degrees of freedom for

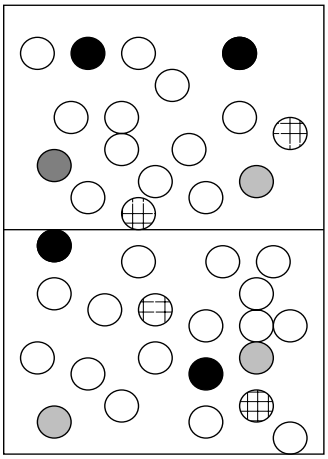


Completely randomized design

Randomly selected for bag limit of 1 walleye/day

Randomly selected for bag limit of 2 walleye/day

Randomly selected for bag limit of 5 walleye/day

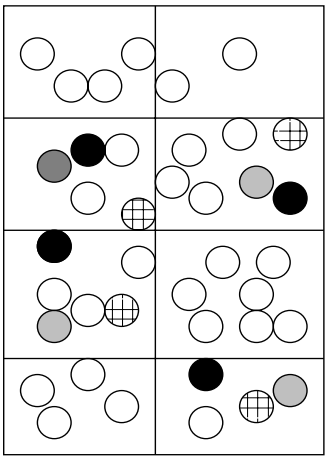


Randomized block design

Randomly selected for bag limit of 1 walleye/day

Randomly selected for bag limit of 2 walleye/day

Randomly selected for bag limit of 5 walleye/day



Random-effect block design

Randomly selected for bag limit of 1 walleye/day

Randomly selected for bag limit of 2 walleye/day

Randomly selected for bag limit of 5 walleye/day

Figure 3.3 Examples of three experimental designs for testing how fish populations in lakes (circles) in a landscape (rectangles) respond to the application of three different daily bag limits in an angling fishery.

Box 3.8 Example of a Completely Randomized Design

The goal of this study was to determine if walleye catch rates differed among Wisconsin lakes with different daily bag limits (Beard et al. 2003). From the thousands of lakes in Wisconsin with walleye populations, six lakes were randomly chosen to have a bag limit of one walleye per day, seven lakes were randomly chosen to have a bag limit of two walleye per day, and nine lakes were randomly chosen to have a bag limit of five walleye per day. For this analysis, the designation of North or South was ignored. A fixed-effects general linear model (GLM, implemented in SAS; SAS 2005) was used for the analysis of these data.

Program

```
   * This data step reads the following information into a data set named
walleye;
   data walleye;
   input lake $ region $ bag_limit catch;
   cards;
   Willow            North          1         2.21
   Mud               North          1         2.32
   Pine              North          1         2.74
   Bass              North          2         2.23
   Perch             North          2         2.25
   Twin              North          2         1.40
   Park              North          2         2.36
   Mendota           North          5         1.78
   Silver            North          5         1.64
   Manistee          North          5         1.97
   Fox               North          5         1.99
   McGee             South          1         2.70
   Deep              South          1         3.63
   Round             South          1         2.82
   Long              South          2         3.09
   Portage           South          2         3.63
   Indian            South          2         2.82
   Wolf              South          5         2.20
   Gull              South          5         1.74
   Black             South          5         2.85
   Goose             South          5         3.01
   Fletcher          South          5         1.72
   ;
   run;
   *These statements call the GLM procedure in SAS, declaring the variable
bag_limit to be a categorical variable, and catch to be the continuous re-
sponse variable. The lsmeans statement requests least-squares means and
standard errors of catch for each level of bag_limit;
   proc glm;
   class bag_limit;
   model catch=bag_limit;
   lsmeans bag_limit/stderr;
   run;
```

(Box continues)

Box 3.8 *(continued)*

Results and Interpretation

Table Results of the GLM procedure in SAS for the dependent variable catch of walleye in 22 lakes with three different bag limits. Abbreviations are given for coefficient of variation (CV), mean square error (MSE), sum of squares (SS), and least-squares means (LSMEAN).

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	1.62317576	0.81158788	2.43	0.1153
Error	19	6.35613333	0.33453333		
Corrected total	21	7.97930909			
R^2	0.203423	Root MSE	0.578389		
CV	23.96337	Catch mean	2.413636		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Bag limit	2	1.62317576	0.81158788	2.43	0.1153

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Bag limit	2	1.62317576	0.81158788	2.43	0.1153

Least-Squares Means

Bag limit	Catch LSMEAN	SE	*P* > |*t*|
1	2.73666667	0.23612614	<0.0001
2	2.54000000	0.21861033	<0.0001
5	2.10000000	0.19279619	<0.0001

The results of the analysis indicate that the daily bag limit had little effect on walleye catch rate. Although the point estimates of mean catch rate (catch LSMEAN) differ somewhat among bag limits, the differences were not greater than would be expected by random chance. The *F*-value for the entire experiment was 2.43 and the resulting *P*-value was 0.1153, which is greater than the 0.05 alpha value commonly used when testing for significant differences among means. Therefore, we would conclude that bag limits had no significant effect on catch rates.

Sums of squares can be computed in several ways. The simplest to understand and the most widely reported in statistical analysis programs are type III SS, which are computed as the difference in SS between two nested models in which one term is left out. Thus, the SS for each term is simply the difference in SS between the full model with all terms present and a reduced model with the term of interest absent. In contrast, SAS also reports type I SS, which are computed as the difference in SS between hierarchical models in which each term is dropped in sequence, beginning with the right-hand term and proceeding to the left. Thus, the SS for each term depend on the order in which the terms are specified by the user. Type III SS are independent of the order in which terms are specified in the model, so are generally preferred over type I SS.

analysis (Cochran and Cox 1957), thus maintaining statistical power when the number of replicates per treatment is small.

The main disadvantage of experiments using a completely randomized design is that the power of these experiments to detect differences among treatments may be relatively low (analogous to simple random sampling, where the precision of point estimates may be low). Randomized designs are most commonly used in laboratory studies, where experimental units are relatively homogeneous, thereby increasing the power of the experiment. In field studies where experimental units vary greatly from unit to unit (Cochran and Cox 1957), variation among units may obscure systematic differences resulting from the treatment. One way to overcome large variation among units is to increase the number of replicates in the experiment, but this comes at additional monetary cost (Brown and Austen 1996).

Completely randomized designs have been used in fisheries management projects mostly where sites were homogenous or where differences among sampling units were not known. For example, Walsh et al. (2002) compared catches from prepositioned area electrofishing and electric seining at 12 randomly selected stream sites. Similarly, Kocik and Taylor (1994) placed brown trout and steelhead in randomly selected sites within an experimental stream to quantify their survival and growth. In both studies, sampling sites were assumed to be relatively homogenous, thereby minimizing variability not accounted for in the experiment. Although completely randomized designs are uncommon in fisheries, they can be useful in small pilot studies that will provide some information about the experimental unit for better design of a full-scale study. Before continuing on with more sophisticated designs, we consider the analysis of this relatively simple design, and discuss some of the critical considerations for data analysis.

3.3.1.1 Analysis of Completely Randomized Design

After the experiment has been conducted and data collected, how do we determine if the treatment(s) led to a response? One tool available is the general linear model (GLM), which contains the familiar analysis of variance (ANOVA) model. General linear model is a term used to refer to an entire class of models that are linear in their parameters (Yandell 1997; Montgomery 2001), which means that no parameter in the model is an exponent or is multiplied or divided by another parameter (Neter et al. 1996). The term general is used because both continuous and categorical variables can be used as predictor variables (Quinn and Keough 2002). In most of these models, we measure a response variable and then determine how this response variable is influenced by one or more predictor variables.

In our creel limit example, the treatments (or predictor variables) are fixed because the bag limits were determined prior to the start of the experiment and then applied according to the completely randomized design. Moreover, we treat the creel limits as categorical variables. This is in contrast to continuous variables (such as lake area), which we will discuss later. When analyzing data from this situation, we use what is called a fixed-effects GLM (Quinn and Keough 2002). The objective of our analysis is to determine whether variation in the means for different treatment levels differs more than would be expected by chance or if

"real" differences in catch rate are related to the bag limit imposed. The GLM for this case can be specified for $i = 1$ to p treatment levels and $j = 1$ to n replicates:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \tag{3.15}$$

where y_{ij} are the observations, μ is the overall population mean of the response variable, α_i is the treatment effect for each level, and ε_{ij} is the unexplained variation among lakes (i.e., statistical error; Quinn and Keough 2002). In our example (Box 3.8), $i = 3$ treatments (bag limits of 1, 2, and 5) and $j = 6$, 7, and 9 replicates. The fixed-effects model can then be used to test the H_0 that all treatment level means (specified as μ_i) are the same:

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_i = \mu.$$

This can also be specified in terms of the test of treatment effects:

$$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_i = 0.$$

An *F*-test is used to compare the variability among groups to the residual variability (*F*-ratio = mean squared error for main effects divided by the mean squared residual error), to determine if the observed differences in group means are greater than would be expected by chance. The observed *F*-ratio is compared to an *F*-distribution with the degrees of freedom in the numerator and denominator being those used for the two mean square errors. If H_0 is true, both group and residual mean square error should estimate the pooled population error term and the *F*-ratio should be 1 (Quinn and Keough 2002). In our bag limit example (Box 3.8), we would compare the *F*-ratio to an *F*-distribution with 2 df (3 treatments; $p - 1$) for the numerator and 19 df (22 observations – 3 treatments) for the denominator.

3.3.1.2 *Assumptions*

The most important assumption when sampling or performing any experiment is that the treatments are randomly applied to the experimental units (Sokal and Rohlf 1995). Failure to select samples at random or to apply treatments at random may result in biased results that are not representative of the true response (Sokal and Rohlf 1995). Applying treatments to lakes or rivers where the investigator suspects they will be most successful is tempting, but the results of the study will not be applicable to any other lakes or rivers. If the fully randomized experimental design is implemented, this insures that the random-selection assumption is satisfied.

Many statistical analyses assume that sample units or the selection of samples are independent (Sokal and Rohlf 1995; Brown and Austen 1996). That is, changes in one sample unit or one sample subject should not affect other sampling units or subjects. Treatments or subjects must therefore be spatially and temporally independent (Sokal and Rohlf 1995; Brown and Austen 1996). In our example

(Box 3.8), this assumption implies that the bag limit imposed on one lake has no effect on the catch rate of walleyes in nearby lakes. This assumption would be violated, for example, if anglers shifted their effort away from lakes where the creel limit was imposed to lakes where the bag limit was not imposed.

Repeated measurements of water bodies over time, which are often used to detect changes in fish populations caused by management actions (stocking, habitat manipulations, or regulations) in particular water bodies (often in relation to water bodies where the management action was not implemented) are not temporally independent. Lack of temporal independence is called time series bias and can lead to problems in estimating parameters (Walters 1985; Caputi 1988; Hilborn and Walters 1992; Myers and Barrowman 1996). To account for a lack of temporal independence among sample units, a repeated-measures design is often used (section 3.3.8). Lack of spatial independence also occurs in fisheries studies (as noted in the previous paragraph; also see examples in telemetry studies, Chapter 14, and watershed analyses, Chapter 18), and a variety of methods have been developed to account for spatial dependency.

Another important assumption of a GLM analysis is that the residual variance must be constant or homoscedastic among observations and treatments (Sokal and Rohlf 1995; Montgomery 2001). Variability among experimental units commonly increases with an increasing mean, thereby leading to heteroscedastic residual variance (Sokal and Rohlf 1995). Inequality of variance is generally diagnosed by using plots of the residuals against the predictor variable and predicted values and either Bartlett's or the modified Levene test (Montgomery 2001). When unequal variance occurs, the data are often transformed to equalize the variance, or the unequal variance is accounted for using a mixed model (section 3.3.4). For example, prior to transformation, catch data are often highly skewed in their distribution, so variance often differs among treatment levels (heteroscedastic residual variance; Figure 3.4). In contrast, after log transformation, catch data may be normally distributed and have equal variance among treatment levels (homoscedastic residual variance; Figure 3.4).

The final assumption necessary when performing many common statistical analyses is that the residual errors are normally distributed (Sokal and Rohlf 1995; Montgomery 2001). Although large departures from normality can significantly affect inferences from a GLM analysis (Montgomery 2001), this is perhaps the least important assumption because the central limit theorem states that with large sample sizes (e.g., greater than 30), estimates of model parameters often approximate a normal distribution regardless of the distribution of the data (Yandell 1997). Departures from normality are tested using normal probability plots and statistical tests such as the Shapiro–Wilk test or Kolmogorov–Smirnov test (Box 3.9; Sokal and Rohlf 1995). Each of these tests examines different aspects of departures from normality, so they sometimes provide conflicting insights. Our preference is for the Shapiro–Wilk test, which is more sensitive to departures in the tails of the distribution, though the Kolmogorov–Smirnov test is also useful because the test statistic, D, is a readily interpretable measure of the maximum difference between the observed and expected cumulative distributions. When

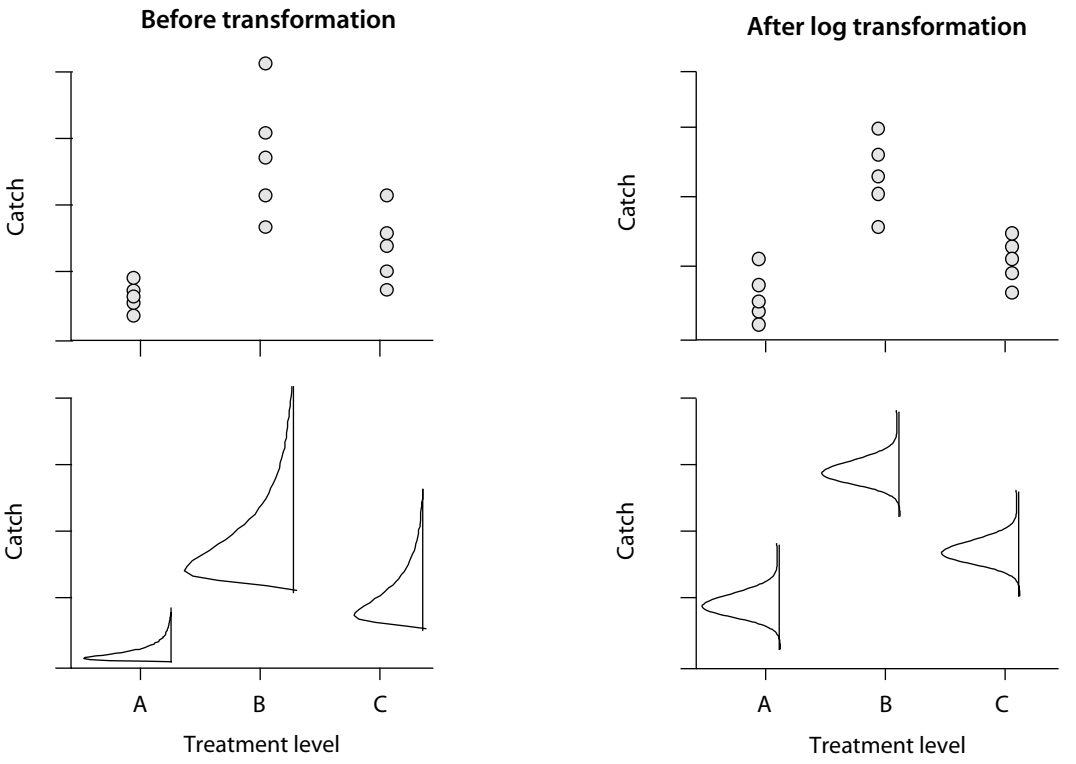


Figure 3.4 Changes to the distribution of catch data following log transformation. The upper panel in each pair of graphs illustrates the distribution of individual data points and the lower panel depicts the hypothetical statistical distribution from which data points were drawn.

examining distributions of residuals for potential violations of the assumption of normal distribution, more concern should be given to distributions with thinner or thicker tails than expected than to distributions that are skewed (Montgomery 2001). Analyses that proceed with nonnormal data will generally lead to fewer significant test results because of reduced power of the test (Montgomery 2001).

One important departure from normality is the presence of outliers, extreme values that lie well outside the distribution of the rest of the data (e.g., more than 3 SDs from the mean) and that are often caused by sampling problems or because the outliers belong to a population that differs from the target population (Montgomery 2001). Outliers can significantly affect the outcome of a statistical analysis and should be examined to determine if they are caused by sampling problems or because they come from another population. If follow-up investigation of an outlier reveals that the outlying datum was caused by a failure of the sampling protocol, the datum can be rejected from the analysis. However, outliers can also be extreme values of the target population so should not be rejected simply because they are outliers.

Box 3.9 Example of How to Test Errors (Residuals) for Normality

In an extension to the example in Box 3.8, the results of the analysis were augmented to examine the normality of residuals.

Program

```
   proc glm;
   class bag_limit;
   model catch=bag_limit;
   lsmeans bag_limit/stderr;
   *The following output request saves a new data set named model_resid,
saving residuals into a variable named resid;
   output out=model_resid r=resid;
   run;
   *These statements call the univariate procedure in SAS, requesting a
normality plot, normality test, and a q-q plot of the variable named resid;
   proc univariate plot normal;
   var resid;
   qqplot resid;
   run;
```

Results

Some output is not shown because it is not critical to this discussion; the pertinent results of this analysis follow.

Table The univariate procedure of SAS was used to evaluate the normality of residuals generated from the analysis in Box 3.8.

Moments

N	22	Sum weights	22
Mean	0	Sum observations	0
SD	0.55015726	Variance	0.30267302
Skewness	0.43350809	Kurtosis	0.01734467

Tests for Normality

Test	Test statistic	Statistic value	*P*-test	*P*-value
Shapiro–Wilk	W	0.933623	$P < W$	0.1460
Kolmogorov–Smirnov	D	0.155156	$P > D$	>0.1500
Cramer–von Mises	W^2	0.120727	$P > W^2$	0.0557
Anderson–Darling	A^2	0.700157	$P > A^2$	0.0602

(Box continues)

Box 3.9 *(continued)*

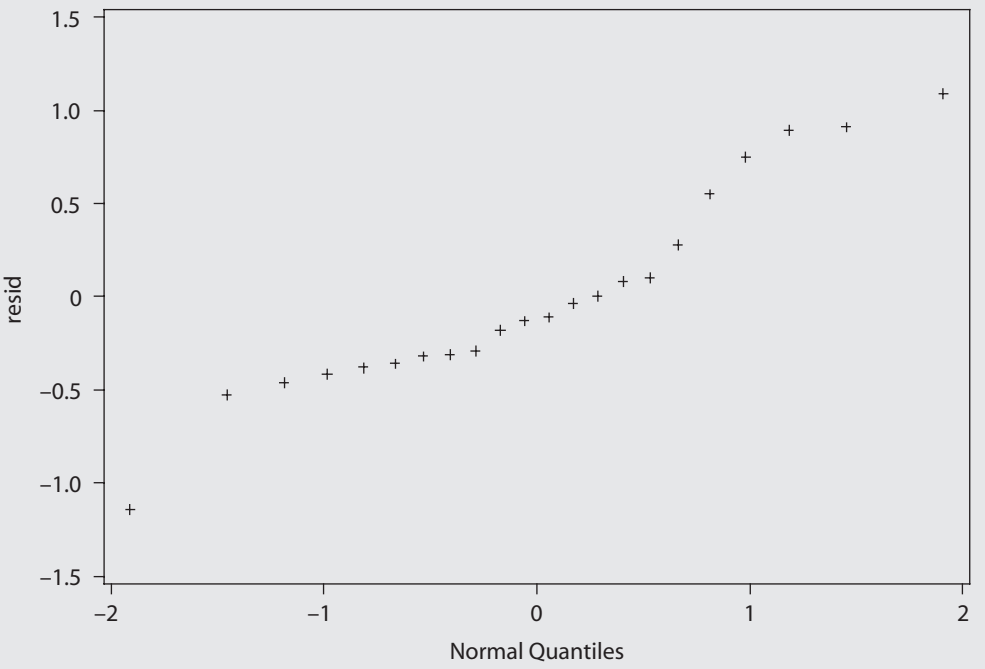


Figure Normal probability plot of residuals versus normal quantiles.

In this example, we used the univariate procedure in SAS to produce a normal probability plot and to provide statistical tests of normality. In a normal probability plot, we are looking for a relatively straight 1–1 line in our plot. In this case, the data show some deviation from a straight line, but none of the normality tests were significant, which indicates that the residuals did not differ significantly from a normal distribution.

When the assumptions of normality or equality of variability are not met, the data can often be transformed into a new scale for which the assumptions are satisfied. However, transformations should be logical and scientifically sound (Yandell 1997). For example, body weight of virtually all organisms will increase as an approximately cubic function of length, so variance in fish weight would likely also increase as an approximately cubic function of length (Brown and Austen 1996). Consequently, the use of a linear model to describe the relationship between weight and length would be incorrect. Transformation of weight and length into their logarithms ($\log_{10}$ or $\log_e$) permits the use of a linear model to estimate parameters of the weight–length relationship and eliminates heteroscedasticity of residual errors. Although transformations may help meet the assumptions of

the GLM, they may also bias parameter estimates (Hayes et al. 1995), so care must be taken in interpreting point estimates.

The selection of an appropriate transformation should rely on an examination of the distribution, mean, and variance of the data. When the mean of a sample is positively correlated with its variance (i.e., variance increases as the mean gets larger), a logarithmic transformation, either base 10 or base *e*, is often appropriate (Sokal and Rohlf 1995). When the mean and variance are similar and do not vary independently, which is often true of count data such as the number of fish caught in a net, the data should be transformed into their square roots to make the variance independent of the mean (Sokal and Rohlf 1995). When the distribution of the data has fewer observations at the mean and at the tails and more observations at intermediate regions than would a normal distribution (platykurtic), which is true of proportion or percentage data such as the percent of lake trout with sea lamprey wounds, the data should be transformed into their arcsines or arcsine square roots (Sokal and Rohlf 1995).

Another approach to transforming the data is to use a nonparametric statistical test. Many nonparametric tests use a rank transformation (i.e., ranking each observation) as a means of reducing the effects of outliers or nonnormality. Coverage of nonparametric methods is beyond the scope of this chapter (see Chapter 1), but they provide a useful suite of methods. As with other transformations, inferences drawn from analyses using a rank transformation do not strictly apply to the arithmetic mean. In the case of rank transformations, inference is generally based on the median, 50th percentile, as a measure of central tendency.

Following transformation, the results of any analysis should be evaluated to determine if the transformation was successful in correcting the observed problem. Also, transforming data to meet statistical assumptions may lead to a model that is not interpretable in the original scale of measurement (Draper and Smith 1981), which may render the model useless for its original purpose. For example, if data for a two-variable model are transformed to meet the assumption of normality, but data for the two variables are each transformed with a different transformation, the resulting statistical model cannot be back-transformed into the original measurement scales for interpretation. Thus, data should be transformed only in the context of an understandable model and its transformation. Sometimes, transformations will still not meet the assumptions necessary for the desired analysis. In such instances, a transformation that achieves approximately equal variances among samples is usually sufficient for analysis, even if the data are slightly nonnormal (Yandell 1997).

3.3.2 Randomized Block Design

One of the disadvantages of a completely randomized design is that natural variation among experimental units obscures the effect of the treatments, thereby reducing the statistical power to detect real differences when they occur. Although more replicates can be taken to overcome low power, a commonly used strategy is to subdivide the population of interest into more homogeneous groups or blocks.

For example, we may already know or suspect that the catch rate of walleye tends to be higher in northern Wisconsin than in southern Wisconsin (Box 3.10). Thus, if we take into account the location of the lake within the state, we can reduce the variability among units within a block. This is directly analogous to the increased precision of stratified random sampling over simple random sampling.

For many fisheries experiments, blocks are often created across time or space (Quinn and Keough 2002). The purpose of blocking is to reduce variability within each group (Cochran and Cox 1957) to estimate means more precisely and to increase power of tests of treatment effects (Quinn and Keough 2002). In a randomized block design, any number of treatments and replicates may be included in the design, and the statistical analysis is straightforward (Cochran and Cox 1957). For a randomized block design to be favored over a completely random design, the precision gained by blocking the treatments must offset the degrees of freedom lost when blocks are used (Yandell 1997). Randomized block designs can be used when blocks are missing, but completely randomized designs are usually better for testing treatment effects if the number of missing blocks is large (Cochran and Cox 1957).

The analysis of randomized block designs is similar to the fully randomized design, except that the effects of the blocking factor are included as an additional effect (Box 3.10). The statistical model for the randomized block design is

$$y_{ijk} = \mu + \alpha_i + \beta_k + \varepsilon_{ijk}, \tag{3.16}$$

here y_{ijk} are the observations, μ is the population mean of the response variable, α_i is the treatment effect for each level, β_k is the effect for each level of the blocking variable, and ε_{ijk} is the unexplained variation among experimental units. This model can then be used to test the null hypothesis that all treatment level means (specified as μ_i) are the same, after taking account for the effect of the blocking variable(s):

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_i = \mu.$$

Randomized block designs are often used in fisheries. For example, Wilderbuer et al. (1998) compared catch per unit effort of various fish species collected by two different types of trawls (Wilderbuer et al. 1998). In this experiment, the two trawl types were simultaneously hauled, and each paired haul was considered as a block because of variability in catches between trawl runs (Wilderbuer et al. 1998). Similarly, Sammons and Bettoli (1999) examined variation in catch and mean length of largemouth bass caught by electrofishing, blocked by transects sampled. Whalen and LaBar (1994) used stream sections as blocks to compare survival and growth of Atlantic salmon stocked at different densities.

3.3.3 Analysis of Covariance

In the randomized block design, variability due to the effects of the categorical blocking variable(s), such as the designation of north and south in the example shown in Box 3.10, is used to remove the confounding effect of this variability.

Box 3.10 Example of a Randomized Block Design

In an extension to the example in Box 3.8, lakes were first blocked into northern and southern Wisconsin lakes, and then treatments were randomly assigned to lakes in each block. A randomized block design should include the blocking factor during the randomization process. The SAS program for this analysis is similar to a completely randomized design, except that block and an interaction term are included in the model. For brevity, the data are not repeated here.

Program

```
   *The following call to the GLM procedure indicates that bag_limit and
region are categorical predictor variables and catch is a continuous response
variable;
   proc glm;
   class bag_limit region;
   model catch=bag_limit region region*bag_limit;
   lsmeans bag_limit region/stderr;
   run;
```

Results and Interpretation

Results of the above analysis are as follows.

Table The GLM procedure for a randomized block design (blocks being northern versus southern lakes) with the dependent variable catch. This analysis is based on the data presented in Box 3.8.

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	5	4.83082242	0.96616448	4.91	0.0065
Error	16	3.14848667	0.19678042		
Corrected total	21	7.97930909			
R^2	0.605419	Root MSE	0.443599		
CV	18.37888	Catch mean	2.413636		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Bag_limit	2	1.93517601	0.96758800	4.92	0.0216
Block	1	2.86174438	2.86174438	14.54	0.0015
Bag_limit*block	2	0.43851616	0.21925808	1.11	0.3523

Least-Squares Means

Bag_limit	Catch LSMEAN	SE	*P* > |*t*|
1	2.73666667	0.18109869	<0.0001
2	2.62000000	0.16940231	<0.0001
5	2.07450000	0.14878776	<0.0001

Block	Catch LSMEAN	SE	*P* > |*t*|
North	2.10944444	0.13498300	<0.0001
South	2.84466667	0.13765619	<0.0001

(Box continues)

Box 3.10 *(continued)*

In this analysis, the bag limit ($F = 4.92$; $P = 0.0216$) and block ($F = 14.54$; $P = 0.0015$) both appear to have an effect on angler catch rates. The interaction between these two factors does not appear to be significant ($F = 1.11$; $P = 0.3523$), which suggests that the effect of bag limits was similar in northern and southern Wisconsin lakes (blocks). Therefore, we can re-run the analysis without the interaction term in the model.

```
proc glm;
class bag_limit region;
model catch=bag_limit region ;
lsmeans bag_limit region/stderr;
run;
```

Results of the analysis without the interaction term for region*bag limit are similar to those for the model with the interaction term.

Table The GLM procedure for a randomized block design (blocks being northern versus southern lakes) with the dependent variable catch. This analysis does not include the interaction term of bag limits*region.

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	4.39230627	1.46410209	7.35	0.0020
Error	18	3.58700282	0.19927793		
Corrected total	21	7.97930909			
R^2	0.550462	Root MSE	0.446406		
CV	18.49515	Catch mean	2.413636		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Bag_limit	2	1.95674263	0.97837132	4.91	0.0199
Block	1	2.76913051	2.76913051	13.90	0.0015

Least-Squares Means

Bag_limit	Catch LSMEAN	SE	*P* > |*t*|
1	2.73666667	0.18224431	<0.0001
2	2.59097810	0.16927875	<0.0001
5	2.06035036	0.14918152	<0.0001

Block	Catch LSMEAN	SE	*P* > |*t*|
North	2.10581833	0.13522076	<0.0001
South	2.81951176	0.13664751	<0.0001

Results of this analysis suggest that walleye catch rates differed significantly among daily bag limits ($F = 4.91$; $P = 0.0199$) and between northern and southern Wisconsin lakes ($F = 13.90$; $P = 0.015$). Importantly, differences in walleye catch rates between northern and southern Wisconsin lakes obscured the effect of daily bag limits when the data were analyzed using a fully randomized design (Box 3.8).

Often, however, potential blocking variables are not categorical but are measured on a continuous scale. For example, lakes can be arbitrarily categorized as "small" and "large" based on their surface area but could also be measured in surface area on a continuous scale. The GLM treats these variables as the independent variable in a regression, and inferences based on these models evaluate the effects of a treatment on the response variable, after accounting for the effect of continuous variable(s) on the response variable (Quinn and Keough 2002). This particular application of a GLM is often termed analysis of covariance (ANCOVA).

As in a randomized block design, one of the main advantages of using an ANCOVA design is that unexplained variability in the response variable is reduced, thereby providing greater statistical power to detect and estimate the effects of treatments. A further advantage is that the results provide insight into the effects of covariates and potential interactions with the treatment variable. Thus, a greater understanding of the experimental system can be obtained with ANCOVA. Another advantage of an ANCOVA design is that the value of the independent covariate is not known a priori but is determined at the time of the experiment, thereby reducing the need to have the entire sampling frame sorted into blocks prior to the experiment.

When using ANCOVA, several assumptions must be met. First, a linear relationship must exist between the response variable (e.g., weight) and the independent variable used as a covariate (e.g., length; Montgomery 2001; Quinn and Keough 2002). We must also assume that the covariate values are similar among treatments (Quinn and Keough 2002). The important implication of this assumption is that ANCOVA should not be used to correct for different values of the covariate in each group (Quinn and Keough 2002). For example, if initial lengths of fish in an experiment were different, we should not include initial lengths to correct for this difference (Quinn and Keough 2002). In regression, we must assume that the covariate, x, is fixed and measured without error (Quinn and Keough 2002).

Because continuous factors are included in ANCOVA, the statistical model looks somewhat different than the model for fully randomized or randomized block designs. Multiple expressions of the ANCOVA model are available, but we prefer

$$y_{ij} = \beta_0 + \alpha_i + \beta x_{ij} + \varepsilon_{ij}, \tag{3.17}$$

where y_{ij} are the observations, β_0 is the intercept for the regression between x and y for the population as a whole, α_i is the treatment effect for each level, β is the common slope for the regression between x_{ij} and y_{ij}, and ε_{ij} is the unexplained variation among experimental units. This model can then be use to test the H_0 that all treatment effects (α_i) are the same:

$$H_0 : \alpha_1 = \alpha_2 = \ldots = 0.$$

An equivalent way of thinking about the H_0 is that all of the regressions between the response variable (y) and the covariate (x) have the same intercept. However, before testing the H_0, we must first test whether the slopes of the regression lines for all treatments are the same (i.e., all regression lines are parallel; Neter et al.

1996; Quinn and Keough 2002). To test the hypothesis that slopes are equal for all treatment levels, the interaction between the fixed effect and the continuous variable (covariate) is evaluated using the slope heterogeneity test (Box 3.11; Neter et al. 1996; Quinn and Keough 2002). If the interaction is significant, then the ANCOVA model (equation [3.17]) does not apply and separate regression models should be fit to each treatment level and then compared (Neter et al. 1996). If the interaction is not significant, then slopes are assumed to be equal, and the ANCOVA model (equation [3.17]) is estimated.

The use of ANCOVA is common in the fisheries literature. Many analyses include the effect of a covariate that is important for understanding the effects of one or more fixed treatments. This is especially common when the treatment effect may be influenced by growth in either length or weight. The other common use of ANCOVA in fisheries is to determine how data can be grouped. In many instances, an analyst will be uncertain whether the covariate in question affects the results of an outcome. For instance, Beard et al. (1997) used ANCOVA when building a predictive model of angler catch rate from walleye density to determine if walleye density differed among length-limit regulation categories and years sampled. When length category and years sampled were not significant, walleye densities were grouped together regardless of length regulation and year (Beard et al. 1997).

3.3.4 Random Effects and Mixed Models

In a randomized block design, the entire population of interest is broken into subgroups (blocks) from which units are selected for treatment. In Box 3.10, for example, all lakes in Wisconsin were designated as coming from the northern or southern part of the state. Thus, a randomized block design is analogous to a stratified random sampling design. In many fisheries investigations, blocking is used to reduce variability, but samples are not collected from all blocks within the population. For example, in an experiment to evaluate the effects of a herbicide application on density of age-0 bluegill, we might randomly select five lakes to receive a herbicide treatment and five lakes to receive no treatment. This could be repeated for 4 years, resulting in 20 treatment lakes and 20 control lakes (Box 3.12). We suspect that recruitment of age-0 bluegill may vary annually because of factors such as weather. Thus, we could use year as a blocking factor. However, our interest lies not just in the years selected for study but also in future years. Thus, the blocks (i.e., year) constitute only a sample of all possible years of interest. Such factors are analogous to clusters in cluster sampling (Figure 3.2). When this is the case, the blocking factor is appropriately treated as a random effect in the statistical model. A model that includes both random effects (years, in this example) and fixed treatments (also known as fixed effects; herbicide treatments, in this example) is referred to as a mixed model.

The statistical model for simple mixed models (i.e., with only a single fixed and a single random effect) is similar to that for the randomized block:

$$y_{ijk} = \mu + \alpha_i + \beta_k + \varepsilon_{ijk}, \tag{3.18}$$

Box 3.11 Example of an Analysis of Covariance Design

The goal of this study was to determine how substrate size affected early growth of brook trout eggs. In a lab experiment, a fisheries scientist placed individual brook trout eggs into containers with different substrates. The investigator also believed that egg diameter would affect early growth, so egg size was measured as a continuous covariate. An analysis of covariance model with egg diameter as the continuous variable and substrate as the categorical treatment variable follows.

Program

```
   *The following data step creates a data set named growth containing the
data that follow;
   data growth;
   input id substrate $ egg_diameter growth;
   cards;
    1          Cobble            8.3           20.0
    2          Cobble            8.5           23.5
    3          Cobble           11.2           24.7
    4          Cobble           10.7           29.5
    5          Cobble            9.6           24.3
    6          Cobble           11.8           31.7
    7          Cobble            9.6           22.1
    8          Cobble            8.9           19.0
    9          Cobble           11.2           17.3
   10          Cobble            8.9           23.3
    1          Gravel           10.3           36.4
    2          Gravel            9.5           25.7
    3          Gravel            8.5           13.6
    4          Gravel            9.9           33.9
    5          Gravel            8.6           17.1
    6          Gravel            8.9           22.6
    7          Gravel           10.4           32.0
    8          Gravel           10.8           40.2
    9          Gravel            9.9           26.6
   10          Gravel           10.1           32.9
    1          Sand              9.3           20.4
    2          Sand              8.8           15.3
    3          Sand              9.2           21.6
    4          Sand             10.0           22.9
    5          Sand             10.5           21.2
    6          Sand             10.2           17.4
    7          Sand              9.4           12.4
    9          Sand             10.7           21.8
   10          Sand             11.8           25.0
   ;
   run;
   *These statements call the GLM procedure in SAS, declaring the variable
substrate to be a categorical predictor variable and growth to be the
continuous response variable. By default, the variable egg_diameter is
treated as a continuous predictor variable;
   proc glm;
   class substrate;
   model growth=substrate egg_diameter egg_diameter*substrate;
   run;
```

(Box continues)

Box 3.11 *(continued)*

Results and Interpretation

The results of the analysis follow.

Table The GLM procedure for the dependent variable growth of brook trout. The variable substrate is a categorical predictor variable and the variable egg diameter is treated as a continuous predictor variable.

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	5	1025.104031	205.020806	17.64	<0.0001
Error	23	267.387693	11.625552		
Corrected total	28	1292.491724			
R^2	0.793122	Root MSE	3.409626		
CV	14.23951	Growth mean	23.94483		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Substrate	2	266.1171709	133.0585855	11.45	0.0004
Egg_diameter	1	557.0652164	557.0652164	47.92	<0.0001
Egg_diameter*substrate	2	311.4589773	155.7294887	13.40	0.0001

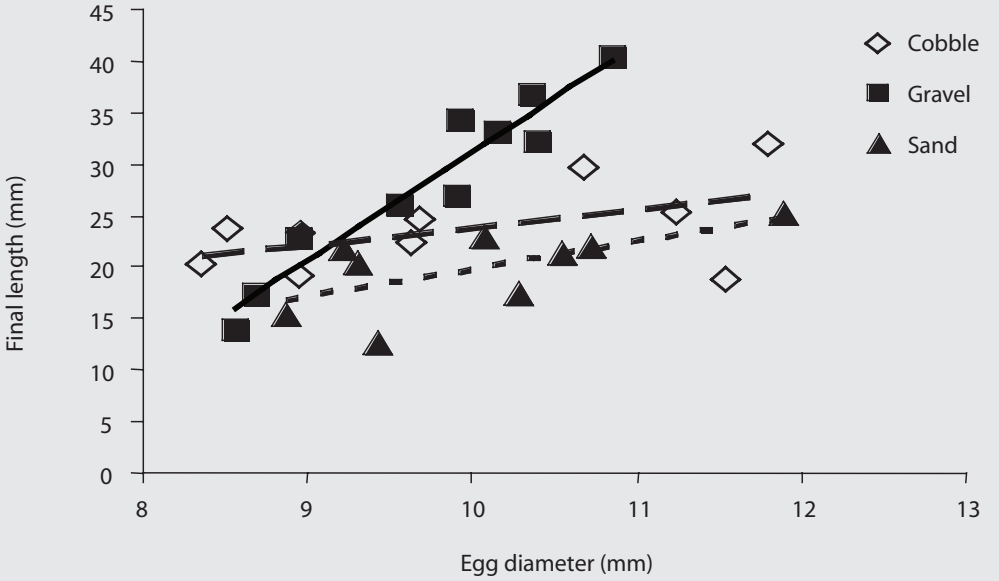
Results of the analysis indicate a significant interaction between egg diameter and substrate (F = 13.40; P = 0.0001), which indicates that egg diameter did not influence final length the same for all substrate classes. Therefore, significance of the two main effects, substrate (F = 11.45; P = 0.0004) and egg diameter (F = 47.92; P < 0.0001), cannot be interpreted because the ANCOVA model (equation [3.17]) does not apply. If the interaction is significant, separate regression models should be fit to each treatment level and then compared.

where y_{ijk} are the observations, μ is the population mean of the response variable, α_i is the treatment effect for each level, β_k is the effect for each level of the random effect variable, and ε_{ijk} is the unexplained variation among experimental units. This model can then be use to test the H_0 that all treatment level means (specified as α_i) are the same, taking into account the effect of the random variable(s):

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_i = \mu.$$

Mixed models have been used occasionally in fisheries investigations but have also not been used when they would be appropriate. The most common mistake is to treat a random factor as a fixed effect, with the consequence that type I errors are underestimated. Buynak and Mitchell (2002) provide an example where a mixed model was applied in a fishery experiment. The study was designed

With a significant interaction between a continuous and categorical variable, the best way to interpret the results is graphically. For these data, growth generally increased with egg diameter, but the increase was higher in gravel substrate (steeper slope) than it was in cobble or sand substrate (shallower slope). In cobble and sand substrates, the relationship between egg diameter and growth was consistent (similar slopes). In addition, growth was higher in gravel than in sand or cobble for egg diameters greater than about 9.5 mm but lower in gravel than in sand or cobble for smaller egg diameters.



to determine the effects of a slot size limit on smallmouth bass populations. In this study, Buynak and Mitchell (2002) set up a mixed-effects model that tested for differences in density between length limit treatment sites (slot size limit versus no slot size limit) and across years. Year was considered a random variable in this model, because Buynak and Mitchell were interested in determining if the effect of years was the same for all years or differed among years (Buynak and Mitchell 2002).

3.3.5 Factorial Design

The factorial design is used when an investigator wants to investigate the effects of more than one factor on the response variable. In a factorial design, each complete trial of the experiment explores all possible combinations of the levels of

Box 3.12 Example of a Mixed-Model Design

The goal of this study was to determine the effect of herbicide treatment on the abundance of age-0 bluegill in lakes. In theory, treatment with herbicide will create greater access to food resources, so abundance of age-0 bluegill should increase. Funds were available for treating and sampling only four lakes each year, along with sampling an equivalent number of untreated control lakes. To increase the sample size available for the experiment, the fisheries scientists treated lakes over 4 years but were concerned that year-to-year variation in weather could obscure the real effect of treatment.

Program

```
   *This data step creates a data set named herb that contains the following
data;
   data herb;
   input year herbicide $ 13-22  lake_id bluegill_yoy;
   cards;
    2001        Treatment        988          86
    2001        Treatment        116         100
    2001        Treatment        375         163
    2001        Treatment         17         135
    2001        Control          592          62
    2001        Control          677          69
    2001        Control          850          56
    2001        Control          566          50
    2002        Treatment        814         172
    2002        Treatment        397         200
    2002        Treatment        175         204
    2002        Treatment        867         153
    2002        Control          557          51
    2002        Control          106         122
    2002        Control          770          42
    2002        Control          111         127
    2003        Treatment        291         117
    2003        Treatment         76         125
    2003        Treatment         35         153
    2003        Treatment        997         123
    2003        Control          385          89
    2003        Control          712         106
    2003        Control          551          34
    2003        Control          567         197
    2004        Treatment        532          83
    2004        Treatment        424          65
    2004        Treatment        908          59
    2004        Treatment        369          69
    2004        Control          192         137
    2004        Control          371          66
    2004        Control          623          28
    2004        Control          515          23
   ;
   run;
```

```
   *These statements call the MIXED procedure in SAS, declaring herbicide and
year to be categorical predictor variables, and bluegill_yoy (age-0) to be a
continuous response variable. The model statement indicates that the Kenward-
Roger method should be used for computing the degrees of freedom. The random
statement identifies year as a random effect, and the lsmeans statement
requests least-squares means for bluegill density for the different levels of
herbicide treatment;
   proc mixed  covtest;
   class herbicide year;
   model bluegill_yoy = herbicide / ddfm=kenwardroger;
   random year/solution;
   lsmeans herbicide;
   run;
```

Results and Interpretation

Results of this analysis are as follow.

Table The mixed procedure of SAS. Herbicide treatment (fixed effect) and year (random effect) are predictor variables, and age-0 bluegill density is the continuous response variable. The convergence criteria were met. Abbreviations are given for –2 · residual log likelihood (–2Res log like); Akaike's Information Criteria (AIC); small sample corrected AIC (AICc); and Bayesian Information Criteria (BIC). Note a smaller value is better for the information criteria indices.

Iteration History

Iteration	Evaluations	–2Res log like	Criterion
0	1	321.73349015	
1	1	317.52062502	0.00000000

Covariance Parameter Estimates

Covariance parameter	Estimate	SE	*Z*-value	*P* > *Z*
Year	689.45	734.59	0.94	0.1740
Residual	1660.62	451.96	3.67	0.0001

Fit Statistics

–2Res log like	317.5
AIC	321.5
AICc	322.0
BIC	320.3

(Box continues)

Box 3.12 *(continued)*

Solution for Random Effects

Effect	Year	SE estimate	Prediction	*df*	*t*-value	$P > \|t\|$
Year	2001	–9.1751	18.5847	3.77	–0.49	0.6489
Year	2002	24.4509	18.5847	3.77	1.32	0.2626
Year	2003	12.2495	18.5847	3.77	0.66	0.5479
Year	2004	–27.5253	18.5847	3.77	–1.48	0.2169

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	$P > F$
Herbicide	1	27	10.53	0.0031

Least-Squares Means

Effect	Herbicide	Estimate	SE	*df*	*t*-value	$P > \|t\|$
Herbicide	Control	78.6875	16.6178	4.52	4.74	0.0067
Herbicide	Treatment	125.44	16.6178	4.52	7.55	0.0010

Results of the analysis suggest that application of herbicide significantly increased the relative abundance of age-0 bluegill ($F = 10.53$; $P = 0.0031$). In control lakes, the mean catch of age-0 bluegill was 78.7 with a SE of 16.6, whereas in treated lakes the mean catch of age-0 bluegill was 125.4 with a SE of 16.6. The effect of the random year effect was not so clear because the covariance estimate for the year effect was 689.45, but the covariance had a SE of 734.59 and a *P*-value of 0.1740. Although this *P*-value is greater than the often-used 0.05, accounting for the potential effects of years is likely an important structural component of the design, and therefore, year should still be included in the model.

factors investigated (Montgomery 2001). For example, an experiment with a levels of factor A and b levels of factor B (where A and B are main effects) includes $a \times b$ treatment combinations. In factorial designs, main effects are generally of primary interest, and if no interactions are present between or among main effects, main effects are simple averages of the effects found for each treatment level (Cochran and Cox 1957; Montgomery 2001). In factorial designs, the factors are considered to be fixed effects (Quinn and Keough 2002).

As an example of a factorial design, the fisheries scientist of an aquaculture facility may be interested in exploring how stocking density and different feeding levels affect the yield of channel catfish in rearing ponds. The fisheries scientist could use only the lowest and highest stocking densities and three feeding levels for the fish, for six possible treatment combinations (Figure 3.5). The fisheries scientist randomly assigns ponds to each treatment combination and runs the experiment. If stocking levels and feeding levels do not interact, the interpretation of

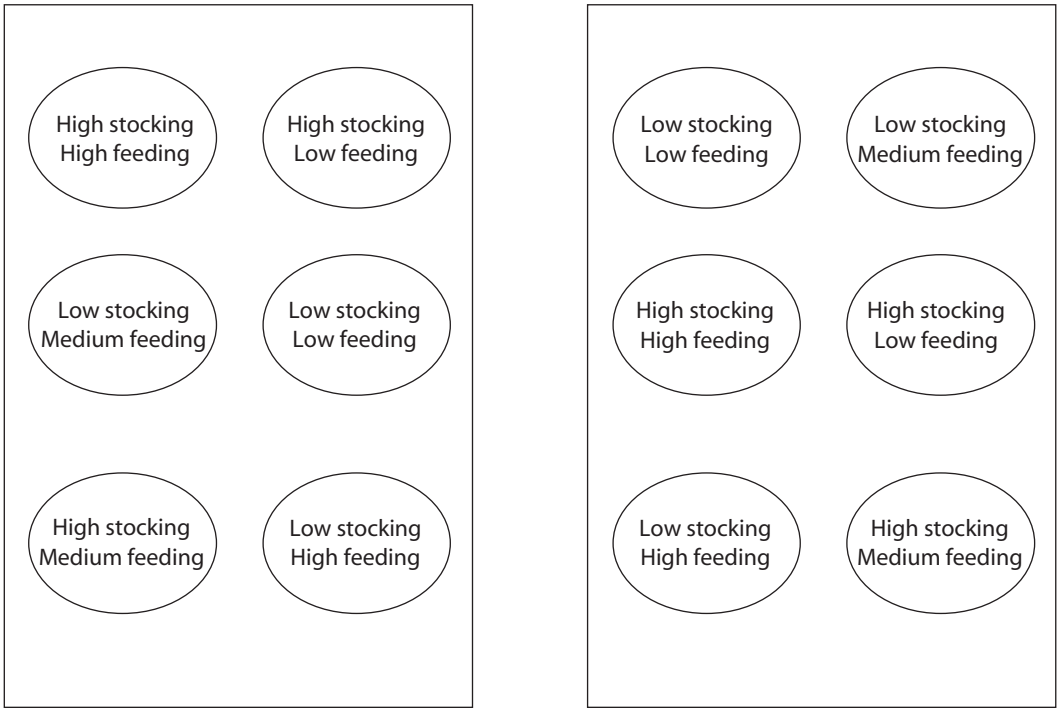


Figure 3.5 Example of a fully randomized factorial design for testing the effect of two stocking densities and three feeding rates. The entire experiment has two replicates for each combination of the factors.

the main effects is straightforward. For instance, the main effect of feeding level, given stocking densities, would be simply the difference between the averages of the results from high and low feeding levels, regardless of the stocking density. That is, we calculate the average for high feeding levels (across all units) and the average for low feeding levels (across all units) and subtract the mean of the low feeding levels from the high feeding levels. The effect is then interpreted as increasing factor *A* from a low to high feeding level causes an average effect equal to the difference between the means (Montgomery 2001). In the absence of interaction, main effects are additive so they are simple to calculate and interpret (Quinn and Keough 2002).

The factorial design helps to understand if and how the effects of each factor interact (Montgomery 2001). Failure to use a factorial design may lead to misinterpretation of results or failure to ascribe results to proper effects. With the presence of an interaction, the effect of factor *A* depends on the level of factor *B* (Montgomery 2001). For example, in our hatchery experiment, at a low stocking density fish may grow at a similar rate regardless of feeding level, but at a high stocking rate feeding rate may affect growth. Interactions are very common in fisheries science because main effects may have synergistic or antagonistic effects (Quinn and Keough 2002). Interactions can make interpretation of main effects difficult and often are easier to interpret when main effects are plotted.

Factorial designs have many advantages over other designs. Factorial experiments are especially useful when the goal of the experiment is to obtain a broad picture of the effects of the factors (Cochran and Cox 1957). If the factors are independent of one another, the factorial experiment can save considerable time and expense (Cochran and Cox 1957; Montgomery 2001). The factorial experiment is most often used in manipulative experiments and in exploratory work, where the factor effects are explored over a range of values (Cochran and Cox 1957; Quinn and Keough 2002).

The statistical model for a simple factorial design with two factors is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \alpha\tau_{ij} + \varepsilon_{ijk}, \tag{3.19}$$

where y_{ijk} are the observations, μ is the population mean of the response variable, α_i is the treatment effect for each level of the first factor, τ_j is the effect for each level of the second factor, $\alpha\tau_{ij}$ is the interaction between main effects, and ε_{ijk} is the unexplained variation among experimental units. This model can then be use to test the H_0 that the means for each level of each factor are the same and that the interaction between the factors is 0.

$$\begin{aligned} H_0 &: \alpha_1 = \alpha_2 = \ldots = \alpha_i; \\ &\tau_1 = \tau_2 = \ldots = \tau_j; \text{ and} \\ &\alpha_1\tau_1 = \alpha_1\tau_2 = \alpha_2\tau_1 = \alpha_2\tau_2 = \ldots = \alpha_i\tau_j. \end{aligned}$$

Factorial designs are commonly used in studies of fisheries management and ecology (Box 3.13). For example, Nowlin and Drenner (2000) used mesocosms to examine the effects of the presence or absence of a planktivore in conjunction with the presence or absence of a fish assemblage on zooplankton densities. Similarly, Dahl (1998) used a factorial design to evaluate the effects of benthivory on benthic assemblages by enclosing standard lengths of stream and then examining the invertebrate assemblage in streams sections with no fish, bullheads, brown trout, and brown trout plus bullheads. In a more complex design, Drenner et al. (1998) examined the effects of nutrient loading, levels of omnivory, and levels of clay on phytoplankton biomass present in mesocosms. In all of these experiments, interactions between factors were suspected, so factorial designs were necessary to understand the effects.

Factorial designs are also beneficial in other types of experiments. For example, Aas et al. (2000) used a factorial design on results of a mail survey to produce hypothetical profiles of fishing opportunities that were based on fishing regulations and expectations of anglers who fished certain waters. Factorial experiments can also be used in computer modeling. For example, Sampson and Yin (1998) used computer simulations of a fractioned factorial design to examine the effects of natural mortality, fishing mortality, and recruitment on the demographic history of a fishery. Factorial designs are common in fisheries, although investigators may not refer to their designs as factorial. If multiple

Box 3.13 Example of a Factorial Design

The goal of this study was to determine how size and stocking location of fingerling Chinook salmon affected survival and subsequent return to the Snake River. Bugert and Mendel (1997) used a 2 × 2 factorial design in which size (subyearling versus yearling) and location of release (on-station versus off-station) were compared to see how these factors affected survival. For this example, we have included only years when all treatment combinations were implemented.

Program

```
data chinook;
input year size$ release$ survival;
cards;
1987 Sub          On    .058
1987 Sub         Off    .155
1987 Yearling     On    .406
1987 Yearling    Off    .319
1988 Sub          On    .058
1988 Sub         Off    .004
1988 Yearling     On    .350
1988 Yearling    Off   1.376
1989 Sub          On    .014
1989 Sub         Off    .008
1989 Yearling     On    .092
1989 Yearling    Off    .320
1990 Sub          On    .047
1990 Sub         Off    .044
1990 Yearling     On    .599
1990 Yearling    Off   3.048
;
run;
```

Because survival was expressed as a percentage, the data were first transformed using the arcsine transformation. The program used to analyze these data follows.

```
data chinook1;
set chinook;
arcsurv=arsin(survival/100);
run;

proc glm;
class size release;
model arcsurv=size release size*release;
lsmeans size release size*release/stderr;
run;
```

Results and Interpretation

Results of this analysis are as follow.

(Box continues)

Box 3.13 *(continued)*

Table The GLM procedure for a 2 × 2 factorial design to assess fingerling Chinook salmon survival with size (subyearling versus yearling) and location of release (on-station versus off-station) as factors (based on Bugert and Mendel 1997). The dependent variable is the arcsin transformation of the percent survival (arcsurv); the number of observations is 16.

Class Level Information

Class	Levels	Values
Size	2	Sub Yearling
Release	2	Off On

General Linear Model

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	0.00039779	0.00013260	3.10	0.0672
Error	12	0.00051284	0.00004274		
Corrected total	15	0.00091063			
R^2	0.436827	Root MSE	0.006537		
CV	151.6227	Arcsurv mean	0.004312		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Size	1	0.00023428	0.00023428	5.48	0.0373
Release	1	0.00008329	0.00008329	1.95	0.1880
Size*release	1	0.00008021	0.00008021	1.88	0.1958

Least-Squares Means

Size	Arcsurv LSMEAN	SE	*P* > |*t*|
Sub	0.00048500	0.00231130	0.8373
Yearling	0.00813815	0.00231130	0.0042

Release	Arcsurv LSMEAN	SE	*P* > |*t*|
Off	0.00659315	0.00231130	0.0146
On	0.00203001	0.00231130	0.3970

Size	Release	Arcsurv LSMEAN	SE	*P* > |*t*|
Sub	Off	0.00052750	0.00326866	0.8745
Sub	On	0.00044250	0.00326866	0.8946
Yearling	Off	0.01265879	0.00326866	0.0022
Yearling	On	0.00361751	0.00326866	0.2901

Results of the analysis suggest that size at stocking significantly affected survival of juvenile Chinook salmon ($F = 5.48$; $P = 0.0373$), but that release location did not significantly affect survival ($F = 1.95$; $P = 0.1880$). Further, the interaction between release location and size at stocking was not significant ($F = 1.88$; $P = 0.1958$).

factors and interactions are included in the study design, a factorial design is very likely the basis for the experiment.

3.3.6 Nested Design

Nested designs often occur when subsamples are taken from the experimental units included in a study. A common type of nesting in fisheries research is to have individual sample sites nested within lakes or streams. In such studies, individual lakes or streams are experimental units, but we need to account for variation among sites within each lake or stream. Similarly, individual fish sampled from a lake are generally not true replicates but should be treated as a nested subsample. For example, in Figure 3.6, fish are nested subsamples within lakes,

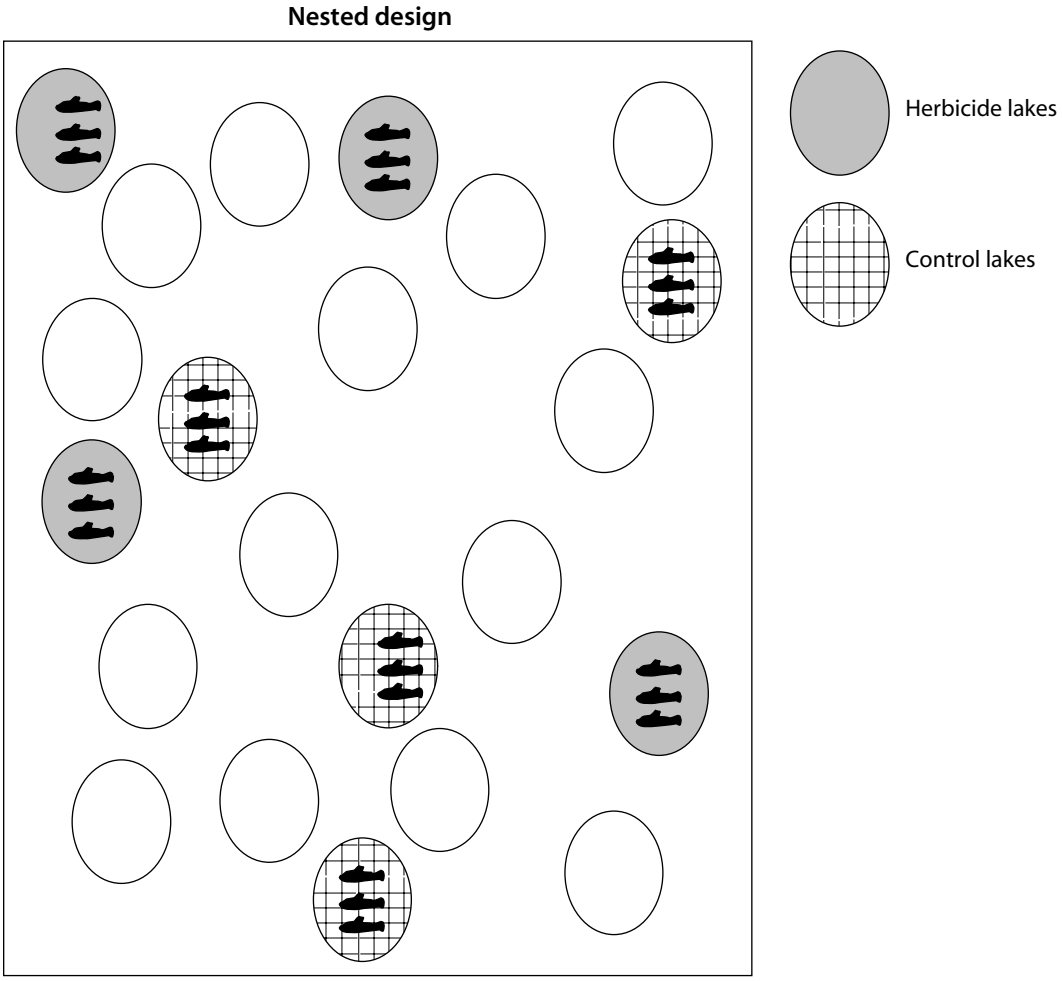


Figure 3.6 Example of a nested design in which lakes are the main experimental unit and individual fish are nested subsamples within each lake. Open circles indicate lakes that were not included in the study as either herbicide lakes or control lakes.

which were randomly selected for application of herbicide or held as control lakes. To account for the experimental design properly, a nested effect needs to be included in the statistical model. The nested effect variable is generally a categorical identifier for each experimental unit. In nested designs, the effects can be either random or fixed, but in the biological sciences the main effect (e.g., the treatment applied) is often fixed and the nested effect (e.g., individual lake identifier) is often random (Quinn and Keough 2002).

The experimental units can sometimes be difficult to identify properly in nested designs (Yandell 1997). For example, in a design in which sampling transects are nested within habitat types in a lake, the primary experimental units are habitat patches, not sampling transects or the lake. Difficulties in identifying primary experimental units in nested designs can also lead to pseudoreplication (Hurlbert 1984), where subsamples (e.g., fish in nets) are confused with truly replicated experimental units (e.g., nets in lakes).

The statistical model for nested designs is similar to that for mixed models (section 3.3.4); an example of a nested design is given in Box 3.14. Nested designs are common in fisheries. The design discussed above, with sampling transects or locations nested in streams or lakes, is appropriate when comparing effects across water bodies. For example, to determine how various benthic taxa varied at different spatial scales, Boyero and Bailey (2001) used a nested design with sampling points nested within riffles nested within streams. Boyero and Bailey (2001) were able to attribute the variation in taxa to these different spatial scales. Using a similar approach, Cole (2001) nested sample cells of different sizes to assess spatial variability in the abundance of clams. Pierce et al. (2001) used a nested design to examine differences in species richness in relation to diel sampling period, sampling gears, and sites, all nested within each lake sampled. Using a random-effects model, Radomski and Goeman (2001) nested developed and undeveloped lakeshore plots within lake development classes to quantify differences in vegetative abundance among lakes and between shoreline types. Toepfer et al. (1999) nested individual leopard darter results within separate trials to separate individual variation in burst speed and numbers when assessing overall swimming performance. Conover et al. (1997) used a nested design to attribute variance in growth rate of young striped bass to individual mothers, nested within the latitude from which they came, to separate genetic and physiological effects of each mother from the effect of latitude.

3.3.7 Split-Plot Design

In a split-plot design, the main experimental units are divided into two or more parts (Cochran and Cox 1957). Different levels of treatments are then applied to a subunit within the main experimental unit. This type of design is similar to a randomized block design, except in the randomized block, the treatment combinations are assigned randomly, not randomly within each main plot (Cochran and Cox 1957). For example, consider a hatchery experiment with two levels of stocking density (high and low) and two feeding levels (high and low). In a split-plot

Box 3.14 Example of a Nested Design

For the example in Box 3.12, where the effect of herbicide treatment on age-0 bluegill density was investigated, we may also be interested in how herbicide treatment affects mean length of age-0 bluegill at the end of the growing season (for this example, assume that length of individual bluegill from each lake in the study was measured). In a nested design, the primary experimental unit is a lake, so each bluegill is not an independent replicate but rather is a subsample from the lake. For brevity, only the lakes sampled in 2001 from Box 3.12 are used in this example.

Table Hypothetical data on lengths of age-0 bluegills from lakes treated with herbicide and control lakes that were not treated with herbicide (an extension of Box 3.12 data).

	Length of age-0 bluegills							
	Treatment lakes				Control lakes			
Summary statistic	988	116	375	17	592	677	850	566
	103	88	97	116	70	83	102	79
	90	95	94	94	79	85	89	72
	98	82	103	112	78	92	82	67
	90	100	94	111	85	85	86	78
	96	84	83	96	65	84	99	83
	88	92	93	111	68	83	88	68
	97	94	90	91	93	79	99	87
	100	79	107	116	80	90	80	75
	89	103	94	109	89	77	81	79
	108	81	86	110	65	85	93	90
Mean	95.9	89.8	94.1	106.6	77.2	84.3	89.9	77.8

Program

The SAS program used to analyze these data follows.

```
  data bluegill;
  input herbicide $ 1-9 lake length;
  cards;
  Treatment        988     103
  Treatment        988      90
  (input data)
  ;
  run;

  * This call to the MIXED procedure is much like in Box 3.12, except that
the random statement is used to indicate that individual bluegills within a
lake are a subsample from a herbicide treatment class;
  proc mixed covtest;
  class lake herbicide;
  model length=herbicide;
  random lake(herbicide);
  lsmeans herbicide;
  run;
```

(Box continues)

Box 3.14 *(continued)*

Results and Interpretation

Results of the analysis follow.

Table The effect of herbicide treatment on age-0 bluegill length. The mixed procedure with the random statement is used to indicate that individual bluegills within a lake are a subsample from an herbicide treatment class. Convergence criteria were met. The estimation method was restricted maximum likelihood (REML).

Model Information	
Data set	WORK.BLUEGILL
Dependent variable	Length
Covariance structure	Variance components
Estimation method	REML
Residual variance method	Profile
Fixed effects SE method	Model-based
Degrees of freedom method	Containment

Class Level Information

Class	Levels	Values
Lake	8	17 116 375 566 592 677 850 988
Herbicide	2	Control treatment

Dimensions	
Covariance parameters	2
Columns in X	3
Columns in Z	8
Subjects	1
Maximum observations per subject	80
Observations used	80
Observations not used	0
Total observations	80

experiment, the investigator would randomly select a stocking density for each of four ponds, divide the ponds in half with barriers, and randomly select a feeding rate to apply to each half of each pond (Figure 3.7). The sample size for feeding rate increased from two, using a factorial design with four ponds, to four, using the split-plot design.

When performing a split-plot experiment, the B effect and $A \times B$ interaction (the feeding rate and feeding rate × stocking density effects in the hatchery experiment) are estimated more precisely than are the A effects (stocking density; Cochran and Cox 1957). As described with the hatchery experiment, the degrees of freedom are smaller for the whole unit than for the subunit comparisons. The

Iteration History

Iteration	Evaluations	–2Res log like	Criterion
0	1	579.91706840	
1	1	561.75492286	0.00000000

Covariance Parameter Estimates

Covariance parameter	Estimate	SE	*Z*-value	*P* > *Z*
Lake (herbicide)	37.3500	25.1356	1.49	0.0686
Residual	61.5000	10.2500	6.00	<0.0001

Fit Statistics

–2Res log like	561.8
AIC	565.8
AICc	565.9
BIC	565.9

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	*P* > *F*
Herbicide	1	6	9.40	0.0220

Least-Squares Means

Effect	Herbicide	Estimate	SE	*df*	*t*-value	*P* > |*t*|
Herbicide	Control	82.3000	3.2977	6	24.96	<0.0001
Herbicide	Treatment	96.6000	3.2977	6	29.29	<0.0001

Results of the analysis indicate that age-0 bluegill in control lakes were significantly shorter (82.3 mm) than in herbicide-treated lakes (96.6 mm; $F = 9.40$; $P = 0.0220$).

primary advantage of the split-plot design is realized when the B and $A \times B$ effects are of greater interest than is the A effect, or when the A effect cannot be tested on small experimental units because the cost or size of the A experimental unit is prohibitive (Cochran and Cox 1957; Montgomery 2001). For example, in the hatchery example the primary interest was in the feeding rate and feeding rate × stocking density effects, whereas the secondary interest was in the stocking density effect. However, the increase in precision of estimating B effects can often lead to results where the effect of factor B is significant and the effect of factor A is not significant (Cochran and Cox 1957). Analysis of data collected using a split-plot design are often complicated and require detailed coding of data to assure analysis

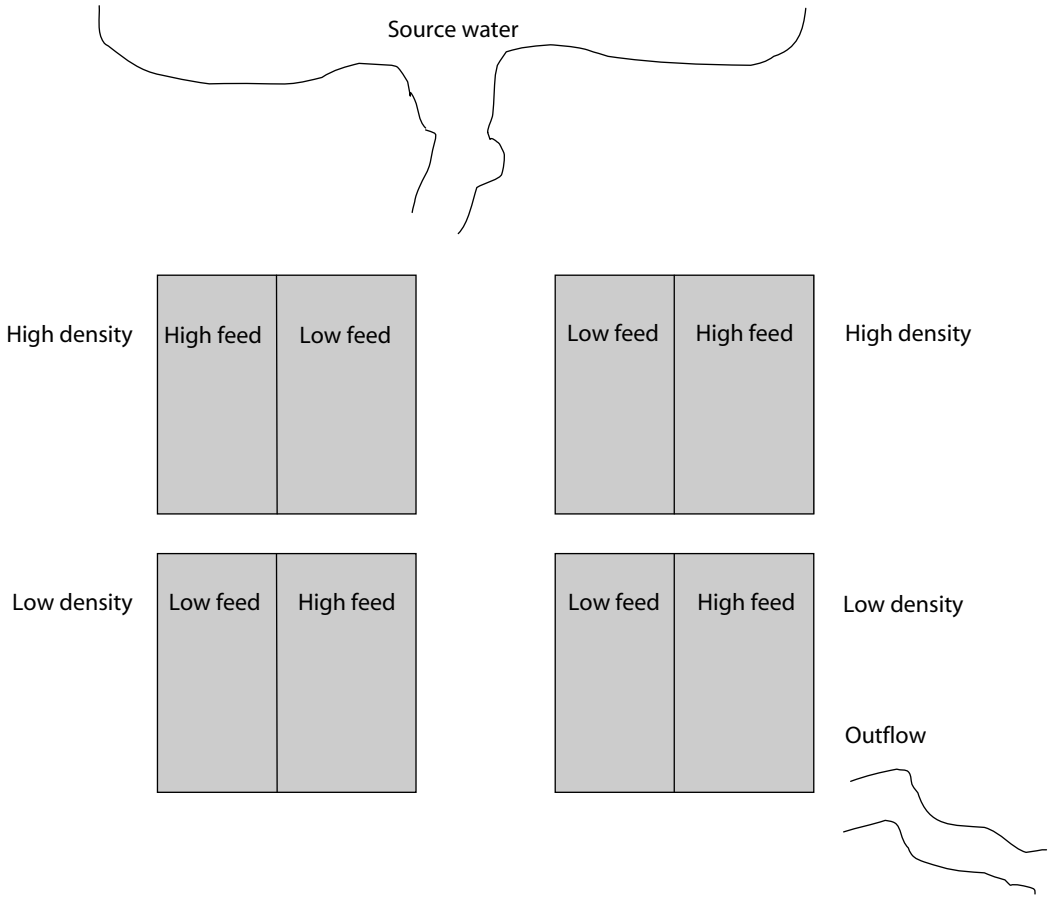


Figure 3.7 Example of a split-plot design to test the effects of feeding and stocking density on growth rates of muskellunge in hatchery ponds. In this experiment, four ponds are each divided in half and two ponds each are randomly selected for high and low stocking density. Within each pond, each side is randomly selected for high or low feeding rates. The overall design includes two replicates for stocking density and four replicates for feeding rate.

programs work correctly. Split-plot designs often contain a mixture of random (e.g., ponds) and fixed effects (e.g., feeding or stocking rates), which may further complicate analysis (Quinn and Keough 2002).

Split-plot designs are uncommon in fisheries, though the repeated-measures split-plot design has been widely applied (Box 3.15; Maceina et al. 1994). More details of that design will be covered in the next section because it combines the aspects of the split-plot and repeated-measures design. An example of a split-plot design by Secor et al. (2000) tested differences in growth performance between anadromous and nonanadromous strains of striped bass; a split-plot design was used to separate growth and salinity effects at three levels of growth and salinity.

3.3.8 Repeated-Measures Design

The repeated-measures design generally refers to experiments in which individual experimental units are observed more than once (Quinn and Keough 2002). In some cases, a single treatment is applied and the experimental unit is observed over time, but in other cases multiple treatments are applied and the experimental unit is observed multiple times. When the same experimental units are observed multiple times, the observations on the response to a treatment are potentially correlated because the same experimental unit is used (Quinn and Keough 2002). Thus, observations are not necessarily independent, and the design and analysis should take this into account. In such experiments, the treatment is typically considered to be a fixed effect and the subject is often a random effect (Montgomery 2001). As with split-plot designs, repeated-measures designs are often complex, and the analysis depends on the details of the situation (e.g., how many times the units are observed and how observations are correlated in time).

Repeated-measures designs are not commonly used in fisheries. A repeated-measures design was used to estimate the retention rate of coded wire tags in paddlefish, which were marked in four locations with coded wire tags. Each individual fish then was examined monthly to determine if tags had been retained or lost in each location, and total tag retention rate was estimated (Fries 2001). In an experiment to train grass carp to respond to different types of sound, Willis et al. (2002) used sound at different frequencies to determine if response varied with frequency. Because individual response types were measured on individual grass carp, the type of sound needed to be corrected for the measurement from individual grass carp to assure that individual grass carp behavior was taken into account in analysis of response to the type of sound (Willis et al. 2002).

In fisheries, repeated-measures designs often refer to a specialized version of the split-plot design in which the repeated measures are taken from the same set of sites (Box 3.15; Maceina et al. 1994). The sites selected are usually thought of as random effects in such designs, so repeated-measure designs are essentially split-plot designs that allow for correlation within each nested random effect (Yandell 1997). The repeated-measures design or the repeated-measures split-plot design often assigns treatments to experimental units, which are then measured over different time intervals (which become the plots). The main difference between a split-plot design and a repeated-measures split-plot design is that the split-plot design allocates within-plot treatments to subunits within each plot, whereas the repeated-measures split-plot design allocates within-subjects treatments sequentially to each subject (Quinn and Keough 2002). In a fisheries experiment that uses a repeated-measures split-plot design, sampling stations are often fixed, so treatments are measured repeatedly at the same site (correlation is present) with interactions between site, treatment, and time (Maceina et al. 1994).

The use of repeated-measures split-plot designs has become common in fisheries because of interest in time period effects of sampling at fixed sites (Maceina et al. 1994). Maceina et al. (1994) were the first to advocate use of repeated-measures split-plot designs in fisheries. In one experiment, Maceina et al. (1994) quantified

Box 3.15 Example of Repeated-Measures Split-Plot Design

The goal of this study was to determine the effects of vegetation removal by grass carp on fish biomass. Maceina et al. (1994) sampled the same six coves twice before and twice after treatment. Main plot *A* included cove, treatment, and cove*treatment interaction effects, and subplot *B* included time and time*treatment interaction effects. Maceina et al. (1994) popularized the use of repeated-measures split-plot designs in fisheries, which is appropriate for analyzing data collected through time at fixed stations. The analysis relies on standard analysis of variance techniques.

Program

```
data cove;
input year treat$ time cove area biomass;
cards;
1980 PRE  1 1 1.51 13854
1980 PRE  1 2  .67  4091
1980 PRE  1 3 2.19 17195
1980 PRE  1 4  .63  5138
1980 PRE  1 5  .64  5148
1980 PRE  1 6  .45  2971
1981 PRE  2 1 1.60  6374
1981 PRE  2 3 1.97 21441
1981 PRE  2 4  .74 17830
1981 PRE  2 5  .66  3577
1981 PRE  2 6  .32  2678
1985 POST 1 1 1.83 3209
1985 POST 1 3 2.39 11556
1985 POST 1 4  .88  8132
1985 POST 1 5  .70  5094
1985 POST 1 6  .49  1973
1986 POST 2 1 1.83 10643
1986 POST 2 2  .43   479
1986 POST 2 3 2.39 11103
1986 POST 2 4  .88  2852
1986 POST 2 5  .70  2489
1986 POST 2 6  .49  8898
 ;
data cove;
set cove;
logbio=log10(biomass);
run;

proc glm;
class cove treat time;
model logbio=cove treat treat*cove time treat*time;
      test h=treat e=treat*cove;
      test h=cove e=treat*cove;
run;
```

Results and Interpretation

The main fixed effects are cove and treatment, and main plot is split into time effects. The interactions were estimated to see if any spatial (treat*cove) or temporal (treat*time) correlations affected the results.

Table The GLM procedure to determine the effects of vegetation removal by grass carp on fish biomass. Six coves were sampled twice before (PRE) and twice after (POST) treatment. Main plot *A* included cove, treatment, and cove*treatment interaction effects, and subplot *B* included time and time*treatment interaction effects (based on Maceina et al. 1994). The dependent variable is $\log_{10}$biomass of fishes (logbio), and the number of observations was 22.

Class Level Information

Class	Levels	Values
Cove	6	1 2 3 4 5 6
Treatment	2	POST PRE
Time	2	1 2

The GLM Procedure

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	13	2.46565668	0.18966590	2.15	0.1406
Error	8	0.70628807	0.08828601		
Corrected total	21	3.17194474			
R^2	0.777333	Root MSE	0.297130		
CV	7.943952	Logbio mean	3.740325		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Cove	5	1.76323921	0.35264784	3.99	0.0409
Treat	1	0.36593102	0.36593102	4.14	0.0762
Cove*treat	5	0.43968044	0.08793609	1.00	0.4767
Time	1	0.01186994	0.01186994	0.13	0.7234
Treat*time	1	0.00436853	0.00436853	0.05	0.8295

Tests of Hypotheses with Type III MS for Cove*Treat as Error Term

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Treat	1	0.36593102	0.36593102	4.16	0.0969
Cove	5	1.76323921	0.35264784	4.01	0.0768

The type III SS, which are properly calculated using the cove*treat interaction MSE, indicate that the main fixed effects of coves ($F = 4.01$; $P = 0.0768$) and vegetation removal by grass carp treatments ($F = 4.16$; $P = 0.0969$) were significant at an alpha of 0.10, which suggests that the treatment affected fish biomass. Time ($F = 0.13$; $P = 0.7234$) and time*treatment ($F = 0.05$; $P = 0.8295$) effects were not important in explaining differences in fish biomass.

the abundance between years of age-0 black crappie and white crappie that were collected in trap nets at fixed stations over 2 d. The repeated-measures split-plot analysis treated stations as replicates, year as the main treatment effect, and day as the plot effect, along with interactions between years and stations and between days and years (Maceina et al. 1994). Using a similar approach, Pierce et al. (2001) determined the effect of the number of marked fish, station, year, and month on capture efficiency of beach seines for various species of fish. In this experiment, the replicate was the station, the fixed effect was the year, and the plot was the month (Pierce et al. 2001). Year and month were significant in explaining capture efficiency (Pierce et al. 2001). Jackson and Hightower (2001) used individual movement data from striped bass (the plot) to separate variance of individual fish from sex and season. They determined how sex and season affected site fidelity of striped bass (Jackson and Hightower 2001). Finally, to quantify spawning substrate preferences of yellow perch in Lake Michigan, Robillard and Marsden (2001) used a repeated-measures split-plot design that treated stations as replicates, year as the main effect, and substrate types as plots.

3.4 CONCLUSION

One point of potential confusion in experimental design is how to distinguish between fixed factors, fixed blocking effects, and random effects. In the hatchery example above, stocking density and feeding rate were factors assigned randomly to individual ponds because we were interested in understanding the response of the system to these factors. Thus, both of these factors are fixed effects. In contrast, the designation of lakes as being northern or southern in Box 3.10 is not something under the control of the investigator; all lakes are assigned to one of these two groupings before the start of the experiment. As such, this is an example of a fixed blocking factor. However, we may be interested in making predictions for lakes in the northern or southern part of the state, even if they were not present in our sample. We can appropriately make predictions for such a case because unsampled lakes must belong to either the northern or southern block. The trickiest situation is for random effects. Imagine, for example, that the state of Wisconsin had been subdivided into 20 different blocks, but we only selected lakes in three of the 20 blocks. We could still use region as a blocking factor, but the three selected blocks do not constitute the entire population (lakes) of 20 blocks, so we could not make predictions for a lake in one of the 17 blocks that were not included in the experiment. Treating the regional designation as a random effect appropriately allows us to take into account block-to-block variation, thereby enabling predictions about lakes in all 20 blocks.

Throughout this discussion, we have focused on true experiments where levels of treatment can be assigned at random by the investigator. However, in many fisheries studies, we are interested in how naturally varying factors affect fish populations, habitat, or anglers. Strictly speaking, such observational studies do not really fall into the category of an experimental design because we cannot infer cause and effect relationships from such studies. Observational

studies are common in fisheries science and yield insight into the dynamics of fishery systems. Many of the methods we have presented in this section are useful for the analysis of observational studies, but we caution the reader to recognize that the conclusions reached from such analyses are akin to correlation and do not imply causality.

As a final comment, the experimental designs presented here represent only a simple subset of the experimental designs used in practice. Elements of several designs are often used to achieve the goals of an experiment. For example, nested designs are frequently used with a factorial design. This occurs because our unit of measurement (e.g., individual fish) is often part of a larger experimental unit (e.g., lake or pond). Elements of repeated-measure designs are also frequently combined with other experimental design components to allow us to determine how experimental units vary over time in response to treatment. Because of the complexity of many experimental designs and analyses in fisheries, we recommend that you consult with a professional statistician before an experiment is started. This will assure that the proper experimental design is used and that the correct analyses techniques are considered and used. The analysis of data from more complex designs needs to be carefully considered but provides much deeper insights into the biology of fisheries systems.

3.5 REFERENCES

Aas, O. E., W. Haider, and L. Hunt. 2000. Angler responses to potential harvest regulations in a Norwegian sport fishery: a conjoint-based choice modeling approach. North American Journal of Fisheries Management 20:940–950.

Bates, D. M., and D. G. Watts. 1988. Nonlinear regression analysis & its applications. Wiley, New York.

Beard, T. D., Jr., S. P. Cox, and S. R. Carpenter. 2003. Impacts of bag limit reductions on angler effort in Wisconsin walleye lakes. North American Journal of Fisheries Management 23:1283–1293.

Beard, T. D., Jr., S. W. Hewett, Q. Yang, R. M. King, and S. J. Gilbert. 1997. Prediction of angler catch rates based on walleye population density. North American Journal of Fisheries Management 17:621–627.

Boyero, L., and R. C. Bailey. 2001. Organization of macroinvertebrate communities at a hierarchy of spatial scales in a tropical stream. Hydrobiologia 464:219–225.

Brown, M. L., and D. J. Austen. 1996. Data management and statistical techniques. Pages 17–62 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Bugert, R. M., and G. W. Mendel. 1997. Adult returns of subyearling and yearling fall chinook salmon released from a Snake River hatchery or transported downstream. North American Journal of Fisheries Management 17:638–651.

Buynak, G. L., and B. Mitchell. 2002. Response of smallmouth bass to regulatory and environmental changes in Elkhorn Creek, Kentucky. North American Journal of Fisheries Management 22:500–508.

Caputi, N. 1988. Factors affecting the time series bias in stock-recruitment relationships and the interaction between time series and measurement error bias. Canadian Journal of Fisheries and Aquatic Sciences 45:178–184.

Cochran, W. G. 1977. Sampling techniques, 3rd edition. Wiley, New York.

Cochran, W. G., and G. M. Cox. 1957. Experimental designs. Wiley, New York.

Cole, R. G. 2001. Staggered nested designs to identify hierarchical scales of variability. New Zealand Journal of Marine and Freshwater Research 35:891–896.

Conover, D. O., J. J. Brown, and A. Ehtisham. 1997. Countergradient variation in growth of young striped bass *(Morone saxatilis)* from different latitudes. Canadian Journal of Fisheries and Aquatic Sciences 54:2401–2409.

Dahl, J. 1998. Effects of a benthivorous and a drift-feeding fish on benthic stream assemblage. Oecologia 116:426–432.

Draper, N. R., and H. Smith. 1981. Applied regression analysis, 2nd edition. Wiley, New York.

Drenner, R. W., K. L. Gallo, R. M. Baca, and J. D. Smith. 1998. Synergistic effects of nutrient loading and omnivorous fish on phytoplankton biomass. Canadian Journal of Fisheries and Aquatic Sciences 55:2087–2096.

Efron, B., and R. J. Tibshirami. 1998. An introduction to the bootstrap. CRC Press LLC, Boca Raton, Florida.

Fries, J. N. 2001. Retention of coded wire tags in four locations in juvenile paddlefish. North American Journal of Fisheries Management 21:962–966.

Hartman, K. J., and S. B. Brandt. 1995. Estimating energy density of fish. Transactions of the American Fisheries Society 124:347–355.

Hayes, D. B., J. K. T. Brodziak, and J. B. O'Gorman. 1995. Efficiency and bias of estimators and sampling designs for determining length–weight relationships of fish. Canadian Journal of Fisheries and Aquatic Sciences 52:84–92.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Jackson, J. R., and J. E. Hightower. 2001. Reservoir striped bass movements and site fidelity in relation to seasonal patterns in habitat quality. North American Journal of Fisheries Management 21:34–45.

Kocik, J. F., and W. W. Taylor. 1994. Summer survival and growth of brown trout with and without steelhead under equal total salmonine densities in an artificial stream. Transaction of the American Fisheries Society 123:931–938.

Lohr, S. L. 1999. Sampling: design and analysis. Duxbury Press, Pacific Grove, California.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot ANOVA design for repeated-measures fishery data. Fisheries (Bethesda) 19(3):14–20.

Montgomery, D. C. 2001. Design and analysis of experiments. Wiley, New York.

Myers, R. A., and N. J. Barrowman. 1996. Is fish recruitment related to spawner abundance? U.S. National Marine Fisheries Service Fishery Bulletin 94:707–724.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. McGraw-Hill/Irwin, Homewood, Illinois.

Nowlin, W. H., and R. W. Drenner. 2000. Context-dependent effects of bluegills in experimental mesocosm communities. Oecologia 122:421–426.

Pierce, C. L., A. M. Corcoran, A. N. Gronbach, S. Hsia, B. J. Mullarkey, and A. J. Schwartzoff. 2001. Influence of diel period on electrofishing and beach seining assessments of littoral fish assemblages. North American Journal of Fisheries Management 21:918–926.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge, UK.

Radomski, P., and T. J. Goeman. 2001. Consequences of human lakeshore development on emergent and floating-leaf vegetation abundance. North American Journal of Fisheries Management 21:46–61.

Robillard, S. R., and J. E. Marsden. 2001. Spawning substrate preferences of yellow perch along a sand–cobble shoreline in southwestern Lake Michigan. North American Journal of Fisheries Management 21:208–215.

SAS (Statistical Analysis System). 2003. SAS/STAT user's guide, version 9.1. SAS Institute, Cary, North Carolina.

Sammons, S. M., and P. W. Bettoli. 1999. Spatial and temporal variation in electrofishing catch rates of three species of black bass (*Micropterus* spp.) from Normandy Reservoir, Tennessee. North American Journal of Fisheries Management 19:454–461.

Sampson, D. B., and Y. Yin. 1998. A Monte Carlo evaluation of the stock synthesis assessment program. Pages 315–338 *in* F. Funk and seven coeditors. Fishery stock assessment models. American Fisheries Society, Bethesda, Maryland.

Satterthwaite, F. E. 1946. An approximate distribution of estimates of variance components. Biometrics 2:110–114.

Seber, G. A. F., and C. J. Wild. 1989. Nonlinear regression. Wiley, New York.

Secor, D. H., T. E. Gunderson, and K. Karlsson. 2000. Effect of temperature and salinity on growth performance in anadromous (Chesapeake Bay) and nonanadromous (Santee–Cooper) strains of striped bass *Morone saxatilis*. Copeia 2000:291–296.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry, 3rd edition. Freeman, New York.

Thompson, S. K. 1992. Sampling. Wiley, New York.

Toepfer, C. S., W. L. Fisher, and J. A. Haubelt. 1999. Swimming performance of the threatened leopard darter in relation to road culverts. Transactions of the American Fisheries Society 128:155–161.

Walsh, M. G., D. B. Fenner, and D. L. Winkelman. 2002. Comparison of an electric seine and prepositioned area electrofishers for sampling stream fish communities. North American Journal of Fisheries Management 22:77–85.

Walters, C. J. 1985. Bias in estimation of functional relationships from time series data. Canadian Journal of Fisheries and Aquatic Sciences 42:147–149.

Whalen, K. G., and G. W. LaBar. 1994. Survival and growth of Atlantic salmon *(Salmo salar)* fry stocked at varying densities in the White River, Vermont. Canadian Journal of Fisheries and Aquatic Sciences 51:2164–2169.

Wilde, G. R., and W. L. Fisher. 1996. Reservoir fisheries sampling and experimental design. Pages 397–407 *in* L. E. Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society, Symposium 16, Bethesda, Maryland.

Wilderbuer, T. K., R. F. Kappenman, and D. R. Gunderson. 1998. Analysis of fishing power correction factor estimates from a trawl comparison experiment. North American Journal of Fisheries Management 18:11–18.

Willis, D. J., M. V. Hoyer, D. E. Canfield, Jr., and W. J. Lindberg. 2002. Training grass carp to respond to sound for potential lake management uses. North American Journal of Fisheries Management 22:208–212.

Yandell, B. S. 1997. Practical data analysis for designed experiments. Chapman and Hall, New York.

Zar, J. H. 1999. Biostatistical analysis, 4th edition. Prentice Hall, Upper Saddle River, New Jersey.

4 Recruitment

Michael J. Maceina and Donald L. Pereira

4.1 INTRODUCTION

4.1.1 Recruitment Assessment in Freshwater Fisheries

Recruitment of young fish into catchable, harvestable, or adult size is necessary to sustain any population and fishery. Recruitment failure, due to overfishing, habitat alteration, or abiotic or biotic events, can lead to reduced adult abundance and reduced angler catch rates. If severe, recruitment failure can ultimately result in severe population declines and collapse of a fishery. However, larval, juvenile, and even adult fish can be stocked to augment a fishery or a population if natural recruitment is low or nil. Conversely, if recruitment is high, then catchable-size abundance and fishing success should be greater if density-dependent mortality and growth reduction are not excessive. Recruitment is typically the strongest determinate influencing populations among the three major factors affecting populations, that is, growth, recruitment, and mortality (Carline et al. 1984; Allen and Pine 2000).

Recruitment success typically varies from year to year in most populations due to a number of factors. Some species from certain water bodies may display fairly constant recruitment each year, whereas other species or populations display highly variable recruitment that will cause wide fluctuations in the number of fish reaching a certain age or length. In marine systems, recruitment rates tend to be log-normally distributed with many average and below average years interspersed with periodic strong year-classes (Hennemuth et al. 1980). Although a similar review has not been conducted for North American freshwater fishes, we suspect a similar pattern. Fecundity in fishes is typically high, and recruitment variability is often caused by density-independent factors, but density-dependent regulation can stabilize recruitment (Cowan et al. 2000). The processes and mechanisms that cause recruitment variation have been intensively investigated for many years (reviewed by Cowan et al. 2000). In this chapter, analysis of the effects of environmental factors and parental abundance on recruitment will be explored.

Protection of adult fish stocks from overexploitation to prevent recruitment overfishing has primarily been addressed for commercially important marine fisheries

(Musick 1999). Recruitment overfishing occurs when a fishing rate is maintained over a long time period that results in low yields due to reduced recruitment of fish to adult or catchable size. These concepts of recruitment overfishing can also be important to the conservation of freshwater fishes. Loss of critical habitat and migration barriers in conjunction with overfishing have been cited as the cause for the decline of many Pacific salmon stocks in the northwestern USA (Stouder et al. 1997). Striped bass recovery in the Chesapeake Bay coincided with protection of mature females from exploitation, which increased juvenile abundance and recruitment (Richards and Rago 1999). Conservation of sturgeons has focused on protection of adults (long-lived species with late time to maturity) and habitat due to the low intrinsic rate of population increase via recruitment (Secor and Waldman 1999; Musick et al. 2000).

Fisheries scientists often evaluate the response of a population and fishery to habitat manipulations or regulation changes. However, an accurate result may not be evident due to recruitment variability, particularly if the evaluation time period is short (Allen and Pine 2000). For example, Bettoli et al. (1992) attempted to determine the effects of the complete removal of all submersed aquatic plants by grass carp on harvestable-size (>25 cm in total length [TL]) largemouth bass. Following vegetation removal, density of age-1 largemouth bass declined but growth rates increased. Although 3 years of post-vegetation-removal data were collected, the long-term effects on the adult population were difficult to detect because largemouth bass longevity was greater than 10 years, and obviously recruitment success or failure may take a number of years to detect (Maceina et al. 1994).

Recruitment is typically defined in terms of age or size, and this definition needs is to be assigned by the fishery scientist. Terms such as age-0 or age-1 recruits refers to the age at the time of collection. For example, catch rates of age-1 crappies collected with trap nets in fall (fish about 18 months old) were used by Maceina and Stimpert (1998) as an index of recruitment. Hansen et al. (1998) estimated density of age-0 walleye in September over a 39-year period, and thus fish were about 5 months old.

The number of fish reaching a certain length can also be used to define recruitment (e.g., number of fish greater than 100 mm), but length categorization is mostly used either as some minimum length of fish that can be caught by anglers and possibly harvested or as some regulated minimum length. For example, number of recruits entering the fishery at a particular length can be used.

4.1.2 Data Required to Assess Recruitment

Fisheries scientists can collect either long-term monitoring data or specific research data to quantify recruitment in a population or among populations. For a particular species or population, the fisheries scientist must determine what time in the early life of a fish confers recruitment to adult size or the fishery. For example, Sammons and Bettoli (1998) showed that low and high larval abundance of white bass, white crappie, and black crappie were associated with weak and strong year-classes, respectively. Pitlo (1997) reported an increase in the commer-

cial catch of channel catfish in the upper Mississippi River was associated with increased abundance of age-0 fish. Buynak et al. (1999) found the abundance of cohorts (year-classes) of age-5 largemouth bass was correlated to electrofishing catch of these same cohorts at age 1.

Either to document recruitment abundance and variability fully or relate some environmental factor to recruitment variation directly, long-term data collection is necessary. Typically, any sampling method that measures density or catch per effort (C/f) can by used to estimate recruitment (see Chapters 7 and 8). These samples should be taken about the same time each year and ideally under similar environmental conditions. Either fixed or random stations serve as replicates samples taken each year (Chapter 3). Conducting statistical analyses when random sampling is used is easier, but spatial variability in recruits may be greater than year-to-year variation, and low or zero catches can complicate the analysis. Thus, more replicate samples may be needed when random sampling is employed. With fixed-station sampling, repeated-measures analysis of variance (ANOVA) can be used to detect temporal or spatial differences in recruits (Maceina et al. 1994), but the analysis is more complicated than simple ANOVA.

Evaluation of recruitment enhancement from habitat manipulation or stocking requires a specific study design. Typically, these investigations are shorter in duration and may involve a few years of pre- and postmanipulation data collection. In addition, the frequency of sample collection may be greater than once per year, particularly if the fisheries scientist wants to investigate the effects of manipulation on successful recruitment.

Annual estimates of recruitment can be generated with monitoring data routinely collected by many state, federal, and provincial natural resource agencies in conjunction with population assessments. Many agencies conduct standardized electrofishing, gill netting, trap netting, seining, and sometimes rotenone sampling of age-0 fishes. Sampling bias and differences in catchability either with gear or location can occur and should be assessed (see Murphy and Willis 1996). Specific investigations to address a priori hypotheses pertaining to recruitment may require more sampling but typically can take place over a shorter period of time.

To determine if a relation exists between recruit and spawner (or parental abundance), long-term data collection of both these variables is necessary. Madenjian et al. (1996) used as little as 13 years of walleye data, whereas Myers et al. (1994) limited their analysis of 72 finfish populations to those with at least 20 years of data. Sources for recruitment–spawner data may include direct estimates from mark–recapture experiments, C/f derived from indices (Tyler and Crawford 1991; Myers et al. 1997), or estimates from stock assessment modeling derived from some form of sequential population analysis, virtual population assessment, or catch at age. We urge caution in any recruit–spawner analysis with fewer than 20 observations.

4.1.3 Freshwater versus Marine Recruitment Assessment

Similar to freshwater systems, recruitment of marine fishes can be highly variable and is regulated by both density independent and dependent factors. In addition,

low parental abundance due to either overexploitation or natural population cycles can drastically reduce reproductive output, which can ultimately result in a decline of recruits into the population and confound detection of environmental variables related to variable recruitment (Walters and Collie 1988). The areal extent of marine ecosystems, the sometimes longer interval between juvenile and adult harvestable ages, the difficulty of sampling certain life stages, and the greater difficulty in sampling recruits can confound recruitment assessment in marine environments as compared with freshwater environments.

Walters and Collie (1988) questioned the use of public funds to support correlative approaches to explain recruitment variation due to environmental variables, particularly for variables that cannot be controlled in marine fisheries. Subsequent predictions can fail due to the short-term period or window of data collection "at the frustrating rate of one observation per year" that will not encompass an even prohibitively longer period of recruitment variability (Walters and Collie 1988). In addition, correlative relations between environment variables and recruitment may be spurious (Walters and Collie 1988; Myers 1998). Myers (1998) examined numerous studies that had reexamined recruitment patterns primarily of marine fishes and found after retesting that many of the previously derived correlates failed to predict recruitment after additional data had been collected.

An early focus in marine stock assessment focused on recruit–spawner relations. In general, these described the number of recruits as a function of parental spawners either as (1) an asymptotic relation whereby recruitment will not increase once a certain number of spawners occur in the population (Beverton–Holt, with compensation) or (2) a domed-shaped relation (Ricker, with overcompensation) whereby the greatest numbers of recruits will be produced at some intermediate abundance of spawners. With a greater number of spawners, recruits will decline due to density dependence (Ricker 1975). In freshwater, Ricker recruit–spawner relations have been shown, for example, for walleye (Madenjian et al. 1996; Hansen et al. 1998) and lake trout (Hansen et al. 1996).

Typically, attempts to define a critical abundance of spawning adults necessary for adequate recruit production based on Ricker or Beverton–Holt equations have been wrought with high variability, confounding effects of environmental factors that affect recruitment, and the necessity to collect long-term data (Hilborn and Walters 1992; Goodyear 1993; Hansen et al. 1998). To circumvent the problem of defining recruit–spawner relations, marine fisheries scientists have attempted to address recruitment overfishing quantitatively by means of the spawning potential ratio (Goodyear 1993). Slipke et al. (2002) introduced the use of the spawning potential ratio to address recruitment overfishing for a freshwater commercial fishery.

4.2 ESTIMATION OF RECRUITMENT

A variety of gears can be used to sample young fish and estimate recruitment. In freshwater, electrofishing gear, gill nets, seines, trap nets, fyke nets, push nets,

bottom and midwater trawls, and rotenone have been used to sample recruits (see Chapters 5–10 in Murphy and Willis 1996). Gear selectivity can cause bias in estimating recruitment (Jackson and Noble 1995) and should be thoroughly evaluated to ensure representative sizes of all members of a cohort are sampled.

4.2.1 Estimates of Recruitment from Population Estimates

The density and biomass of recruits can be estimated with single or multiple mark–recapture methods, depletion methods, and toxicants (Chapter 8). Serns (1982) used electrofishing to conduct multiple mark–recapture procedures to estimate the population of age-0 smallmouth bass in the fall from 1974 to 1981 in Lake Nebish, Wisconsin, and estimates varied from 1,174 to 7,764 fish. Rider et al. (1994) blocked off 0.1–0.2-ha coves that contained submersed vegetation or open water to estimate the density of age-0 largemouth bass by means of an electrofishing catch depletion technique. Typically, four to seven 10-min passes were required to deplete the number of individuals in order to compute a Leslie and Davis (1939) linear regression line (Rider et al. 1994).

Long-term (>20 years) annual cove rotenone samples were used to estimate the biomass of age-0 black basses in Bulls Shoals Reservoir, Arkansas (Ploskey et al. 1996) and age-0 crappie density in four Mississippi reservoirs (Allen and Miranda 1998). Hoyer and Canfield (1996) used 0.1-ha block nets and rotenone applied in limnetic and littoral regions of Florida lakes to estimate annual density of age-0 largemouth bass. Allen et al. (1999) used a shoreline rotenone technique to relate age-0 largemouth bass density to chlorophyll-*a* concentrations and larval gizzard shad and threadfin shad densities. Fisher and Zale (1993) used a 12.2-m-long by 1.8-m-deep bag seine with 4.8-mm mesh and captured age-0 largemouth bass from quadrants of a known area. Average annual catch rates varied from 0.42 to 3.12 fish/100 m^2, and density was related to water-level fluctuations in conjunction with the implementation of a new water-level regulation schedule on Grand Reservoir, Oklahoma (Fisher and Zale 1993).

4.2.2. Estimates of Recruitment from Indices

In many instances, estimating the density and biomass of recruits (either at age 0 or age 1) is not feasible or too costly. Sampling with electrofishing gear, gill nets, seines, trawls, and trap nets can provide standardized units of effort for either time or distance and can be used to index recruit abundance. Willis and Stephen (1987) and Sammons and Bettoli (2000) used experimental monofilament gill nets to estimate catch per net-night of age-0 walleye and age-1 white bass; recruits of both species varied over two orders of magnitude, and this variation was related to hydrologic variables.

Maceina and Stimpert (1998) used age-1 catch rates (*N*/net-night) of black crappie and white crappie in trap nets as an index of recruitment in Alabama reservoirs and related the variation in recruitment to reservoir hydrologic variables. Sammons and Bettoli (2000) used DC electrofishing in the spring to capture age-1

largemouth bass along 40 randomly chosen 100-m transects throughout Normandy Reservoir, Tennessee, and average annual catch rates varied from about 0.1 to about 1.7 fish/100 m. Jackson and Noble (2000) used a handheld electrofishing apparatus from a boat to collect age-0 largemouth bass during 3-week intervals from June to October from 1988 to 1998; peak average catch rates among years varied from 3.8 to 46.7 age-0 fish/20 min.

Bronte et al. (1993) used a semiballoon bottom trawl with a 11.9-m headrope, a 15.6-m footrope, and a 12.7-mm-mesh cod end to sample all sizes of yellow perch, and the total catch of age-2 fish was used as an index of recruitment and year-class strength. Similarly, Madenjian et al. (2000), using 8-m and 11-m bottom trawls, documented a reduction in catch of age-0 white bass over a 20-year period from Lake Erie that was related to the subsequent decline in the fishery. Counihan et al. (1999) used a 6.2-m high-rise bottom trawl to assess abundance of age-0 white sturgeons. Highly variable catches were due to patchy distributions and were not normally distributed (Counihan et al. 1999). These authors recommended that indices of presence and absence and *C/f* both be used to assess recruitment levels of age-0 white sturgeons in the Columbia River. Beach seines were used to collect age-0 striped bass from the Chesapeake Bay from 1954 to 1996 to develop a quantitative juvenile index, which was later used to determine that recruitment overfishing caused the collapse of this valuable fishery (Richards and Rago 1999).

4.2.3 Use of Marks and Tags to Assess Recruitment

Coded wire tags, dyes, chemical marking, morphological marks, and genetic tags can be employed to determine the success of stocking and may also be used to examine recruitment processes of wild fish. To evaluate stocking success, otoliths of juvenile fishes can be marked with alizarin complexone, calcein, or oxytetracycline, and these chemicals can be applied by immersion, injection, or orally through prepared foods in a hatchery (Thomas et al. 1995).

Isermann et al. (2002) successfully used oxytetracycline immersion to form marks on young crappie otoliths and suggested this technique can be used effectively to identify stocked crappies up to about 2 years old. Paragamian et al. (1992) established the reliability of stress checks on otoliths of hatchery-reared kokanees to distinguish these fish from wild fish in Lake Pend Oreille, Idaho. Counts of daily increments in relation to stress checks allowed for correct identification of fish from several co-occurring release groups that had been stocked at different times in the same season (Paragamian et al. 1992).

Buynak and Mitchell (1999) used alternating pectoral fin clips of stocked age-0 largemouth bass (about 11 cm TL) over a 5-year stocking period to evaluate contribution of stocked fish to naturally produced fish in a 1,200-ha Kentucky reservoir. Over time, stocked fish contributed 25% of the total electrofishing catch (Buynak and Mitchell 1999).

Ryan et al. (1998) stocked genetically distinct Florida largemouth bass adults that contained a unique allele (*sIDHP*109*) expressed in the allozyme locus for

isocitrate dehydrogenase (IUBMB [1992] number 1.1.1.42) that was different than the *sIDHP*100* and *sIDHP*122* alleles found, respectively, in the northern largemouth bass and Florida largemouth bass population in Lake Galdwater, Texas. Offspring homozygous for the *sIDHP*109* allele were produced, grown out in nursery ponds, and stocked at rate of 8.8 fish/ha in summer (Ryan et al. 1998). To assess stocking success, age-0 and age-1 largemouth bass were sampled in fall and spring with DC electrofishing, and recruits were identified using electrophoresis (Ryan et al. 1998). Murphy et al. (1983) identified allele frequencies at the malate dehydrogenase locus (*mMDH-2**; IUBMB [1992] number 1.1.1.37) of hatchery-raised walleye, and the success of supplemental stocking was evaluated by quantifying shifts in cohort allele frequencies due to the stocking of juvenile fish with allele frequencies different from resident-hatched fish.

Ludsin and DeVries (1997) used three different color dyes that were injected into small (<100 mm TL), medium (100–150 mm TL), and large (>150 mm TL) age-0 largemouth bass in the fall to assess overwinter size-dependent mortality. In the spring at age 1, a higher proportion of the larger-size individuals were collected, which indicated that size in the fall influenced recruitment to age 1. Parsons and Pereira (2001) used coded wire tags to evaluate walleye stockings and estimated the extent of natural reproduction in three Minnesota lakes. About 95,000 hatchery-reared striped bass were individually marked with coded wire tags and released into Delaware Bay, and a total population estimate of age-0 striped bass was derived from recaptures of both tagged and wild-produced fish (Burton and Weisberg 1994).

4.2.4 Otolith Microstructure Analysis to Assess Recruitment

The analysis of daily increments on the otoliths of fishes (Pannella 1971) can provide fisheries scientists with insights on early life history aspects of fish population dynamics, including recruitment. Research has sometimes shown that early hatched cohorts, identified by enumerating daily growth rings, not only have a size advantage compared with later-hatched cohorts but grow faster and are more likely to recruit to the population due to increased survival (Ludsin and DeVries 1997). Conversely, early hatching in spring could be detrimental to larval fish survival and subsequent recruitment due to unstable climatic conditions such as low water temperatures, high variation in air and water temperatures, or windy conditions (Kramer and Smith 1962; Summerfelt 1975; Crecco and Savoy 1987; Rice et al. 1987).

Isely et al. (1987) and Maceina et al. (1988) used incremental counts of daily growth rings to assess temporal spawning patterns, growth, and recruitment potential of mixed populations of northern and Florida largemouth bass. Crecco and Savoy (1987) identified 5-d cohorts of American shad, estimated cohort mortality, and found that recruitment was influenced by density-dependent processes and strongly mediated by hydrologic and climatic conditions.

4.3 RECRUITMENT VARIABILITY AND FACTORS RELATED TO YEAR-CLASS STRENGTH

4.3.1 Temporal Variation in Recruitment

Abundance of recruits can be relatively stable or highly variable over time. Fisheries scientists collect long-term monitoring or research data to assess temporal variation in recruitment. Generally, data should be collected at about the same time each year from a random, systematic, or stratified sampling design (Chapter 3).

Allen and Pine (2000) reviewed data on recruitment variability in white crappie and black crappie populations and largemouth bass populations based on age-0 and age-1 abundances, which were assessed using electrofishing, trap nets, and rotenone sampling. Coefficients of variation (CV; 100 · SD/mean) for recruits averaged 82% (55–124%) for crappies and 66% (11–189%) for largemouth bass. In Lake Escanaba, Wisconsin, mark–recapture population estimates based on electrofishing were conducted for age-0 walleye each September–October from 1958 to 1996 (Hansen et al. 1998). In this lake, age-0 walleye density averaged 99 fish/ha, varied from 5 to 299 fish/ha, and had a CV of 78% (Hansen et al. 1998). This long-term database is useful for understanding the dynamics of walleye recruitment in Lake Escanaba but represented a tremendous amount of sampling effort over a long time period that obviously cannot be achieved for every system.

High CV values in recruitment will cause population characteristics and associated angler catches to fluctuate. In short-lived populations (i.e., less than 8 years), three to four successive weak year-classes can cause a population to decline drastically. Software programs such as GIFSIM (Taylor 1981), MOCPOP (Beamesderfer 1991), and FAST (Slipke and Maceina 2000) can be used to simulate the response of a fish population over time to stochastic recruitment. Fisheries scientists may be interested in examining some mean or median level of recruitment and associated variance and incorporating this variation (SD, CV, or range) into modeling or other types of analyses.

Kimura (1988) presented a two-way ANOVA technique based on log-linear models that can be applied to catch data of different age fish to test for differences among year-class abundances. In many instances with sampling, collection gears may be positively or negatively biased for a certain age. This bias was evident in the crappie data presented in Table 4.1; age-1 catch was higher than age-0 catch for the 1993 year-class. If fish are collected over time from a number of different age-groups, then potential age-selective bias in the sample can be ameliorated by considering multiple catches of different age fish from the same year-class (Box 4.1).

Over time, recruitment will vary and may show an increasing, decreasing, or stable pattern. Long-terms changes in recruitment may be a function of variation in water quality, habitat, climatic factors, introduction of a competitive species, or excessive exploitation. Bettoli et al. (1992) found a significant correlation (r = 0.80; $P < 0.05$) between density of age-1 largemouth bass and macrophyte coverage. A reduction in age-0 channel catfish and striped bass abundance was associated with high exploitation of adults for both of these species, and a subsequent

Table 4.1 Age-0, age-1, and age-2 black crappie and white crappie catch rates (species catch per unit effort, *C/f,* were pooled) for 11 year-classes collected with trap nets from Weiss Lake, Alabama, from 1989 to1999. Three reservoir hydrologic variables are also provided (partial data set presented in Maceina and Stimpert 1998). Water levels in Weiss Lake are regulated for flood control and power generation. Mean winter stage was the average daily stage between 1 January and 31 March (prior to crappie spawning). Mean winter retention was derived by dividing average daily volume by discharge, which was computed from average daily readings for Weiss Lake. Full summer pool is normally obtained around 15 April each year at an elevation of 171.95 m above mean sea level (msl). Mean spring stage was computed from average daily stages between 1 April and 31 May and coincided with crappie spawning (Travnichek et al. 1996).

Year-class	Catch per unit effort			Mean winter stage (m msl)	Mean winter retention (d)	Mean spring stage (m msl)
	Age-0	Age-1	Age-2			
1989		3.12	0.49	170.85	7.2	171.79
1990	8.03	5.32	2.43	171.76	4.2	171.75
1991	0.47	0.39	0.39	170.73	7.4	171.75
1992	0.61	0.97	0.61	170.67	6.6	171.82
1993	1.38	3.59	1.32	170.99	6.2	171.77
1994	2.73	2.62	0.92	170.93	5.9	171.87
1995	1.66	0.57	0.47	170.88	6.8	171.80
1996	9.89	8.63	2.11	171.39	5.5	171.85
1997	1.86	0.93	0.21	170.83	6.1	171.90
1998	3.72	1.17		171.04	5.6	171.83
1999	2.18			170.77	9.7	171.82

reduction in exploitation resulted in increased recruitment (Pitlo 1997; Richards and Rago 1999). White bass decreased over time in Lake Erie, and this trend was related to a declining temporal trend in the abundance of age-0 fish (Madenjian et al. 2000; Box 4.2).

Fisheries scientists can also examine temporal differences in recruitment from changes in habitat features or by manipulating habitat characteristics. Fisher and Zale (1993) examined abundance of age-0 largemouth bass during a 12-year period for which data were collected prior to and after a change in the reservoir-regulated water levels. In many instances, a manipulation is conducted in a single area or water body and pre- and postmanipulation responses in recruitment are measured. If young fish are collected over time, then one can use one of the ANOVA designs presented by Hubert and Fabrizo (Chapter 7).

4.3.2 Spatial Variation in Recruitment

Fish recruitment can vary among water bodies, within water bodies, and among different habitats within a single water body (Allen and Pine 2000). For example, Wrenn et al. (1996) found that density of age-0 largemouth bass was greater in areas of Lake Guntersville, Alabama, that contained Eurasian water milfoil compared with areas that were devoid of aquatic vegetation. Sammons and Bettoli (2000) collected age-0 largemouth bass for a 6-year period to examine the relation of reservoir hydrology and largemouth bass recruitment. Four distinct areas of the

Box 4.1 Log-Linear Model to Test for Year-Class Abundance Differences

Below we conduct a test for year-class abundance differences among the 1990 to 1997 year-classes (YEARCL) based on catch rates of age-0, age-1, and age-2 crappies (AGE in years) from Weiss Lake (Table 4.1). Trap-net catch rates are transformed to natural log values (LCATCH) to homogenize variances as recommended by Kimura (1988) for log-linear analysis. The data in Table 4.1 were rearranged to conduct the analysis. Year of collection (YEARCOL) was included in the data file, and the following SAS (2001) program was written to conduct the analysis.

Program

```
DATA WECRA;
INPUT YEARCOL YEARCL AGE CATCH;
LCATCH=LOG(CATCH);
LINES;
1990    1990    0    8.03
1991    1990    1    5.32
1992    1990    2    2.43
1991    1991    0    0.47
1992    1991    1    0.39
1993    1991    2    0.39
1992    1992    0    0.61
(continue data input)
;
PROC GLM; CLASS YEARCL AGE;
MODEL LCATCH=YEARCL AGE;
LSMEANS YEARCL/T PDIFF STDERR;
MEANS YEARCL/LSD LINES ALPHA=0.001786; RUN;
```

Results

Table Output for two-way analysis of variance (ANOVA) and comparison of least-squares means for catch (dependent variable LCATCH). There were 24 observations in the data set. Abbreviations are given for coefficient of variation (CV), mean square error (MSE), sum of squares (SS), and least-squares mean (LSMEAN).

Class Level Information

Class	Levels	Values
YEARCL	8	1990 1991 1992 1993 1994 1995 1996 1997
AGE	3	0 1 2

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	9	23.11022510	2.56780279	11.67	0.0001
Error	14	3.07916486	0.21994035		
Corrected total	23	26.18938997			
R^2	0.882427	Root MSE	0.46897798		
CV	138.6533	LCATCH mean	0.33823792		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
YEARCL	7	19.01504754	2.71643536	12.35	0.0001
AGE	2	4.09517756	2.04758878	9.31	0.0027

Least-Squares Means for H_0 LSMEAN = 0

YEARCL	LCATCH LSMEAN	SE	*P* > |*t*|	LSMEAN number
1990	1.54751636	0.27076456	0.0001	1
1991	–0.87941322	0.27076456	0.0058	2
1992	–0.33968395	0.27076456	0.2302	3
1993	0.62595581	0.27076456	0.0365	4
1994	0.62803144	0.27076456	0.0360	5
1995	–0.27010797	0.27076456	0.3354	6
1996	1.73115220	0.27076456	0.0001	7
1997	–0.33754732	0.27076456	0.2330	8

Least-Squares Means for H_0 LSMEAN *I* = LSMEAN *j*

LSMEAN number (*j*)	LSMEAN number (*I*)							
t-test, and *P*[a]	1	2	3	4	5	6	7	8
1								
t-value		6.337973	4.928459	2.406673	2.401252	4.74676	–0.47957	4.922879
P		0.0001	0.0002	0.0305	0.0308	0.0003	0.6389	0.0002
2								
t-value	–6.33797		–1.40951	–3.9313	–3.93672	–1.59121	–6.81754	–1.41509
P	0.0001		0.1805	0.0015	0.0015	0.1339	0.0001	0.1789
3								
t-value	–4.92846	1.409513		–2.52179	–2.52721	–0.1817	–5.40803	–0.00558
P	0.0002	0.1805		0.0244	0.0242	0.8584	0.0001	0.9956
4								
t-value	–2.40667	3.9313	2.521787		–0.00542	2.340088	–2.88624	2.516207
P	0.0305	0.0015	0.0244		0.9958	0.0346	0.0120	0.0247
5								
t-value	–2.40125	3.93672	2.527207	0.005421		2.345508	–2.88082	2.521627
P	0.0308	0.0015	0.0242	0.9958		0.0343	0.0121	0.0244
6								
t-value	–4.74676	1.591212	0.181699	–2.34009	–2.34551		–5.22633	0.176119
P	0.0003	0.1339	0.8584	0.0346	0.0343		0.0001	0.8627
7								
t-value	0.479568	6.817541	5.408028	2.886241	2.880821	5.226329		5.402448
P	0.6389	0.0001	0.0001	0.0120	0.0121	0.0001		0.0001
8								
t-value	–4.92288	1.415093	0.00558	–2.51621	–2.52163	–0.17612	–5.40245	
P	0.0002	0.1789	0.9956	0.0247	0.0244	0.8627	10.0001	

(Box continues)

Box 4.1 *(continued)*

Least-Significant-Difference Test for Variable LCATCH[b]			
T grouping	Mean	*N*	YEARCL
A	1.7312	3	1996
A			
A	1.5475	3	1990
A			
B A	0.6280	3	1994
B A			
B A	0.6260	3	1993
B			
B C	–0.2701	3	1995
B C			
B C	–0.3375	3	1997
B C			
B C	–0.3397	3	1992
C			
C	–0.8794	3	1991

[a] To ensure overall protection level, only probabilities associated with preplanned comparisons should be used.

[b] Means with the same letter are not significantly different. Alpha = 0.001786, *df* = 14, MSE = 0.21994, critical value of *t* = 3.84, and least significant difference = 1.4723. This test controls the type I comparisonwise error rate not the experimentwise error rate.

Interpretation

The two-way ANOVA indicated that both age and year-class were significant ($P < 0.01$) class variables related to crappie *C/f*. Inspection of the data in Table 4.1 suggested abundance of a year-class decreased at older ages, and this was supported by the analysis ($F = 9.31$; $df = 2, 14$). Accounting for the effects of age, *C/f* also varied by year-class ($F = 12.35$; $df = 7, 14$).

reservoir were chosen each year, and replicate electrofishing transects were conducted to collect fish in August and September each year. A split-plot repeated-measures design (Maceina et al. 1994) was used to test for spatial differences in *C/f* among these four areas of the reservoir for data collected in 1992 because the same six electrofishing transects were sampled repeatedly in August and September for that year (Table 4.2; Box 4.3).

4.3.3 Use of Adult Age-Structure Data to Estimate Recruitment Variability

Total annual mortality in a fish population can be obtained from catch-curve regression analysis (see Chapter 6) by which a sample of fish that has recruited to the fishery is collected and aged and the natural log of the number at age (*y*-variable) is regressed against age (*x*-variable). When a single sample of fish is collected that represents a number of cohorts or year-classes, highly variable

Below the ANOVA is a table that presents the least-squares means (LSMEANS) for $\log_e(C/f)$, LCATCH, the associated SE, and a probability value from a *t*-test that the least-squares mean is not equal to 0 (null hypothesis, H_0: LSMEAN = 0). A number assigned to each least-squares mean represents each year-class. The LSMEANS procedure in SAS creates least-squares means for the class variables. These are also referred to as adjusted means (SAS 2001). For one-way ANOVA, or in this example, a balanced two-way ANOVA, least-squares means are computed as arithmetic means. In this example, each year-class is weighted by sample size or the number of age-groups ($N = 3$).

Next, a matrix table is presented that shows pairwise t-tests among all eight year-classes and corresponding probability levels. These comparisons allow the fisheries scientists either to accept or reject the H_0 that the least-squares means or year-class abundance estimates are the same for the two year-classes being compared. A caution statement at the end of the table warns the analyst only to make preplanned comparisons (SAS 2001). If many comparisons are being made, some statistical differences may be detected due to random chance and are not true differences.

Because 28 comparisons were made among these eight year-classes (7 + 6 + 5 + 4 + 3 + 2 + 1), a Bonferroni correction can be applied to an alpha level to reduce the probability of making a type I error. For this example, if we set alpha at 0.05 and divide by 28, the new alpha level of 0.001786 can be specified in the SAS program to perform Fischer's least-significance-difference multiple- range test, which is analogous to multiple pairwise *t*-tests. Thus, fewer statistical differences would be evident compared with a standard alpha value of 0.05. The fisheries scientist needs to decide what level of a type I error is acceptable when making multiple comparisons. In this example, the Bonferroni correction is clearly highly conservative and would become more restrictive in rejecting the H_0 as the number of years of data collection increases. If a priori assignment of year-class groups can be assigned, then a Bonferroni correction for preplanned comparisons would be less or contrast statements can be set up to test for preplanned comparisons. From these results, the 1990 and 1996 year-classes were more abundant than were the 1991, 1992, 1995, and 1997 year-classes. Abundance of the 1993 and 1994 year-classes was intermediate and statistically similar to some of these weaker and more abundant year-classes.

recruitment will cause the relation between number at age and age to vary (Ricker 1975; Maceina 1997).

Guy and Willis (1995) introduced and used the recruitment variability index (RVI) to assess black crappie reproductive success in South Dakota. To compute the RVI, the cumulative relative frequency (CRF) distribution (the same as presented for the Kolmogorov–Smirnov one-sample test or PROC CHART in SAS [2001]) is used to describe the magnitude and distribution of the frequency-of-occurrence-at-age data. This index is sensitive to year-classes that are completely missing from the sample. The RVI is computed as

$$\text{RVI} = [\text{CRF}/(N_m + N_p)] - N_m/N_p, \tag{4.1}$$

where N_m is equal to the number of missing year-classes (no fish were collected), N_p is equal to the number of year-classes present in the sample, and the N_p must

Box 4.2 Evaluation of Time Series Trends in Recruit Abundance

The following program presents a plot and computes the Pearson correlation coefficient between age-0 *C/f* (AGE0CPE) of white bass and year and the Kendall tau-*b* nonparametric correlation coefficient for ranks between these two variables (data published in Mandenjian et al. 2000). In addition, the simple linear regression between *C/f* and year was computed along with the Durbin–Watson statistic (DW) to determine temporal autocorrelation. Finally, the residuals from the regression were plotted against year by means of the following SAS program.

Program

```
DATA WHBASS;
INPUT YEAR AGE0CPE @@;
LINES;
1972 24.38 1973 4.29 1974 10.06 1975 18.16 1976 23.44 1977 20.38 1978 8.06
1979 11.36 1980 25.24 1981 20.49 1982 4.88 1983 2.1 1984 7.68
1985 4.52 1986 3.14 1987 0.57 1988 4.25 1989 1.35 1990 8.42 1991 2.04
1992 3.66 1993 2.84 1994 2.49 1995 0.6 1996 4.85 1997 3.14
;
PROC PLOT; PLOT AGE0CPE*YEAR;
PROC CORR; VAR AGE0CPE YEAR;
PROC CORR KENDALL; VAR AGE0CPE YEAR;
PROC REG; MODEL AGE0CPE=YEAR/DW;
OUTPUT OUT=A R=RES;
PROC PLOT; PLOT RES*YEAR/VFRE=0; RUN;
```

Results

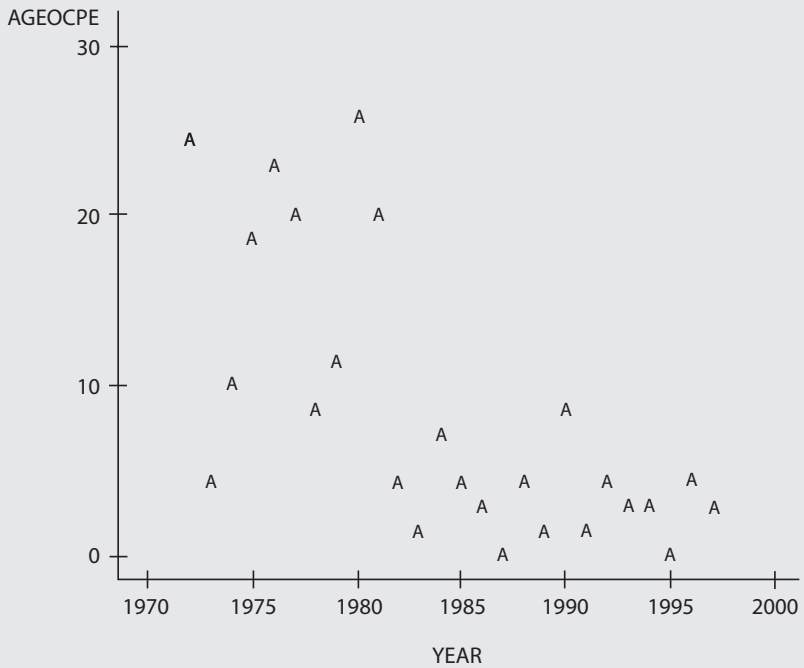


Figure Age-0 *C/f* (AGE0CPE) of white bass versus year, in which A represents one observation.

Table Output for Pearson and Kendall tau-*b* correlation coefficients. Computed are the Pearson correlation coefficient between AGE0CPE of white bass and year and the Kendall tau-*b* nonparametric correlation coefficient for ranks between these two variables. Both correlation coefficients test for $P > |R|$ under the H_0 that R (rho) = 0 and $N = 26$.

Simple Statistics

Variable	*N*	Mean	SD	Sum	Minimum	Maximum
AGE0CPE	26	8.5535	8.0782	222.3900	0.5700	25.2400
YEAR	26	1985	7.6485	51597	1972	1997

Pearson Correlation Coefficients

	AGE0CPE	YEAR
AGE0CPE	1.00000	–0.67212
	0.0	0.0002
YEAR	–0.67212	1.00000
	0.0002	0.0

Kendall's Tau-*b* Correlation Coefficient

	AGE0CPE	YEAR
AGE0CPE	1.00000	–0.48690
	0.0	0.0005
YEAR	–0.48690	1.00000
	0.0005	0.0

Time Series Regression and Test for Autocorrelation

Source	*df*	SS	Mean square	*F*-value	$P > F$
Model	1	736.99067	736.99067	19.775	0.0002
Error	24	894.43972	37.26832		
Corrected total	25	1631.43039			

R^2	0.4517	Root MSE	6.10478
Adjusted R^2	0.4289	AGE0CPE mean	8.55346
CV	71.37202		

Variable	*df*	Parameter estimate	SE	*t*-value	$P > \|t\|$
Intercept	1	1417.304215	316.79344310	4.474	0.0002
YEAR	1	–0.709877	0.15963274	–4.447	0.0002

Durbin–Watson statistic (DW)	1.500
Number of observations	26
First-order autocorrelation	0.216

(Box continues)

Box 4.2 *(continued)*

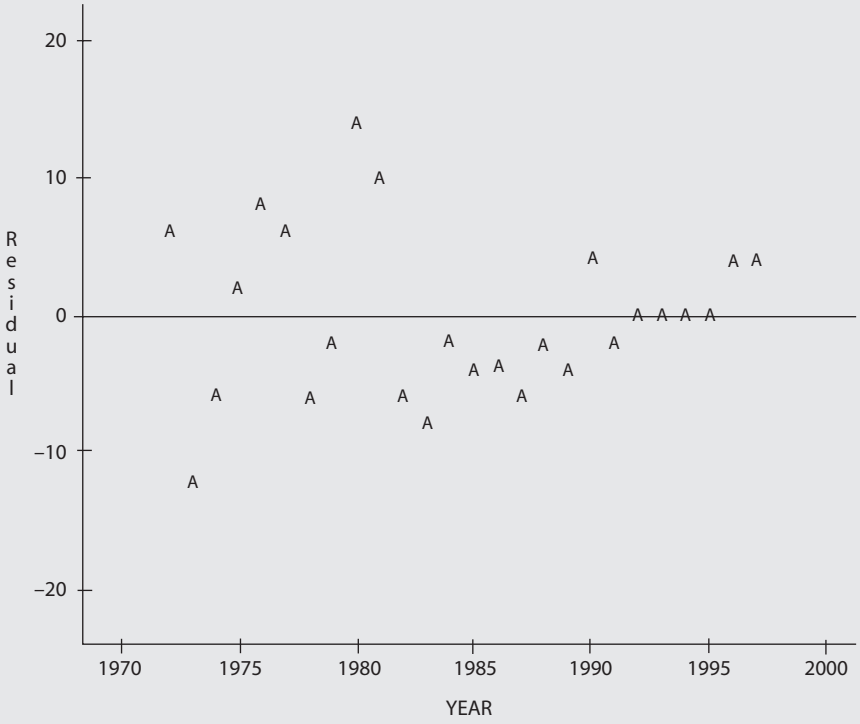


Figure Residuals from the time series regression of age-0 white bass *C/f* versus year, in which A represents one observation.

Interpretation

The SAS plot that examines the relationship between age-0 *C/f* and year clearly showed a decline in white bass recruitment over time. Correlation and regression analyses indicated a significant ($P < 0.01$) decrease in *C/f* of age-0 white bass over time. The Kendall tau-*b* correlation coefficient is a nonparametric test that numerically ranked the years and age-0 *C/f* and then computed the

exceed the N_m. The RVI varies from –1 to 1, and increases in RVI indicate less recruitment variability. The index assumes that fish are fully recruited to the sampling gear, that catch at age is a valid representation of year-class strength, and that there are not year-classes beyond the last age-group represented in the sample (Guy and Willis 1995). In addition, Guy (1993) recommended that the RVI should not be computed when less than three year-classes are present.

Maceina (1997) built upon the RVI concept and developed a quantitative index of recruitment variability based on the residuals or errors associated with

association between these two ranks. Similar to Pearson correlation coefficients, Kendall tau-*b* correlation coefficients can vary from –1 to 1, and these values would be computed if ranks completely matched. The value of Kendall tau-*b* correlation coefficient was less than that of the Pearson coefficient, but both tests indicated a strong, significant ($P \leq 0.01$) decreasing temporal trend in age-0 *C/f*. The Kendall tau-*b* correlation coefficient is useful for time series data because one or a few extremely high or low values can be so influential when using linear methods that the results would be biased.

Autocorrelation can be troublesome with time series analysis because errors or residuals may not be independent. For example, when measuring the response of a variable to some factor, the same factor also influenced that variable of interest some time in the past. The same processes could be evident with recruitment data, for example, due to long-term temporal patterns in climate or recruitment overfishing. If adults have been overexploited to the extent that not enough recruits are being produced to replace adults, then a negative feedback loop occurs; fewer adults confer fewer recruits to become adults, and then there are fewer recruits in the next generation. Although autocorrelation is somewhat troublesome statistically with time series analysis, this should not preclude the inferences drawn from primary trends.

The SAS plot of the residuals from the regression of age-0 *C/f* versus year showed a somewhat even cyclic pattern of high and low residuals from 1972 to the early 1980s, negative residuals during the rest of the 1980s, followed by residual values around 0 or greater in the 1990s. These residuals did not appear randomly scattered, particularly from 1972 to about 1985, and suggested a weak cyclic pattern in white bass recruitment. The Durbin–Watson statistic (DW) is a test for the existence of a first-order autoregressive process. The personal computer version of SAS (2001) does not provide a statistical probability that tests for a first-order autocorrelation, but probability distribution tables for the DW values can be found in Montgomery and Peck (1982). For this example, the computed DW was 1.5, which exceeded the critical DW value of 1.45 ($P = 0.05$). Thus, errors were autocorrelated based on the DW statistic. The first-order correlation is the actual correlation between adjacent residuals. We computed a first-order autocorrelation of 0.22, which is moderately low, but for time series analysis, the number of years of data was relatively high ($N = 26$). Values approaching –1 or 1 show a high degree of autocorrelation. See Montgomery and Peck (1982) and Freund and Littell (1991) for more information. High values for first-order correlations and significant DW values suggest a cyclic pattern and dictate that the fisheries scientist should investigate this phenomenon in more detail.

catch curves and subsequently verified the index (Maceina 2004). An assumption of this analysis is that positive and negative residuals associated with catch-curve regressions represent strong and weak year-classes. Thus, variation about the catch-curve regression is primarily associated with recruitment variability, though in some instances density-dependent mortality among adult fish may also influence the relation between number at age and age. Maceina (2004) more thoroughly explained the use of this approach for quantifying recruitment, and an example is provided (Box 4.4).

Table 4.2 Catch (*N*/100 m) of age-0 largemouth bass along six 100-m transects (Rep) in four different regions of Lake Normandy for three successive time periods (Time) spaced 2 weeks apart in August (Aug) and September (Sep) 1992 (partial data set from Sammons and Bettoli 2000). The four areas were the Lower Basin (LB), Riley Creek (RC), the Upper Basin (UB), and Carroll Creek (CC).

Month	Area	Rep	Time	Catch	Month	Area	Rep	Time	Catch
Aug	CC	1	1	0	Aug	RC	1	2	2
Aug	CC	2	1	3	Aug	RC	2	2	0
Aug	CC	3	1	0	Aug	RC	3	2	2
Aug	CC	4	1	0	Aug	RC	4	2	1
Aug	CC	5	1	2	Aug	RC	5	2	4
Aug	CC	6	1	1	Aug	RC	6	2	1
Aug	LB	1	1	4	Aug	UB	1	2	0
Aug	LB	2	1	2	Aug	UB	2	2	1
Aug	LB	3	1	11	Aug	UB	3	2	0
Aug	LB	4	1	3	Aug	UB	4	2	0
Aug	LB	5	1	6	Aug	UB	5	2	1
Aug	LB	6	1	3	Aug	UB	6	2	0
Aug	RC	1	1	3	Sep	CC	1	3	0
Aug	RC	2	1	2	Sep	CC	2	3	3
Aug	RC	3	1	0	Sep	CC	3	3	1
Aug	RC	4	1	5	Sep	CC	4	3	0
Aug	RC	5	1	3	Sep	CC	5	3	1
Aug	RC	6	1	0	Sep	CC	6	3	1
Aug	UB	1	1	2	Sep	LB	1	3	2
Aug	UB	2	1	1	Sep	LB	2	3	3
Aug	UB	3	1	4	Sep	LB	3	3	0
Aug	UB	4	1	0	Sep	LB	4	3	2
Aug	UB	5	1	3	Sep	LB	5	3	3
Aug	UB	6	1	0	Sep	LB	6	3	4
Aug	CC	1	2	0	Sep	RC	1	3	0
Aug	CC	2	2	4	Sep	RC	2	3	0
Aug	CC	4	2	1	Sep	RC	4	3	1
Aug	CC	5	2	0	Sep	RC	5	3	0
Aug	CC	6	2	1	Sep	RC	6	3	0
Aug	LB	1	2	4	Sep	UB	1	3	1
Aug	LB	2	2	1	Sep	UB	2	3	0
Aug	LB	3	2	3	Sep	UB	3	3	3
Aug	LB	4	2	1	Sep	UB	4	3	0
Aug	LB	5	2	5	Sep	UB	5	3	0
Aug	LB	6	2	3	Sep	UB	6	3	0

4.3.4 Examination of the Influence of Environmental Factors on Recruitment

Correlation, simple and multiple linear regression, and nonlinear regression techniques are commonly used to explain and predict variation in recruitment because biotic and abiotic variables that influence recruitment typically vary from year to year. For example, Busch et al. (1975), Kallemeyn (1987), and Hansen et al. (1998) found that adverse climatic conditions in spring during walleye spawning

Box 4.3 Evaluation of Spatial Differences in Recruit Abundance

Table 4.2 contains a data set to test for spatial differences in age-0 largemouth bass catch in Lake Normandy, Tennessee (data from Sammons and Bettoli 2000). In this example, four distinct areas of the reservoir (Lower Basin [LB], Riley Creek [RC], Upper Basin [UB], and Carroll Creek [CC]) were chosen to examine spatial variation in abundance of fish along 100-m shoreline electrofishing transects. A handheld DC electrofishing unit was used at night. Six fixed sites, or replicate transects, were chosen within each area and sampled three times at 2-week intervals starting the second week of August 1992 and ending the second week of September 1992. Thus, 24 transects were conducted over three time intervals for a total of 72 transects, or observations.

Because replicate samples were collected at fixed locations over the three time periods within each of the same areas, a split-plot repeated-measures ANOVA was used to test for differences in number of fish among areas (Maceina et al. 1994). In addition, this analysis also tested for differences in catch over time and examined the time × area interaction. The program and analysis were divided into main-plot A, which included the class variables area, replicates (REP), and the area × replicate interaction, and subplot B, which contained the time and the time × area interaction effects. The mean square error (MSE, or type III sums of squares) of the area × replication term was used as the error term in the denominator and the MSE for area as the numerator of an *F*-test for statistical differences in the number caught among the four areas in main-plot A. The MSE generated from the entire ANOVA was used in the denominator of the *F*-test to determine if statistical differences in catch occurred over the three time periods (subplot B), as well as for testing for any interaction between time periods and areas (subplot B).

The following SAS (2001) program provides output to test for differences in catch among areas.

Program

```
DATA NORM_LMB;
INPUT YEAR MONTH AREA $ REP TIME CATCH;
LINES;
92     8     CC      1       1        0
92     8     CC      2       1        3
92     8     CC      3       1        0
(continue data input)
;
PROC GLM; CLASS TIME AREA REP;
MODEL COUNT = AREA REP  REP*AREA TIME  TIME*AREA;
TEST H = AREA E = REP*AREA;
MEANS AREA/SNK E=REP*AREA;
MEANS TIME/SNK;
RUN;
```

(Box continues)

Box 4.3 *(continued)*

Results

Table Output for split-plot repeated-measures ANOVA (type 1 SS omitted) for which the dependent variable is catch. Four areas in analysis are the Lower Basin (LB), Riley Creek (RC), the Upper Basin (UB), and Carroll Creek (CC). The Student–Newman–Keuls' (SNK) multiple-range test compares the variable catch among these sites.

Class Level Information

Class	Levels	Values
TIME	3	1 2 3
AREA	4	CC LB RC UB
REP	6	1 2 3 4 5 6

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	31	163.2083333	5.2647849	2.05	0.0166
Error	40	102.7777778	2.5694444		
Corrected total	71	265.9861111			
R^2	0.613597	Root MSE	1.602949		
CV	92.32984	COUNT mean	1.736111		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
AREA	3	65.26388889	21.75462963	8.47	0.0002
REP	5	21.40277778	4.28055556	1.67	0.1652
AREA*REP	15	43.31944444	2.88796296	1.12	0.3678
TIME	2	18.36111111	9.18055556	3.57	0.0374
TIME*AREA	6	14.86111111	2.47685185	0.96	0.4618

Test of Hypotheses Using Type III MSE for AREA*REP as an Error Term

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
AREA	3	65.26388889	21.75462963	7.53	0.0026

Student-Newman-Keuls' Test for CATCH by AREA[a]			
Number of means	2	3	4
Critical range	1.2073957	1.4713811	1.6326421
SNK Grouping	Mean	*N*	AREA
A	3.3333	18	LB
B	1.5556	18	RC
B			
B	1.1667	18	CC
B			
B	0.8889	18	UB

Student-Newman-Keuls' Test for CATCH by Time[b]			
Number of means	2	3	
Critical range	0.9352266	1.1262514	
SNK Grouping	Mean	*N*	TIME
A	2.4167	24	1
A			
B A	1.5833	24	2
B			
B	1.2083	24	3

[a] This test controls the type I experimentwise error rate under the complete H_0 but not under partial H_0s. Alpha = 0.05, *df* = 15, and MSE = 2.887963. Means with the same letter are not significantly different.

[b] This test controls the type I experimentwise error rate under the complete H_0 but not under partial H_0s. Alpha = 0.05, *df* = 40, and MSE = 2.569444. Means with the same letter are not significantly different.

The main-plot A test detected a significant ($P < 0.01$) difference in catch among areas. Student–Newman–Keuls' (SNK) multiple-range test based on the proper variance term (MSE of the area × replicate interaction) indicated that catch of fish in the Lower Basin was greater than that in the other three areas, and no differences in catch were evident among Riley Creek, Carroll Creek, and the Upper Basin. The Student–Newman–Keuls' test represents one of many multiple-range tests offered by SAS (2001). Other multiple-ranges tests that statistically separate mean values among treatments can be more or less likely to control the type I error rate.

In addition, the split-plot repeated-measures ANOVA indicated time, a temporal change in catch, was a significant ($P < 0.05$) term in the model. The multiple-range test showed that for all four areas combined, catch was highest during the first sampling time period and declined during the second and third sampling time period. This might be expected as young fish abundance would be expected to decline over time due to natural mortality. An interaction between area and time was not evident ($P = 0.46$) and, thus, did not confound interpretation of temporal effects due to differences in catch among areas for different time periods.

Box 4.4 The Use of Catch-Curve Regression to Identify Weak and Strong Year-Class Formation

This example contains a data set (data published in Maceina and Bettoli 1998) that uses catch-curve regression to detect strong and weak year-class formation in a largemouth bass population. In addition, a reservoir hydrologic variable is included that will be used later (see section 4.3.4) to examine the association between year-class strength and an environmental variable. In spring 1993, 653 age-2 to age-11 largemouth bass were collected using DC electrofishing. Age–length keys (Bettoli and Miranda 2001) were used to estimate the age structure for the entire sample from examination of 190 otoliths.

The SAS program below first computed the regression between the natural log of number at age (LNUM) against age and used the predicted values for the natural log of number at age (PLNUM) as weighting factors when the catch-curve analysis was recomputed. Thus, the second catch-curve regression computed the least-squares fit using the predicted values from the first fit as weights. From this regression, the residuals were computed and printed with the year-class (YEARCL) and age identified. For this analysis, we assumed all fish age 2 and older were fully recruited to the electrofishing gear and the fishery.

Program

```
DATA GUN_LMB;
INPUT YEARCL AGE NUM MEANRET @@;
LNUM=LOG(NUM + 1);
LMEANRET=LOG10(MEANRET);
LINES;
91 2  175  13.7 90 3  273  16.9 89 4   28   9.6 88 5   79  47.7
87 6   18  19.5 86 7   49  49.5 85 8   21  31.0 84 9    8   9.6
83 10   0  10.5 82 11   2  23.2
;
PROC REG NOPRINT; MODEL LNUM=AGE/R; ID YEARCL AGE;
OUTPUT OUT=A P=PLNUM;
DATA B; SET A;
W=PLNUM;
PROC PRINT; VAR YEAR AGE NUM LNUM W;
PROC REG; WEIGHT W; MODEL LNUM= YEARCL AGE/R; RUN;
```

Results

Table Data for 653 age-2 to age-11 largemouth bass collected using DC electrofishing. The number at age (NUM) is given with its associated weighting factor, LNUM (=$\log_e$[NUM + 1]). Weight is the predicted value for the natural log of number at age.

Observation	YEARCL	AGE	NUM	LNUM	Weight
1	1991	2	175	5.17048	5.48748
2	1990	3	273	5.61313	4.97418
3	1989	4	28	3.36730	4.46088
4	1988	5	79	4.38203	3.94758
5	1987	6	18	2.94444	3.43428
6	1986	7	49	3.91202	2.92098
7	1985	8	21	3.09104	2.40768
8	1984	9	8	2.19722	1.89438
9	1983	10	0	0.00000	1.38108
10	1982	11	2	1.09861	0.86778

Table Catch-curve regression for LNUM versus AGE.

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	47.50049	47.50049	23.584	0.0013
Error	8	16.11255	2.01407		
Corrected total	9	63.61304			
R^2	0.7467	Root MSE	1.41918		
Adjusted R^2	0.7150	LNUM mean	3.86169		
CV	36.75024				

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	6.344692	0.56991090	11.133	0.0001
AGE	1	–0.480520	0.09894628	–4.856	0.0013

Predicted LNUM (PLNUM) and Residuals from Weighted Regression

Observation	AGE	YEARCL	Weight	LNUM	PLNUM	Predicted SE	Residual
1	2	1991	5.4875	5.1705	5.3837	0.402	–0.2132
2	3	1990	4.9742	5.6131	4.9031	0.331	0.7100
3	4	1989	4.4609	3.3673	4.4226	0.277	–1.0553
4	5	1988	3.9476	4.3820	3.9421	0.252	0.4399
5	6	1987	3.4343	2.9444	3.4616	0.265	–0.5171
6	7	1986	2.9210	3.9120	2.9811	0.310	0.9310
7	8	1985	2.4077	3.0910	2.5005	0.377	0.5905
8	9	1984	1.8944	2.1972	2.0200	0.455	0.1772
9	10	1983	1.3811	0	1.5395	0.540	–1.5395
10	11	1982	0.8678	1.0986	1.0590	0.630	0.0396

Residuals and Associated Outlier Statistics from Weighted Regression

Observation	AGE	YEARCL	SE residual	Student residual	–2–1–0 1 2[a]	Cook's *D*
1	2	1991	0.453	–0.470	| | |	0.087
2	3	1990	0.544	1.306	| |** |	0.316
3	4	1989	0.612	–1.724	| *** | |	0.304
4	5	1988	0.668	0.658	| |* |	0.031
5	6	1987	0.719	–0.720	| * | |	0.035
6	7	1986	0.770	1.209	| |** |	0.119
7	8	1985	0.833	0.709	| |* |	0.051
8	9	1984	0.925	0.192	| | |	0.004
9	10	1983	1.080	–1.426	| ** | |	0.254
10	11	1982	1.387	0.029	| | |	0.000

(Box continues)

Box 4.4 *(continued)*

Residuals and Associated Outlier Statistics from Weighted Regression (*continued*)	
Sum of residuals	0
Sum of squared residuals	16.1125
Predicted residual SS (Press)	24.7446

[a] Graphical representation of Student residuals.

Using weighted regression for catch-curve analysis deflates the importance of rare and older fish when computing regression coefficients by proportioning the contribution to each product and cross product by the corresponding PLNUM–age data point. Thus, residuals for older and rarer cohorts of fish such as the 1983 year-class will be less with weighted than with unweighted regression. Generally, older and rarer year-classes are less likely to be accurately represented if a small to moderate sample of fish is collected, and this weighted regression procedure is recommended. Larger samples of fish will more accurately represent all year-classes, hence residuals from unweighted catch-curves regressions can be used.

From the analysis, the 1985 and 1988 year-classes could be considered moderately strong whereas the 1986 and 1990 year-classes were relatively more abundant and represented greater year-class strength. Conversely, the 1987 and 1991 year-classes were moderately weak, and even poorer year-class formation was evident for the 1983 and 1989 year-classes. The graphic below illustrates the number at age versus year-class (not a plot from SAS). The solid line represents the least-square fit to the data using weighted regression and the dashed lines are the residuals.

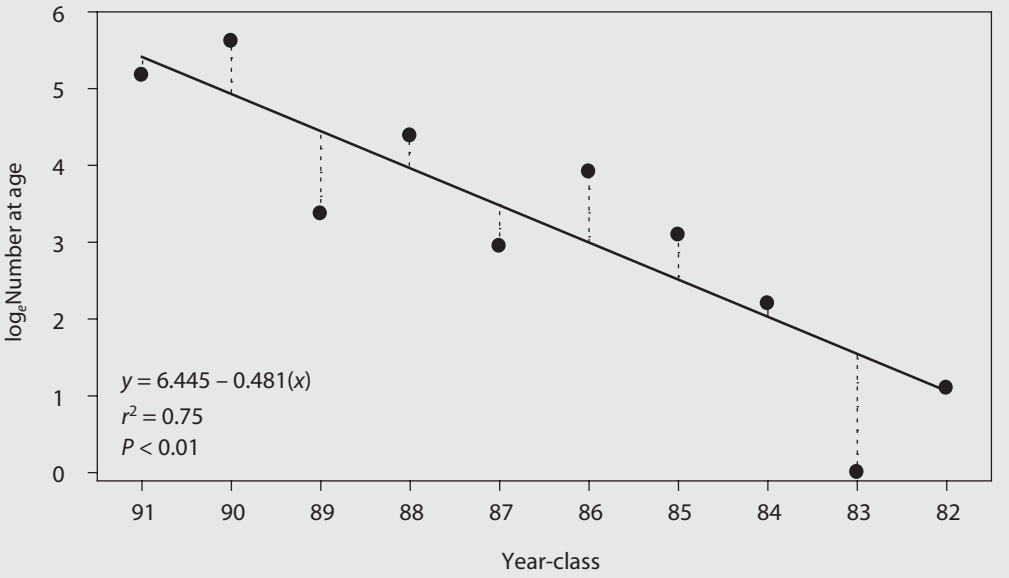


Figure The number at age versus year-class. The solid line represents the least-square fit to the data using weighted regression and the dashed lines are the residuals.

activity inhibited successful recruitment. Allen et al. (1999) found a positive correlation ($r = 0.61$) between chlorophyll-*a* concentrations and density of age-0 largemouth bass. Using linear regression, Reinert et al. (1997) used a host of reservoir hydrologic variables to explain 58–99% of the variation in electrofishing catch rates of age-0 and age-1 largemouth bass and spotted bass. Serns (1982) used linear regression and found that mean water temperature from June to August explained 74% of the variation in age-0 smallmouth bass density. Maceina and Stimpert (1998) found winter (January–March) retention in reservoirs prior to spawning and post-winter (April–November) retention were negatively and positively related to, respectively, black crappie and white crappie recruitment ($R^2 = 0.62$). An example of using correlative and regression techniques to explain recruitment variation is given in Box 4.5

The collection of long-term data to document recruitment variation and relate this variation to abiotic and biotic variables is desirable for fisheries scientists attempting to explain fluctuations in recruitment. In the absence of long-term data on recruitment levels, the residuals or errors associated with catch-curve regressions can be used as an index of recruitment variability and compared to biotic and abiotic variables (Maceina 1997). Maceina (1997) expanded the use of simple linear catch-curve regression to incorporate an additional independent environmental variable(s) (ENVIR) that was measured when fish were age 0. The generalized equation is

$$\log_e \text{number} = b_0 - b_1(\text{age}) \pm b_2(\text{ENVIR}). \tag{4.2}$$

For this equation, weighted regression is used to deflate the influence of rarer and older fish in the analysis similar to the procedures in Box 4.5. This technique has been used to explain environmental factors related to variation in fish recruitment for a number of species (Maceina and Bettoli 1998; Slipke et al. 1998; DiCenzo and Duval 2002; Maceina 2003), and an example is shown in Box 4.6. In addition, residuals can be pooled among water bodies and different years of collections and examined in relation to environmental variables to explain recruitment variation (Maceina and Bettoli 1998).

In regression analysis, transforming independent and dependent variables to natural log, common log, or inverse values can improve fit, reduce heteroscedastic variances, and sometimes explain better nonlinear fit between variables. Nonlinear regression can be a useful tool to explain and show graphically, for example, that progressively higher levels of some independent variable will result in an increase (or decrease) in some measure of recruitment before eventually reaching an asymptotic level.

4.4 RECRUIT–SPAWNER RELATIONSHIPS

An important component of fisheries management is to determine if a relationship exists between recruitment and spawner abundance. A quantitative understanding of the amount of recruitment that is necessary to sustain a fishery is

Box 4.5 Use of Correlation, Simple Regression, and Multiple Regression Analyses to Explain Recruitment Variation

From the data presented in Table 4.1, the relations between *C/f* of age-0 crappies (CPE0) and reservoir hydrologic conditions were determined. The respective year-classes (YEARCL) were also noted. The following SAS (2001) program plots bivariate relations between *C/f* of age-0 fish and hydrologic variables, computes the Pearson product moment correlation coefficients among age-0 catch and the reservoir hydrologic terms, and finally computes multiple regressions to describe and predict age-0 catch from these hydrologic variables.

Program

```
DATA WECRA;
INPUT YEARCL CPE0   WINSTAGE WINRET SPRSTAGE;
LINES;
1989      .         170.85    7.2    171.79
1990     8.03       171.76    4.2    171.75
(continue data input)
;
PROC PLOT; PLOT CPE0*WINSTAGE; PLOT CPE0*WINRET; PLOT CPE0*SPRSTAGE;
*/ plots not presented but should be examined by the fisheries scientist;
PROC CORR; VAR CPE0 WINSTAGE WINRET SPRSTAGE;
PROC REG; MODEL CPE0=WINSTAGE WINRET SPRSTAGE/SS1 SS2 SCORR1 PCORR2 VIF
COLLINOINT;
PROC REG; MODEL CPE0=WINSTAGE SPRSTAGE/SS1 SS2 SCORR1 PCORR2 VIF
COLLINOINT;
PROC REG; MODEL CPE0=WINSTAGE; RUN;
```

Results

Table Output for correlation analysis among *C/f* of age-0 crappies (CPE0) and the three hydrologic variables, mean winter stage (WINSTAGE), mean winter retention (WINRET), and mean spring stage (SPRSTAGE) (see Table 4.1). The Pearson correlation coefficient tests for $P > |R|$ under the H_0 that R (rho) = 0.

Simple Statistics						
Variable	*N*	Mean	SD	Sum	Minimum	Maximum
CPE0	10	3.25300	3.18380	32.53000	0.47000	9.89000
WINSTAGE	11	170.98545	0.32163	1880.84000	170.67000	171.76000
WINRET	11	6.47273	1.39075	71.20000	4.20000	9.70000
SPRSTAGE	11	171.82273	0.05781	1890.05000	171.75000	171.92000

Pearson Correlation Coefficients				
	CPE0	WINSTAGE	WINRET	SPRSTAGE
CPE0				
R	1.00000	0.89313	–0.56302	–0.03217
P	0.0	0.0005	0.0902	0.9297
N	10	10	10	10
WINSTAGE				
R	0.89313	1.00000	–0.71949	–0.29292
P	0.0005	0.0	0.0126	0.3820
N	10	11	11	11
WINRET				
R	–0.56302	–0.71949	1.00000	0.38038
P	0.0902	0.0126	0.0	0.2485
N	10	11	11	11
SPRSTAGE				
R	–0.03217	–0.29292	0.38038	1.00000
P	0.9297	0.3820	0.2485	0.0
N	10	11	11	11

Table Multiple and linear regression analyses for the dependent variable CPE0.

Analysis of Variance with Three Hydrologic Variables					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	79.79125	26.59708	13.952	0.0041
Error	6	11.43796	1.90633		
Corrected total	9	91.22921			
R^2	0.8746	Root MSE	1.38070		
Adjusted R^2	0.8119	CPE0 mean	3.25300		
CV	42.44383				

Parameter Estimates						
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|	Type I SS
Intercept	1	–4237.774608	1511.6915952	–2.803	0.0310	105.820090
WINSTAGE	1	9.614666	1.95648369	4.914	0.0027	72.771053
WINRET	1	0.084910	0.47521332	0.179	0.8641	1.008889
SPRSTAGE	1	15.110553	8.50931361	1.776	0.1261	6.011308

(Box continues)

Box 4.5 *(continued)*

Variable	*df*	Type II SS	Squared semi-partial correlation	Squared partial correlation	Variance inflation
Intercept	1	14.981201		0.00000000	
WINSTAGE	1	46.037663	0.79767273	0.80099459	2.03665416
WINRET	1	0.060861	0.01105884	0.00529277	2.22235140
SPRSTAGE	1	6.011308	0.06589236	0.34450202	1.22457891

Colinearity Diagnostics (Intercept Adjusted)

Number	Eigenvalue	Condition index	Variable proportion		
			WINSTAGE	WINRET	SPRSTAGE
1	1.99908	1.00000	0.0912	0.0913	0.0911
2	0.72319	1.66261	0.1262	0.0340	0.8576
3	0.27773	2.68288	0.7826	0.8747	0.0512

Analysis of Variance with Two Hydrologic Variables

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	79.73039	39.86519	24.268	0.0007
Error	7	11.49882	1.64269		
Corrected total	9	91.22921			
R^2	0.8740	Root MSE	1.28167		
Adjusted R^2	0.8379	CPE0 mean	3.25300		
CV	39.39976				

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|	Type I SS
Intercept	1	–4273.240691	1391.1235060	–3.072	0.0180	105.820090
WINSTAGE	1	9.380227	1.34721346	6.963	0.0002	72.771053
SPRSTAGE	1	15.553433	7.55648750	2.058	0.0786	6.959337

Variable	*df*	Type II SS	Squared semi-partial correlation type I	Squared partial correlation type 2	Variance inflation
Intercept	1	15.500246		0.00000000	
WINSTAGE	1	79.635947	0.79767273	0.87382619	1.12067509
SPRSTAGE	1	6.959337	0.07628408	0.37703313	1.12067509

Colinearity Diagnostics (Intercept Adjusted)				
			Variable proportion	
Number	Eigenvalue	Condition index	WINSTAGE	SPRSTAGE
1	1.32815	1.00000	0.3359	0.3359
2	0.67185	1.40600	0.6641	0.6641

Analysis of Variance with One Hydrologic Variable					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	72.77105	72.77105	31.540	0.0005
Error	8	18.45816	2.30727		
Corrected total	9	91.22921			
R^2	0.7977	Root MSE	1.51897		
Adjusted R^2	0.7724	CPE0 mean	3.25300		
CV	46.69443				

Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	–1445.158045	257.90658352	–5.603	0.0005
WINSTAGE	1	8.470290	1.50823184	5.616	0.0005

Based on *C/f* of age-0 crappies as a recruitment index, this variable was positively correlated to winter water stage prior to spawning ($R = 0.89$; $P < 0.01$) and weakly, but negatively, correlated to corresponding winter retention ($R = –0.56$; $P = 0.09$). No relation was evident between spring stage ($R = –0.03$; $P = 0.92$) and age-0 *C/f*. As expected, winter stage and retention were inversely related ($R = –0.72$; $P < 0.05$).

Of the three hydrologic terms used as regressors in multiple regression analysis of *C/f* of age-0 crappies, winter retention was the weakest independent variable. The overall model was highly significant ($P < 0.01$), and the three hydrologic terms explained about 87% of the variation in age-0 *C/f*. However, diagnostics used to detect multicolinearity among independent variables showed winter stage and winter retention covaried. The condition index was elevated (2.683), and the colinearity diagnostics (variable proportion) associated with this condition index was 0.783 and 0.875 for winter stage and retention, respectively, and showed multicolinearity existed in the model. In addition, winter retention was not a significant ($P = 0.86$) regressor in the model. Hence, winter retention was dropped from subsequent analyses, and age-0 *C/f* was regressed against winter stage and spring stage. We decided for this example to keep spring stage in the analysis even though winter stage was obviously the most influential regressor. In this multiple regression,

(Box continues)

Box 4.5 *(continued)*

the partial regression coefficient for spring stage was modestly significant ($P = 0.08$) and positive, which suggested after accounting for the effects of winter stage, slightly higher water levels in spring may enhance crappie recruitment. With the use of squared partial correlation coefficients (pr^2, squared partial correlation type II in table above), winter stage provided the greatest contribution ($pr^2 = 0.87$) to age-0 *C/f* compared with spring stage ($pr^2 = 0.38$). The squared semi-partial correlations (squared semi-partial correlation type I) were fitted to the independent variables in the order that they were entered into the model and showed that after first accounting for the effects of winter stage (79.8%), spring stage explained an additional 7.6% of the variance for age-0 *C/f*. The sum of the squared semi-partial correlations will equal the coefficient of determination for the entire model. Finally, a simple linear model was computed that regressed age-0 *C/f* to winter stage (see figure below; not a plot from SAS 2001).

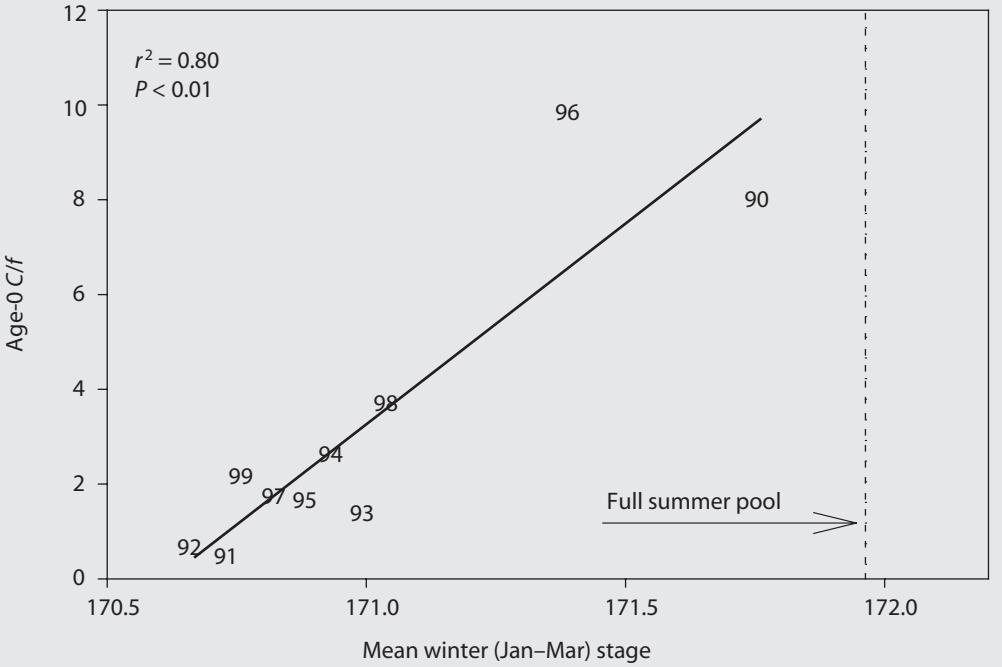


Figure Linear model of *C/f* of age-0 crappies versus winter stage. Numeric values along regression line refer to year-classes.

extremely useful information. Some fisheries scientists have argued that a quantifiable relation between recruits and spawners does not exist and that abiotic and biotic processes influence recruitment independently of spawner abundance (Van Den Avyle and Hayward 1999). However, Myers and Barrowman (1996) provided clear evidence for a positive relationship between recruits and spawners, though the results were more compelling for marine than for freshwater populations.

4.4.1 Types of Recruit–Spawner Relations

Several recruit–spawner models are commonly used, and detecting density dependence, or compensation, is of primary importance in fitting these models. Recruitment in wild populations will be limited by environmental constraints at relatively high densities, and therefore the rate of recruitment (i.e., the number of recruits produced per unit of spawners) may decrease at high levels of spawner abundance. Two common curves have two coefficients with similar functions. One coefficient (α) represents density-independent recruitment and is often referred to as the productivity coefficient. This is the rate of recruitment in the absence of any environmental constraints, and represents the slope of the stock–recruitment curve at the origin. The second coefficient (β) arises from density-dependent processes. At relatively high spawning stock levels various ecological processes (e.g., rate of predation, habitat or food limitations) will result in compensation in the survival of recruits, and recruitment rate will decline with an increase in spawner abundance.

Box 4.6 Incorporation of an Environmental Term into a Catch-Curve Regression to Explain Fluctuations in Recruitment

From the data presented in the SAS program in Box 4.4 and the program below, April–July retention will first be plotted against the residuals from the weighted catch-curve regression for largemouth bass. Then, this term will be added to the simple linear catch-curve regression to compute a multiple regression. The mean retention (MEANRET) between April–July corresponds to the hatching and post-hatching time period for each year-class when fish were age 0 (Maceina et al. 1995). The variables YEARCL, AGE, NUM, and LNUM are defined in Box 4.4.

Program

```
DATA GUN_LMB;
INPUT YEARCL  AGE NUM MEANRET;
LNUM=LOG(NUM + 1);
LMEANRET=LOG10(MEANRET);
LINES;
91 2  175  13.7 90 3  273  16.9 89 4   28   9.6 88 5   79  47.7
(continue data input)
;
PROC REG NOPRINT; MODEL LNUM=AGE/R; ID YEARCL AGE;
OUTPUT OUT=A  P=PLNUM;
DATA B; SET A; W=PLNUM;
PROC REG NOPRINT; WEIGHT W; MODEL LNUM=AGE/R;
OUTPUT OUT=C R=RES;
PROC PLOT; PLOT RES*MEANRET/VREF=0;
PROC CORR; VAR RES MEANRET;
PROC REG; WEIGHT W; MODEL LNUM=AGE LMEANRET/SS1 SS2 PCORR2; RUN;
```

(Box continues)

Box 4.6 *(continued)*

Results

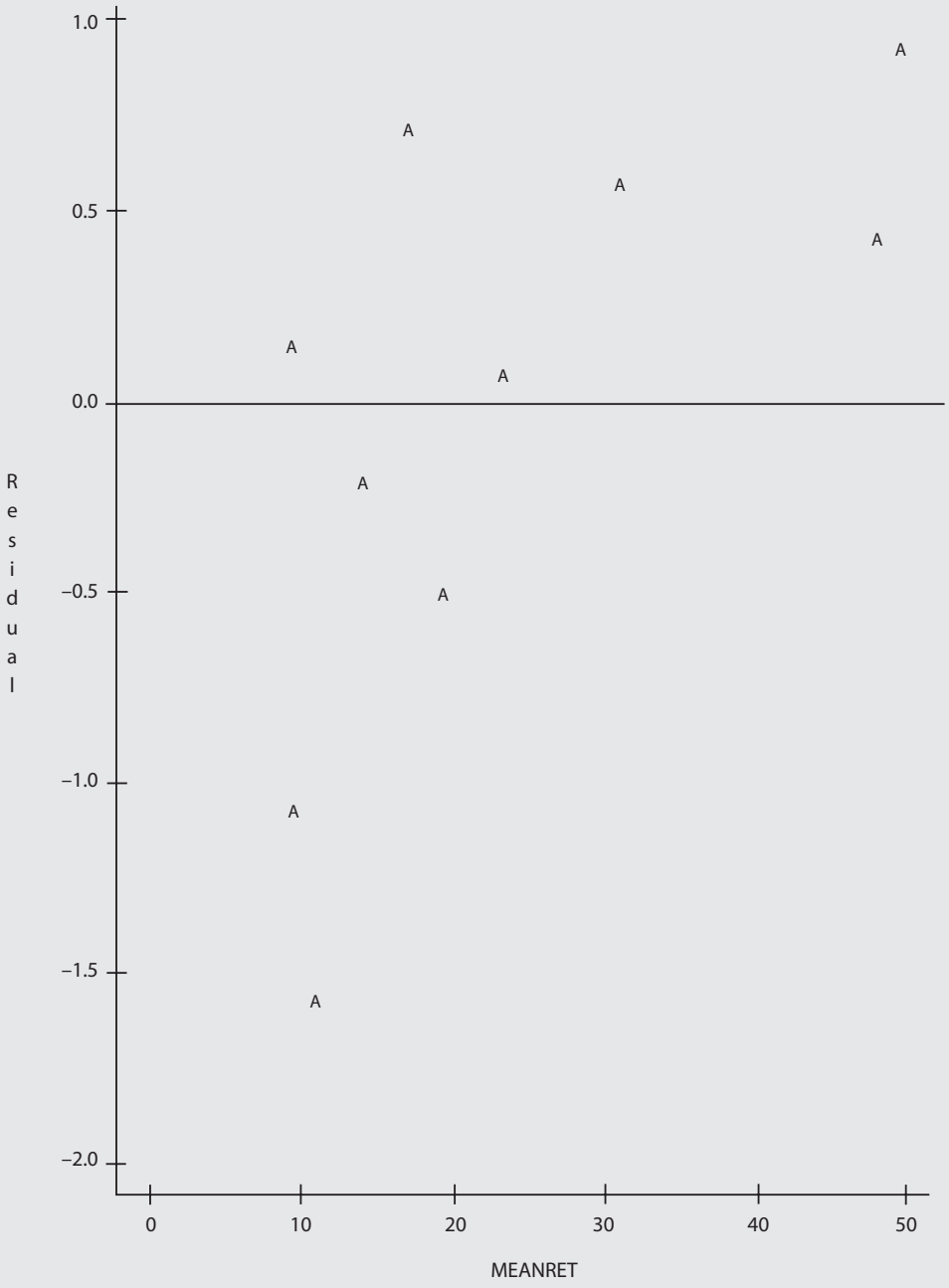


Figure Plot of residuals from the weighted catch-curve regression for largemouth bass versus mean retention, in which A represents one observation.

Table Output for correlation analysis between residuals (RES) from the weighted catch-curve regression for largemouth bass and the mean retention (MEANRET) in the reservoir between April–July.

Simple Statistics

Variable	*N*	Mean	SD	Sum	Minimum	Maximum	Label
RES	10	–0.0437	0.7975	–0.4368	–1.5395	0.9310	Residual
MEANRET	10	23.1200	15.0095	231.2	9.6000	49.5000	

Pearson Correlation Coefficients

	RES	MEANRET
RES		
R	1.00000	0.65716
P	0.0	0.0390
MEANRET		
R	0.65716	1.00000
P	0.0390	0.0

Table Multiple regression analysis for the dependent variable LNUM ($\log_e$[NUM + 1]) of largemouth bass. Log_{10}(mean retention) is given by LMEANRET.

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	55.57442	27.78721	24.197	0.0007
Error	7	8.03862	1.14837		
Corrected total	9	63.61304			
R^2	0.8736	Root MSE	1.07162		
Adjusted R^2	0.8375	LNUM mean	3.86169		
CV	27.75011				

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	3.950567	1.00022232	3.950	0.0055
AGE	1	–0.526061	0.07666302	–6.862	0.0002
LMEANRET	1	2.048881	0.77270892	2.652	0.0329

(Box continues)

Box 4.6 *(continued)*

Variable	*df*	Type I SS	Type II SS	Squared partial correlation type II
Intercept	1	473.867469	17.914685	.
AGE	1	47.500491	54.073342	0.87057856
LMEANRET	1	8.073930	8.073930	0.50109573

The SAS plot of residuals of the weighted catch-curve regression against April–July retention (MEANRET) showed higher retention was associated with progressively higher residuals, or stronger year-class formation for largemouth bass, whereas lower retention (<15 d) was associated with lower recruitment for those year-classes, and the relation was not linear. The significant correlation ($R = 0.66$; $P < 0.05$) was computed between catch-curve residuals and April–July retention, and thus the variation about the catch-curve plot in Box 4.4 was related to this hydrologic term. The plot of catch-curve residuals against April–July appeared nonlinear, hence retention was transformed to $\log_{10}$ values and the multiple regression equation computed. The addition of the retention term explained an additional 13% of the variation in the catch-curve above that explained by the simple catch-curve regression (Box 4.4). Based on the squared partial correlation coefficient ($pr^2 = 0.50$), retention was an important variable in explaining the variation in number at age beyond that explained by age alone (see Maceina and Bettoli 1998 for further analysis).

4.4.1.1 *Beverton–Holt Recruit–Spawner Curve*

The recruitment curve developed by Beverton and Holt (1957) assumes that competition among early life stages for any limited resource (e.g., food or space) will cause recruits (R) to increase initially, then to decline to an asymptotic value as spawner abundance (S) increases (Figure 4.1). One form of the Beverton–Holt curve is

$$R = \frac{\alpha S}{1 + \beta S}, \tag{4.3}$$

where α is the productivity or density-independent coefficient and is the maximum recruitment rate (R/S) at low spawner abundance (i.e., the initial slope), and β determines the level of density dependence. Maximum recruitment represented by the asymptote is equal to α/β. An example of estimating a Beverton–Holt recruit–spawner curve is given in Box 4.7.

4.4.1.2 *Ricker Recruit–Spawner Curve*

In some fish populations, the recruit–spawner relation may be dome shaped, with the number of recruits declining at higher levels of spawner abundance due to overcompensation (Ricker 1954). Overcompensation may arise from such obvious processes as cannibalism but more importantly can be induced by predation on

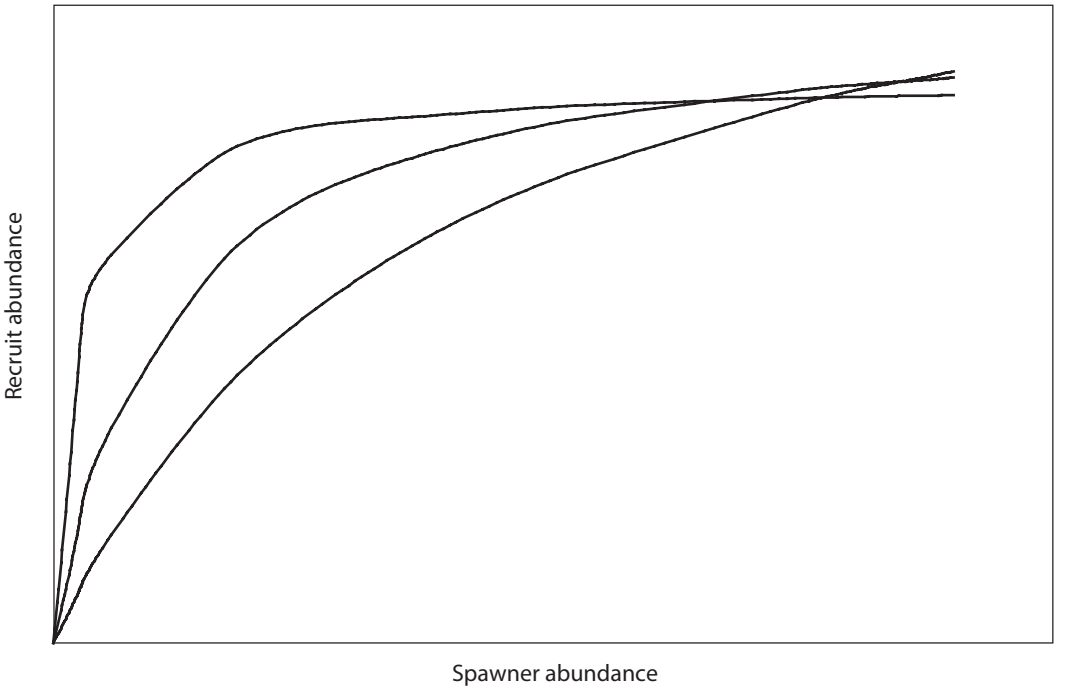


Figure 4.1 Beverton–Holt recruit–spawner curves computed from equation (4.3). The three curves show different productivity coefficients (α) but approximately the same maximum level of density dependence (β), showed by the asymptote.

prerecruits, including predation by other species. Specifically, if a predator responds to increased prey (i.e., the potential recruits in our stock–recruit relationship) either by increasing its own abundance or predatory effectiveness, then overcompensation may occur. Thus, while cannibalism may be an attractive explanation for overcompensation, predation by a variety of species may also result in a domed recruit–spawner curve. Ricker (1954) also suggested that overcompensation may be a prevalent condition in lake ecosystems that are more confining and potentially habitat limited in comparison to large marine systems. Hilborn and Walters (1992) reported that overcompensation will also arise when growth of young fish is density dependent and predation is size dependent. Therefore, fisheries scientists should examine their data for a dome shaped recruitment curve even for species that are not expected to be cannibalistic. In semelparous salmon species, overcompensation may arise from superpositioning of spawning redds and disease outbreaks affecting egg mortality at high spawner densities (Ricker 1954) .

The Ricker curve is defined by

$$R = Se^{\alpha - \beta S}, \tag{4.4}$$

where α is the density-independent or productivity coefficient, representing the rate of recruitment at low spawner abundance (i.e., the slope at the origin), and β

Box 4.7 Computation of the Beverton–Holt Recruit-Spawner Curve

From 1991 to 1996, crappies were collected from three Alabama reservoirs (LAKE) that displayed similar hydrologic conditions (data from Ozen 1997); 16 to 20 trap nets were used as described in Box 4.1. Fish were collected in the fall of each year, aged, and weighed (1 g). The variable SPAWNER was determined by dividing total weight of all age-2 and older crappies (assumed to be adults) by the number of net-nights of effort and RECRUIT was determined by dividing the total number of age-0 crappies by the number of net-nights of effort. The program below plots the relation between recruits and spawners, then describes the relations between recruits and spawners using nonlinear regression for untransformed and $\log_e$ transformed data (equations [4.3] and [4.5], respectively). From the last nonlinear regression, predicted recruits (PRE_LREC) was regressed against observed recruits to provide additional statistical inference. The predicted number of recruits and associated residuals from the last nonlinear regression were derived and printed. In the nonlinear procedure in SAS (PROC NLIN), the parameters (PARMS) statement refers to approximate coefficients for α (A in SAS) and β (B in SAS) in the nonlinear regression that are provided by the fisheries scientist to initiate the analysis. Hougaard's skewness values for α and β were computed for each nonlinear regression. Finally, residual values from the last nonlinear regression were summed.

Program

```
DATA REC_SPA;
INPUT LAKE $ YEAR SPAWNER RECRUIT @@;
*/ SPAWNER = INDEX OF ADULT BIOMASS OF AGE-2 AND OLDER CRAPPIE;
*/RECRUIT = TRAP NET CATCH RATE OF AGE-0 CRAPPIE;
LRECRUIT=LOG(RECRUIT);
LINES;
AL 91 340 5.41 AL 92 907 3.00 AL 93 171 2.41 AL 94 1040 2.25
AL 95 55 0.41 AL 96 524 8.71
DE 92 213  1.13 DE 93 1034 4.66 DE 94 457 1.94 DE 95 200 7.28
DE 96  669 10.56
JB 90 372 9.33 JB 91 386 2.19 JB 92 585 6.75 JB 93 660 13.85
JB 94 337 3.58 JB 95 396 3.48 JB 96 620 23.70
;
PROC PLOT; PLOT RECRUIT*SPAWNER; */ plot not shown;
PROC NLIN HOUGAARD; PARMS A=0.03 B=0.002;
MODEL RECRUIT=(A*SPAWNER)/(1 + B*SPAWNER);
PROC NLIN HOUGAARD; PARMS A=0.01 B=0.002;
MODEL LRECRUIT=LOG((A*SPAWNER)/(1 + B*SPAWNER));
OUTPUT OUT=A P=PRE_LREC R=RES;
PROC REG; MODEL PRE_LREC=LRECRUIT;
DATA B; SET A;
PRE_REC=EXP(PRE_LREC);
PROC PRINT; VAR LAKE YEAR RECRUIT SPAWNER PRE_LREC RES;
PROC MEANS NOPRINT; VAR RES;
OUTPUT  OUT=B SUM=SUMRES;
PROC PRINT; VAR SUMRES; RUN;
```

Results

Table Nonlinear regression (NLIN) of RECRUIT (total number of age-0 crappies divided by number of net-nights of effort) versus SPAWNER (total weight of age-2 and older crappies divided by number of net-nights of effort). In the estimation summary, *R*, PPC(B), and RPC(B) are measures and diagnostics of the degree of convergence of the model; smaller values represent better model fit. An intercept was not specified for this model.

Iterative Phase

Iteration	A	B	SS
0	0.0300	0.00200	502.9
1	0.0418	0.00388	486.2
.			
.			
.			
17	0.0386	0.00369	484.5
18	0.0386	0.00369	484.5

Estimation Summary

Method	Gauss-Newton
Iterations	18
R	9.328×10^{-6}
PPC(B)	0.000068
RPC(B)	0.000103
Object	6.74×10^{-11}
Objective	484.4791
Observations read	18
Observations used	18
Observations missing	0

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	Approximate *P* > *F*
Regression	2	747.8	373.9	12.35	0.0006
Residual	16	484.5	30.2799		
Uncorrected total	18	1232.3			
Corrected total	17	552.2			

Parameter Estimates

Parameter	Estimate	Approximate SE	Approximate 95% confidence limits		Hougaard's skewness
A	0.0386	0.0474	–0.0619	0.1391	5.6771
B	0.00369	0.00675	–0.0106	0.0180	6.1140

(Box continues)

Box 4.7 *(continued)*

Approximate Correlation Matrix		
	A	B
A	1.0000000	0.9864924
B	0.9864924	1.0000000

Table Nonlinear regression of $\log_e$RECRUIT (LRECRUIT) versus SPAWNER, with right side of equation (4.5) transformed using a natural log. This is followed by a linear regression between predicted (PRE_LREC) and observed recruits. Nonlinear regression model details as above.

Iterative Phase			
Iteration	A	B	SS
0	0.0100	0.00200	18.3365
1	0.0147	0.00110	10.2141
.			
.			
.			
9	0.0170	0.00142	10.1303
10	0.0170	0.00142	10.1303

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	Approximate *P* > *F*
Regression	2	42.5153	21.2576	9.38	0.0074
Residual	16	10.1303	0.6331		
Uncorrected total	18	52.6455			
Corrected total	17	16.0679			

Parameter Estimates					
Parameter	Estimate	Approximate SE	Approximate 95% confidence limits		Hougaard's skewness
A	0.0170	0.00934	–0.00282	0.0368	2.6078
B	0.00142	0.00192	–0.00264	0.00549	2.9081

Linear Regression of Predicted versus Observed Recruits					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	1.66803	1.66803	9.71	0.0067
Error	16	2.74841	0.17178		
Corrected total	17	4.41644			
R^2	0.3777	Root MSE	0.41446		
Adjusted R^2	0.3388	PRE_LREC mean	1.42551		
CV	29.07430				

Table Predicted number of recruits (PRE_LREC) and associated residuals (RES) and the sum of the residuals (SUMRES). Also given are the variables SPAWNER, RECRUIT, LAKE (from Ozen 1997), and YEARCL.

Observation and sum	LAKE	YEARCL	RECRUIT	SPAWNER	PRE_LREC	RES
1	AL	1991	5.41	340	3.89039	0.32974
2	AL	1992	3.00	907	6.72230	–0.80682
3	AL	1993	2.41	171	2.33517	0.03154
4	AL	1994	2.25	1040	7.11974	–1.15194
5	AL	1995	0.41	55	0.86608	–0.74782
6	AL	1996	8.71	524	5.09635	0.53595
7	DE	1992	1.13	213	2.77528	–0.89854
8	DE	1993	4.66	1034	7.10312	–0.42152
9	DE	1994	1.94	457	4.70153	–0.88520
10	DE	1995	7.28	200	2.64343	1.01305
11	DE	1996	10.56	669	5.81874	0.59599
12	JB	1990	9.33	372	4.12979	0.81501
13	JB	1991	2.19	386	4.23010	–0.65832
14	JB	1992	6.75	585	5.42008	0.21943
15	JB	1993	13.85	660	5.77837	0.87416
16	JB	1994	3.58	337	3.86719	–0.07717
17	JB	1995	3.48	396	4.30019	–0.21163
18	JB	1996	23.70	620	5.59234	1.44408
SUMRES						-2.02×10^{-9}

Interpretation

For 18 iterations, an optimal solution (convergence criteria met) was found that minimized the residual error for this recruit–spawner data. The analysis indicated that the nonlinear regression was highly significant ($F = 12.35$; $df = 2, 16$; $P < 0.01$), but this test is highly suspect (see section 4.4.3). The nonlinear regression procedure in SAS (2001) does not compute a coefficient of determination, but this can be approximated by subtracting the residual SS from the corrected total SS and then dividing this by the corrected SS. For this example, $r^2 = (552.2 - 484.5)/552.2 = 0.12$. The spawner abundance did not explain a high proportion in variation in crappie recruits, although statistically significant. The coefficients for α (0.0386) and β (0.00369) are given along with approximate SEs and 95% confidence limits for these coefficients. Thus from equation (4.3),

$$R = 0.0386 \times S/(1 + 0.00369 \times S) .$$

Hougaard's skewness values for α and β were 5.7 and 6.1, respectively, which were high and indicated these parameters were not normally distributed and were potentially biased. Thus, the equation may be inaccurate. Finally a correlation matrix was presented that estimates the relation between α and β; correlations typically will also be high as these coefficients will covary when the least-squares solution is computed through iteration.

(Box continues)

Box 4.7 *(continued)*

The second model includes lognormal error structure by taking the natural logarithms of both sides of the nonlinear Beverton–Holt recruit–spawner equation An optimal solution was found (convergence criteria met), and the output suggested that the regression was significant ($P < 0.01$). The approximate r^2 value was 0.37 ([16.07 – 10.13]/16.07). Next, the coefficients are given for α (0.0170) and β (0.00142) with approximate SEs and confidence intervals. Hougaard's skewness values were still high for α and β (2.6 and 2.9) but lower than those computed for normal error structure. The next analysis presents the linear regression between predicted and observed recruits ($\log_e$ transformed). The *F*-statistic (9.71) and r^2 value (0.38) were very similar to those derived from the previous nonlinear regression.

The last table contains a print of predicted crappie recruits (PRE_LREC) and residuals (RES) for a given level of crappie spawners computed from the last nonlinear regression equation. Observed and predicted recruits and spawners can be used to construct a bivariate plot, and the nonlinear regression line (not a plot from SAS) of the relation between crappie recruits and spawners is shown below. The sum of the residuals (SUMRES) was approximately 0, which suggested an optimal least-squares fit to the data.

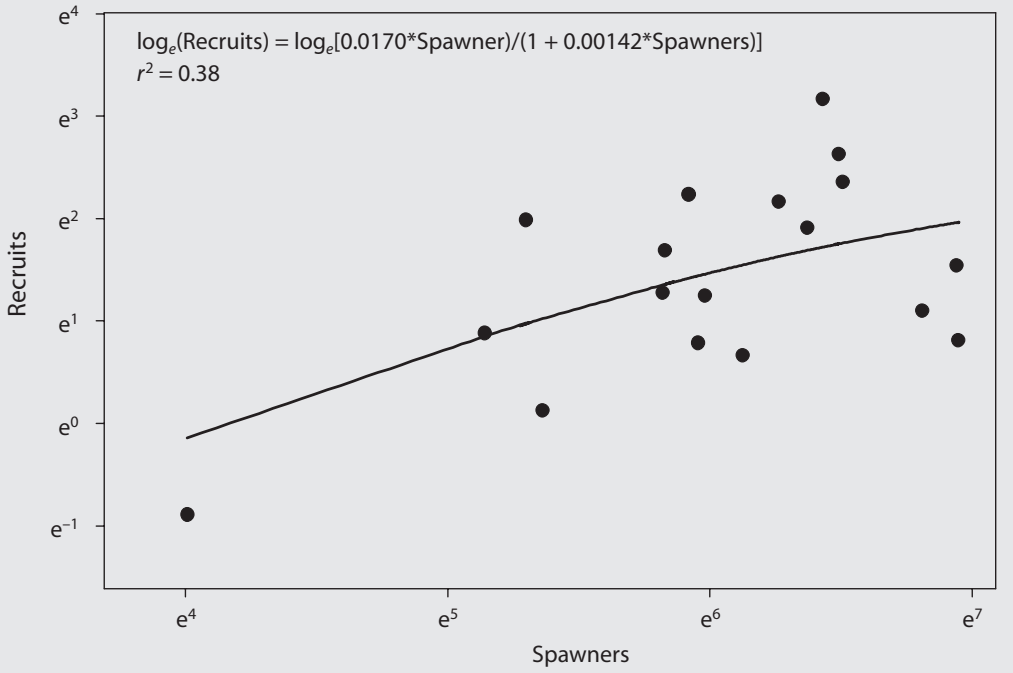


Figure Plot of the raw data ($\log_e$ scale) and the nonlinear regression of the relation between crappie recruits and spawners.

is the density-dependent coefficient, with the curve reaching a maximum at β^1 units of spawners before declining (Figure 4.2). An example of computing a Ricker recruit–spawner curve is given in Box 4.8.

4.4.1.3 *Additional Recruit–Spawner Curves*

Fisheries scientists may attempt to construct an alternative recruit–spawner curve that differs from the Beverton–Holt and Ricker forms. Alternate curves may integrate other processes that affect the early life history of a species. The Beverton–Holt and Ricker recruit–spawner curves typically will accommodate most of these conditions. However, a variety of other mathematical models for describing these relationships can be found in Cushing (1971, 1973), Deriso (1980), Shepherd (1982), with further elaboration by Schnute (1985), and Reish et al. (1985).

4.4.2 Estimation of Recruit–Spawner Coefficients

There are several methods available to fisheries scientists for fitting models to recruit–spawner data, and these include both linear and nonlinear regression procedures. Extreme data points, especially those associated with abnormally high

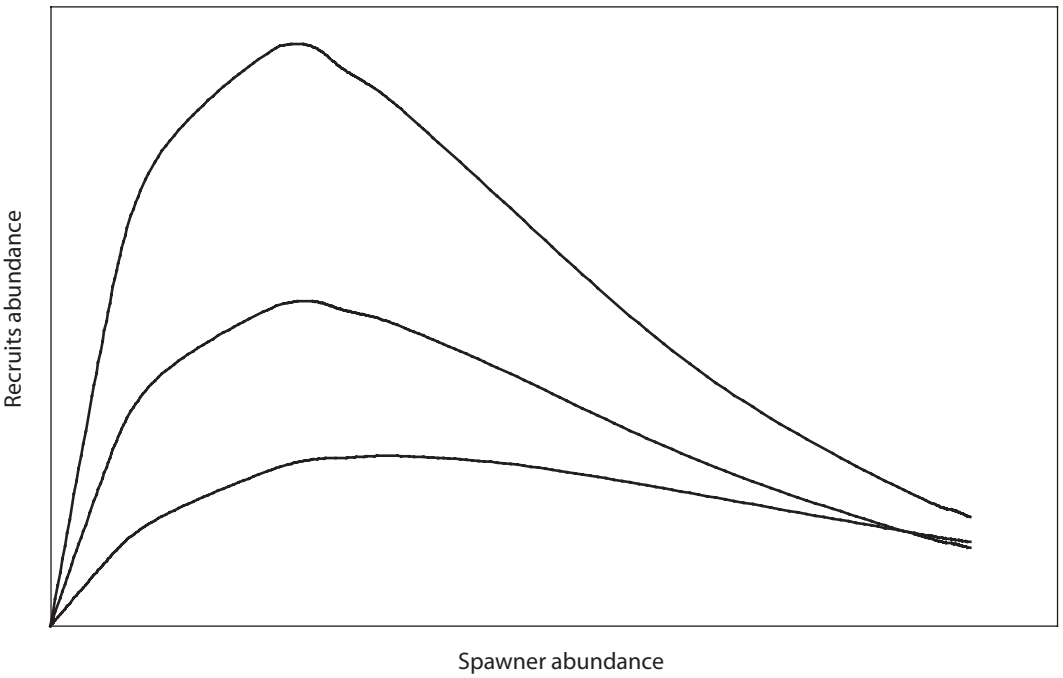


Figure 4.2 Ricker recruit–spawner curves computed using equation (4.4). The three curves have different values for the density-independent coefficient (α) but display approximately the same values for the density-dependent coefficient (β). This results in variation in the maximum recruitment, but the maximum occurs at approximately the same abundance of spawners.

Box 4.8 Computation of Ricker Recruit–Spawner Curves with the Inclusion of an Environmental Term to Explain Recruit Variation

Population estimates for age-5 and older adult walleye (SPAWNER) and age-0 walleye (RECRUIT) were made in Escanaba Lake, Wisconsin, from 1958 to 1991 (data presented in Hansen et al. 1998; see Table 4.3). The following SAS (2001) program computes a nonlinear regression to describe the relation between recruits and spawners assuming lognormal error structure (equation [4.6]). From this regression, predicted recruits are regressed against observed recruits to provide additional statistical inference. Next the program computes the Ricker recruit–spawner relation (equation [4.7]) using linear regression. The corrected coefficient of determination and associated *F*-statistic was given by regressing predicted recruits against observed recruits. Finally, the program also computes the nonlinear regression with lognormal error structure in the recruit–spawner relation to include the variation in May air temperature (MTEMPCV) as an additional regressor of walleye recruits (equation [4.9] modified to include lognormal error structure).

Program

```
DATA WALLEYE;
NPUT YEAR RECRUIT SPAWNER MTEMPCV;
*/AGE0 AND AGE5 IS THE NUMBER OF WALLEYE IN ESCANABA LAKE, WI;
*/ MAYTEMP IS THE CV FOR MAY AIR TEMPERATURES;
RATIO=RECRUIT/SPAWNER;
LRATIO=LOG(RATIO);
LRECRUIT=LOG(RECRUIT);
LINES;
1958    4532    775     0.24125
1959    22996   2310    0.16319
.
.
.
1990    35607   735     0.19356
1991    4876    1261    0.32032
;
PROC NLIN DATA=WALLEYE HOUGAARD; PARMS A=4 B=0;
MODEL LRECRUIT=LOG(SPAWNER*EXP(A + B*SPAWNER));
OUTPUT OUT=A P=PRE;
PROC REG; MODEL PRE=LRECRUIT;
PROC REG DATA=WALLEYE; MODEL LRATIO=SPAWNER;
OUTPUT OUT=B P=P R=RESIDUAL;
DATA C; SET B;
PRATIO= EXP(P);
PRECRUIT=PRATIO*SPAWNER;
LPREC=LOG(PRECRUIT);
PROC REG; MODEL LPREC=LRECRUIT;
PROC NLIN DATA=WALLEYE HOUGAARD; PARMS A=4 B=0 C=-7.0;
MODEL LRECRUIT=LOG(SPAWNER*EXP(A + B*SPAWNER + C*MTEMPCV));
RUN;
```

Results

Table Ricker recruit–spawner curve using nonlinear regression and accounting for lognormal error structure. The Gauss-Newton method is employed and convergence criterion was met.

Iterative Phase			
Iteration	A	B	SS
0	4.0000	0	194.0
1	3.3916	–0.00118	33.7931

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	Approximate $P > F$
Regression	2	2724.0	1362.0	2.13	0.1540
Residual	32	33.7931	1.0560		
Uncorrected total	34	2757.8			
Corrected total	33	36.0449			

Parameter Estimates					
Parameter	Estimate	Approximate SE	Approximate 95% confidence limits		Hougaard's skewness
A	3.3916	0.4118	2.5529	4.2303	-771×10^{-20}
B	–0.00118	0.000302	–0.00179	–0.00056	-292×10^{-19}

Table Linear regression between predicted $\log_e$RECRUIT (PRE; from the previous nonlinear regression) and observed recruits ($\log_e$ transformed).

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	$P > F$
Model	1	0.15590	0.15590	2.14	0.1534
Error	32	2.33335	0.07292		
Corrected total	33	2.48925			
R^2	0.0626	Root MSE	0.27003		
Adjusted R^2	0.0333	PRE mean	8.94712		
CV	3.01809				

(Box continues)

Box 4.8 *(continued)*

Table Linear regression between $\log_e$(RECRUIT/SPAWNER) (LRATIO) versus SPAWNER and linear regression between re-predicted (LPREC) and observed recruits.

Analysis of Variance of LRATIO versus SPAWNER					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	16.04272	16.04272	15.19	0.0005
Error	32	33.79315	1.05604		
Corrected total	33	49.83587			
R^2	0.3219	Root MSE	1.02764		
Adjusted R^2	0.3007	LRATIO mean	1.94113		
CV	52.93998				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	3.39157	0.41176	8.24	<0.0001
SPAWNER	1	–0.00118	0.00030179	–3.90	0.0005

Analysis of Variance of LPREC versus Observed Recruits					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	0.15590	0.15590	2.14	0.1534
Error	32	2.33335	0.07292		
Corrected total	33	2.48925			
R^2	0.0626	Root MSE	0.27003		
Adjusted R^2	0.0333	LPREC mean	8.94712		
CV	3.01809				

Table Ricker recruit–spawner curve from nonlinear regression using lognormal error structure and including the variation in May air temperature as a environmental predictor (coefficient C) of recruitment variation (LRECRUIT). The Gauss-Newton method is employed and convergence criterion was met.

Iterative Phase				
Iter	A	B	C	SS
0	4.0000	0	–7.0000	32.5687
1	4.7915	–0.00073	–7.8388	21.9990

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Regression	3	2735.8	911.9	9.90	0.0005
Residual	31	21.9990	0.7096		
Uncorrected total	34	2757.8			
Corrected total	33	36.0449			

Parameter Estimates					
Parameter	Estimate	Approximate SE	Approximate 95% confidence limits		Hougaard's skewness
A	4.7915	0.4815	3.8095	5.7736	2.23×10^{-16}
B	–0.00073	0.000271	–0.00128	–0.00018	1.06×10^{-16}
C	–7.8388	1.9228	–11.7604	–3.9173	-485×10^{-19}

Interpretation

The nonlinear regression for the Ricker recruit–spawner relation assuming lognormal error structure converged quickly to find an optimal least-squares fit to the data. However, based on this nonlinear regression and the linear regression between predicted and observed recruits, the relation was not significant ($F = 2.13 - 2.14$; $P = 0.15$), and spawners only explained about 6% of the variation in recruits. The confidence intervals for α and β were positive and negative, respectively, and did not overlap with 0, which suggested that adult walleye abundance explained only a small percentage of the total variation in walleye recruits. Hougaard's skewness values approximated 0 and showed the parameter coefficients were normally distributed and potentially not biased. Walleye recruitment in Escanaba Lake was weakly explained as

$$\log_e R = \log_e(Se^{3.392 - 0.001885}).$$

Similar to the Beverton–Holt recruit–spawner relation, α and β were just approximated in the PARMS statement, and SAS (2001) estimates the optimal solution for these two coefficients.

For the liner regression method (equation [4.7]) between the $\log_e$ of the ratio of recruit to spawner (LRATIO) against spawner, identical values for α (3.3916) and β (–0.00118) were derived by this method as compared with the nonlinear regression, but the computed coefficient of determination and F-statistic were much higher. Note that the errors, or residual SS, were identical (33.79) for both the linear and nonlinear regression, but the corrected total SS varied. However, the linear regression of predicted versus observed recruits for the linear method (equation [4.7]) computed nearly an identical coefficient of determination (0.06) and *F*-statistic ($F = 2.14$) to that derived in the previous analysis using nonlinear regression.

(Box continues)

Box 4.8 *(continued)*

Finally, the variation in May air temperature (MTEMP) that occurred during walleye hatch in Escanaba Lake was obviously a highly significant regressor of recruits when included in the recruit–spawner curve and improved the approximate coefficient of determination from 0.062 to 0.390. The equation was

$$\log_e R = \log_e[Se^{4.792 - 0.00730S - 7.839(\text{MTEMP})}].$$

The slope coefficient for the variation in May air temperature was negative and indicated that greater fluctuations in air temperatures in May was related to reduced production of walleye recruits. Hougaard's skewness values approximated 0 and indicated the three parameters included in the nonlinear regression were normally distributed. Inclusion of additional regressors can be tested using full and reduced model techniques presented in Montgomery and Peck (1982), which test for the reduction in the residual SS or error in the full model. Fisheries scientists should be cautious and conservative when adding additional predictors to stock–recruitment models. A plot (not a SAS 2001 plot) of the recruit–spawner relation and the nonlinear regressions for normal and lognormal error structure is presented below for walleye from Lake Escanaba.

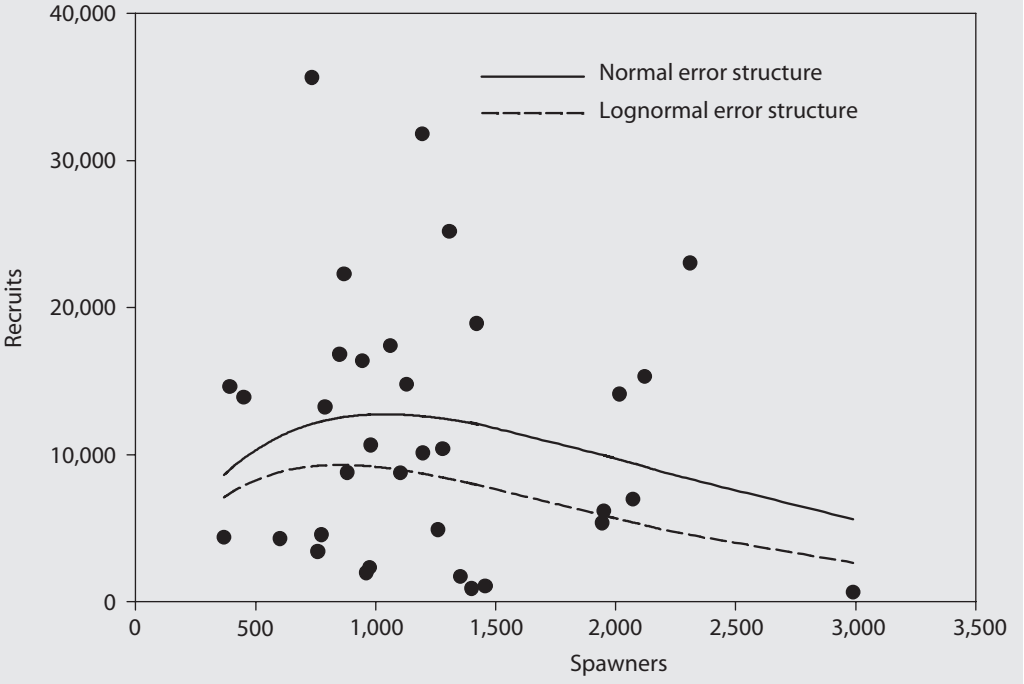


Figure Recruit–spawner relation and the nonlinear regressions for normal and lognormal error structure is presented below for walleye from Lake Escanaba, Wisconsin.

number of recruits, will influence the curve, and thus robust-fitting algorithms may also be appropriate. An important initial consideration in fitting recruit–spawner curves is the error distribution for the model. Many parameter estimation methods have the common assumption that the model residuals are normally distributed, but Peterman (1981) showed that models with lognormal errors were more appropriate to recruit–spawner data. Two common observations of recruit–spawner data are that (1) distributions are skewed to the left and display a long right tail, and (2) the amount of variation will be proportional to the average recruitment expected at a given spawner abundance, and recruitment will appear more variable at high spawner abundance when compared with low abundance. Thus, lognormal errors typically are apparent, and data transformation is usually necessary.

Nonlinear regression algorithms can be used to fit the Beverton–Holt (equation [4.3]) and Ricker (equation [4.4]) models that will provide for lognormal errors by taking the logarithm of both sides of these equations. Expressed in this manner, the Beverton–Holt (equation [4.5]) and Ricker (equation [4.6]) models are

$$\log_e(R) = \log_e[\alpha S/(1 + \beta S)], \tag{4.5}$$

and

$$\log_e(R) = \log_e(S^{\alpha - \beta S}). \tag{4.6}$$

These equations can be fitted using nonlinear procedures in SAS (2001; see Boxes 4.7 and 4.8). Equation (4.6) can also be expressed in a linear form as

$$\log_e(R/S) = \alpha - \beta S, \tag{4.7}$$

which retains the lognormal equation structure contained in equation (4.6) (see Box 4.8). Thus, the intercept will provide an estimate of productivity parameter, α, with the density-dependent term, β, estimated by the slope of equation (4.7). Equations (4.6) and (4.7) provide identical estimates of α and β (see Box 4.8), but statistical properties and associated inferences usually vary between these two equations (see section 4.4.3).

4.4.3 Statistical Properties and Inferences Associated with Recruit–Spawner Models

Recruit–spawner models can be computed using nonlinear regression techniques, but some of the properties that exist for linear regression do not apply for nonlinear least-squares estimates (Neter et al. 1996). For nonlinear regression, the sum of the residuals may not equal 0 and nonlinear models have no intercept, hence the corrected sum of squares (SS) may have no meaning (Freund and Littell 1991). The error or residual SS and the regression SS may not necessarily sum to the total SS for nonlinear models (Neter et al. 1996). Thus, the computation of the coefficient of determination and associated mean squares used in *F*-tests will

likely be incorrect. Finally, the linear method (equation [4.7]) to estimate coefficients of α and β for the Ricker recruit–spawner model has spawner abundance on both sides of the equation, and a spurious test for β (density-dependent term) will typically arise. The error or residual SS are identical when computing the nonlinear (equation [4.6]) and linear (equation [4.7]) equations, but the model and total SS vary. Thus, coefficient of determination and corresponding *F*-statistic computed from the mean squares will typically be different between nonlinear and linear computations.

In our review of published papers, most fisheries scientists report the coefficient of determination and sometimes the probabilities associated with *F*-statistics for recruit–spawner relations. We recommend fisheries scientists use caution and careful evaluation when making statistical inferences pertaining to recruit–spawner relations. To assist fisheries scientists in interpreting the statistical strength of the recruit–spawner relations and other nonlinear regressions, we recommend the following procedures be used. Some of these were recommended by Neter et al. (1996).

1. Although not statistical, a plot of the normal and lognormal error structure of the recruit–spawner relation should always be conducted to help interpret the shape (Ricker, Beverton–Holt, or other) and relative fit or strength of the relation.
2. For nonlinear regression, the closer the sum of the residuals is to 0, the more likely a better fit to the model has occurred (see Box 4.7).
3. For most computational programs such as SAS (2001), numerous iterations are conducted that minimize the residual SS to produce a least-square fit to the data, hence producing the "optimal" model and associated parameters coefficients. If the number of iterations to solve the equation is high (>10–20), then the results may be suspect. In SAS, users are required to provide an initial estimate of the specified regression coefficients, such as α and β for the Beverton–Holt and Ricker recruit–spawner relations (see Boxes 4.7 and 4.8). If the number of iterations is high, then the fisheries scientist can adjust the initial estimate of the regression coefficients to reduce the number of iterations required to obtain the optimal least-squares fit.
4. Hougaard (1985) presented a method to measure the skewness of each of the parameter coefficients generated from nonlinear regression. Coefficient estimates that are more normally distributed, hence are less biased, display absolute skewness values close to 0 (see Boxes 4.7 and Box 4.8).
5. Bootstrap parameter estimation can provide a method to examine if the estimates of the parameter coefficients are approximately normal and whether the bias in each of the parameter coefficients is relatively small (see section 4.4.5 and Box 4.9).
6. Fisheries scientists can predict the number of recruits from nonlinear regression (normal or lognormal error structure) and regress this value against the observed number of recruits (see Boxes 4.7 and 4.8). The coefficient of determination and associated *F*-statistic to test that a recruit–spawner relation exists

Box 4.9 Computation of Bootstrapped Parameter Estimates for the Ricker Recruit–Spawner Curve

The SAS program below conducts bootstrapped parameter estimation for walleye recruit–spawner data (recruits given as *R,* spawners given as *S*) listed in Table 4.3. The program uses the nonlinear form of the Ricker recruit–spawner relation and incorporates lognormal error structure (equation [4.6]). In total, 500 estimates were generated. The program includes information that provides an explanation of the computations that each statement is doing (noted by */).

Program

```
*/INPUT DATA;
DATA ALLDATA;
INPUT J YEARCLASS R S @@;
CARDS;
1 1958 4532 775
2 1959 22996 2310
(continue data input)
33 1990 35607 735
34 1991 48761261
;
*/transform data;
DATA TRANSDATA;
SET ALLDATA;
LOGR=LOG(R);
LOGS=LOG(S);
RUN;

*/ FIT CURVE TO ORIGINAL DATA SET USING EQN. 4.6;
PROC NLIN DATA=TRANSDATA MAXITER=60 METHOD=MARQUARDT;
PARMS ALPHA=1  BETA=.001;
MODEL LOGR=LOGS+ ALPHA - BETA*S;
OUTPUT OUT=PREDOUT PREDICTED=PRCT;
RUN;

*/creates data set that contains 34 residuals for logR;
DATA NEWDATA;
SET PREDOUT;
RESRCT=LOGR-PRCT;
DROP J YEARCLASS R S LOGR LOGS PRCT;
RUN;
*/define bootstrap macro;
%MACRO BOOT;
%DO I=1 %TO 500;

*/create bootstrap data set;
*/creates data set of 34 random residuals;
DATA TEMP;
CHOICE=INT(RANUNI(23456+&I)*N)+1;
SET NEWDATA POINT=CHOICE NOBS=N;
J+1;
IF J>N THEN STOP;
RUN;
```

(Box continues)

Box 4.9 *(continued)*

```
   */creates data set containing logS, logR, predicted logR and predicted logR
+ random residual;
   data analysis;
   SET PREDOUT;
   SET TEMP;
   */ADDS RANDOM RESIDUAL TO PREDICTED LOGR;
   BSRCT=PRCT+RESRCT;
   RUN;

   */fit curve to bootstrap data set;
    PROC NLIN DATA=ANALYSIS MAXITER=60 METHOD=MARQUARDT NOPRINT;
   PARMS ALPHA=1  BETA=.001;
   MODEL BSRCT=LOGS+ ALPHA - BETA*S;
   IF _ITER_=60 THEN CONVERGE=0;
   IF _ITER_<60 THEN CONVERGE=1;
   ID CONVERGE;
   OUTPUT OUT=BOOTOUT PARMS=ALPHA BETA CONVERGE;
   RUN;

   */ delete unnecessary data;
   DATA TEMPBOOT;
   SET BOOTOUT;
   IF J<34 THEN DELETE;
   RUN;

   */save parameter estimates from each bootstrap run;
   PROC APPEND BASE=ALLBOOT DATA=TEMPBOOT;
   RUN;

   %END;
   %MEND;
   */end macro define;

   */run bootstrap macro;
   %BOOT;

   */summarize bootstrap results;
   PROC UNIVARIATE DATA=ALLBOOT;
   VAR ALPHA BETA;
   OUTPUT OUT=BOOTSUM P10=P10 P90=P90;
   RUN;
   QUIT;
```

Results

Table Output for bootstrapped estimation of Ricker recruit–spawner curve assuming lognormal error structure. The Marquardt method is employed. The dependent variable is $\log_e R$ (logR).

Iterative Phase			
Iteration	A	B	SS
0	1.0000	0.00100	194.9
1	3.3916	0.00118	33.7931

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	Approximate $P > F$
Regression	2	2724.0	1362.0	2.13	0.1540
Residual	32	33.7931	1.0560		
Uncorrected total	34	2757.8			
Corrected total	33	36.0449			

Approximate Correlation Matrix		
	A	B
A	1.0000000	–0.9037713
B	–0.9037713	1.0000000

Table Distribution patterns of α for the Ricker recruit–spawner curve as described by the UNIVARIATE procedure.

Moments			
N	500	Sum weights	500
Mean	3.38337872	Sum observations	1691.68936
SD	0.40222105	Variance	0.16178178
Skewness	–0.2077502	Kurtosis	–0.0742384
Uncorrected SS	5804.35489	Corrected SS	80.7291065
CV	11.8881475	SE Mean	0.01798787

Basic Statistical Measures			
Location		Variability	
Mean	3.383379	SD	0.40222
Median	3.390297	Variance	0.16178
Mode		Range	2.17523
		Interquartile range	0.53847

(Box continues)

Box 4.9 *(continued)*

Tests for Location: $Mu_0 = 0$				
Test	Statistic symbol	Statistic value	Comparison	*P*-value
Student's *t*	*t*	188.0922	$P > \|t\|$	<0.0001
Sign	*M*	250	$P \geq \|M\|$	<0.0001
Signed rank	*S*	62625	$P \geq \|S\|$	<0.0001

Quantiles	
Quantile	Estimate
100% Maximum	4.30214
99%	4.22918
95%	4.04890
90%	3.91125
75% Q3	3.66141
50% Median	3.39030
25% Q1	3.12294
10%	2.88759
5%	2.68240
1%	2.36726
0% Minimum	2.12691

Table Distribution patterns of β for the Ricker recruit–spawner curve as described by the UNIVARIATE procedure.

Moments			
N	500	Sum weights	500
Mean	0.00117151	Sum observations	0.58575379
SD	0.00029632	Variance	8.78072×10^{-8}
Skewness	–0.0636916	Kurtosis	–0.0124084
Uncorrected SS	0.00073003	Corrected SS	0.00004382
CV	25.2941464	SE mean	0.00001325

Basic Statistical Measures			
Location		Variability	
Mean	0.001172	SD	0.0002963
Median	0.001169	Variance	8.78072×10^{-8}
Mode		Range	0.00168
		Interquartile range	0.0003932

Tests for Location: $Mu_0 = 0$				
Test	Statistic symbol	Statistic value	Comparison	*P*-value
Student's *t*	*t*	88.40259	$P > \|t\|$	<0.0001
Sign	*M*	250	$P \geq \|M\|$	<0.0001
Signed rank	*S*	62625	$P \geq \|S\|$	<0.0001

Quantiles	
Quantile	Estimate
100% Maximum	0.002026458
99%	0.001804193
95%	0.001680089
90%	0.001553971
75% Q3	0.001367037
50% Median	0.001169344
25% Q1	0.000973827
10%	0.000807687
5%	0.000676374
1%	0.000422876
0% Minimum	0.000343403

From the 500 estimates of the Ricker recruit–spawner, α and β were positively correlated ($r = 0.91$), similar to the *r*-value of –0.90 from the approximate correlation matrix from the single nonlinear equation computed at the beginning of the SAS output. High correlation between nonlinear parameters typically occurs as these coefficients are simultaneously determined and are not independent regressors of the dependent variable.

The output for 500 randomized Ricker recruit–spawner curves and coefficients showed that the mean and median values for α and β were very similar to the empirical estimates generated from the nonlinear equation ($\alpha = 3.392$ and $\beta = -0.00188$). Student *t*-tests, sign, and sign ranks tests for both α and β indicated these coefficients were not equal to 0 for these 500 estimates. The minimum value for β was not less than 0, and thus this bootstrap method provided evidence that weak density dependence occurred in the recruit–spawner relation for walleye in Escanaba Lake. Finally, the UNIVARIATE procedure in SAS showed α and β were approximately normally distributed, similar to the results from Hougaard's skewness values computed in Box 4.8.

using this method can provide more accurate and correct statistical inference information. In some instances, both the nonlinear and the linear regression between predicted and observed recruits will compute nearly identical coefficients of determination and *F*-statistics, and conducting both analyses can provide a level of confidence to the statistical properties of the recruit–spawner relation. The degrees of freedom for the linear model associated with the predicted versus observed recruits is equal to 1 (spawners are the only independent variable), which is correct, compared with 2 *df* for the Beverton–Holt and Ricker recruit–spawner models, where the α and β coefficients are generated for nonlinear regression.

7. For equation (4.7), the linear regression between predicted and observed recruits will compute a nearly identical coefficient of determination and *F*-statistic as equation (4.6) for the lognormal error structure for the Ricker recruit–spawner relation (see Box 4.8). Slight differences are due to computational differences between linear and nonlinear methods. If equation (4.7) is used to estimate α and β, then the linear regression of predicted versus observed recruits is recommended to make statistical inferences.

4.4.4 Incorporation of Environmental Terms to Explain Additional Variation in Recruit–Spawner Models

A broad suite of both abiotic and biotic factors such as climate and prey abundance may explain recruitment variation above that explained by spawners. Additional explanatory variables can easily be added to the traditional recruit–spawner models. The Beverton–Holt model (equation [4.3]) can be modified as

$$R = \frac{\alpha S}{1 + \beta S} \left(e^{c_1 x_1 + \ldots + c_n x_n}\right), \tag{4.8}$$

where x_1 to x_n are n additional independent variables, and c_1 to c_n are the respective estimated coefficients. Hilborn and Walters (1992) suggested expressing the x-values as deviations from a mean value. A similar modification can be applied to the Ricker model for equation (4.4) as

$$R = Se^{\alpha - \beta S + c_1 x_1 + \ldots + c_n x_n}, \tag{4.9}$$

where the additional terms are defined as in equation (4.8). An example of adding an environmental term to a Ricker recruit–spawner curve with lognormal error structure is shown in Box 4.8.

The addition of external variables to recruit–spawner models has been extensively debated and should be used with caution (Walters and Collie 1988). Robust analysis of environmental factors should include testing the integrity of any relations over a relatively long period of time. Myers (1998) reanalyzed a large number of previous studies of recruitment–environmental correlates and found that few relations persisted over time. Tyler (1992) argued in support of research on

environmental factors while acknowledging the criticisms of such work and specifically cautioned against "data dredging," whereby a fisheries scientist assembles an extensive list of environmental factors and tests for correlation with a recruitment time series. This analytical approach can be highly vulnerable to spurious correlation that may arise simply at random or may be due to the selection of an improper error rate for hypothesis testing. Tyler (1992) suggested that data suitable for correlative studies should include several time periods of both increasing and decreasing trends in recruitment. An iterative, operational approach that combines mechanistic simulation with additional analyses from natural experiments will help reject some of the alternate hypotheses developed from the conceptual phase of the study. Finally, we urge caution in the interpretation of statistical output based on either linear or nonlinear methods when environmental variables are added to recruit–spawner models (see section 4.4.3).

4.4.5 Estimates of Uncertainty in Recruit–Spawner Curves

To obtain reliable estimates of uncertainty for recruit–spawner coefficients is difficult for a variety of reasons (Hilborn and Walters 1992). One problem with recruit–spawner data is that model errors are often autocorrelated and thus result in time series bias of parameter estimates. Thus, the assumption of independent errors that is necessary for standard parametric statistical inference is violated. Equations (4.5) and (4.6) provide relatively reliable estimates of the recruit–spawner model parameters, as these equations incorporate lognormal errors (Hilborn and Walters 1992).

Hilborn and Walters (1992) recommended jack-knife and bootstrap methods for producing reliable confidence intervals about recruit–spawner coefficients. Although the bootstrap method is computationally more intensive than is the jack-knife method, the former will provide frequency distributions of parameter estimates, and reasonable confidence intervals can be extracted even when the distribution is asymmetrical. The bootstrap involves resampling with replacement either the original data pairs or residuals from the model fit. For regression models, Efron and Tibshirani (1998) recommend bootstrapping residuals due to strong assumptions that must be made regarding linear models when bootstrapping data pairs. One iteration would involve drawing n random residuals with replacement and adding these to the original observations of the y-variable. The parameters are reestimated from this resampled data set. This process is repeated from 100 to 1,000 times to obtain a frequency distribution of the estimated parameters from which we can then estimate the variance of this distribution and bias corrected and accelerated confidence intervals (Efron and Tibshirani 1998). An example of utilizing bootstrapped methods for a Ricker recruit–spawner curve is presented in Box 4.9.

4.4.6 Sources of Bias in Recruit–Spawner Relations

Two primary sources of bias in estimating recruit–spawner coefficients are time series bias and measurement error bias. In wild populations, spawner abundance

fluctuates and quite often, mortality from juvenile life stages to the time of maturation (i.e., recruitment to the spawning population) may or may not be relatively constant over time. Thus, variation in spawner abundance will not be independent of the process errors that impart variation in recruitment. Large recruitment events will therefore lead to increased spawner abundance in the future, and vice versa. Under this condition, the errors in recruit–spawner models are not independent but are autocorrelated (i.e., a good year is likely to be followed by a good year, and vice versa), and we therefore violate a key assumption of parametric statistics resulting in potential bias of parameter estimates.

For walleye from Escanaba Lake (Hansen et al. 1998; see Table 4.3), the largest observed spawner abundance (2,990) was about eight times greater than the smallest (369). Hilborn and Walters (1992) suggested that biases can be ignored if the smallest stocks are less than 10% the size of the largest; in Escanaba Lake this ratio was about 12%. Walters and Ludwig (1981) suggested that if spawners are estimated with ± 30% error or better, then bias from measurement errors is probably not severe. In Escanaba Lake, spawner abundance was estimated with mark–recapture methods, and standard deviations of these estimates varied from 5.1% to 19.2% and averaged 10% of the mean from 1959 to 1991 (Carpenter et al. 1994). Thus, measurement error of spawner abundance may not impose serious bias on the estimated recruit–spawner coefficients for walleye in Escanaba Lake.

4.5 USE OF ADULT SPAWNER DATA TO ASSESS RECRUITMENT OVERFISHING

In some instances in freshwater fisheries, recreational or commercial exploitation (or both) can so severely deplete the number of adults in the population at such a high rate that recruitment is reduced. Evidence presented by Davidoff et al. (1973) and Walker et al. (1993) for lake whitefish, Chevalier (1977) and Anthony and Jorgensen (1977) for walleye, and Eshenroder (1977) for yellow perch strongly suggested that recruitment overfishing was associated with a decline in catch and yield in these freshwater fisheries. Rieman and Beamesderfer (1990) found the recruit–spawner relation exerted the greatest influence on the dynamics of a white sturgeon population, which indicated these long-lived, slow-growing fish were vulnerable to recruitment overfishing. Secor and Waldman (1999) found that high exploitation caused the Atlantic sturgeon population in Delaware Bay to collapse in the early 1900s due to recruitment overfishing. Slipke et al. (2002) predicted the reduction in young channel catfish was associated with overharvest of adults in the upper Mississippi River.

In the 1980s, marine fisheries scientists attempted to address the problem of recruitment overfishing quantitatively and developed a simple index termed the spawning potential ratio (SPR; Goodyear 1993). Typically, attempts to use Ricker or Beverton–Holt equations of recruit–spawner relations to define a critical abundance of spawning adults (Hilborn and Walters 1992) have been wrought with high variability, confounding effects of environmental factors that affect recruitment, and the lack of long-term data collection (Goodyear 1993; Hansen et al. 1998).

Table 4.3 Recruit (age-0) and spawner (age-5 and older) walleye data for Escanaba Lake, Wisconsin (from Hansen et al. 1998). Abundances of recruits and spawners were determined from mark–recapture population estimates, and May temp CV is the coefficient of variation in May air temperature, which is when walleye were spawning.

	Number		
Year-class	Age-0	Age-5 and older	May temp CV
1958	4,532	775	0.24125
1959	22,996	2,310	0.16319
1960	628	2,990	0.46056
1961	879	1,400	0.33028
1962	14,747	1,130	0.22618
1963	13,205	790	0.20596
1964	31,793	1,195	0.19229
1965	10,621	981	0.20363
1966	22,271	870	0.3452
1967	8,736	1,104	0.27511
1968	8,761	883	0.10884
1969	18,885	1,421	0.17799
1970	10,098	1,198	0.2106
1971	3,394	760	0.22098
1972	1,697	1,354	0.39461
1973	25,159	1,308	0.19696
1974	14,093	2,016	0.20992
1975	1,932	962	0.33459
1976	2,292	976	0.24803
1977	17,386	1,062	0.19815
1978	5,334	1,945	0.32837
1979	6,957	2,073	0.4162
1980	1,036	1,458	0.26409
1981	16,345	946	0.25728
1982	6,149	1,952	0.27111
1983	10,366	1,280	0.18882
1984	16,795	851	0.28661
1985	14,599	394	0.12269
1986	15,299	2,121	0.18605
1987	13,882	452	0.14723
1988	4,351	369	0.18968
1989	4,262	603	0.34298
1990	35,607	735	0.19356
1991	4,876	1,261	0.32032

The SPR is the number of mature eggs produced at a certain level of exploitation for a given population divided by total number of eggs produced in the population if no fish were exploited. Goodyear (1993) defined potential recruit fecundity (P) as the number of mature eggs that could be produced by an average recruit in a population where density-dependent growth and survival did not occur. This represents the actual average lifetime production of mature eggs per recruit at equilibrium population densities in the absence of any density-dependent

suppression of maturation or fecundity at age. Potential recruit fecundity (P) is determined (Goodyear 1993) from

$$P = \sum_{i=1}^{n} E_i \prod_{j=0}^{i=1} S_{ij} \tag{4.10}$$

where

n = number of ages in the population;
E_i = mean fecundity of females of age i in the absence of density-dependent growth;
S_{ij} = $e^{(-[Fij + Mij])}$, the density-independent annual survival probabilities of females of age i when age j;
F_{ij} = the fishing mortality rate of females of age i when age j; and
M_{ij} = the natural mortality rate of females of age i when age j.

Exponential functions for fishing mortality (F) and natural mortality (M) are incorporated into this integral equation similar to predicting cohort abundance.

The SPR is defined as

$$\text{SPR} = P_{\text{fished}} / P_{\text{unfished}} . \tag{4.11}$$

The SPR has a maximum value of 1.00 (unity) and declines toward 0 as fishing mortality increases. The software developed by Slipke and Maceina (2000) can compute SPR values for fish displaying a wide variety of different life history traits and different rates of fishing and natural mortality.

Goodyear (1993) recommends SPR targets of no less than 20–30% based on observations of pelagic marine species. Slipke et al. (2002) estimated the critical SPR to maintain adequate recruitment of channel catfish in the upper Mississippi River was 10% based on the response of C/f of age-0 fish and subsequent harvest of adult fish. Quist et al. (2002) recommended a maximum conservative SPR target of 40–50% to protect overexploitation of shovelnose sturgeon in the Missouri River. Other than the work of Slipke et al. (2002), critical values for SPR have not been defined nor used to evaluate freshwater sport or commercial fisheries, but exploring the utility of SPR for these fisheries warrants investigation. Target SPRs are achieved by protecting mature females by means of harvest regulations. The SPR is used as a management criterion to maintain adequate females in the population to prevent recruitment overfishing. Typically, mature ova production increases exponentially with fish length or linearly with weight. In some instances, larger females can produce one to two orders of magnitude more eggs than can smaller sexually mature fish. For example, management strategies to maintain white sturgeon in the Columbia River include protecting older mature females that can be caught and released using hook-and-line gear but allowing harvest of a slot length (92–183 cm TL) of smaller fish (Rieman and Beamesderfer 1990). The use of SPR critical values should be analyzed with caution as a direct relationship between fecundity and subsequent recruits may not occur, and from year to year environmental variables can also influence reproductive success.

4.6 SUMMARY

A wide variety of parametric statistical procedures can be used by fisheries scientists to examine spatial and temporal fluctuations in recruit abundance. These tools can also be used to examine the effects of manipulations or biotic and abiotic impacts on recruitment success. The relation between recruits and spawners can be investigated, but typically these relations require long-term data collection, accurate estimates of spawner abundance are difficult to obtain, and recruitment variation is likely to be more influenced by environmental conditions. For most freshwater fishes, fluctuations in recruitment exert a very strong influence on population dynamics, and obtaining information on recruitment variation is paramount to understanding and managing fisheries. Typically, most recreational and commercial fisheries display wide variation in recruits, and this variation should be considered when sampling designs are considered, planned, and executed. Although recruit variation can be high, adequate replicates can be taken to provide the fisheries scientist with enough statistical power to examine and test hypotheses related to fish recruitment.

4.7 REFERENCES

Allen, M. S., J. C. Greene, F. R. Snow, M. J. Maceina, and D. R. DeVries. 1999. Recruitment of largemouth bass in Alabama reservoirs: relations to trophic state and larval shad occurrence. North American Journal of Fisheries Management 19:67–77.

Allen, M. S., and L. E. Miranda. 1998. An age-structure model for erratic crappie fisheries. Ecological Modelling 107:289–303.

Allen, M. S., and W. E. Pine III. 2000. Detecting fish population responses to a minimum length limit: effects of variable recruitment and duration of evaluation. North American Journal of Fisheries Management 20:672–682.

Anthony, D. D., and C. R. Jorgensen. 1977. Factors in the declining contributions of walleye *(Stizostedion vitreum vitreum)* to the fishery of Lake Nipissing, Ontario, 1960–76. Journal of the Fisheries Research Board of Canada 34:1703–1709.

Beamesderfer, R. C. 1991. MOCPOP 2.0: a flexible system for simulation of age-structure populations and stock-related functions. Oregon Department of Fish and Wildlife Information Report 91–4, Clackamas.

Bettoli, P. W., M. J. Maceina, R. L. Noble, and R. K. Betsill. 1992. Piscivory in largemouth bass as a function of aquatic vegetation abundance. North American Journal of Fisheries Management 12:509–516.

Bettoli, P. W., and L. E. Miranda. 2001. Cautionary note about estimating mean length at age with subsampled data. North American Journal of Fisheries Management 21:425–428.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London

Bronte, C. R., J. H. Selgeby, and D. V. Swedberg. 1993. Dynamics of a yellow perch population in western Lake Superior. North American Journal of Fisheries Management 13:511–523.

Burton, W. H., and S. B. Weisberg. 1994. Estimating abundance of age-0 striped bass in the Delaware River by using marked hatchery fish. North American Journal of Fisheries Management 14:347–354.

Busch, W. D., R. L. Scholl, and W. L. Hartman. 1975. Environmental factors affecting the strength of walleye *(Stizostedion vitreum vitreum)* year-classes in western Lake Erie, 1960–70. Journal of the Fisheries Research Board of Canada 32:1733–1743.

Buynak, G. L., and B. Mitchell. 1999. Contribution of stocked advanced-fingerling largemouth bass to the population and fishery at Taylorsville Lake, Kentucky. North American Journal of Fisheries Management 19:494–503.

Buynak, G. L., B. Mitchell, D. Bunnell, B. McLemore, and P. Rister. 1999. Management of largemouth bass at Kentucky and Barkely lakes, Kentucky. North American Journal of Fisheries Management 19:59–66.

Carline, R. F., B. L. Johnson, and T. J. Hall. 1984. Estimations and interpretation of proportional stock density for fish populations in Ohio impoundments. North American Journal of Fisheries Management 4:139–154.

Carpenter, S. R., A. Munoz-Del-Rio, S. Newman, P. W. Rasmussen, and B. M. Johnson. 1994. Interactions of anglers and walleyes in Escanaba Lake, Wisconsin. Ecological Applications 4:822–832.

Chevalier, J. R. 1977. Changes in walleye *(Stizostedion vitreum vitreum)* populations in Rainy Lake and factors in abundance, 1924–75. Journal of the Fisheries Research Board of Canada 34:1696–1702.

Counihan, T. D., A. I. Miller, and M. J. Parsley. 1999. Indexing the relative abundance of age-0 white sturgeons in an impoundment of the lower Columbia River from highly skewed trawling data. North American Journal of Fisheries Management 19:520–529.

Cowan, J. H., K. A. Rose, and D. R. DeVries. 2000. Is density-dependent growth in young-of-the-year fishes a question of critical weight? Reviews in Fish Biology and Fisheries 10:61–89.

Crecco, V. A., and T. F. Savoy. 1987. Review of recruitment mechanisms of the American shad: the critical period and match–mismatch hypothesis reexamined. Pages 455–468 *in* M. J. Dadswell, R. J. Klauda, C. M. Moffitt, R. L. Saunders, R. A. Rulifson, and J. E. Cooper, editors. Common strategies of anadromous and catadromous fishes. American Fisheries Society Symposium 1, Bethesda, Maryland.

Cushing, D. H. 1971. Dependence of recruitment on parent stock in different groups of fishes. Journal du Conseil, Conseil International pour l'Exploration de la Mer 33: 340–362.

Cushing, D. H. 1973. Dependence of recruitment on parent stock. Journal of the Fisheries Research Board of Canada 30:1965–1976.

Davidoff, E. B., R. W. Rybicki, and K. H. Doan. 1973. Changes in the population of lake whitefish *(Coregonus clupeaformis)* in Lake Winnipeg from 1944 to 1969. Journal of the Fisheries Research Board of Canada 30:1667–1682.

Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences 37:268–282.

DiCenzo, V. J., and M. C. Duval. 2002. Importance of reservoir inflow in determining white bass year-class strength in three Virginia reservoirs. North American Journal of Fisheries Management 22:620–626.

Efron, B., and R. J. Tibshirani. 1998. An introduction to the bootstrap. CRC Press, Boca Raton, Florida.

Eshenroder, R. L. 1977. Effects of intensified fishing, species change, and spring water temperatures on yellow perch, *Perca flavescens*, in Saginaw Bay. Journal of the Fisheries Research Board of Canada 34:1830–1838.

Fisher, W. L., and A. V. Zale. 1993. Effect of water level fluctuations on abundance of young- of-the-year largemouth bass in a hydropower reservoir. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 45(1991):422–431.

Freund, R. J., and R. C. Littell. 1991. SAS system for regression. SAS Institute, Cary, North Carolina.

Goodyear, C. P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. Pages 67–71 *in* S. J. Smith, J. J. Hunt, and D. Rivard, editors. Risk evaluation and biological reference points for fisheries management. Canadian Journal of Fisheries and Aquatic Sciences, Special Publication 120.

Guy, C. S. 1993. Structure, dynamics, and movement patterns for crappie in South Dakota waters. Doctoral dissertation. South Dakato State University, Brookings.

Guy, C. S., and D. M. Willis. 1995. Population characteristics of black crappie in South Dakota waters: a case for ecosystem-specific management. North American Journal of Fisheries Management 15:754–765.

Hansen, M. J., M. A. Bozek, J. R. Newby, S. P. Newman, and M. D. Staggs. 1998. Factors affecting recruitment of walleyes in Escanaba Lake, Wisconsin, 1958–1996. North American Journal of Fisheries Management 18:764–774.

Hansen, M. J., and six coauthors. 1996. Causes of declining survival of lake trout stocked in U.S. waters of Lake Superior in 1963–1986. Transactions of the American Fisheries Society 125:831–843.

Hennemuth, R. C., J. E. Palmer, and B. E. Brown. 1980. A statistical description of recruitment in eighteen selected fish stocks. Journal of Northwest Atlantic Fisheries Science 1:101–110.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hougaard, P. 1985. The appropriateness of the asymptotic distribution in a nonlinear regression model in relation to curvature. Journal of the Royal Statistical Society 47: 103–114.

Hoyer, M. V., and D. E. Canfield Jr. 1996. Largemouth bass abundance and aquatic vegetation in Florida lakes: an empirical analysis. Journal of Aquatic Plant Management 34:23–32.

Isely, J. J., R. L. Noble, J. B. Koppelman, and D. P. Philipp. 1987. Spawning period and first-year growth of northern, Florida, and intergrade stocks of largemouth bass. Transactions of the American Fisheries Society 116:757–762.

Isermann, D. A., P. W. Bettoli, S. M. Sammons, and T. N. Churchill. 2002. Initial postsocking mortality, oxytetracycline marking, and year-class contribution of black-nosed crappies stocked into Tennessee reservoirs. North American Journal of Fisheries Management 22:1399–1408.

IUBMB (International Union of Biochemistry and Molecular Biology). 1992. Enzyme nomenclature. Academic Press, San Diego, California.

Jackson, J. R., and R. L. Noble. 1995. Selectivity of sampling methods for juvenile largemouth bass in assessments of recruitment processes. North American Journal of Fisheries Management 15:408–418.

Jackson, J. R., and R. L. Noble. 2000. Relationships between annual variations in reservoir conditions and age-0 largemouth bass year-class strength. Transactions of the American Fisheries Society 129:699–715.

Kallemeyn, L. W. 1987. Correlations of regulated lake levels and climatic factors with abundance of young-of-the-year walleye and yellow perch in four lakes in Voyageurs National Park. North American Journal of Fisheries Management 7:513–521.

Kimura, D. L. 1988. Analyzing relative abundance indices with log-linear models. Transactions of the American Fisheries Society 8:175–180.

Kramer, R. H., and L. H. Smith. 1962. Formation of year classes in largemouth bass. Transactions of the American Fisheries Society 91:29–41.

Leslie, P. H., and D. H. Davis. 1939. An attempt to determine the absolute number of rats on a given area. Journal of Animal Ecology 8:94–113.

Ludsin, S. A., and D. R. DeVries. 1997. First-year recruitment of largemouth bass: the interdependency of early life stages. Ecological Applications 7:1024–1038.

Maceina, M. J. 1997. Simple application of using residuals from catch-curve regressions to assess year-class strength in fish. Fisheries Research 32:115–121.

Maceina, M. J. 2003. Verification of the influence of hydrologic factors on crappie recruitment in Alabama reservoirs. North American Journal of Fisheries Management 23:470–480.

Maceina, M. J. 2004. Verifying residuals from catch curves to detect recruitment variation in largemouth bass and crappie. North American Journal of Fisheries Management 24:231–236.

Maceina, M. J., and P. W. Bettoli. 1998. Variation in largemouth bass recruitment in four mainstream impoundments of the Tennessee River. North American Journal of Fisheries Management 18:998–1003.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot analysis of variance design for repeated-measures fishery data. Fisheries (Bethesda) 19(3):14–20.

Maceina, M. J., B. R. Murphy, and J. J. Isely. 1988. Factors regulating Florida largemouth bass stocking success and hybridization with northern largemouth bass in Aquilla Lake, Texas. Transactions of the American Fisheries Society 117:221–231.

Maceina, M. J., S. J. Rider, and S. T. Szedlmayer. 1995. Density, temporal spawning patterns, and growth of age-0 and age-1 largemouth bass *(Micropterus salmoides)* in vegetated and unvegetated areas of Lake Guntersville, Alabama. Pages 497–511 *in* D. C. Secor, J. M. Dean, and S. E. Campana, editors. Recent developments in fish otolith research. University of South Carolina Press, Columbia.

Maceina, M. J., and M. R. Stimpert. 1998. Relations between reservoir hydrology and crappie recruitment in Alabama. North American Journal of Fisheries Management 18:104–113.

Madenjian, C. P., R. L. Knight, M. T. Bur, and J. L. Forney. 2000. Reduction in recruitment of white bass in Lake Erie after invasion of white perch. Transactions of the American Fisheries Society 129:1340–1353.

Madenjian, C. P., J. T. Tyson, R. L. Knight, M. W. Kershner, and M. J. Hansen. 1996. First-year growth, recruitment, and maturity of walleyes in western Lake Erie. Transactions of the American Fisheries Society 125:821–830.

Montgomery, D. C., and E. A. Peck. 1982. Introduction to linear regression analysis. Wiley, New York.

Murphy, B. R., L. A. Nielsen, and B. J. Turner. 1983. Use of genetic tags to evaluate stocking success for reservoir walleyes. Transactions of the American Fisheries Society 112:457–463.

Murphy, B. R., and D. W. Willis, editors. 1996. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Musick, J. A. 1999. Ecology and conservation of long-lived marine animals. Pages 1–10 *in* J. A. Musick, editor. Life in the slow lane: ecology and conservation of long-lived marine animals. American Fisheries Society Symposium 23, Bethesda, Maryland.

Musick, J. A., and seventeen coauthors. 2000. Marine, estuarine, and diadromous fish stocks at risk of extinction in North America (exclusive of the Pacific salmonids). Fisheries (Bethesda) 25(11):6–30.

Myers, R. A. 1998. When do environment–recruitment correlations work? Reviews in Fish Biology and Fisheries 8:285–305.

Myers, R. A., and N. J. Barrowman. 1996. Is fish recruitment related to spawner abundance? U.S. National Marine Fisheries Service Fishery Bulletin 94:707–724.

Myers, R. A., G. Mertz, and J. Bridson. 1997. Spatial scales of interannual recruitment variations of marine, anadromous, and freshwater fish. Canadian Journal of Fisheries and Aquatic Sciences 54:1400–1407.

Myers, R. A., A. A. Rosenberg, P. M. Mace, N. Barrowman, and V. R. Restrepo. 1994. In search of thresholds for recruitment overfishing. ICES Journal of Marine Science 51: 191–205.

Neter, J., M. H. Kutner, C. J. Nachtsheeim, and W. Wasserman. 1996. Applied linear statistical models, 4th edition. McGraw-Hill, Boston.

Ozen, O. 1997. Crappie population characteristics in six Alabama reservoirs. Master's thesis. Auburn University, Auburn, Alabama.

Pannella, G. 1971. Fish otoliths: daily growth layers and periodic patterns. Science 173: 1124–1127.

Paragamian, V. L., E. C. Bowles, and B. Hoelscher. 1992. Use of daily growth increments on otoliths to assess stockings of hatchery-reared kokanees. Transactions of the American Fisheries Society 121:785–791.

Parsons, B. G., and D. L. Pereira. 2001. Relationship between walleye stocking and year-class strength in three Minnesota lakes. North American Journal of Fisheries Management 21:801–808.

Peterman, R. M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Canadian Journal of Fisheries and Aquatic Sciences 38:1113–1119.

Pitlo, J., Jr. 1997. Response of upper Mississippi River channel catfish populations to changes in commercial harvest regulations. North American Journal of Fisheries Management 17:848–859.

Ploskey, G. R., J. M. Nestler, and W. M. Bivin. 1996. Predicting black bass reproductive success from Bull Shoals Reservoir hydrology. Pages 422–441 *in* L. E. Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society Symposium 16, Bethesda, Maryland.

Quist, M. C., C. S. Guy, M. A. Pegg, P. J. Braaten, C. L. Pierce, and V. T. Travnichek. 2002. Potential influence of harvest of shovelnose sturgeon populations in the Missouri River system. North American Journal of Fisheries Management 22:537–549.

Reinert, T. R., G. R. Ploskey, and M. J. Van Den Avyle. 1997. Effects of hydrology on black bass reproductive success in four southeastern reservoirs. Proceedings of the Annual Conference Southeastern Association Fish and Wildlife Agencies 49(1995):47–57.

Reish, R. L., R. B. Deriso, D. Ruppert, and R. J. Carroll. 1985. An investigation of the population of dynamics of Atlantic menhaden *(Brevoortia tyrannus)*. Canadian Journal of Fisheries and Aquatic Sciences 42(Supplement 1):147–157.

Rice, J. A., L. B. Crowder, and M. E. Holey. 1987. Exploration of mechanisms regulating larval survival in Lake Michigan bloater: a recruitment analysis based on characteristics of individual larvae. Transactions of the American Fisheries Society 116:703–718.

Richards, R. A., and P. J. Rago. 1999. A case history of effective fishery management: Chesapeake Bay striped bass. North American Journal of Fisheries Management 19:356–375.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada 11:559–623.

Ricker, W. E. 1975. Computation and interpretation of biological statistics in fish populations. Fisheries Research Board of Canada Bulletin 191.

Rider, S. J., M. J. Maceina, and D. R. Lowery. 1994. Comparisons of cove rotenone and electrofishing catch-depletion estimates to determine abundance of age-0 largemouth bass in vegetated and unvegetated areas. Journal of Freshwater Ecology 9:19–27.

Rieman, B. E., and R. C. Beamesderfer. 1990. White sturgeon in the lower Columbia River: is the stock overexploited? North American Journal of Fisheries Management 10: 388–396.

Ryan, M. J., M. A. Webb, and L. T. Fries. 1998. Contribution of largemouth bass reared in nursery ponds to year classes in two Texas reservoirs. Proceedings Annual Conference Southeastern Association Fish and Wildlife Agencies 50(1996):131–138.

Sammons, S. M., and P. W. Bettoli. 1998. Larval sampling as a fisheries management tool: early detection of year-class strength. North American Journal of Fisheries Management 18:137–143.

Sammons, S. M., and P. W. Bettoli. 2000. Population dynamics of a reservoir sport fish community in response to hydrology. North American Journal of Fisheries Management 20:791–800.

SAS. 2001. SAS, version 8.2. SAS Institute, Cary, North Carolina.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences 42:414–429.

Secor, D. H., and J. R. Waldman. 1999. Historical abundance of Delaware Bay Atlantic sturgeon and potential rate of recovery. Pages 203–216 *in* J. A. Musick, editor. Life in the slow lane: ecology and conservation of long-lived marine animals. American Fisheries Society Symposium 23, Bethesda, Maryland.

Serns, S. L. 1982. Relation of temperature and population density to first-year recruitment and growth of smallmouth bass in a Wisconsin lake. Transactions of the American Fisheries Society 111:570–574.

Shepherd, J. G. 1982. A versatile new stock–recruitment relationship for fisheries, and the construction of sustainable yield curves. Journal du Conseil International pour l'Exploration de la Mer 40:67–75.

Slipke, J. W., and M. J. Maceina. 2000. Fishery analyses and simulation tools (FAST). Auburn University, Auburn, Alabama.

Slipke, J. W., M. J. Maceina, V. H. Travnichek, and K. C. Weathers. 1998. Effects of a 356-mm minimum-length limit on the population characteristics and sport fishery of smallmouth bass in the Shoals Reach of the Tennessee River, Alabama. North American Journal of Fisheries Management 18:76–84.

Slipke, J. W., A. D. Martin, J. Pitlo Jr., and M. J. Maceina. 2002. Use of the spawning potential ratio in freshwater fisheries management: the upper Mississippi River channel catfish fishery. North American Journal of Fisheries Management 22:1295–1300.

Stouder, D. J., P. A. Bisson, and R. J. Naiman, editors. 1997. Pacific salmon and their ecoystems: status and future options. Chapman and Hall, New York.

Summerfelt, R, C. 1975. Relationship between weather and year-class strength of largemouth bass. Pages 166–174 *in* R. H. Stroud and H. Clepper, editor. Black bass biology and management. Sport Fishing Institute, Washington, D.C.

Taylor, C. J. 1981. A generalized inland fishery simulator for management biologists. North American Journal of Fisheries Management 1:60–72.

Thomas, L. M., S. A. Holt, and C. R. Arnold. 1995. Chemical marking techniques for larval and juvenile red drum *(Sciaenops ocellatus)* otoliths. Pages 703–718 *in* D. C. Secor, J. M. Dean, and S. E. Campana, editors. Recent developments in fish otolith research. University of South Carolina Press, Columbia.

Travnichek, V. H., M. J. Maceina, and R. A. Dunham. 1996. Hatching time and early growth of black crappie, white crappie, and their naturally produced F_1 hybrid in Weiss Lake, Alabama. Transactions of the American Fisheries Society 125:334-338.

Tyler, A. V. 1992. A context for recruitment correlations: why marine fisheries biologists should still look for them. Fisheries Oceanography 1:97–106.

Tyler, A. V., and W. R. Crawford. 1991. Modeling of recruitment patterns in Pacific cod *(Gadus macrocephalus)* in Hecate Strait, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 48:2240–2249.

Van Den Avyle, M. J., and R. S. Hayward. 1999. Dyamics of exploited fish populations. Pages 127–166 *in* C. C. Kohler, and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Walker, S. H., M. W. Prout, W. W. Taylor, and S. R. Winterstein. 1993. Population dynamics and management of lake whitefish stocks in Grand Traverse Bay, Lake Michigan. North American Journal of Fisheries Management 13:73–85.

Walters, C. J., and J. S. Collie. 1988. Is research on environmental factors useful to fisheries management? Canadian Journal of Fisheries and Aquatic Sciences 45:1848–1854.

Walters, C. J., and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock–recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences 38:704–710.

Willis, D. W., and J. L. Stephen. 1987. Relationships between storage ratio and population density, natural recruitment, and stocking success of walleye in Kansas reservoirs. North American Journal of Fisheries Management 7:279–282.

Wrenn, W. B., D. R. Lowery, M. J. Maceina, and W. C. Reeves. 1996. Largemouth bass and aquatic plant abundance in Guntersville Reservoir, Alabama. Pages 382–393 *in* L. E. Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society Symposium 16, Bethesda, Maryland.

5 Age and Growth

J. Jeffery Isely and Timothy B. Grabowski

5.1 INTRODUCTION

The ability to determine ages of fishes without bias is critical to effective management and research. Accurate age information can provide valuable insights into critical life history events. Often, migrations related to spawning or ontogenetic changes in environmental requirements are also age dependent. Age data can be coupled with numbers of individuals to produce an age-frequency distribution, from which patterns in mortality can be determined. Similarly, deviations in expected numbers at age can provide insights into year-class strength variability and the effects of environment on survival.

When age and size information are combined, we can evaluate growth. Growth provides us with some indication of resource utilization and the effectiveness of our management strategies. Our ability to model growth and to understand variables that affect growth both within and among populations is critical to our ability to manage fisheries effectively. When we evaluate age, growth, and mortality (see Chapter 6) in combination, we begin to understand the complex relationship between population size and biomass (see Chapter 8). This understanding is the basis of modern fisheries resource allocation and management.

5.2 AGE DETERMINATION AND VALIDATION

Primary methods employed by fisheries scientists to estimate ages of fishes are recovery of known-age fish, evaluation of length-frequency distributions, and interpretation of calcified structures. Under unique circumstances, additional methods employed by researchers include evaluations of isotope decay rates and chemical microanalysis.

5.2.1 Use of Known-Age Fish

The most direct method of determining age is by the recovery of known-age fish. Although costly and time consuming, the method is most useful to validate ages determined by other methods. In this method, fish of known age are reared under

natural conditions or marked and released into the wild to be recaptured at a later time. Dyes and stains that are incorporated in hard parts, such as oxytetracycline or alizarin complexone, have been used to validate annual and daily growth-increment formation (Brothers 1990). Individuals are either immersed into a bath containing a dye or injected with a chemical that is incorporated into the aging structure. After a minimum of one annual growth cycle (or several daily growth cycles), the fish is recaptured and the structure is examined. Although the relationship between the number of annuli between the mark and the margin of the structure is used to validate the annual deposition of increments, the technique validates annual increment formation during only the time period of the study. It is then inferred that all rings are similarly formed. It is important to examine a variety of sizes and ages when employing this technique (Campana 2001).

5.2.2 Length-Frequency Method

Because fishes in temperate climates generally spawn over a relatively short period each year, but grow over a relatively long period each year, there are natural discontinuities in the length-frequency distribution between age-classes within a population (Macdonald and Pitcher 1979; Macdonald 1987). At any given time, the length-frequency distribution of a population is composed of a variety of age-classes. In theory, each year-class forms a unique length distribution resulting in a separate mode in the cumulative distribution. The method of estimating age by separating overlapping length distributions has been used since the late 1800s.

Although this method works well to separate early age-classes, the decrease in annual growth in length as individuals age, combined with natural variability in growth among individuals, results in increasing overlap in age-specific length distributions with older cohorts (Figure 5.1). In most species, only the youngest two or three cohorts are readily distinguished using this method. The method also has several other disadvantages. Environmental conditions often result in disjunct spawning or survival within a single spawning season, resulting in multi-modal length-frequency distributions within year-classes. Geographic differences in environmental quality, density dependency, or other factors may also result in differential growth between groups spawned within the same time period. Differences in year-class strength may result in an underrepresentation of one or more year-classes that are masked by a more dominant cohort. Schooling species often associate by size, resulting in little within-school variability in size across age-classes. Most sampling and fishing gear are size selective and collect samples that are biased by size or growth rates of individuals. As a result, most samples represent a subset of the population length distribution that contains only the fastest-growing younger fish and slowest-growing older fish along with normally growing fish of intermediate age. The success of the length-frequency method of age estimation requires a large sample drawn at random from the population. Although it is useful in fast-growing, short-lived species, the technique is most useful to corroborate age distributions derived from some other method. Separating the overlap in length-frequency distributions requires an iterative statistical procedure. Fournier

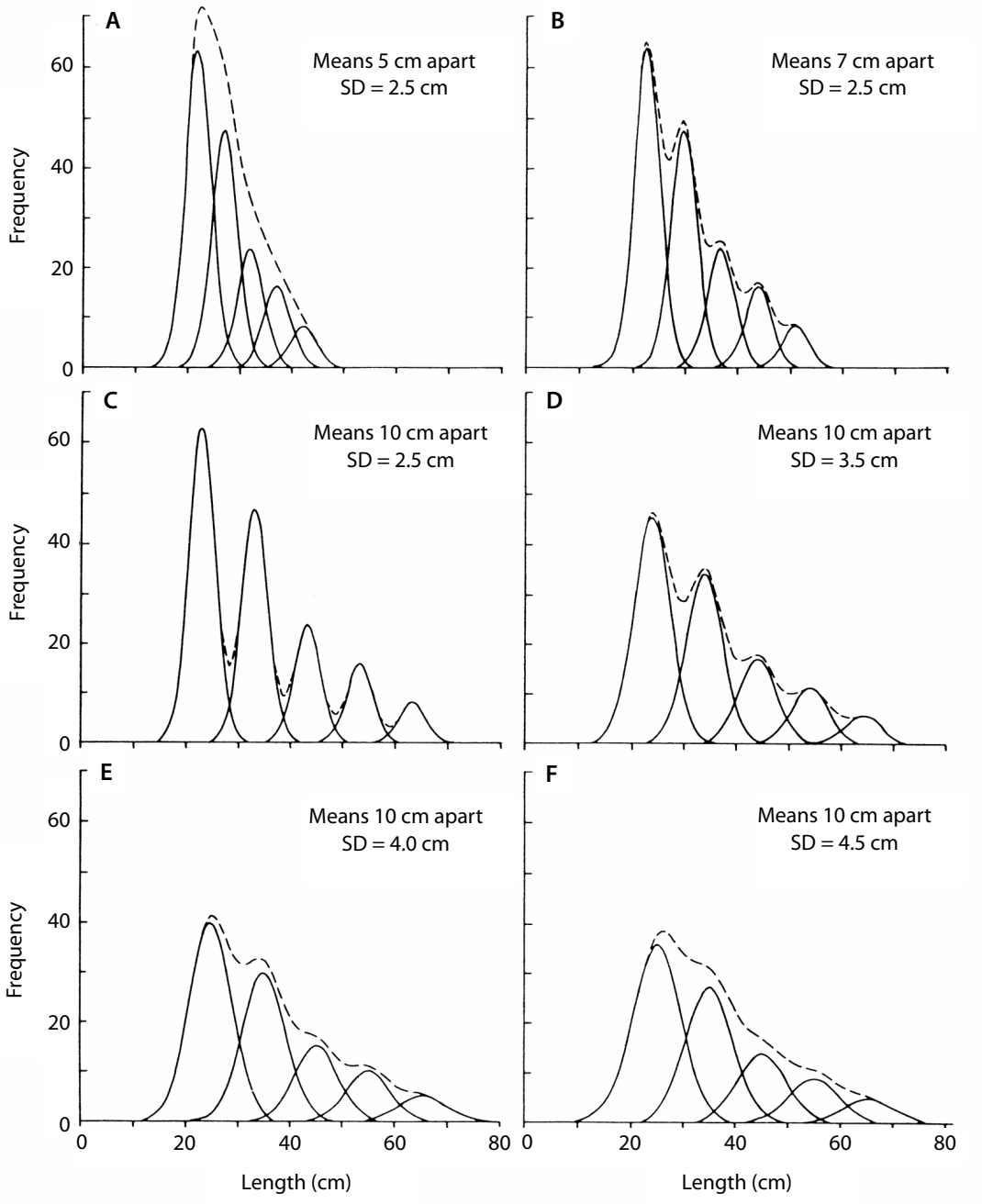


Figure 5.1 Illustration of the effect of changes in the mean, standard deviation, and relative sizes of cohorts upon a length-frequency distribution. The dashed lines represent the frequency distribution of a hypothetical population and solid lines indicate cohorts. The contribution to the population by each cohort is constant. Reprinted from Macdonald and Pitcher (1979) with permission.

et al. (1998) developed a length-based, age-structured model, MULTIFAN–CL, which provides an integrated method of estimating age composition and other parameters from length-frequency data. The method incorporates Bayesian parameter estimation and procedures for hypothesis testing to assist model development. The reader is referred to Macdonald and Pitcher (1979) and Macdonald (1987) for further information on the use of length-frequency distributions to assign individuals to age-groups.

5.2.3 Interpretation of Calcified Structures

Intra-annual variability in environmental variables such as temperature, salinity, dissolved oxygen, and productivity often produce a seasonal cycle in fish growth. This seasonal cycle is recorded as discernible increments in calcified structures in fishes because of the differential deposition of calcium and protein in relation to growth. Structures commonly used to age fish include scales, otoliths, spines, fin rays, vertebrae, and other bony structures (DeVries and Frie 1996). The successful interpretation of calcified structures to age individuals relies on the ability to recognize patterns in the layered deposition of material. As annual growth increments decrease with age, the spacing and distinctness of growth increments in calcified structures also decreases, often resulting in a negative bias in age estimates of older fishes.

5.2.3.1 Scales

Historically, scales were the most popular structure used to estimate age. Despite their limitations, they remain an important and commonly employed tool for assessing age and growth in many species. Scales were first recognized to contain age information as early as 1890. The technique was commonly used in assessments of European marine fisheries at the turn of the century but was not applied widely in North America until the 1920s (Carlander 1987). A review by Van Oosten (1929) described the methods and established guidelines for scale interpretation that led to widespread employment of the method beginning in the 1930s. With a few modifications, the techniques described by Van Oosten (1929) are still used today. Although it is relatively easy and inexpensive to collect and prepare scales, the identification of annuli requires skill and experience. The identification of false annuli can be a critical component of age and growth studies utilizing scales. Additionally, scales do not develop at hatching and may not appear until the fish is at a relatively advanced stage of development (Ward and Leonard 1954; Sire and Arnulf 1990; Sire et al. 1997).

The major advantage of the use of scales for aging is that the fish need not be sacrificed for data collection. This is of particular importance in studies focusing on endangered or threatened species as well as in situations where the removal of fish from the study area would bias study results. However, the potential bias toward underestimation of age is a major disadvantage.

5.2.3.2 Otoliths

Otoliths, or ear stones, are acellular structures formed by the crystallization of calcium carbonate in a protein matrix. This process is growth dependent and occurs throughout the life of the fish (Popper and Lu 2000). Material is permanently deposited and is not resorbed as in scales. Consequently, otoliths constitute a permanent record of growth for a fish to the extent that fossil otoliths can be used to reconstruct the life histories of ancient fishes (Woydack and Morales-Nin 2001). Otoliths were first used to determine fish age in the late 1800s (see review by Van Oosten 1929). However, the discovery in the 1970s that otoliths form daily increments has become an important advancement in fisheries science (Pannella 1971). Daily ages determined from age-0 fish have led to the incorporation of early life history information, such as cohort-specific growth and mortality, into stock assessments and the evaluation of the effects of environmental conditions on growth and mortality over short temporal scales. The analysis of daily otolith increments in fisheries research has been reviewed by Campana and Neilson (1985). Although otolith increments are usually easier to interpret than are scale increments, the recognition of otolith annuli and daily growth increments still requires skill and experience. There are two disadvantages to the use of otoliths in age determination: sacrifice of the fish is necessary for otolith removal and a large investment of time is required to prepare them for reading. These factors should be considered when designing an age and growth study.

5.2.3.3 Spines and Fin Rays

In cases in which sacrifice is impractical and scales are inadequate or nonexistent, spines or fin rays may be used to determine age. This method is most commonly employed to age catfishes and sturgeons but can be applied to a wide range of species (Beamish 1981). The reader is referred to Boyko (1946) and Sneed (1951) for details regarding preparation and reading. A major disadvantage of this technique is that spines contain a central lumen, which expands as the fish grows. The expanding lumen erodes early annuli and can cause age and growth estimates to be biased (Nash and Irwin 1999; Buckmeier et al. 2002). Annuli can be somewhat more irregular than those in scales and otoliths, potentially rendering them useless for back-calculating growth. However, the use of spines or rays in conjunction with other structures or in cases where other structures fail to produce reliable results warrants consideration (Beamish 1981).

5.2.3.4 Vertebrae and Other Bony Structures

Historically, other structures such as opercular bones (Bardach 1955) and cleithra have been used to determine the ages of fishes. Studies utilizing these structures are not common, despite yielding age estimates that are comparable to those from scales and otoliths (Baker and Timmons 1991; Baker and McComish 1998). The usual methods of age determination for bony fish do not work for cartilaginous fish. However, some structures such as vertebrae contain mineralized calcium

phosphate, which is deposited in proportion to size and has proven useful in determining age (Stevens 1975; Clement 1992; Natanson et al. 2001). Other structures such as the thorns of skates and rays (Gallagher and Nolan 1999) and the spines of dogfishes (McFarlane and Beamish 1987a) also have been employed for determining age and growth in elasmobranchs.

5.2.4 Validation of Age Estimates from Calcified Structures

Common assumptions of estimating age from hard parts is that increments are formed annually or daily and that all marks are readily identifiable. Although generally correct, these assumptions are not always valid (Beamish and McFarlane 1983; Campana 2001). Variability in growth resulting from environmental extremes, spawning, disease, or injury may result in marks that appear similar in structure to annual increments (Mugiya and Uchimura 1989; Morales-Nin 2000). Allometric growth and the slowing of growth with increasing age may render annual or daily marks difficult to distinguish. In recent years, the potentially large effects of underestimates in age on management decisions related to harvest and growth have re-emphasized the importance of validating age estimation procedures (Beamish and McFarlane 1983; Campana 2001).

5.2.4.1 Natural Marks

In some cases natural marks have been used as a method of age validation. Occasionally a natural or anthropogenic event will create a reference mark on calcified structures. For example, the eruption of Mount Pinatubo in 1992 resulted in reduced productivity in lakes throughout the northeastern United States and Canada. This reduced productivity has manifested in slow growth of fishes during that year, which can be seen across age-classes (King et al. 1999a; 1999b). This natural mark has been used to validate the ages of fishes that were living in 1992. Another widespread mark that has been used to validate ages in long-lived fishes is the incorporation of radioactive carbon (^{14}C) from nuclear bomb tests in the 1950s into tissues (Kalish 1993; Kalish et al. 1997; Campana et al. 2002). Typically, the nucleus of the otolith is used to determine the year of birth of the individual. The technique can be used broadly to separate fish into individuals born prior to and after nuclear weapons testing or to validate specific ages when used in conjunction with a reference chronology of atmospheric ^{14}C levels. Other events such as El Niño–La Niña events (Woodbury 1999) and oil spills (Gallego et al. 1995) that are associated with a specific date have the potential to be used as marks for independent validation of age estimates. However, to date, this method has not been widely investigated.

5.2.4.2 Radiochemical Dating

The ages of fishes determined from otoliths can also be validated using radiochemical dating. This process takes advantage of the decay of radioactive trace elements deposited in the otoliths during their formation. The ratio of parent to daughter isotopes in the nucleus of the otolith can be used to estimate the time of

its formation. This technique is dependent upon removal of the nucleus and can be extremely sensitive to the removal of excess material. Additionally, the resolving power of the technique renders it suitable only for long-lived species (Francis 2003). The reader is referred to Bennett et al. (1982) and Andrews et al. (1999) for detailed descriptions of this technique.

5.2.4.3 *Marginal Increment Analysis*

Marginal increment analysis (MIA) is a commonly used method for evaluating both annulus and daily increment formation to validate age estimates. It tests the assumption that a growth increment formed on an annual or daily cycle will fit a saw-toothed pattern when the average state of completion of that increment in the population is plotted against time (Figure 5.2). Marginal increment analysis is popular because it is easy and cost effective relative to other validation techniques. It uses repeated sampling at regular intervals through time to determine when an annulus or increment is deposited. The application of the technique has come under question, as reviewed by Beckman and Wilson (1995) and Campana (2001). There are severe technological limitations in measuring a growth increment along the increasingly thin and curved edge of an otolith. These limitations often lead

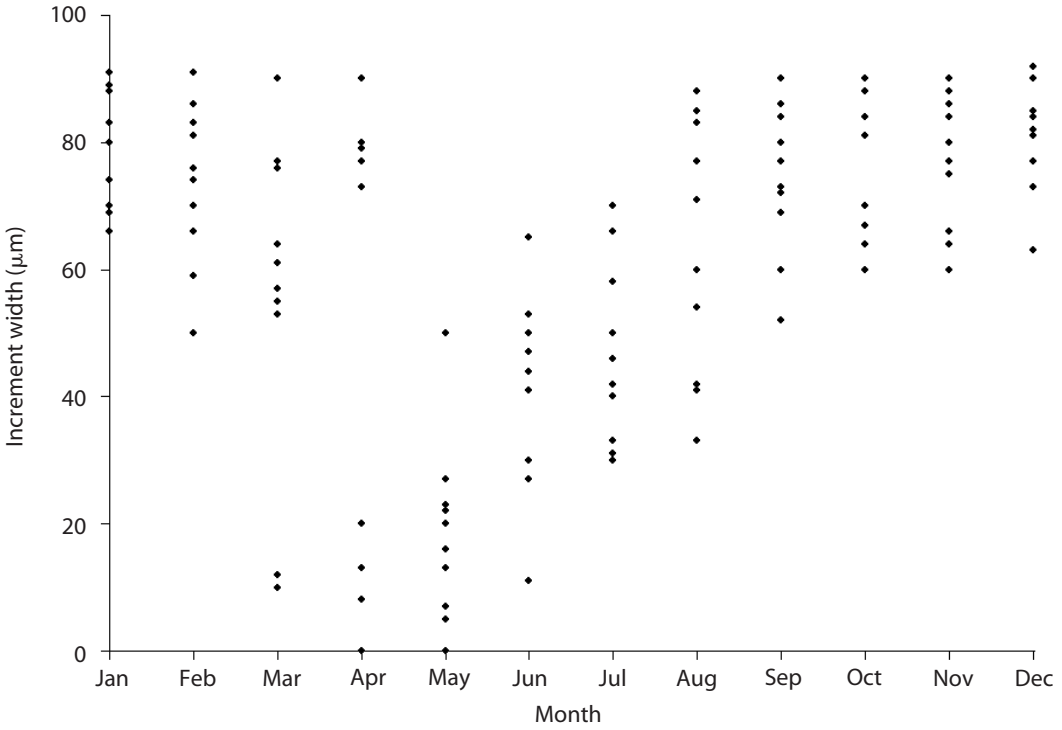


Figure 5.2 Plot of marginal increment, the amount of translucent material between the last increment and the otolith margin, for largemouth bass sampled monthly in a southeastern reservoir. Increment width increases until another annulus is deposited, in this case, between March and May.

to subjective interpretation of results. There are also problems with applying the results of MIA from a younger, faster-growing age-class to older cohorts within the same population (Campana 2001). This tends to result in bias toward underestimating age in older individuals. There also have been unexplained instances where inconsistencies in the timing of increment formation among years and locations have been observed (Beckman and Wilson 1995).

Campana (2001) outlined several aspects of a well-designed validation study using MIA. The most important point is to interpret the results objectively by means of an appropriate statistical analysis. It is essential that only a limited number of age-classes are used in a MIA and that samples from these age-classes be randomized before reading to avoid subjectivity. Finally, at least two complete annual or daily growth cycles should be examined during the course of MIA.

5.2.4.4 Date-Specific Marking

Validation of both annual and in some cases daily growth increments can be accomplished through the recapture of physically or chemically marked fish. The use of chemical marks is perhaps the most powerful of the validation tools, but it also carries the same drawbacks as any mark–recapture study (see DeVries and Frie 1996 for details). Otoliths and other hard parts will incorporate chemicals such as oxytetracycline, alizarin complexone, calcein, and strontium. These chemicals bind to calcium, resulting in a mark on the growth increment forming at the time that will fluoresce under ultraviolet light (Weber and Ridgway 1962, 1967; Rahn and Perrin 1970; Hettler 1984; Wilson et al. 1987). For a review of the use of these chemicals the reader is referred to McFarlane and Beamish (1987b) and Brothers (1990). The marks have a high retention rate on internal structures (Reinert et al. 1998) but may degrade on external structures, such as scales and fin rays, which are exposed to sunlight.

Traditional marks such as externally visible marks or electronic tags such as passive integrated transponder (PIT) tags have been used. With this technique, known-age fish are released and recaptured some time later. By comparing a reference sample collected at the time of release with samples collected from recaptured marked fish, annual deposition can be validated.

5.2.4.5 Captive Rearing

Individuals held and reared in captivity can be used to validate daily growth increment formation. This technique generally is not considered suitable to validate annulus formation because laboratory conditions cannot fully recreate the natural environment. Even daily increments differ in appearance in captive-reared individuals. However, their frequency of formation is rarely influenced because of the endogenous control of the process (Geffen 1987; Morales-Nin 2000).

5.2.5 Applications of Age Data

In addition to its use in estimating growth (see section 5.3), age data can be used in several other applications. Data collected from otoliths and other hard parts

are utilized to construct age- or cohort-specific models of mortality and survivorship. Otoliths are increasingly being used as biological data recorders of temperature and salinity regimes (Campana 2005). However despite these advances, age–length keys and hatch date analysis remain important tools frequently used to evaluate population structure and events that are not easily observed, such as spawning and migration.

5.2.5.1 *Age–Length Keys*

The relationship between age and length is relatively stable within a population. Consequently, age can account for a large amount of the variability in length. Given a sample of fish that has been aged, we can produce a probability matrix of the proportion of individuals within a certain length-class having a certain age (Fridriksson 1934; Ketchen 1949; Isermann and Knight 2005). This table is often referred to as an age–length key (Box 5.1). The age–length key can then be used to estimate the age of fish of a given length so that length frequency from a much larger sample can be converted to age frequency (Isermann and Knight 2005). The use of ages estimated from age–length keys can significantly reduce the time or cost associated with aging large numbers of fish. The method is particularly valuable when applied to rare or endangered species, for which the collection of tissues used in aging may be problematic. When applied to early life stages, age-frequency information can provide insight into spawning and migration not available from length information alone. It is important to note that the usefulness of age–length keys is generally restricted in time and space. Variability in growth among years and geographic locations (Westrheim and Ricker 1978; Terceiro and Ross 1993; Bettoli and Miranda 2001) may bias the results obtained from using age–length keys developed from other times or places.

5.2.5.2 *Hatch Date Analysis*

By using information obtained from otolith daily growth increments, it is possible to determine the hatch date of larval and juvenile fishes. In early life stages, hatch date distributions can be used to glean information on the importance of density dependent and independent factors on spawning, growth, and survival. This technique has numerous applications including identifying the periodicity of spawning events, locating spawning habitats, and examining cohort-specific patterns of mortality. Although similar to age-frequency analysis, hatch date analysis uses age to back-calculate hatch date; then adjustments are made for the effects of cumulative mortality on the numbers produced at each date. By incorporating mortality information, scientists are better able to estimate egg production and other variables important in assessing stocks.

5.3 GROWTH

Growth is the addition of biomass to either a population or an individual. In fisheries management, we attempt to optimize the efficiency of harvest by balancing individual growth, population biomass, and mortality. If we harvest young fish, we

Box 5.1 Creating an Age–Length Key

Fisheries scientists often collect length data on large samples, but age data, because of the large amount of effort involved, are generally collected on smaller samples (i.e., subsamples). In some cases, we wish to convert our length data to age data. We do this through the use of an age–length key. We start with a data set containing individual length and age data. By dividing length data into a series of discrete intervals, we can determine the frequency of ages within each interval. These frequencies are transformed into probabilities, which are later used to convert numbers at length to numbers at age. In this example, we have age and length (tl) data for adult spotted sucker. We create a series of length intervals and create a new variable (tlint) that is a discrete representation of the length data. In this case, we develop 2 cm (20 mm) length-groups and name each group by the low end of the interval. We then determine cell frequencies and calculate cell probabilities using Proc Freq in SAS (SAS 2004). By adding some options to the tables statement, we can suppress the printing of the frequencies and percentages we don't need.

Program

```
data spotted;
input tl age;

if 90<= tl < 100 then tlint = 90;
else if 100<= tl < 120 then tlint = 100;
else if 120<= tl < 140 then tlint = 120;
else if 140<= tl < 160 then tlint = 140;
else if 160<= tl < 180 then tlint = 160;
else if 180<= tl < 200 then tlint = 180;
else if 200<= tl < 220 then tlint = 200;
else if 220<= tl < 240 then tlint = 220;
else if 240<= tl < 260 then tlint = 240;
else if 260<= tl < 280 then tlint = 260;
else if 280<= tl < 300 then tlint = 280;
else if 300<= tl < 320 then tlint = 300;
else if 320<= tl < 340 then tlint = 320;
else if 340<= tl < 360 then tlint = 340;
else if 360<= tl < 380 then tlint = 360;
else if 380<= tl < 400 then tlint = 380;
else if 400<= tl < 420 then tlint = 400;
else if 420<= tl < 440 then tlint = 420;
else if 440<= tl < 460 then tlint = 440;
else if 460<= tl < 480 then tlint = 460;
else if 480<= tl < 500 then tlint = 480;
else if 500<= tl < 520 then tlint = 500;
else if 520<= tl < 540 then tlint = 520;

datalines;
100                1
111                1
114                1
384                4
(input remaining data)
;

proc freq;
tables tlint*age / nocol nofreq nocum nopercent;
run;
```

Program Output

The output consists of a table containing the row percent, which is equal to the probability that a fish within a certain size interval is a certain age.

Table Output from the frequency procedure. Given is the probability that a fish within a given length interval (tlint) is a certain age.

tlint and total number fish	Age										Total number
	1	2	3	4	5	6	7	8	9	10	
90	16.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
100	66.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
120	16.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
240	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
300	0.00	0.00	9.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
320	0.00	0.00	18.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
340	0.00	0.00	27.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
360	0.00	0.00	27.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
380	0.00	0.00	18.18	30.77	0.00	0.00	0.00	0.00	0.00	0.00	
400	0.00	0.00	0.00	46.15	0.00	0.00	0.00	0.00	0.00	0.00	
420	0.00	0.00	0.00	23.08	60.00	0.00	0.00	0.00	0.00	0.00	
440	0.00	0.00	0.00	0.00	40.00	100.00	0.00	100.00	40.00	0.00	
460	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	40.00	0.00	
480	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	100.00	
Total number	6	4	11	13	10	5	2	3	5	2	61

Program

Once an age–length key is generated, the length-frequency distribution from the larger sample is put into the same interval format as the aged sample. The cell frequencies are then multiplied by the frequencies from the age–length key to estimate the age distribution of the sample.

```
data spotall;
input tl;

if 90<= tl < 100 then tlint = 90;
else if 100<= tl < 120 then tlint = 100;
else if 120<= tl < 140 then tlint = 120;
else if 140<= tl < 160 then tlint = 140;
else if 160<= tl < 180 then tlint = 160;
else if 180<= tl < 200 then tlint = 180;
else if 200<= tl < 220 then tlint = 200;
else if 220<= tl < 240 then tlint = 220;
else if 240<= tl < 260 then tlint = 240;
else if 260<= tl < 280 then tlint = 260;
else if 280<= tl < 300 then tlint = 280;
else if 300<= tl < 320 then tlint = 300;
else if 320<= tl < 340 then tlint = 320;
else if 340<= tl < 360 then tlint = 340;
```

(Box continues)

Box 5.1 *(continued)*

```
else if 360<= tl < 380 then tlint = 360;
else if 380<= tl < 400 then tlint = 380;
else if 400<= tl < 420 then tlint = 400;
else if 420<= tl < 440 then tlint = 420;
else if 440<= tl < 460 then tlint = 440;
else if 460<= tl < 480 then tlint = 460;
else if 480<= tl < 500 then tlint = 480;
else if 500<= tl < 520 then tlint = 500;
else if 520<= tl < 540 then tlint = 520;

datalines;
336
336
336
395
395
395
395
386
386
386
386
416
416
416
416
452
452
(input remaining data)
;

proc means mean;
class tlint;
run;
```

Program Output

The above program will produce a summary table of the number of fish per length interval.

Table Output from the means procedure. Summary statistics for the variable length (tl) for each length intervals (tlint).

tlint	*N*	Mean	SD	Minimum	Maximum
300	3	318.0000000	0	318.0000000	318.0000000
320	6	335.5000000	0.5477226	335.0000000	336.0000000
340	12	350.7500000	6.3263518	344.0000000	359.0000000
360	12	372.2500000	7.8985039	360.0000000	379.0000000
380	30	392.3333333	6.1941760	382.0000000	399.0000000
400	28	413.0000000	4.4886689	405.0000000	418.0000000
420	48	432.5625000	5.5039879	420.0000000	438.0000000
440	51	449.7450980	4.5380310	443.0000000	459.0000000
460	61	466.9836066	4.1412229	462.0000000	474.0000000
480	83	492.4216867	6.2627830	480.0000000	499.0000000
500	29	512.2068966	2.0244807	510.0000000	514.0000000
520	36	528.4444444	1.6978044	526.0000000	530.0000000

Program

By using the information from this table as a summary data set, we create a data set for each age-group and then merge the data sets to create an aged sample.

```
data spotfreq;
input tlint num;
datalines;
300                3
350                6
340               12
360               12
380               30
400               28
420               48
440               51
460               61
480               83
500               29
520              36;
run;

data spotage1;
set spotfreq;
if tlint = 90 then age = 1;
else if tlint = 100 then age = 1;
else if tlint = 120 then age = 1;
if tlint = 90 then nage = (num* 100)/100;
else if tlint = 100 then nage = (num* 100)/100;
else if tlint = 120 then nage = (num* 100)/100;
if nage = . then delete;
run;

data spotage2;
set spotfreq;
if 240 then age = 2;
if 240 then nage = (num* 100)/100;
if nage = . then delete;
run;

data spotage3;
set spotfreq;
if tlint = 300 then age = 3;
else if tlint = 320 then age = 3;
else if tlint = 340 then age = 3;
else if tlint = 360 then age = 3;
else if tlint = 380 then age = 1;
if tlint = 300 then nage = (num* 100)/100;
else if tlint = 320 then nage = (num* 100)/100;
else if tlint = 340 then nage = (num* 100)/100;
else if tlint = 360 then nage = (num* 100)/100;
else if tlint = 380 then nage = (num* 33.33)/100;
if nage = . then delete;
run;
```

(Box continues)

Box 5.1 *(continued)*

```
data spotage4;
set spotfreq;
if tlint = 380 then age = 4;
else if tlint = 400 then age = 4;
else if tlint = 420 then age = 4;
if tlint = 380 then nage = (num* 66.67)/100;
else if tlint = 400 then nage = (num* 100)/100;
else if tlint = 420 then nage = (num* 33.33)/100;
if nage = . then delete;
run;

data spotage5;
set spotfreq;
if tlint = 420 then age = 5;
else if tlint = 440 then age = 5;
if tlint = 420 then nage = (num* 66.67)/100;
else if tlint = 440 then nage = (num* 28.57)/100;
if nage = . then delete;
run;

data spotage6;
set spotfreq;
if tlint = 440 then age = 6;
if tlint = 440 then nage = (num* 35.71)/100;
if nage = . then delete;
run;

data spotage7;
set spotfreq;
if tlint = 460 then age = 7;
if tlint = 460 then nage = (num* 50.00)/100;
if nage = . then delete;
run;

data spotage8;
set spotfreq;
if tlint = 440 then age = 8;
else if tlint = 480 then age = 8;
if tlint = 440 then nage = (num* 21.43)/100;
else if tlint = 480 then nage = (num* 33.33)/100;
if nage = . then delete;
run;

data spotage9;
set spotfreq;
if tlint = 440 then age = 9;
else if tlint = 460 then age = 9;
else if tlint = 480 then age = 9;
if tlint = 440 then nage = (num* 14.29)/100;
else if tlint = 460 then nage = (num* 50.00)/100;
else if tlint = 480 then nage = (num* 66.67)/100;
if nage = . then delete;
run;

data spotage;
set spotage1 spotage2 spotage3 spotage4 spotage5 spotage6 spotage7 spotage8
spotage9;
run;

proc print;
run;
```

Program Output

The resulting data set contains the number of fish in each age-group (nage) by length category.

Table Number of fish in each age-group (nage) by length category (tlint) for the larger sample.

tlint	Number	Age	nage
300	3	2	3.0000
350	6	2	6.0000
340	12	2	12.0000
360	12	2	12.0000
380	30	2	30.0000
400	28	2	28.0000
420	48	2	48.0000
440	51	2	51.0000
460	61	2	61.0000
480	83	2	83.0000
500	29	2	29.0000
520	36	2	36.0000
300	3	3	3.0000
340	12	3	12.0000
360	12	3	12.0000
380	30	1	9.9990
380	30	4	20.0010
400	28	4	28.0000
420	48	4	15.9984
420	48	5	32.0016
440	51	5	14.5707
440	51	6	18.2121
460	61	7	30.5000
440	51	8	10.9293
480	83	8	27.6639
440	51	9	7.2879
460	61	9	30.5000
480	83	9	55.3361

may optimize numbers but lose biomass because we have not allowed individuals to reproduce. Alternatively, if we harvest older fish, individual biomass may be maximized, but a large portion of the population will be lost to natural mortality. The interplay between growth and mortality is, therefore, critical in determining management strategies. Growth is also an important component in understanding the ecology of a species at both the individual and population level, as it is a convenient method for assessing the quality of a habitat and tracing life histories.

On an individual basis, change in length is proportional to change in weight. We can relate length and weight using the equation

$$W = aL^b, \tag{5.1}$$

where W is weight, L is length, and a and b are constants. This relationship can be expressed in linear form with the equation

$$\log_e W = a + b \log_e L\,. \tag{5.2}$$

When change in all three dimensions is similar across all sizes, we consider growth to be isometric. This results in the special case in which the exponent $b = 3$. In most species, body shape does not change with age; therefore, most species grow isometrically. In species for which individuals change shape with age either through metamorphosis, development of secondary sex characteristics, or senescence, growth is said to be allometric, and $b \neq 3$.

For a variety of reasons including ease, we often measure length rather than weight. One representation of growth can be obtained by simply comparing the change in modal lengths through time within a population when the modes are considered to represent distinct age-classes. Similarly, the progression of modes through time can be used to estimate growth (Figure 5.3). This method of growth determination assumes that the sample is drawn at random with respect to size and that growth across age-classes is similar through time. Variation in growth within and among years may result in biased estimates when different year-classes are compared.

There are several different methods to express growth numerically (Ricker 1975; Busacker et al. 1990). If growth is reported as the change in length or weight over a given time interval, then it is termed absolute growth and expressed as

$$\Delta L_{\text{absolute}} = L_2 - L_1\,, \tag{5.3}$$

where $\Delta L_{\text{absolute}}$ is absolute growth, L_1 is initial length, and L_2 is final length. However, growth can also be expressed as a percent increase in length or weight relative to an initial value. This is referred to as relative growth and is generally expressed as

$$\Delta L_{\text{relative}} = \frac{L_2 - L_1}{L_1}\,100\,, \tag{5.4}$$

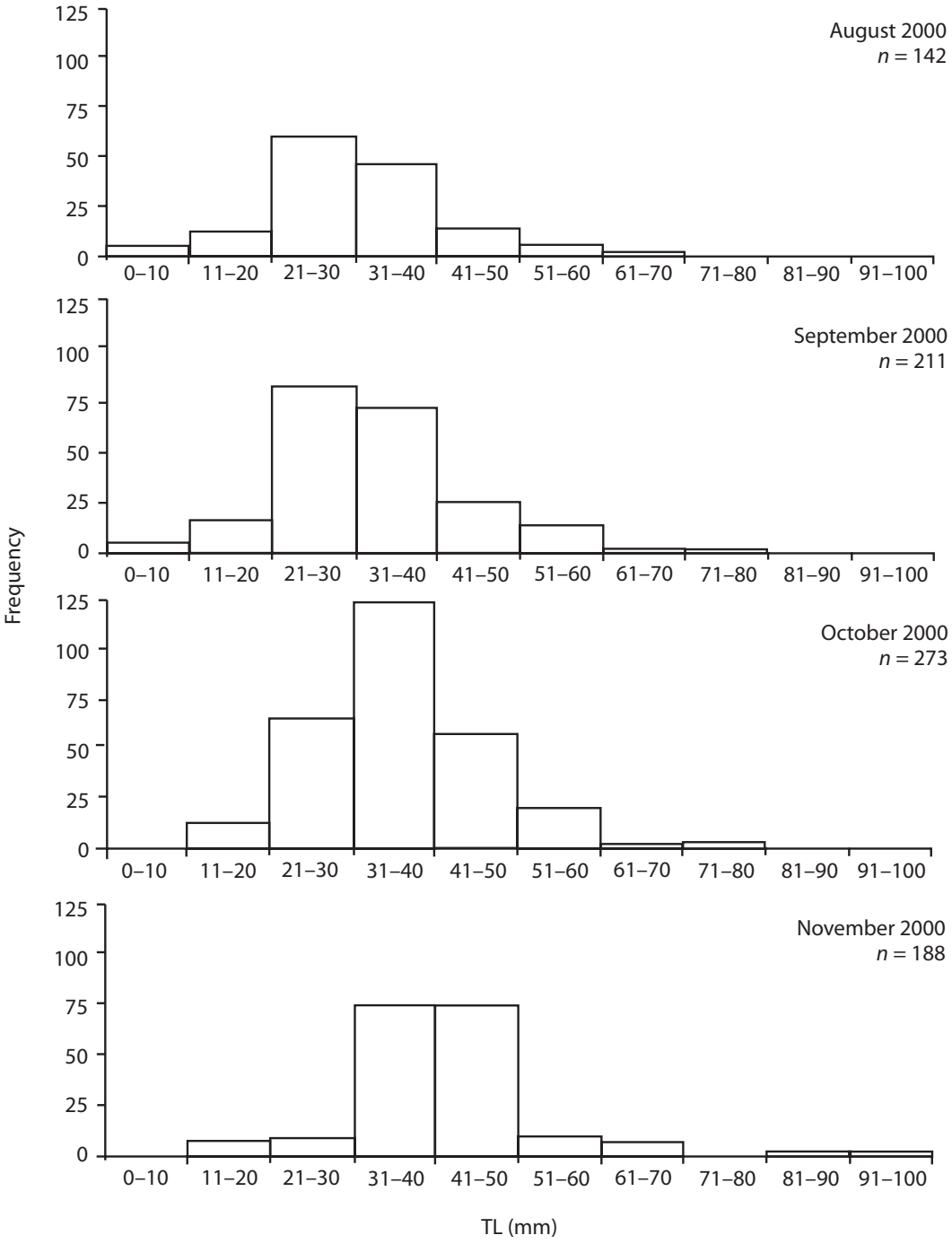


Figure 5.3 Length-frequency distributions of molly miller from August through November 2000 showing an increase in total length (TL) through time of age-0 fish represented in the samples.

where $\Delta L_{\text{relative}}$ is relative growth. Both absolute and relative growth also can be expressed as a rate in terms of growth per unit time:

$$G_{\text{absolute}} = \frac{L_2 - L_1}{t_2 - t_1}, \text{ and} \tag{5.5}$$

$$G_{\text{relative}} = \frac{L_2 - L_1}{L_1(t_2 - t_1)}, \tag{5.6}$$

where G is growth rate either absolute or relative, t_1 is initial time, t_2 is final time, and L_1 and L_2 are the corresponding lengths for those times. In the case that growth is exponential over a short period of time (<1 year) it is best reported as an instantaneous rate:

$$G = \frac{\log_e L_2 - \log_e L_1}{t_2 - t_1}. \tag{5.7}$$

These calculations of growth rates yield an estimate of growth that is appropriate over short time scales (days to months). Growth over longer time periods tends to deviate from these simple, linear, or exponential estimates and requires more complex models that will be discussed later.

5.3.1 Back-Calculation of Length from Calcified Structures

If we assume that the growth of calcified structures is proportional to overall fish growth, a simple ratio or direct proportion method can be used to back-calculate size at annulus formation (Box 5.2). If we know the length of the fish, the radius of the calcified structure, and the radius to each annulus, we can use the equation

$$\frac{L_i}{L_c} = \frac{S_i}{S_c}, \text{ or } L_i = \frac{S_i L_c}{S_c}, \tag{5.8}$$

where S_i is the radius at annulus formation, S_c is the overall radius, L_i is the length at annulus formation, and L_c is the fish length at capture (Box 5.2). Although this relationship generally holds true, fisheries scientists have noticed that it often results in an underestimation of length when scales are used. A tacit assumption of proportionality is that scales are formed early in development. For many species, this is not true. Scales may not form in some species until the individual reaches lengths of 5 cm or greater. Work done by Fraser (1916) and Lee (1920) suggested that a correction factor (a) be added to the equation to account for the delay in scale formation. The resulting equation,

Box 5.2 Determining Mean Back-Calculated Length at Age

In addition to providing estimates of age, hard parts are often used to back-calculate length at younger ages. To demonstrate how this is accomplished, we will be using a data set determined from scales and describing the age and growth of spotted sucker from the Savannah River. For each fish, our data set contains an identification number (ID), sex, total length at capture (L_c), year of capture (date), age, radius of ageing structure (scale) at capture (S_c), annulus *i* (inc), and scale radius at each annulus *i* (S_i) for each individual annulus.

Dahl–Lea Method

We start with the simple case in which the growth of the structure used for ageing is directly proportional to the growth of the fish. This method is generally referred to as the Dahl–Lea method (Dahl 1907; Lea 1910) and allows one to back-calculate length at age for individual fish. The formula is

$$L_i = L_c(S_i / S_c),$$

where L_i is back-calculated length at annulus *i*, L_c is length at capture, S_i is ageing-structure radius at annulus *i*, and S_c is ageing-structure radius at capture. Using the SAS code below, we can generate back-calculated total lengths (L_i) and calculate mean length at age for the spotted sucker population.

Program

```
data sucker;
input ID$ sex$ Lc date age Sc inc Si;
Li = LC * (Si/Sc);
cards;
07447              M     336    2004       3    16.3       1       5
07447              M     336    2004       3    16.3       2    12.9
07447              M     336    2004       3    16.3       3    16.3
35334              F     395    2004       4    18.4       1     4.8
35334              F     395    2004       4    18.4       2     9.9
35334              F     395    2004       4    18.4       3    16.5
35334              F     395    2004       4    18.4       4    18.4
44736              F     386    2004       4    18.6       1     4.9
44736              F     386    2004       4    18.6       2     8.5
44736              F     386    2004       4    18.6       3    13.6
44736              F     386    2004       4    18.6       4    18.6
(input remaining data)
;
run; quit;

proc means data=sucker mean stderr std;
title 'Mean back-calculated TL at age for spotted sucker';
class inc;
var Li;
run; quit;
```

(Box continues)

Box 5.2 *(continued)*

Program Output

The above SAS program will yield the following output for our spotted sucker data set.

Table Mean back-calculated total length (L_i) at age for spotted sucker generated the means procedure.

Annulus *i*	Number of observations	Back-calculated total length		
		Mean	SE	SD
1	65	93.2168665	2.3905126	19.2729288
2	65	217.2709364	5.0262681	40.5230692
3	65	328.9349467	5.6120787	45.2460251
4	52	373.7611035	6.1277514	44.1878437
5	35	388.8813533	7.3027812	43.2038363
6	27	404.2197909	7.3382603	38.1307192
7	21	413.5260280	6.5852539	30.1774246
8	18	430.0907986	6.5060450	27.6028112
9	18	452.1169315	6.6510322	28.2179397
10	14	464.0779850	7.4847926	28.0055297
11	11	478.1245135	9.1952021	30.4970351
12	9	488.7881206	8.7617199	26.2851596
13	6	498.4247331	10.1487390	24.8592322
14	4	501.2899579	10.1750810	20.3501620
15	3	516.9213162	17.0469489	29.5261816
16	2	540.9583333	26.9583333	38.1248406
17	1	580.0000000		

Fraser–Lee Model

In some cases, structures such as scales may take some time to form after hatch or metamorphosis. Consequently, early length estimates are biased. The Fraser–Lee model (Fraser 1916; Lee 1920) accounts for this bias by including a biological intercept in the model. The model is

$$L_i = a + (L_c - a)(S_i/S_c).$$

The variable a is the intercept determined from the ageing-structure radius and fish length relationship and the other variables are previously defined.

Because we are using scales to back-calculate length at age, we will likely require a correction factor. Because we did not collect empirical data or find information in the literature regarding the length of scale formation in spotted sucker, then we must estimate this parameter by modeling the known relationship between ageing-structure radius and fish-length at capture from our spotted sucker data set. Even had we found this information in the literature, performing the below calculations is another good way to check one's data.

Program

```
proc glm data=sucker;
title 'Estimate of biological intercept';
model Li = Si;
run; quit;
```

We use regression analysis to determine the relationship between scale radius at age and back-calculated total length at age. In a larger data set, it may be possible to use L_c and S_c directly to estimate the Fraser–Lee correction factor. However, our spotted sucker data set does not have any individuals younger than age 3. Using L_c and S_c provides us with unrealistically large estimates.

Another important consideration is that units of measurement for L_i and S_i are the same for this calculation. Whereas this is not a concern during calculations of length at age because the units cancel, it will yield inaccurate estimates for the correction factor. In our spotted sucker data set, scales were magnified 24× and measured in centimeters. A conversion will be necessary as total length was measured in millimeters. Therefore, we will create a new variable containing the converted scale radii with the following statement.

```
data sucker_Si2;
set sucker;
Li = Lc * (Si/Sc);
Si2 = (Si*10)/24;
run; quit;

proc glm data=sucker_Si2;
title 'Estimate of biological intercept';
model Li = Si2;
run; quit;
```

This will convert scale radius from centimeters to millimeters and account for making the measurements under magnification. The program will now yield the following output.

Program Output

Table Estimate of biological intercept for the dependent variable mean back-calculated total length (L_i) based on the general linear model (GLM) procedure. The number of observations used and read was 416. Abbreviations are given for coefficient of variation (CV), mean square error (MSE), and sum of squares (SS).

General linear model

Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	1	6486640.962	6486640.962	3672.42	<0.0001
Error	414	731253.767	1766.313		
Corrected total	415	7217894.729			
R^2	0.898689	Root MSE	42.02753		
CV	13.00363	Li mean	323.1985		

(Box continues)

Box 5.2 *(continued)*

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Si2	1	6486640.962	6486640.962	3672.42	<0.0001

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Si2	1	6486640.962	6486640.962	3672.42	<0.0001

Variable	Estimate	SE	*t*-value	*P* > |*t*|
Intercept	35.49138299	5.17548973	6.86	<0.0001
Si2	50.88336354	0.83965286	60.60	<0.0001

The intercept of this linear model will be an estimate of the Fraser–Lee correction factor. Therefore, we must be able to reject the null hypothesis that the intercept is not different from 0. If we cannot reject this null hypothesis, then a correction factor is not likely to be necessary. In the case of our spotted sucker data, a correction factor is necessary. Now, the correction factor a can be used to calculate mean back-calculated total length (L_i) at age with the Fraser–Lee correction.

Program

```
data sucker_a;
set sucker;
Li = Lc * (Si/Sc);
Si2 = (Si*10)/24;
a = 35.5;
Li_corrected = a + (Lc - a) * (Si/Sc);
run; quit;

proc means data=sucker_a mean stderr std;
title 'Corrected mean back-calculated TL at age for spotted sucker';
class inc;
var Li_corrected;
run; quit;
```

$$L_i = \frac{L_c - a}{S_c} S_i + a\,, \tag{5.9}$$

where a is the size of the individual at the time of scale formation, provides an unbiased estimate in length when scales are used and is referred to as the Fraser–Lee or intercept-corrected direct proportion model. While this formula is widely used, it may not be the most precise estimate of length at age. There can also be differences in the precision of back-calculation depending upon the structure,

Program Output

The above program produces the following output.

Table Corrected mean back-calculated total length (L_i) at age for spotted sucker generated by the means procedure.

Annulus *i*	Number of observations	Back-calculated total length		
		Mean	SE	SD
1	65	120.8801223	2.1454817	17.2974263
2	65	234.5165149	4.4996539	36.2773695
3	65	336.8015953	4.9458073	39.8743735
4	52	379.4639509	5.3958210	38.9098182
5	35	394.8724181	6.4986305	38.4464168
6	27	409.7792454	6.5739649	34.1593235
7	21	419.0105377	5.9427822	27.2332491
8	18	434.8327817	5.9630989	25.2992861
9	18	455.2899399	6.1182144	25.9573855
10	14	466.9561266	7.0110838	26.2330736
11	11	480.5614840	8.7472611	29.0113831
12	9	490.6474731	8.3865848	25.1597543
13	6	499.9279199	9.9833260	24.4540545
14	4	502.6128845	10.6974006	21.3948012
15	3	517.8177039	17.4957374	30.3035062
16	2	541.3281250	27.3281250	38.6478050
17	1	580.0000000		

This method is very useful when sample sizes are small and additional growth information is needed. It can also be used to develop data to test for size-selective mortality (or Lee's phenomenon), a common occurrence in commercial fisheries. Growth histories from specific year-classes can be compared or data can be converted to year-specific growth to compare inter-annual variations.

necessitating careful selection of the model used (Campana 1990; Klumb et al. 2001). The reader is referred to Francis (1990) for a review of alternative methods. Once calculated, size-at-age information between sexes and populations can be compared (Box 5.3).

Often times, back-calculated lengths fall below the mean of observed lengths from the same population. This apparent change in growth over time was first described by Lee in 1920 and is discussed by Ricker (1975) and others to a greater extent. Interestingly, Lee's phenomenon can be related to (1) failure to use the

Box 5.3 Assessing Differences in Length at Age between Groups

Now that we have corrected back-calculated length at age, we can test for differences between groups. For example, we commonly want to test for a sex effect on length at age. We can use our previous example to evaluate differences between sexes by means of an analysis of covariance (ANCOVA) approach. We start with our spotted sucker data set containing fish identification number (ID), sex, total length at capture (L_c), year of capture (date), age, radius of aging structure (scale) at capture (S_c), annulus increment number (inc), and radius of aging structure at inc (S_i). We calculate the length at each increment using a direct proportion method and incorporate the Fraser–Lee correction factor calculated in Box 5.2. Given that growth has a curvilinear component, we create a dummy variable (incsq) to be incorporated into the model.

Program

```
data sucker;
input ID$ sex$ Lc date age Sc inc Si;
a = 35.5;
Li_corrected = a + (Lc - a) * (Si/Sc);
incsq = inc*inc;
cards;
07447              M     336    2004       3    16.3       1       5
07447              M     336    2004       3    16.3       2    12.9
07447              M     336    2004       3    16.3       3    16.3
35334              F     395    2004       4    18.4       1     4.8
35334              F     395    2004       4    18.4       2     9.9
35334              F     395    2004       4    18.4       3    16.5
35334              F     395    2004       4    18.4       4    18.4
44736              F     386    2004       4    18.6       1     4.9
44736              F     386    2004       4    18.6       2     8.5
44736              F     386    2004       4    18.6       3    13.6
44736              F     386    2004       4    18.6       4    18.6
(input remaining data)
;
run; quit;
```

Once the data are entered we can run the GLM procedure to test the null hypothesis that there is no difference between males and females in the slope of the length at age regressions.

```
proc glm data=sucker;
title 'Testing for equal slopes between males and females';
class sex;
model Li_corrected= sex inc incsq sex*inc sex*incsq;
run; quit;
```

This program will yield the following output.

Program Output

Table Test of the assumption of equal slopes for male and female spotted suckers (ANCOVA) by means of the GLM procedure with Li-corrected (back-calculated total length, L_i, with the Fraser–Lee correction factor) as the dependent variable.

General linear model

Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	5	5486250.292	1097250.058	589.40	<0.0001
Error	410	763266.344	1861.625		
Corrected total	415	6249516.636			
R^2	0.877868	Root MSE	43.14656		
CV	12.92255	Li-corrected mean	333.8859		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Sex	1	116871.182	116871.182	62.78	<0.0001
lnc	1	4286871.155	4286871.155	2302.76	<0.0001
lncsq	1	937786.528	937786.528	503.75	<0.0001
lnc*sex	1	9742.781	9742.781	5.23	0.0227
lncsq*sex	1	134978.646	134978.646	72.51	<0.0001

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Sex	1	49896.978	49896.978	26.80	<0.0001
lnc	1	1974734.921	1974734.921	1060.76	<0.0001
lncsq	1	688068.591	688068.591	369.61	<0.0001
lnc*sex	1	143885.275	143885.275	77.29	<0.0001
lncsq*sex	1	134978.646	134978.646	72.51	<0.0001

The value of interest here is the *P*-value for the interaction terms inc*sex and incsq*sex. This tests the null hypothesis that the slopes are equal between males and females. In this case, it appears that males and females have different slopes and thus grow at different rates. We would use this information to justify modeling the growth of the two sexes separately.

Similarly, differences in growth rates between populations or other treatment variables can be evaluated by placing a population or treatment identifier in the data set and substituting it for sex in the analysis.

corrected body–scale relationship, (2) a bias resulting from size-selective sampling or harvest, or (3) variation in mortality rates as a function of growth. Size selectivity may be the most common bias, as fish tend to be sampled or harvested by size-selective gear rather than by age-selective gear, so the fastest-growing individuals are sampled or harvested first. Consequently, individuals that live the longest tend to be the slowest-growing individuals in the population, resulting in smaller back-calculated sizes at younger ages.

5.3.2 Growth in Weight

Theoretically, in an unlimited environment growth is exponential and can be modeled using the equation

$$w_t = w_0 e^{gt}, \tag{5.10}$$

where w_t is weight at time t, w_0 is initial weight, e is the base of the natural logarithm, and g is a growth coefficient. Although this equation could be used to estimate population growth as well as individual growth, it is seldom applicable for either over long periods. This model assumes no limitations on growth, and this is rarely the case. The model is useful, however, to estimate production of growth within a single growing season or early in development. As previously mentioned, growth in weight is not used as commonly as growth in length. However, weight can be substituted for length in the growth models presented below and will maintain the same form. Coefficients estimated for the resulting equations, obviously, will be different.

5.3.3 Growth in Length

The weight model presented above is not useful to represent growth in length. Early in life, length and weight both increase very rapidly. However, as fish age, small changes in length can result in large changes in weight (equation [5.1]). Although fish are thought to exhibit indeterminate growth, length often approaches an asymptote. A number of models have been used to model length, but the model developed by von Bertalanffy (1938) generally fits fish length data well. It has become a standard among fisheries scientists. The model is represented as

$$l_t = L_\infty(1 - e^{-K(t - t_0)}), \tag{5.11}$$

where l_t is length at time t, L_∞ is the asymptotic length, K is a growth coefficient, and t_0 is a time coefficient at which length would theoretically be 0.

Unlike the simple exponential model, obtaining estimates of L_∞, t_0, and K requires an iterative solution. Most statistical and graphics software packages now contain programs that calculate maximum likelihood estimates (see Chapter 8 for explanation of maximum likelihood) of nonlinear regression parameters such as those in the von Bertalanffy growth equation. Historically, these parameters

were estimated using a graphical solution. Walford (1946) observed that when length at age $t + 1$ was plotted against length at age t, the slope of the line was equal to e^{-K}, where K is the same growth coefficient as in the von Bertalanffy model (Figure 5.4). If a line with a slope of 1 (i.e., a 45° line) is drawn through the

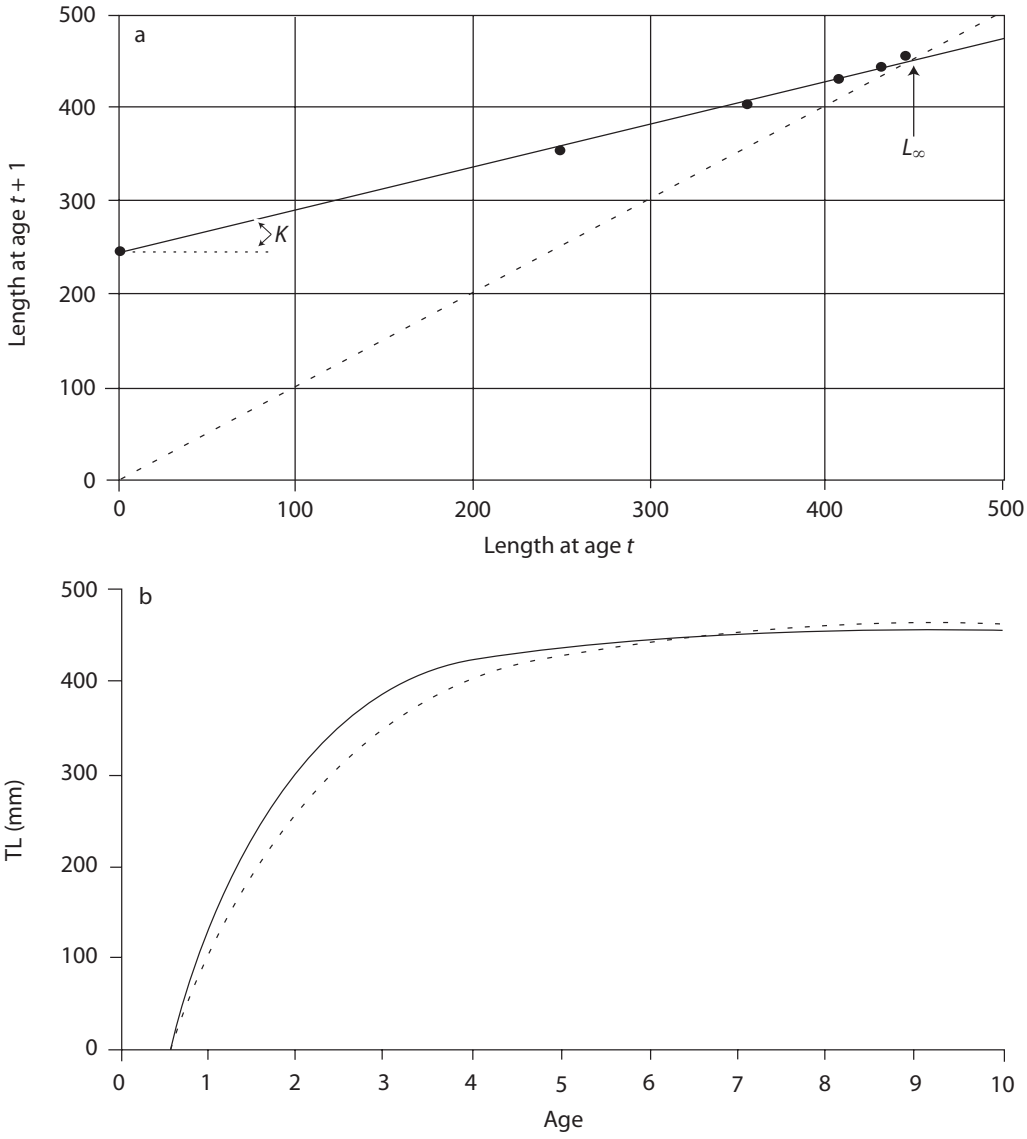


Figure 5.4 (a) Walford plot (solid line) of spotted sucker and the resulting von Bertalanffy growth curve parameters (L_∞ = asymptotic length and K = growth coefficient) estimated from the Walford plot. Dashed line with a slope of 1 is drawn through the origin to provide an estimate of L_∞ at its intersection with the Walford plot. (b) The von Bertalanffy growth curve generated from a Walford plot (solid line) is compared with one generated using iterative procedures illustrated in Box 5.4 (dotted line).

origin, the intersection of the two lines indicates the size at which change in length is theoretically 0 and provides an estimate of L_∞. The parameter t_0 may be estimated by substituting known values for length and age, and the estimates of K and L_∞ derived from the Walford plot may be placed into the von Bertalanffy equation. Although this method produces reasonable estimates, mathematical solutions are more precise (Box 5.4).

The von Bertalanffy model often works well across the entire life history of a fish, meaning it can be applied to a single sample integrated across year-classes or can be developed for individual year-classes. Data also may be stratified by sex or geographic area, modeled independently, and compared using analysis of covariance (ANCOVA) (Box 5.5).

In some mark–recapture studies, we may know size at capture and recapture and time at large, but we may not know age. Fabens (1965) proposed a modification of the von Bertalanffy equation to model growth under this unique circumstance. This model is useful for work on relatively rare or endangered species for which collection of materials on which to base age is impractical or on marine reptiles and other organisms for which a method to determine age has not been identified. The Fabens model is

$$R_i = M_i + (L_\infty - M_i)(1 - e^{-K\Delta t_i}) \tag{5.12}$$

where R_i is the length at recapture of the ith individual, M_i is the length at marking (or first capture) of the ith individual, L_∞ and K are parameters of the von Bertalanffy growth equation, and Δt_i is the time at large. Model parameters can be estimated using a maximum likelihood estimator, or nonlinear fit program, as with the von Bertalanffy model (Box 5.6). Note that this method does not provide and estimate for the time at zero length, t_0, which must be estimated through some other method, by using empirical early growth data, or by substituting known age and length values and parameter estimates into the equation as above.

Although the von Bertalanffy model has become the method of choice for modeling growth in length, other growth models may be more appropriate depending upon the species of interest and the specific circumstances (Ricker 1975). Other commonly applied growth models are included below.

Richards (1959):

$$l_t = D + (L_\infty - D)(1 + He^{-k(t-t_0)})^{-1/H}; \tag{5.13}$$

Gompertz (1825):

$$l_t = L_\infty e^{-ke^{-gt}} \tag{5.14}$$

and the logistic (Verhulst 1838, 1845):

$$l_t = \frac{L_\infty A}{A + (L_\infty - A)e^{-kt}}. \tag{5.15}$$

Box 5.4 Fitting a von Bertalanffy Growth Curve

The length-at-age data on spotted sucker illustrated in Figure 5.4 and Box 5.2 will be used here. For each individual in the data set, we have entered total length (tl) and age. Therefore, each fish represents a single degree of freedom in the analysis. To minimize bias, similar numbers of fish from each year-class should be included in the model. If older or younger age-classes are not well represented in the analysis, confidence limits at the extremes of the curve may expand dramatically or the model will fail to converge. Parameters for the growth curve can now be estimated iteratively using a nonlinear regression approach with the following SAS program.

Program

```
data spotage;
input tl age;
cards;
388                4
418                4
438                4
428                5
539               10
432                4
444                7
421                4
438                4
(input remaining data)
;
run;

proc nlin data = spotage;
model tl = linf*(1-EXP(-k*(age-t0)));
parameters linf = 1000 k = 0.1 t0 = 0.1;
output out = explen p = extl;
run;
```

The model statement expresses the von Bertalanffy model in SAS format. Other models may be substituted. Here are some examples of common growth models expressed in SAS format.

Richards:

model Lt = D+(Lmax–D)*(1+ He**(–k*(t–t0)))**(–1/H) ;

Gompertz:

model Lt = Lmax*exp(–k*exp(–g*t)) ; and

logistic:

model Lt = (Lmax*A)/(A+(Lmax –A)*exp(–k*t) .

The parameters statement provides initial parameter estimates. These values can be estimated from traditional methods such as the Walford plot or by using reasonable values obtained from the literature or from similar species. For example, the asymptotic length, L_∞ (linf), can be estimated as the average length of the oldest age-group. The output statement creates a data set with expected (predicted) values for length at each age, which can then be plotted or analyzed further.

(Box continues)

Box 5.4 *(continued)*

Program Output

The output consists of the results of the iterations, the associated sums of squares, a regression analysis containing the statistical significance of the model, the parameter estimates and associated confidence limits, and a correlation matrix for the parameter estimates.

Table Nonlinear regression analysis (NLIN procedure) of the total length (tl) of spotted sucker and estimates of the von Bertalanffy growth model parameters: asymptotic length (linf), growth coefficient (k), and time coefficient (t0) where length would theoretically be 0. Iterations based on the based on the Gauss-Newton method; convergence criterion was met. The acronyms PPC (prospective parameter change measure) and RPC (retrospective parameter change measure) refer to how well the model met its convergence criteria; the reader is advised to see SAS for details of definitions and procedures.

Iterative phase				
Iteration	linf	k	t0	SS
0	1000.0	0.1000	0.1000	469817
1	776.8	0.1027	–1.5355	326499
2	579.8	0.1096	–4.5622	302753
3	549.8	0.1226	–6.6931	44520.3
4	542.7	0.1382	–6.3290	30574.0
5	535.5	0.1505	–5.8205	30269.8
6	530.3	0.1599	–5.4872	30190.3
7	526.6	0.1670	–5.2450	30171.2
8	523.9	0.1725	–5.0682	30165.6
9	522.0	0.1766	–4.9377	30163.7
10	520.5	0.1798	–4.8406	30163.0
11	519.5	0.1822	–4.7680	30162.6
12	518.7	0.1841	–4.7134	30162.5
13	518.1	0.1855	–4.6723	30162.4
14	517.6	0.1865	–4.6412	30162.3
15	517.3	0.1874	–4.6177	30162.3
16	517.1	0.1880	–4.5999	30162.3
17	516.9	0.1885	–4.5864	30162.3
18	516.7	0.1888	–4.5761	30162.3
19	516.6	0.1891	–4.5683	30162.3
20	516.5	0.1893	–4.5624	30162.3
21	516.5	0.1894	–4.5579	30162.3
22	516.4	0.1896	–4.5544	30162.3
23	516.4	0.1897	–4.5518	30162.3
24	516.3	0.1897	–4.5498	30162.3
25	516.3	0.1898	–4.5483	30162.3
26	516.3	0.1898	–4.5472	30162.3
27	516.3	0.1899	–4.5463	30162.3
28	516.3	0.1899	–4.5457	30162.3
29	516.3	0.1899	–4.5451	30162.3
30	516.3	0.1899	–4.5448	30162.3

Estimation summary	
Method	Gauss-Newton
Iterations	30
Subiterations	1
Average subiterations	0.033333
R	9.028×10^{-6}
PPC(t0)	0.000065
RPC(t0)	0.000085
Object	2.48×10^{-10}
Objective	30162.26
Observations read	95
Observations used	95
Observations missing	0

Regression analysis

Source	*df*	SS	Mean square	Approximate *F*-value	*P* > *F*
Model	2	38328.9	19164.4	58.45	<0.0001
Error	92	30162.3	327.9		
Corrected total	94	68491.2			

Parameter estimates

Parameter	Estimate	Approximate SE	Approximate 95% confidence limits	
linf	516.3	48.8409	419.3	613.3
k	0.1899	0.1194	–0.0473	0.4271
t0	–4.5448	3.4102	–11.3178	2.2283

Approximate correlation matrix

	linf	k	t0
linf	1.0000000	–0.9872305	–0.9616729
k	–0.9872305	1.0000000	0.9927332
t0	–0.9616729	0.9927332	1.0000000

The model indicates that growth of spotted sucker can be estimated using the equation

$$l_t = L_\infty[1 - e^{-K(t - t_0)}],$$

where

$$l_t = 516.3[1 - e^{-0.1899(t + 4.5448)}].$$

Box 5.5 Identifying Environmental Effects on Growth

Often, fisheries scientists are interested in evaluating the effects of some management strategy on growth. Length limits, fertilization, and water level manipulations, for example, may all produce time-specific effects. We cannot simply compare pre-treatment length with post-treatment length. Weisburg and Frie (1987) demonstrated a method of isolating annular growth effects by calculating growth increment and assigning this not only to a specific age but to a specific year. We will use data collected from a population of spotted sucker to test the effects of an extended drought on growth. The drought occurred from 2000 through 2003. Rather than test for the effect of a specific individual year, we group years together by rainfall. Although it would have improved the statistical performance of the model, note that it is not necessary to sample pre-treatment fish length as long as the post-treatment sample contains a representative sample of fish that were alive during the pre-treatment period. In this case, year-classes from normal and drought (dry) years were present in the sample.

Program

```
data spotted_weather;
input id $ sex $ tl w year age bcyear bcage bctl growth;
if bcyear < 2000 then group = "normal";
else if bcyear >1999 then group = "dry";
cards;
04111       F       486     710    2004       9    1995       1      76      76
04111       F       486     710    2004       9    1996       2     181     104
04111       F       486     710    2004       9    1997       3     275      94
(input remaining data)
;
run;

proc glm;
class bcage group;
model growth = bcage group bcage*group;
run;
```

The model statement evaluates annual length increment (growth) as a function of weather conditions (group), age (bcage), and the interaction between age and weather (bcage*group). In cases in which the interaction is significant, results can be interpreted that the treatment affected age-classes differently. For example, we could conclude that during a drought, younger fish might grow slower because of poor habitat but that older fish would grow faster due to a concentration of prey caused by decreased water levels.

Program Output

The output consists of a standard *F* table including model degrees of freedom, the associated model and partial sums of squares, model and partial *F*-values, and significance levels.

Table Evaluation of annual length increment (growth) as a function of weather conditions (group) and age (bcage).

General linear model					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	13	341678.9817	26282.9986	60.18	<0.0001
Error	230	100453.0305	436.7523		
Corrected total	243	442132.0123			
R^2	0.772799	Root MSE	20.89862		
CV	23.71861	Growth mean	88.11066		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
bcage	8	333756.1313	41719.5164	95.52	<0.0001
group	1	6500.5460	6500.5460	14.88	0.0001
bcage*group	4	1422.3044	355.5761	0.81	0.5173

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
bcage	8	210629.6102	26328.7013	60.28	<0.0001
group	1	3025.6062	3025.6062	6.93	0.0091
bcage*group	4	1422.3044	355.5761	0.81	0.5173

The model indicates that both age (bcage) and weather (group) accounted for significant proportions of variation in growth, but that no significant interaction between age and weather was detected. Therefore, we can reduce the model and further evaluate the effects of age and year on growth. Using the same data, we now run the following model.

Program

```
proc glm;
class bcage group;
model growth = bcage group;
means group;
lsmeans group / adjust=dunnett pdiff=control('dry');
run;
```

The reduced model drops the interaction term. We then calculate mean growth for each category of rainfall. The lsmeans statement calculates least-squares means for normal and drought levels and then compares mean values using a *t*-test. The following output is produced.

Program Output

Table Comparison of growth of spotted suckers in dry and normal years. The number of observations read and used was 244. The least-squares means (lsmeans) comparisons are made with the Dunnett–Hsu adjustment for multiple comparisons. The null hypothesis being tested LSMean1 = LSMean2 compares mean growth in dry and normal conditions.

Class level information		
Class	Levels	Values
bcage	9	1 2 3 4 5 6 7 8 9
group	2	dry normal

(Box continues)

Box 5.5 *(continued)*

The GLM procedures

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	9	340256.6774	37806.2975	86.84	<0.0001
Error	234	101875.3349	435.3647		
Corrected total	243	442132.0123			
R^2	0.769582	Root MSE	20.86539		
CV	23.68090	Growth mean	88.11066		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
bcage	8	333756.1313	41719.5164	95.83	<0.0001
group	1	6500.5460	6500.5460	14.93	0.0001

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
bcage	8	340056.3180	42507.0398	97.64	<0.0001
group	1	6500.5460	6500.5460	14.93	0.0001

Group statistics

Level of group	*N*	Growth mean	Growth SD
Dry	195	87.6564103	45.5918683
Normal	49	89.9183673	28.3870839

Least-squares means

Growth	Group lsmean	*P* > |*t*|
Dry	60.7771191	0.0001
Normal	47.1788748	

The model indicates that growth of spotted suckers was higher during drought conditions than during normal conditions. In this case, there were only two groups and the results of the pairwise *t*-test are the same as for the general model. However, the same procedure could be used when more than two groups are present.

Box 5.6 Estimating Growth from Mark and Recapture Data

In this example, data on the carapace length of loggerhead turtles at mark and at recapture will be used to fit a von Bertalanffy growth curve by means of the Fabens (1965) method. For each individual in the data set, time at large (days) has been calculated from the mark and recapture dates. Carapace length at mark (clmark) and at recapture (clrecap) and time at large (timeoutd) has been entered for each individual. To calculate the von Bertalanffy growth parameters in a standard form, time at large has been converted from days to years (timeouty). Each individual, therefore, represents a single degree of freedom in the analysis. If older and younger age-classes are not well represented in the analysis, or if time at large is long with respect to the expected age of the animal, convergence criteria for parameter estimation may not be met. Parameters for the growth curve can now be estimated iteratively using a nonlinear regression approach with the following SAS program.

Program

```
data turlen;
input markd $ clmark recapd $ clrecap timeoutd;
timeouty = timeoutd/365;

cards;
6/28/00          70.3 6/24/03    76.0    1091
7/20/00          60.5 7/15/03    64.7    1090
8/3/00           65.6 7/18/02    67.9     714
8/3/00           61.2 6/25/03    65.2    1056
7/9/01           76.9 6/24/03    79.1     715
7/12/01          64.4  6/9/03    67.5     697
7/18/01          97.4 6/17/02    97.9     335
3/30/98          60.9 6/21/02    67.7    1544
6/27/00          62.5 6/19/03    69.1    1087
5/12/99          67.0 7/12/00    69.5     427
6/18/89          41.0 9/21/91    52.0    1187
8/28/91          60.412/24/94    70.5    1945
7/14/86          19.3 8/28/91    42.0    1869
6/29/90          40.5  2/6/93    49.8     890
5/10/92          28.0 6/23/96    48.0    1503
6/30/92          26.011/15/95    42.0    1230
8/8/96           64.0 6/19/98    69.6     680
;
run;

proc nlin data = turlen;
model clrecap = clmark + (linf - clmark)*(1 - exp(-K*timeouty));
parms linf = 100 k = 0.1;
output out = pturlen p = pclrecap;
run;
```

The model statement expresses the von Bertalanffy model in SAS format without the usual t_0 parameter. The parameter t_0 can be estimated independently or by using a known-age individual to "anchor" the growth curve once L_∞ (linf) and K (k) have been estimated. As with our initial nonlinear fit exercise, the Fabens (1965) approach can be used to estimate parameters of other growth models.

(Box continues)

Box 5.6 *(continued)*

Again, the parameters statement provides initial parameter estimates, and the output statement creates a data set with expected (predicted) values for carapace length at recapture, which can then be plotted or analyzed further.

Program Output

The output consists of the results of the iterations and the associated sums of squares, a regression analysis containing the statistical significance of the model, parameter estimates and associated confidence limits, and a correlation matrix for the parameter estimates.

Table Nonlinear regression analysis (NLIN) of loggerhead turtle carapace length at recapture (clrecap). Estimates of growth model parameters L_∞ (linf) and K (k) are produced; convergence criterion was met. An intercept was not specified for the regression model.

Iterative phase			
Iteration	linf	k	SS
0	100.0	0.1000	375.4
1	89.5614	0.0839	30.3416
2	87.4962	0.0857	28.9656
3	88.0507	0.0847	28.9197
4	88.0477	0.0845	28.9169
5	88.0490	0.0845	28.9169
6	88.0481	0.0845	28.9169

Estimation summary	
Method	DUD
Iterations	6
Object	4.114×10^{-9}
Objective	28.91686
Observations read	17
Observations used	17
Observations missing	0

In these models, l_t is the size at time t, L_∞ is the asymptotic length, t_0 is the time at size 0, k and g are generalized growth parameters that vary slightly in definition between models, and A, D, and H are position parameters used to constrain the inflection point. The logistic model differs from the von Bertalanffy, Richards, and Gompertz models in that it is symmetrical in relation to the inflection point. These models are not commonly employed in fisheries but are frequently used for other organisms.

Regression model					
Source	*df*	SS	Mean square	Approximate *F*-value	*P* > *F*
Regression	2	74210.9	37105.4	19247.7	<0.0001
Residual	15	28.9169	1.9278		
Uncorrected total	17	74239.8			
Corrected total	16	3257.3			

Parameter estimates				
Parameter	Estimate	Approximate SE	Approximate 95% confidence limits	
linf	88.0481	3.5192	80.5471	95.5491
k	0.0845	0.00820	0.0671	0.1020

Approximate correlation matrix		
	linf	k
linf	1.0000000	–0.9079249
k	–0.9079249	1.0000000

The model indicates that growth (as carapace length, *Cl*) of loggerhead turtle can be estimated using the equation

$$Cl_t = L_\infty(1 - e^{Kt}),$$

where

$$Cl_t = 88.0481(1 - e^{-0.0845t}).$$

The model assumes $t_0 = 0$. If t_0 can be estimated independently, then the model can be adjusted accordingly.

5.4 SUMMARY

In this chapter, we learned about the unique properties of fish calcified structures to record growth history and the importance of validating the interpretation of these structures. We learned that unbiased age determination is the backbone of modern stock assessment. Because of the close relationship between age and length, ages determined from a subsample can be used to estimate the age distribution of

the population. If we know age and size, we can determine growth and compare growth attributes between populations. If our age data are frequent relative to the age of the fish, we can model the change in size through time using nonlinear models such as the model proposed by von Bertalanffy (1938). We can also compare current growth with historic or back-calculated growth to evaluate size-selective processes in fisheries. Age and growth data are a critical component in the effective management of fisheries resources. While age and growth analyses are generally straightforward, collection and interpretation requires skill and experience. The application of age and growth data to recruitment, mortality, and other population models can be expensive and time consuming but are critical to the conclusions drawn from these studies. All attempts should be made to incorporate them whenever possible. One of the most important recent advances in the field of age and growth has been the detection of daily increments in otoliths. This discovery has allowed fisheries scientists to apply analytical techniques to age-0 fish that were previously reserved for adult fish. Consequently, we now have a better ability to evaluate factors affecting recruitment and year-class strength formation.

5.5 REFERENCES

Andrews, A. H., K. H. Coale, J. L. Nowicki, C. Lundstrom, Z. Palacz, E. J. Burton, and G. M. Cailliet. 1999. Application of an ion-exchange separation technique and thermal ionization mass spectrometry to ^{226}Ra determination in otoliths for radiometric age determination of long-lived fishes. Canadian Journal of Fisheries and Aquatic Sciences 56:1329–1338.

Baker, E. A., and T. S. McComish. 1998. Precision of ages determined from scales and opercles for yellow perch *Perca flavescens*. Journal of Great Lakes Research 24:658–665.

Baker, T. T., and L. S. Timmons. 1991. Precision of ages estimated from 5 bony structures of Arctic char (*Salvelinus alpinus*) from the Wood River system, Alaska. Canadian Journal of Fisheries and Aquatic Sciences 48:1007–1014.

Bardach, J. E. 1955. The opercular bone of the yellow perch, *Perca flavescens,* as a tool for age and growth studies. Copeia 1955:107–109.

Beamish, R. J. 1981. Use of fin-ray sections to age walleye pollock, Pacific cod, and albacore, and the importance of this method. Transactions of the American Fisheries Society 110:287–299.

Beamish, R. J., and G. A. McFarlane. 1983. The forgotten requirement for age validation in fisheries biology. Transactions of the American Fisheries Society 112:735–743.

Beckman, D. W., and C. A. Wilson. 1995. Seasonal timing of opaque zone formation in fish otoliths. Pages 27–43 *in* D. H. Secor, J. M. Dean, and S. E. Campana, editors. Developments in fish otolith research. University of South Carolina Press, Columbia.

Bennett, J. T., G. W. Boehlert, and K. K. Turekian. 1982. Confirmation of longevity in *Sebastes diploproa* (Pisces: Scorpaenidae) from Pb-210/Ra-226 measurements in otoliths. Marine Biology 71:209–215.

Bettoli, P. W., and L. E. Miranda. 2001. Cautionary note about estimating mean length at age with subsampled data. North American Journal of Fisheries Management 21: 425–428.

Boyko, E. G. 1946. Age determination of fishes based on cross sections of fin rays. Progressive Fish-Culturist 12:47–48.

Brothers, E. B. 1990. Otolith marking. Pages 183–202 *in* N. C. Parker, A. E. Giorgi, R. C. Heidinger, D. B. Jester, Jr., E. D. Prince, and G. A. Winans, editors. Fish-marking techniques. American Fisheries Society, Symposium 7, Bethesda, Maryland.

Buckmeier, D. L., E. R. Irwin, R. K. Betsill, and J. A. Prentice. 2002. Validity of otoliths and pectoral spines for estimating ages of channel catfish. North American Journal of Fisheries Management 22:934–942.

Busacker, G. P., I. A. Adelman, and E. M. Goolish. 1990. Growth. Pages 363–387 *in* C. B. Schreck and P. B. Moyle. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Campana, S. E. 1990. How reliable are growth back-calculations based on otoliths? Canadian Journal of Fisheries and Aquatic Sciences 47:2219–2227.

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation models. Journal of Fish Biology 59:197–242.

Campana, S. E. 2005. Otolith science entering the 21st century. Marine and Freshwater Research 56:485–495.

Campana, S. E., L. J. Natanson, and S. Myklevoll. 2002. Bomb dating and age determination of large pelagic sharks. Canadian Journal of Fisheries and Aquatic Sciences 59: 450–455.

Campana, S. E., and J. D. Neilson. 1985. Microstructure of fish otoliths. Canadian Journal of Fisheries and Aquatic Sciences 42:1014–1032.

Carlander, K. D. 1987. A history of scale age and growth studies of North American freshwater fish. Pages 3–14 *in* R. C. Summerfelt and G. E. Hall, editors. Age and growth of fishes. Iowa State University Press, Ames.

Clement, J. G. 1992. Reexamination of the fine-structure of endoskeletal mineralization in chondrichthyans—implications for growth aging and calcium homeostasis. Australian Journal of Marine and Freshwater Research 43:157–181.

DeVries, D. R., and R. V. Frie. 1996. Determination of age and growth. Pages 483–512 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Fabens, A. J. 1965. Properties and fitting of the von Bertalanffy growth curve. Growth 29:265–289.

Fournier D. A., J. Hampton, and J. R. Sibert. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences 55: 2105–2116.

Francis, R. I. C. C. 1990. Back-calculation of fish length: a critical review. Journal of Fish Biology 36:883–902.

Francis, R. I. C. C. 2003. The precision of otolith radiometric ageing of fish and the effect of within-sample heterogeneity. Canadian Journal of Fisheries and Aquatic Sciences. 60:441–447.

Fraser, C. M. 1916. Growth of the spring salmon. Transactions of the Pacific Fisheries Society 1915:29–39.

Fridriksson, A. 1934. On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale. Rapports Pour-von Reunion de Conseil Internacionale Exploracion du Mer 86:1–14.

Gallagher, M., and C. P. Nolan. 1999. A novel method for the estimation of age and growth in rajids using caudal thorns. Canadian Journal of Fisheries and Aquatic Sciences 56:1590–1599.

Gallego, A., L. H. Cargill, M. R. Heath, S. J. Hay, and T. Knutsen. 1995. An assessment of the immediate effect of the Braer oil-spill on the growth of herring larvae using otolith microstructure analysis. Marine Pollution Bulletin 30:536–542.

Geffen, A. J. 1987. Methods of validating daily increment deposition in otoliths of larval fish. Pages 223–240 *in* R. C. Summerfelt and G. E. Hall, editors. Age and growth of fishes. Iowa State University Press, Ames, Iowa.

Gompertz, B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philosophical Transactions of the Royal Society of London 115:513–585.

Hettler, W. F. 1984. Marking otoliths by immersion of marine fish larvae in tetracycline. Transactions of the American Fisheries Society 113:370–373.

Isermann, D. A., and C. T. Knight. 2005. A computer program for age–length keys incorporating age assignment to individual fish. North American Journal of Fisheries Management 25:1153–1160.

Kalish, J. M. 1993. Pre- and post-bomb radiocarbon in fish otoliths. Earth and Planetary Science Letters 114:549–554.

Kalish, J. M., J. M. Johnston, D. C. Smith, A. K. Morison, and S. G. Robertson. 1997. Use of the bomb radiocarbon chronometer for age validation in the blue grenadier *Macruronus novaezelandiae*. Marine Biology 128:557–563.

Ketchen, K. S. 1949. Stratified subsampling for determining age distributions. Transactions of the American Fisheries Society 79:205–212.

King, J. R., B. J. Shuter, and A. P. Zimmerman. 1999a. Empirical links between thermal habitat, fish growth, and climate change. Transactions of the American Fisheries Society 128:656–665.

King, J. R., B. J. Shuter, and A. P. Zimmerman. 1999b. Signals of climate trends and extreme events in the thermal stratification pattern of multibasin Lake Opeongo, Ontario. Canadian Journal of Fisheries and Aquatic Sciences 56:847–852.

Klumb, R. A., M. A. Bozek, and R. V. Frie. 2001. Validation of three back-calculation models by using multiple oxytetracycline marks formed in the otoliths and scales of bluegill × green sunfish hybrids. Canadian Journal of Fisheries and Aquatic Sciences 58:352–364.

Lee, R. M. 1920. A review of the methods of age and growth determination in fishes by means of scales. Fishery Investigations, Series II, Marine Fisheries, Great Britain Ministry of Agriculture, Fisheries and Food 4(2):1–35.

Macdonald, P. D. M. 1987. Analysis of length-frequency distributions. Pages 371–384 *in* R. C. Summerfelt and G. E. Hall, editors. Age and growth of fishes. Iowa State University Press, Ames.

Macdonald, P. D. M., and T. J. Pitcher. 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. Journal of the Fisheries Research Board of Canada 36:987–1001.

McFarlane, G. A., and R. J. Beamish. 1987a. Validation of the dorsal spine method of age determination for spiny dogfish. Pages 287–300 *in* R. C. Summerfelt and G. E. Hall, editors. Age and growth of fishes. Iowa State University Press, Ames, Iowa.

McFarlane, G. A., and R. J. Beamish. 1987b. Selection of dosages of oxytetracycline for age validation studies. Canadian Journal of Fisheries and Aquatic Sciences 44:905–909.

Morales-Nin, B. 2000. Review of the growth regulation processes of otolith daily increment formation. Fisheries Research 46:53–67.

Mugiya, Y., and T. Uchimura. 1989. Otolith resorption induced by anaerobic stress in the goldfish *Carassius auratus*. Journal of Fish Biology 35:813–818.

Nash, M. K., and E. R. Irwin. 1999. Use of otoliths versus pectoral spines for aging adult flathead catfish. Pages 309–316 *in* E. R. Irwin, W. A. Hubert, C. F. Rabeni, H. L. Schramm, Jr., and T. Coon, editors. Catfish 2000: proceedings of the international ictalurid symposium. American Fisheries Society, Symposium 24, Bethesda, Maryland.

Natanson, L. J., J. J. Mello, and S. E. Campana. 2001. Validated age and growth of the porbeagle shark, *Lamna nasus,* in the western North Atlantic Ocean. U.S. National Marine Fisheries Service Fishery Bulletin 100:266–278.

Pannella, G. 1971. Fish otoliths: daily growth layers and periodical patterns. Science 173:1124–1127.

Popper, A. N., and Z. Lu. 2000. Structure–function relationships in fish otolith organs. Fisheries Research 46:15–25.

Rahn, B. A., and S. M. Perrin. 1970. Calcein blue as a fluorescent label in bone. Experimentia 26:519–520.

Reinert, T. R., J. Wallin, M. C. Griffin, M. J. Conroy, and M. J. Van den Avyle. 1998. Long-term retention and detection of oxytetracycline marks applied to hatchery-reared larval striped bass, *Morone saxatilis*. Canadian Journal of Fisheries and Aquatic Sciences 55:539–543.

Richards, F. J. 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10:290–300.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

SAS Institute. 2004. SAS/STAT 9 user's guide. SAS Institute, Cary, North Carolina.

Sire, J. Y., F. Allizard, O. Babiar, J. Bourguignon, and A. Quilhac. 1997. Scale development in zebrafish (*Danio rerio*). Journal of Anatomy 190:545–561.

Sire, J. Y., and I. Arnulf. 1990. The development of squamation in 4 teleostean fishes with a survey of the literature. Japanese Journal of Ichthyology 37:133–143.

Sneed, K. E. 1951. A method for calculating the growth of channel catfish, *Ictalurus lacustris punctatus*. Transactions of the American Fisheries Society 80:174–183.

Stevens, J. D. 1975. Vertebral rings as a means of age determination in the blue shark (*Prionace glauca L.*). Journal of the Marine Biological Association of the United Kingdom 55:657–665.

Terceiro, M., and J. L. Ross. 1993. A comparison of alternative methods for the estimation of age from length data for Atlantic coast bluefish (*Pomatomus saltatrix*). U.S. National Marine Fisheries Service Fishery Bulletin 91:534–549.

Van Oosten, J. 1929. Life history of the lake herring *Leucichthys artedi* LeSueur of Lake Huron as revealed by its scales, with a critique of the scale method. U.S. Bureau of Fisheries Bulletin 44:265–428.

Verhulst, P.-F. 1838. Notice sur la loi que la population suit dans son accroissement. Correspondances Mathématiques et Physiques 10:113–121.

Verhulst, P.-F. 1845. Recherches mathématiques sur la loi d'accroissement de la population. Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-lettres de Bruxelles, Series 2, 18:3–38.

von Bertalanffy, L. 1938. A quantitative theory of organic growth. Human Biology 10:181–213.

Walford, L. A. 1946. A new graphic method of describing the growth of animals. Biological Bulletin 90:141–147.

Ward, H. C., and E. M. Leonard. 1954. Order of appearance of scales in the black crappie, *Pomoxis nigromaculatus*. Proceedings of the Oklahoma Academy of Science 33:138–140.

Weber, D. D., and G. J. Ridgway. 1962. The deposition of tetracycline drugs in bones and scales of fish and its possible use for marking. Progressive Fish-Culturist 24:150–155.

Weber, D. D., and G. J. Ridgway. 1967. Marking Pacific salmon with tetracycline antibiotics. Journal of the Fisheries Research Board of Canada 24:849–865.

Westrheim, S. J., and W. E. Ricker. 1978. Bias in using an age–length key to estimate age-frequency distributions. Journal of the Fisheries Research Board of Canada 35:184–189.

Wilson, C. A., D. W. Beckman, and J. M. Dean. 1987. Calcein as a fluorescent marker of otoliths of larval and juvenile fish. Transactions of the American Fisheries Society 116:668–670.

Woodbury, D. 1999. Reduction of growth in otoliths of widow and yellowtail rockfish (*Sebastes entomelas* and *S. flavidus*) during the 1983 El Nino. U.S. National Marine Fisheries Service Fishery Bulletin 97:680–689.

Woydack, A., and B. Morales-Nin. 2001. Growth patterns and biological information in fossil fish otoliths. Paleobiology 27:369–378.

6 Mortality

Leandro E. Miranda and Phillip W. Bettoli

6.1 INTRODUCTION

Mortality is a concept that describes the rate at which individuals are lost from a population. This concept is central to understanding the ecology of populations, particularly their dynamics, and is essential to managing fish stocks. Each species has developed mortality patterns, with specific distribution over life stages and age-groups. High mortality is common at the egg or larval stages, largely due to abiotic conditions, but the lethal effects of abiotic conditions usually become minor when the larvae become mobile. In the early stages of external feeding, limited food may directly influence mortality. If the fish survives, limited food becomes only an indirect source of mortality by retarding growth and lengthening the time spent searching for food, which makes the fish more vulnerable to predation. Later in life, fishing may be an important source of mortality. Knowledge about the patterns and causes of mortality helps fisheries scientists understand inter- and intraspecific interactions and interactions between the population and its abiotic environment.

When studying fish populations from a consumptive outlook, mortality has traditionally been separated into natural and fishing sources. Natural mortality combines death by disease, starvation, predation, inadequate environmental conditions, and old age; most of these causes are interdependent, so the distinctions are arbitrary. Fishing mortality combines harvest and any effect directly linked to the fishing process (e.g., bycatch in commercial fishing gear or death after catch and release). Describing and estimating total, natural, and fishing mortalities is often a challenge in natural populations given sampling limitations and inability to meet fully the assumptions of most estimation procedures.

6.2 BASIC CONCEPTS

Mortality represents the number of individuals that die during a certain time interval. If, for instance, N_t individuals are present in a population at the start of an interval of length $\hat{I}_t$, and N_{t+1} survive to the end of the interval, then $(N_t - N_{t+1})/\hat{I}_t$ equals interval absolute mortality. When comparing populations over time or space, interval absolute mortality can be uninformative because population sizes may

differ. A more useful expression is obtained by representing $(N_t - N_{t+1})/\hat{I}_t$ as a fraction of N_t, and as such, interval absolute mortality becomes interval mortality rate (A; $[\{N_t - N_{t+1}\}/\hat{I}_t]/N_t$) and comparable over populations. The interval mortality rate represents the fraction of individuals present at the start of an interval that actually dies during the interval. Traditionally, A has been taken to represent a 1-year $\hat{I}_t$, but may be defined to represent any $\hat{I}_t$ time interval.

Theory and empirical observations suggest that the number of fish in a cohort does not decline linearly through a time interval. Instead, it declines approximately exponentially at a rate proportional to the number alive at any point in time (Figure 6.1). This pattern of decrease indicates that $A = ([N_t - N_{t+1}]/\hat{I}_t)/N_t$ is not constant over time because it is affected by a changing N_t. An alternative method for expressing mortality is the instantaneous mortality rate (Z; Table 6.1), which linearizes the exponential pattern of A through a logarithmic transformation; thus, $(\log_e N_t - \log_e N_{t+1})/\hat{I}_t = Z$ for any $\hat{I}_t$. We note that as $\hat{I}_t$ approaches zero, Z and A converge because Z represents the death rate at an instant, whereas A represents the death rate at the end of an interval; as the interval becomes small and its width approaches zero, an instant and an interval become indistinguishable.

For example, if by the end of a 1-year interval $A_{\text{annual}} = 0.80$ and, thus, $Z_{\text{annual}} = -\log_e(1 - 0.80) = 1.61$, then the instantaneous monthly mortality rate $Z_{\text{month}} = 1.61/12 = 0.134$ and interval monthly mortality rate $A_{\text{month}} = 1 - e^{-Z_{\text{month}}} = 0.125$. Similarly,

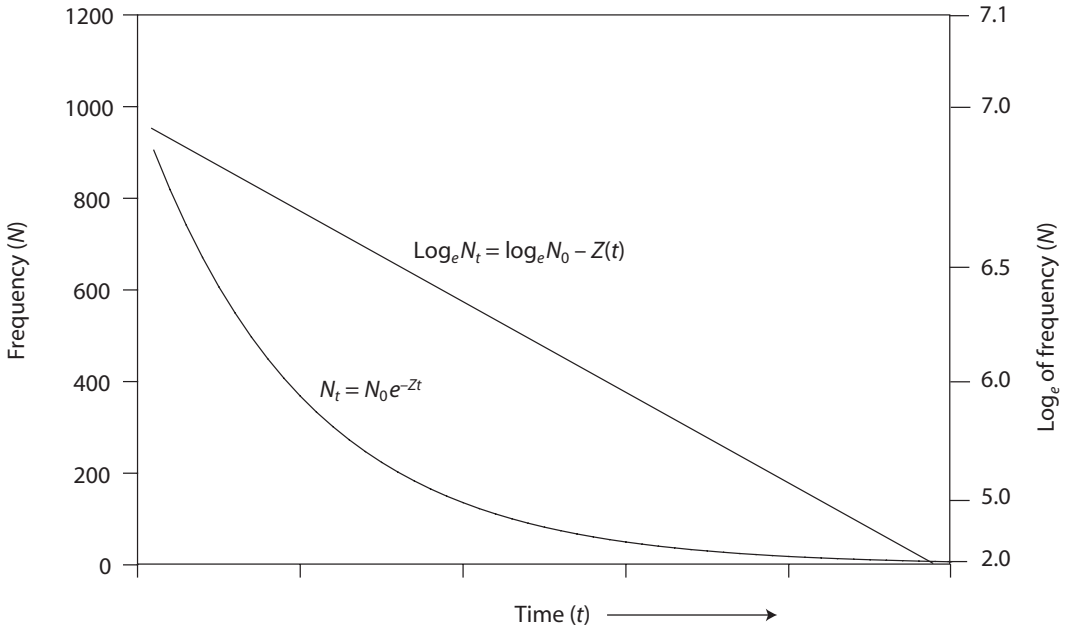


Figure 6.1 Catch curves are based on the assumption that a cohort will decline in frequency at a rate that is proportional to the abundance of the cohort at each instant in time. A $\log_e$ transformation of frequency (N_t) changes an exponential curve into a straight line, which can be described using least-squares regression. The slope of the regression line, Z, represents the instantaneous mortality rate; the intercept (N_0) represents the estimated density at time zero.

Table 6.1 Parameters descriptive of mortality rates and relations among parameters. Symbols are as follows: interval mortality rate (A); instantaneous mortality rate (Z); interval fishing mortality (μ); interval natural mortality (v); instantaneous natural mortality (M); instantaneous fishing mortality (F); conditional natural mortality (n); and conditional fishing mortality (m).

Mortality rates	Total	Fishing	Natural
Interval	$A = \mu + \nu = 1 - e^{-Z}$	$\mu = FA/Z = \nu F/M$	$\nu = MA/Z = \mu M/F$
Instantaneous	$Z = F + M = -\log_e(1 - A)$	$F = \mu Z/A = \mu M/\nu$	$M = \nu Z/A = \nu F/\mu$
Conditional interval	$A = m + n - mn$	$m = 1 - e^{-F}$	$n = 1 - e^{-M}$

$Z_{week} = 1.61/52 = 0.0310$ and $A_{week} = 0.0305$; $Z_{day} = 1.61/365 = 0.00441$ and $A_{day} = 0.00442$. Note that A and Z become alike as interval width decreased from a year to a day. Also, note that the additive property of instantaneous rates allows flexibility to interpolate or predict mortality for intervals other than the ones estimated and to estimate the number of individuals surviving to any point in time (Box 6.1).

Box 6.1 Basic Mortality Computations

Take for example a hypothetical fish population consisting of a single age-group. At the start of a 12-month interval, the age-group consists of 1,000 individuals, and at the end it has been reduced by mortality to 700. For this example,

interval absolute mortality = $N_0 - N_{12} = 1{,}000 - 700 = 300$;
interval mortality rate = $(N_0 - N_{12})/N_0 = (1{,}000 - 700)/1{,}000 = 0.300$; and
instantaneous mortality rate = $Z_{12} = -\log_e(1 - [N_0 - N_{12}]/N_0) = -\log_e(1 - [1{,}000 - 700]/1{,}000) = 0.357$.

Now, suppose we wish to know the fraction of the population remaining, the number of individuals, and the number of deaths at the end of 4 and 8 months intervals. For this, Z_{12} must be partitioned into 4-month (Z_4) and 8-month (Z_8) estimates as

$Z_4 = 4(Z_{12}/12) = 0.119$, and
$Z_8 = 8(Z_{12}/12) = 0.238$.

Interval mortality rates during the 4-month (A_4) and 8-month (A_8) intervals are then calculated as

$A_4 = 1 - e^{-Z_4} = 1 - e^{-0.119} = 0.112$, and
$A_8 = 1 - e^{-Z_8} = 1 - e^{-0.238} = 0.212$.

Numbers of individuals remaining (survival) after 4 months (N_4) and 8 months (N_8) are represented by

$N_4 = N_0 - (N_0 A_4) = 1{,}000 - (1{,}000 \cdot 0.112) = 888$, and
$N_8 = N_0 - (N_0 A_8) = 1{,}000 - (1{,}000 \cdot 0.212) = 788$.

The number of deaths during each interval is therefore 112 in the first 4-months, 100 between month 4 and month 8, and 88 between month 8 and month 12.

Interval and instantaneous mortality rates are also defined for fishing and natural mortalities. The sum of interval natural (ν) and fishing (μ) mortalities adds up to A, whereas the sum of instantaneous natural (M) and fishing (F) mortalities adds up to Z (Table 6.1). Interval mortality A and instantaneous mortality Z are associated as $A = 1 - e^{-Z}$ (Table 6.1); however, $\nu = 1 - e^{-M}$ and $\mu = 1 - e^{-F}$ only when natural and fishing mortalities occur in separate intervals, which is infrequent in freshwater fisheries. When they occur in the same interval, $1 - e^{-M}$ and $1 - e^{-F}$ are also defined as n and m, respectively, and referred to as conditional interval mortality because they estimate potential deaths during the interval had it been the only acting mortality. When n and m occur simultaneously, they compete for the same fish and do not add up to A; instead, $A = m + n - mn$ (Table 6.1).

The effect of harvest on the total mortality of a population can be either additive or compensatory (Nichols et al. 1984). Additive mortality implies that an increment in fishing mortality leads to an equal increment in total mortality. Compensatory mortality implies that an increment in fishing mortality leads to a smaller or no increment in total mortality because natural mortality adjusts downwards to compensate for reduced density. Populations near carrying capacity are more likely to be regulated by density-dependent processes and display compensatory mortality. Hence, a population may exhibit additive mortality at low density and compensatory mortality at high density (section 6.8).

6.3 CATCH-CURVE MODELS

Catch curves and their use in estimating mortality rates of fish populations have a history dating back to C. G. J. Petersen in the late nineteenth century. Most fisheries scientists are familiar with classic catch curves that graphically depict the decline in the number of older fish in a sample; however, the term catch curve applies to any analysis where the change in number of fish over age-classes is considered. This section will discuss common and historical approaches to using catch-at-age data to estimate mortality rates. Catch-curve techniques require several assumptions, including constant recruitment and mortality over years and equal catchability for all ages under consideration. If recruitment is constant, and the analysis is restricted only to ages fully recruited to the gear, then observed declines in abundance of successive age-classes would be due solely to mortality. We will discuss these assumptions and how to deal with situations in which one or more assumptions are not met.

6.3.1 Relative Abundance of Consecutive Age-Classes

In this approach, a single random sample comprising several age-groups is examined. The relative abundance of fish in consecutive age-classes is used to estimate mortality rates. These methods should not be used with catch data from a single sampling season (i.e., a standing age-frequency distribution) unless annual mortality and recruitment are thought to be reasonably constant and all age-groups considered are nearly equally vulnerable to the sampling gear.

6.3.1.1 *Heincke's Method*

In the early twentieth century, fisheries biologists readily took advantage of new techniques for aging marine fishes to examine mortality rates of exploited stocks. If it was assumed that equal numbers of fish were produced each year (i.e., recruitment was constant), then the ratio of the number of fish collected from two consecutive year-classes served as an estimate of interval mortality rate,

$$A = 1 - \frac{N_{t+1}}{N_t} \cdot \tag{6.1}$$

Heincke (1913; cited in Ricker 1975) noted that old fish were less common in a random sample of a population than were young fish, and therefore more weight should be placed on the numbers of young fish when estimating mortality rates. Heincke's method calculated A and its standard error, SE_A, as

$$A = n_0/N, \text{ and} \tag{6.2}$$

$$SE_A = \sqrt{\frac{A(1-A)}{N}}, \tag{6.3}$$

where n_0 was the number of fish in the youngest age considered and N the sum of all fish considered (Box 6.2). Note that it is not important to have accurate ages of fish older than the age-group that serves to start the age series (Ricker 1975). Although this method is used infrequently, it is appropriate when old fish cannot be accurately aged, or when the circumstances prevent the sacrifice of large (likely old) fish to obtain hard bony structures for aging.

6.3.1.2 *Robson and Chapman's Method*

When the age of every fish in a large random sample is known with reasonable certainty, then a simple approach presented by Robson and Chapman (1961), and discussed by Ricker (1975) and Van Den Avyle and Hayward (1999), can be used to estimate survival rate (S) and its standard error (SE_S) as

$$S = \frac{T}{N+T-1}, \text{ and} \tag{6.4}$$

$$SE_S = \sqrt{S\left(S - \frac{T-1}{N+T-2}\right)}, \tag{6.5}$$

where N is the total number of fish fully recruited to the gear and T is derived from the distribution of vulnerable ages in the sample as shown in Box 6.3. Robson

Box 6.2 Heincke's Method of Estimating Annual Mortality

From a reservoir, a large random sample of spotted bass was collected with electrofishing gear, and fish age was determined by inspecting otoliths. The number of fish in each age-class is given below.

Age (years)	1	2	3	4	5	6	7+
Number	257	407	147	32	17	5	4

There was some disagreement over the ages of the four largest and oldest fish, but they were all at least 7 years of age, so the data were coded accordingly. The low catch of age-1 fish relative to age-2 fish suggested that age-1 fish were not fully recruited to the electrofishing gear. When the calculations were limited to age-2 and older fish, annual mortality calculated with equation (6.2) was

$$A = \frac{n_0}{N} = \frac{407}{407 + 147 + 32 + 17 + 5 + 4} = \frac{407}{612} = 67\%,$$

and its standard error was

$$SE_A = \sqrt{\frac{A(1 - A)}{N}} = \sqrt{\frac{0.67(1 - 0.67)}{612}} = 1.9\%$$

Using the same equation, the estimated annual mortality rate for age-3 and older fish was 72%. Alternatively, mortality rates between consecutive years could have been calculated using equation (6.1). For instance, annual mortality between age 2 and age 3 is

$$A_{2-3} = 1 - \frac{N_3}{N_2} = 1 - \frac{147}{407} = 64\%,$$

and between age 3 and age 4 is

$$A_{3-4} = 1 - \frac{N_4}{N_3} = 1 - \frac{32}{147} = 78\%.$$

Both of these approaches are very sensitive to violations of the assumption of constant recruitment. If recruitment is known to vary widely, other mortality estimation techniques should be considered.

and Chapman's method is a discrete-time model (Jensen 1985) that estimates interval survival using maximum-likelihood estimation. The assumptions regarding constant survival, constant recruitment, and equal vulnerability also apply to this method. In fact, Robson and Chapman (1961) stated that the age-frequency distribution from a single sample provides no insight whatsoever into the force of mortality acting on the population unless it can be stated that recruitment and mortality do not vary among years and among ages. Robson and Chapman (1961)

Box 6.3 Robson and Chapman's Maximum-Likelihood Estimate of Survival

Assume that all the fish in a large sample were aged and the numbers of fish in each age-class were tallied, as below. Along with constant (or near constant) recruitment and survival rates, the assumption of equal vulnerability to capture must be met. A cursory examination of the catch-at-age data suggests that the two youngest age-groups were not fully vulnerable, or recruited, to the gear (i.e., the curve does not truly begin to descend until age 3); therefore, the analysis will apply to only age-3 and older fish. The first step is to code each age, starting with zero for the youngest age considered fully recruited.

Age (years)	1	2	3	4	5	6	7	8	9	10
Catch (N_x)	90	164	162	110	55	41	20	14	7	5
Coded age (x)	–	–	0	1	2	3	4	5	6	7

An unbiased estimator of the annual survival rate (S) is

$$S = \frac{T}{N + T - 1},$$

where N is the total number of fish fully recruited to the gear ($N = 162 + 110 + 55 + 41 + 20 + 14 + 7 + 5 = 414$), and T is derived from the distribution of vulnerable ages in the sample, $T = \Sigma\,(xN_x) = 0(162) + 1(110) + \ldots + 7(5) = 570$. Therefore,

$$S = \frac{570}{414 + 570 - 1} = 0.580.$$

The precision of this survival rate estimate is assessed by calculating its standard error, SE_S, as

$$SE_S = \sqrt{S\left(S - \frac{T-1}{N+T-2}\right)} = \sqrt{0.580\left(0.580 - \frac{570-1}{414+570-2}\right)} = 0.018.$$

Note that the precision of this survival estimate is a function of the number of fish examined. Approximate 95% confidence intervals (CI) on the survival rate estimate are

$$CI_{0.95} = S \pm 1.96(SE_S) = S \pm 1.96(0.018) = 0.58 \pm 0.035.$$

Annual mortality rate (A) is $1 - S$, or 0.42 ± 0.035.

provided two useful alternatives to equations (6.4) and (6.5). First, they modified equations (6.4) and (6.5) to allow for estimation when only some of the youngest age-groups are aged and the remaining age-groups are pooled; this procedure sacrifices potential information available from the sample but may enhance accuracy and precision of predictions. Second, they provided a modified equation to apply to catch curves derived using age-length keys: when age-length keys (see Chapter 5) are applied (i.e., the fish that are aged represent subsamples from

fixed length-groups), additional variation is introduced into the survival estimate, and equation (6.5) must be modified to calculate the variance.

6.3.2 Linearized Catch Curves

If fish density declines at a rate proportional to the number of fish present at each point in time, density will decline exponentially (Figure 6.1). Most fish populations exhibit this decline, and this characteristic allows estimation of instantaneous and interval mortality rates. If the $\log_e$ of frequency is plotted in relation to time, the slope of a line fit to those observations will be the instantaneous mortality rate (Z). The instantaneous mortality rate can assume values varying from 0 to slightly over 4, which correspond to interval mortality rates between zero and nearly 100%.

If all fish in a large random sample are aged, and natural logarithms of the catch at each age are taken, the slope of a regression line fit to the descending right limb of the catch curve should represent Z (Figure 6.2). Such plots are widely used by fisheries scientists to estimate mortality rates. Although the mathematics involved in estimating the slope of the catch curve are clear-cut (Box 6.4), there are a number of concerns or assumptions that need to be addressed when using

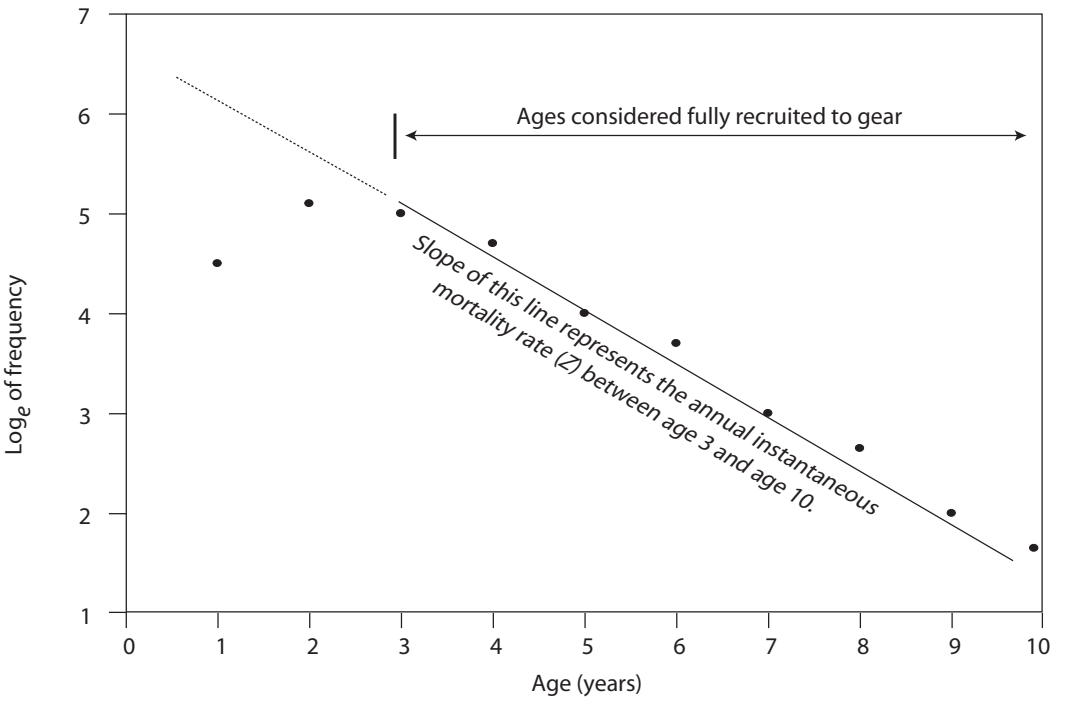


Figure 6.2 Hypothetical catch curve for a freshwater fish population sampled with electrofishing gear. The dashed line represents the expected catch of fish at age 1 and age 2. Catch-curve analysis would be limited to the descending right-hand portion of the curve between ages 3 and 10. See Box 6.4 for further explanation.

catch curves. Partial recruitment of the youngest age-classes to the gear is common, and estimates of mortality must be restricted to those ages considered "fully recruited" to the gear. In Figure 6.2, the catch curve has an ascending left-limb corresponding to lower than expected catches of age-1 and age-2 fish. The low catches of young fish could be due to the bias of electrofishing gear toward larger fish (i.e., only the largest age-1 and age-2 fish were vulnerable to capture). Perhaps young fish were less likely to occupy the shoreline habitat sampled with electrofishing gear. Similar vulnerability issues are also possible when using other gears, such as gill nets, trap nets, and trawls. Alternatively, the ascending left limb of the catch curve could have been caused by the production of weak year-classes in the 2 years preceding the sample. Without repeated sampling in subsequent years, it is impossible to determine which explanation (gear bias or poor recruitment) is most feasible.

Constant recruitment is the exception rather than the rule in many fish populations; however, moderate and random variations in recruitment will not change the general form of a catch curve, and mortality rates can still be estimated (Ricker 1975). In practice, reasonable estimates of annual mortality can usually be derived from catch curves for species such as crappies that often exhibit erratic recruitment (Allen 1999). A common scenario is that depicted by the catch curve in Figure 6.3A, which shows recruitment of largemouth bass varying erratically among years. In these situations, successive years of data can be pooled (Figure 6.3B), and the influence of erratic recruitment can be dampened. Data are combined if it can be assumed that the population is in a state of equilibrium except for random variations in recruitment (Ricker 1975). Pooling several years of data may also resolve the problem of small sample size, particularly for the oldest age-classes. Extreme variation in catch-curve mortality estimates is possible when the few representatives of the oldest age groups are included (Van Den Avyle and Hayward 1999), and it is customary to truncate the analysis at the oldest age-group with at least five representatives. In Figure 6.3A, only 102 fish were collected in the 1992 sample, and the frequency of fish in the oldest age-class (age 7) was less than 5. Pooling data from two consecutive years (Figure 6.3B) reduced the scatter of points around the catch curve and allowed estimation of annual mortality out to age 7. Alternatively, the information provided by each age-group may be weighted according to their representation in the sample (section 6.3.4).

Modest fluctuations in recruitment are acceptable when constructing catch curves, if the fluctuations are random in nature and not serially correlated over time. However, steadily decreasing or increasing recruitment can confound catch-curve analyses. For instance, the introduction of a forage fish to boost prey abundance for piscivores may have the unintended consequence of reducing recruitment of those same piscivores (Johnson and Goettl 1999). Similarly, the phenomenon of reservoir aging may cause long-term shifts in community composition and thereby recruitment (Agostinho et al. 1999). In a population experiencing steadily declining recruitment, a catch curve constructed from a single random sample will underestimate annual mortality. Conversely, steadily increasing recruitment would cause overestimation of annual mortality. Systematic changes

Box 6.4 Mortality Rates from the Slope of Regression Line

The catch-at-age data shown in Box 6.3 and Figure 6.2 are repeated here, along with the natural logarithms of the number at each age.

Age (years)	1	2	3	4	5	6	7	8	9	10
Number	90	164	162	110	55	41	20	14	7	5
Log_enumber	4.50	5.10	5.09	4.70	4.01	3.71	3.00	2.64	1.95	1.61

The catch-curve analysis is limited to those ages considered fully recruited to the gear (age 3 and older). At least five fish in the oldest age-class are present, so the mortality rate will apply to ages 3–10. Using least-squares regression, the slope of the line describing the relation between $\log_e$ of number (*y*-variable) and age (*x*-variable) can be calculated longhand, by means of a spreadsheet, or with the following SAS program:

```
Data A;
Input Age Catch @@;
If Age < 3 then delete;
LogN = Log (catch);
Cards;
1 90  2 164  3 162  4 110  5 55  6 41  7 20  8 14  9 7  10 5;
Proc Reg Data = A; Model LogN = Age;
Run;
```

The SAS output consists of an analysis of variance (ANOVA) table and estimates of the slope, as follows.

Table Regression procedure ($\log_e$ unweighted data) for catch-at-age data. Abbreviations are as follows: mean square error (MSE) and coefficient of variation

$$\text{CV} = 100\left(\frac{\sqrt{MSE}}{\overline{X}}\right), \text{ where } \sqrt{MSE} = \text{Root MSE in SAS output.}$$

Analysis of Variance

Source	*df*	Sum of squares	Mean square	*F*-value	$P > F$
Model	1	10.97660	10.97660	1072.55	< 0.0001
Error	6	0.06140	0.01023		
Corrected total	7	11.03801			
R^2	0.9944	Root MSE	0.10116		
Adjusted R^2	0.9935	Dependent mean	3.33739		
CV	3.03123				

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	$P > \lvert t \rvert$
Intercept	1	6.66033	0.10758	61.91	<0.0001
Age	1	–0.51122	0.01561	–32.75	<0.0001

The slope of the line (–0.51122) is listed under the heading "Parameter Estimates" for the variable "age." The slope of the line represents the instantaneous annual mortality rate, *Z*. The antilog (e^{-Z}) of the instantaneous mortality rate is the annual survival rate (*S*) or 60%, and mortality (*A*) is 1 – *S*, or 40%. The standard error of the slope (SE_z), obtained from the SAS program or equation (6.6), was 0.01561. Thus, the 95% CIs for *Z* are

$$CI_{0.95} \text{ of } Z = Z \pm t_{(0.05,\, n-2)}SE_z .$$

Eight ages were used in the catch curve; therefore, there are 8 – 2 df. Thus,

$CI_{0.95}$ of $Z = 0.511 \pm 2.447(0.0156) = 0.473$ and 0.549, and
$CI_{0.95}$ of $A = 1 - e^{-0.473} = 0.377$ and $1 - e^{-0.549} = 0.422$.

These results suggest that one would be 95% confident that the true mortality rate was between 38% and 42%, and the best estimate would be 40%. Note that the precision of this estimate is a function of the number of age-groups present. When these data were analyzed using the Chapman–Robson method (Box 6.3), the estimated annual mortality rate was slightly higher (42%).

The following SAS program performs weighted regression analysis on the above data, deflating the importance of older, rare fish in the sample. In this example, each observation is weighted by the predicted number of fish in each age-class as suggested by Maceina and Bettoli (1998). The first regression procedure calculates predicted values of $\log_e$(catch) for each age and outputs them to a second data set, where they are used as weights in the second regression procedure.

```
Data A;
Input Age Catch @@;
If Age < 3 then delete;
LogN = Log (catch);
Cards;
1 90  2 164  3 162  4 110  5 55  6 41  7 20  8 14  9 7  10 5
;
Proc Reg Data = A;  Model LogN = Age; Output out = B   Predicted = W;
Proc Reg Data = B;  Model  LogN  =  Age; Weight W;
Run;
```

The SAS output for the first regression procedure is the same as above; the ANOVA results and slope estimate for the weighted regression procedure are as follows:

Table Regression procedure ($\log_e$ weighted data) for catch-at-age data.

Analysis of Variance					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	1	32.14136	32.14136	968.34	<0.0001
Error	6	0.19915	0.03319		
Corrected total	7	32.34051			
R^2	0.9938	Root MSE	0.18219		
Adjusted R^2	0.9928	Dependent Mean	3.74851		
CV	4.86026				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	6.66128	0.10002	66.60	<0.0001
Age	1	–0.51139	0.01643	–31.12	<0.0001

In this example, the slope of the weighted regression (–0.51139) is almost identical to the slope for the unweighted regression line (–0.51122), although that is not always the case. It is usually desirable to estimate mortality rates over the greatest number of age-classes, but the oldest ages are often represented by fewer than five individuals. Use of weighted regression may allow relaxing the "requirement" that the oldest age-class should always be represented by at least five individuals because the influence of the oldest age-classes on the regression line will be reduced.

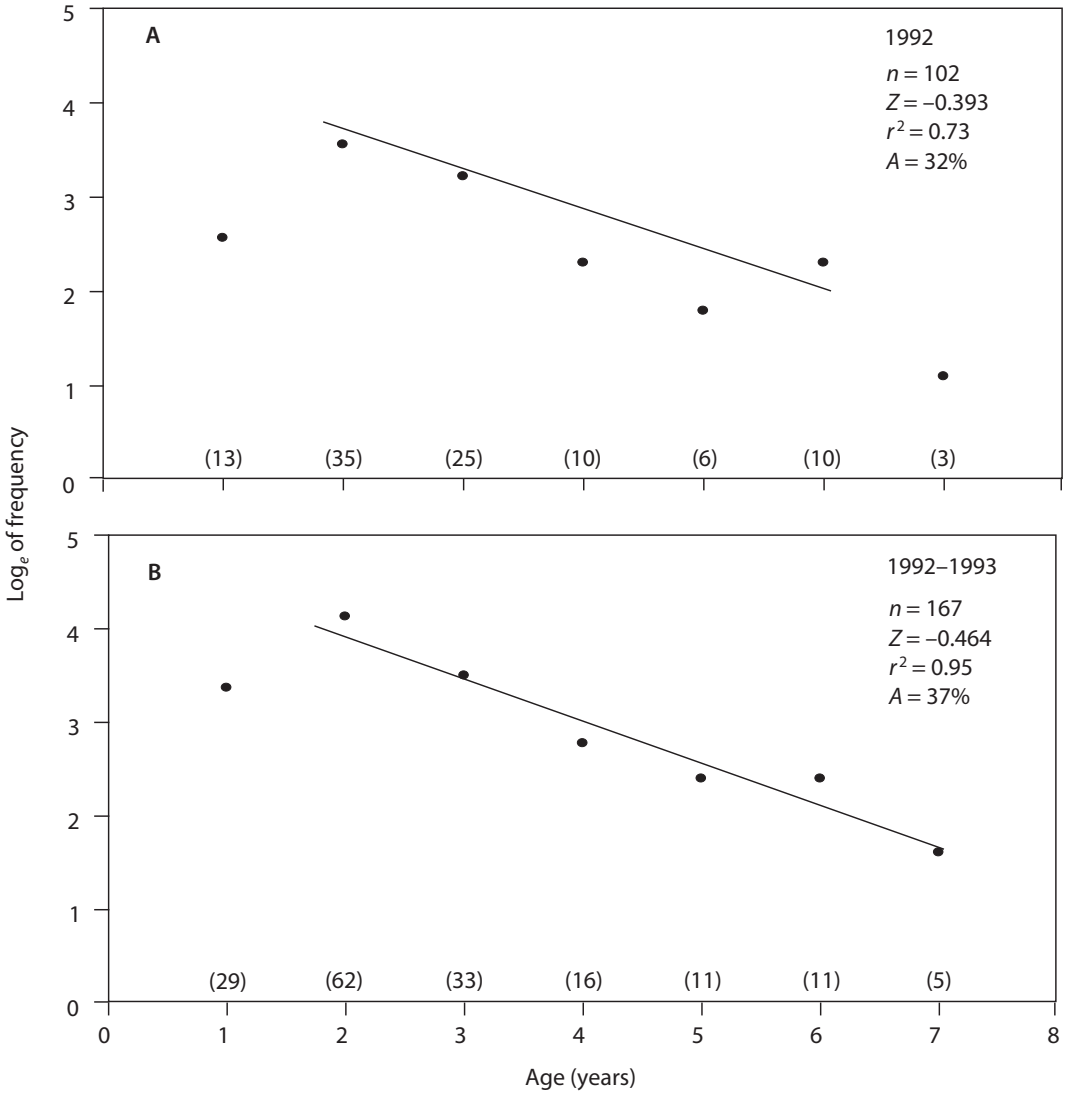


Figure 6.3 Catch curves for largemouth bass collected in (A) one year and (B) two consecutive years in Normandy Reservoir, Tennessee. Numbers in parentheses are the catch at each age. Pooling catch-at-age data over two consecutive years did not appreciably change the estimate of annual mortality; however, the influence of erratic recruitment was dampened and age-7 fish could now be included in the analysis, both of which contributed to an increased r^2. The estimate of interval mortality rate is given as *A*.

(or lack thereof) in recruitment could be identified by examining historical trends in abundance of age-0 fish. For instance, many agencies rely on annual fall trapnetting to index the abundance of age-0 and age-1 crappies, and long-term databases may be readily available (Chapter 4). If trends in recruitment are detected, mortality rates could be estimated after adjusting for variable recruitment (Box 6.5) or by analysis of cohort catch curves (section 6.3.5).

Box 6.5 Adjusting Catch-at-Age Data for Unequal Recruitment

Trap-netting of white crappies in a midwestern reservoir in the spring of 2001 provided information on the standing age structure for the 281 fish collected representing six age-classes. Previous studies indicated that all ages were recruited to the trap nets, so the goal was to calculate instantaneous mortality between ages 1 and 6. The catch-at-age data clearly indicated that recruitment varied widely among years, and therefore one of the assumptions of catch-curve analysis was violated. It was assumed that the trap-net catches accurately indexed recruitment variability, and these data were used to create an index of year-class strength (I_i) for each i year as $I_i = r_i/r_L$, where r_i = number of age-0 fish collected with trap nets in year *i,* and r_L = lowest number of age-0 fish collected during the time series. The index was used to adjust the representation of each year-class N_i in spring 2001 to a constant recruitment as N_i/I_i.

Table Adjusting for non-constant recruitment.

	Year-class					
Metric	2000	1999	1998	1997	1996	1995
Unadjusted standing age distribution estimated by trap-netting in 2001 (N_i)	150	28	5	69	12	17
Number of age-0 fish collected annually with trap nets (r_i) between 1995 and 2000	1,665	556	111	2,330	445	1,220
Index of year-class strength ($I_i = r_i/r_L$)	15	5	1	21	4	11
Adjusted catch (N_i/I_i)	10	5.6	5.0	3.3	3.0	1.5

Plots of the adjusted and unadjusted data relative to age would show that the adjusted data have less scatter around the regression line. In this example, the slope of the adjusted catch-curve line was $Z = -0.34$, translating into $A = 29\%$.

Many freshwater sport fish populations are maintained or augmented by stocking age-0 fish, which can confound catch-curve analysis. In situations where consistent numbers and sizes of fish are stocked annually, and no natural reproduction occurs, the assumption of constant recruitment might be easily met. Conversely, estimating mortality rates using catch curves is confounded when rates, sizes, and frequency of stockings vary, as depicted in Figure 6.4; in such instances, other estimation approaches should be investigated (Box 6.5 and section 6.3.5).

It should be apparent from the comments above that catch curves require fairly large samples of at least several hundred individuals, particularly for long-lived species. Accurate aging in most locales requires the use of otoliths, and if it is important to limit the number of individuals sacrificed, age-length keys can be used to estimate the number of fish at each age from subsampled data (Bettoli and Miranda 2001).

Biases in catch-curve mortality estimates due to unequal recruitment can often be identified and sometimes rectified (e.g., by pooling several years of data). However, variation in mortality rates among age-classes may be difficult to detect and hard to remedy. Ricker (1975) described different shapes, or functional forms,

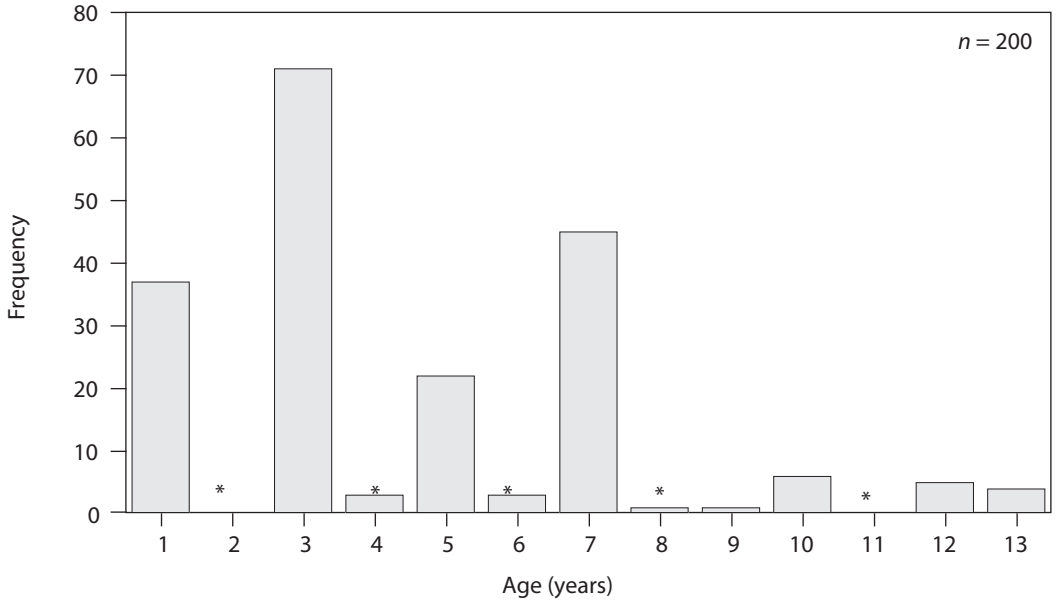


Figure 6.4 Frequency-at-age data for a standing age distribution of walleyes in Dale Hollow Reservoir, Tennessee–Kentucky, December 1999 (Vandergoot and Bettoli 2001). Asterisks denote years when no walleyes were stocked. In years when walleyes were stocked the number of fish stocked varied, which confounds the use of catch curves to estimate annual mortality.

of catch curves derived from empirical data and how forces of natural mortality and fishing mortality may shape the curves. In situations where larger (older) fish are exploited at higher rates than are small fish, the possibility exists that higher rates of fishing mortality are compensated by falling rates of natural mortality (Allen et al. 1998), resulting in no substantial change in total mortality rates over all ages fully recruited to the gear.

When natural mortality is constant, a catch curve for a heavily exploited population with a minimum-length-limit harvest regulation in effect might increase in slope beyond some age due to intense exploitation past the length limit (Figure 6.5). Such biases, caused by violation of the constant-mortality assumption, can be reduced if the catch curve is split into the unexploited and exploited segments and analysis applied to each segment independently. For instance, if a walleye population is being fished under a 40-cm length limit, the catch-curve analysis could be applied separately to those age-classes smaller than 40 cm (Z would represent M if catch–release mortality and illegal harvest were low), and those larger than 40 cm (Z would represent $M + F$ if it was assumed that all legal-sized fish were exploited equally).

6.3.3 Precision of Catch-Curve Mortality Estimates

The precision of the instantaneous mortality rate Z derived from regression of abundance as a function of age is assessed by calculating its variance (S_z^2), which is the variance of the slope of the regression line (Neter et al. 1990):

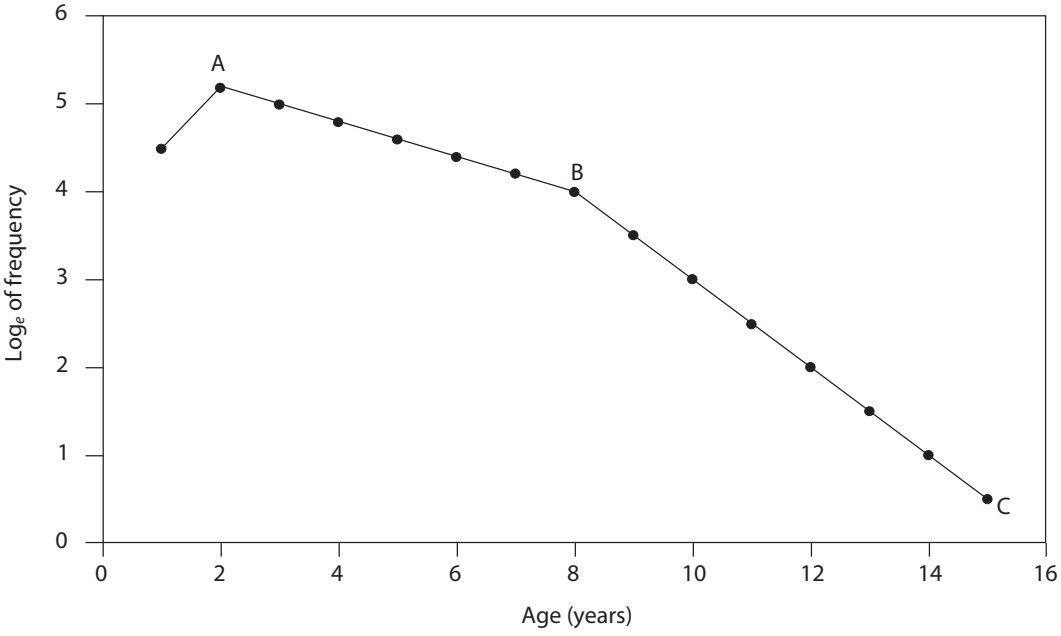


Figure 6.5 Catch curve for a hypothetical population that experiences constant recruitment and constant natural mortality but an increase in fishing mortality past age 8. Separate estimates of Z should be calculated for fish age 2 through age 8 (line A-B) and for fish age 8 through age 15 (line B-C).

$$S_z^2 = \frac{\text{MSE}}{\Sigma X_i^2 - ([\Sigma X_i]^2/n)}, \tag{6.6}$$

where MSE is the mean square error term from the regression model, X_i are the ages used in constructing the catch curve, and n the number of i ages included in regression. The square root of S_z^2 represents the standard error of the slope (S_z), and 95% confidence limits would be

$$\text{CI}_{0.95} \text{ of } Z = Z \pm t_{(0.05,\, n-2)}\, S_z. \tag{6.7}$$

The proper t-value is that for a two-tailed test. The statistical software package SAS (SAS Institute 1998), as well as others, provides estimates of the standard errors of the slopes when performing regression analysis. Note that the precision of Z increases with the number of ages i included in the analysis and decreases as the scatter of points along the regression line increases. In only rare instances will the slope not be declared different from zero; such outcomes should not preclude calculation and reporting of mortality rates (Maceina and Bettoli 1998).

Testing whether two instantaneous mortality rates differ is equivalent to testing for inequality of slopes. Although the mathematics are cumbersome, the null hypothesis that the slopes are equal can be tested using an F-test generated by a SAS program (Box 6.6).

Box 6.6 Comparing Instantaneous Mortality Rates from Catch Curves

Comparing instantaneous mortality rates (*Z*) for two or more populations is equivalent to comparing the slopes of the catch-curve regression lines. Below are catch-at-age data for two populations that fully recruited to the gear at age 2.

Age (years)	1	2	3	4	5	6	7	8	9
Lake 1	433	818	243	67	48	5	30	42	22
Lake 2	305	491	155	100	30	49	16	6	

The SAS program to calculate and compare the slopes of the catch-curve regression lines is given below.

```
Data  A;
Input  Lake  Age  Catch  @@;
LnCatch  =  log(catch);
If  Age <  2  then delete;
Cards;
1 1 433  1 2 818  1 3 243  1 4 67   1 5 48  1 6 5   1 7 30  1 8 42  1 9 22
2 1 305  2 2 491  2 3 155  2 4 100  2 5 30  2 6 49  2 7 16  2 8 6
;
Proc SORT;  By Lake;
Proc  REG;  Model  LnCatch = Age; By Lake;
Proc  GLM;  Class Lake;  Model  LnCatch  =  Age Lake Age*Lake;   Run;
```

The output is given as follows:

Table Catch-curve regression ($\log_e$ catch) for Lake 1.

Analysis of Variance					
Source	*df*	Sum of squares	Mean square	*F*-value	$P > F$
Model	1	8.94246	8.94246	7.02	0.0380
Error	6	7.63908	1.27318		
Corrected total	7	16.58154			
R^2	0.5393	Root MSE	1.12835		
Adjusted R^2	0.4625	Dependent mean	4.01440		
CV	28.10766				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	$P > \|t\|$
Intercept	1	6.55225	1.03737	6.32	0.0007
Age	1	–0.46143	0.17411	–2.65	0.0380

Table Catch-curve regression ($\log_e$ catch) for Lake 2.

Analysis of Variance					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	1	12.18239	12.18239	72.74	0.0004
Error	5	0.83743	0.16749		
Corrected total	6	13.01981			
R^2	0.9357	Root MSE	0.40925		
Adjusted R^2	0.9228	Dependent mean	3.95749		
CV	10.34116				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	7.25554	0.41649	17.42	<0.0001
Age	1	–0.65961	0.07734	–8.53	0.0004

Table The general linear model (GLM) procedure for comparison of regressions ($\log_e$ catch) of lakes 1 and 2 (n = 15). Sum of squares abbreviated as SS.

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	21.13693771	7.04564590	9.14	0.0025
Error	11	8.47650657	0.77059151		
Corrected total	14	29.61344429			
R^2	0.713762	Root MSE	0.877833		
CV	22.01277	Log_ecatch mean	3.987838		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Age	1	20.08252514	20.08252514	26.06	0.0003
Lake	1	0.39457442	0.39457442	0.51	0.4892
Age*Lake	1	0.65983816	0.65983816	0.86	0.3746

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Age	1	21.11299343	21.11299343	27.40	0.0003
Lake	1	0.26295830	0.26295830	0.34	0.5709
Age*Lake	1	0.65983816	0.65983816	0.86	0.3746

The instantaneous mortality rates (i.e., slopes of the catch-curve regression lines) for lakes 1 and 2 were –0.46143 and –0.65961, respectively; thus, annual mortality rates were 37% and 48%. Direct your attention to the Type III sum of squares (SS). The null hypothesis that the two slopes were similar is tested with the *F*-value associated with the Age*Lake interaction term (F = 0.86). At 1 and 11 df, the significance of the test is P = 0.3746. Thus, we fail to reject the hypothesis that the slopes were similar.

6.3.4 Weighted Catch-Curve Analysis

Regression lines fit for catch curves give equal weight to each observation. For example, in the sample of largemouth bass in Figure 6.3B, a frequency of five age-7 fish carried as much weight when fitting the line as a frequency of 62 age-2 fish. However, it is sometimes desirable to weight each observation according to the amount of information it contains (Steel and Torrie 1980). Weighted linear regression will deflate the influence of older and rarer fish (Maceina 1997). A SAS program to perform weighted catch-curve regression is given in Box 6.4. In our largemouth bass example with the 2 years of data (Figure 6.3B), the weighted regression using the predicted $\log_e$(catch) at each age as the weighting factor yielded a slope Z of –0.479, which translates to an A of 38%, in this case similar to the unweighted estimate of 37%.

6.3.5 Cohort Catch Curves

When ancillary data suggest that recruitment, mortality rates, or both are varying enough to render standard catch-curve analysis unreliable, mortality rates can be estimated by following a year-class, or cohort, over time. Although this approach avoids the need for assuming constant recruitment, the assumption of constant mortality is still required if mortality is estimated by regressing catch at age over more than two ages or years.

All of the preceding catch-curve examples have discussed estimating annual mortality rates based on a single, large random sample that represents a standing age structure or a pooling of several annual samples that represents an average standing age structure. However, catch curves can also be constructed to estimate cohort mortality over short time frames by use of multiple samples. For instance, if a cohort of hatchery fish is marked before stocking, subsequent sampling of marked fish should reveal a decline in its abundance over time. If the catch data are $\log_e$ transformed and plotted against days poststocking, the slope of the line will represent the instantaneous daily mortality rate, which can then be expanded to estimate mortality on a weekly, monthly, or annual basis, as in the example of brown trout in Figure 6.6. Similarly, in a study of age-0 largemouth bass mortality, Timmons et al. (1980) collected fish in shoreline rotenone samples weekly through the summer and fall and fit a catch curve to the declining catch per unit effort. Weekly instantaneous mortality rate was –0.226, which translated into a weekly interval mortality rate of 20%.

Correspondingly, annual mortality for individual year-classes can be estimated by examining declining abundance in annual samples (Box 6.7). In many situations in which routine monitoring efforts provide samples of fish that are subsequently aged, long-term databases are available to perform these analyses.

6.4 LENGTH-BASED MODELS

Length-based models do not use estimates of age directly; instead they use growth parameters such as the L_∞ (asymptotic length) and K (rate at which L_∞ is approached) parameters from the von Bertalanffy or other growth models (Chapter

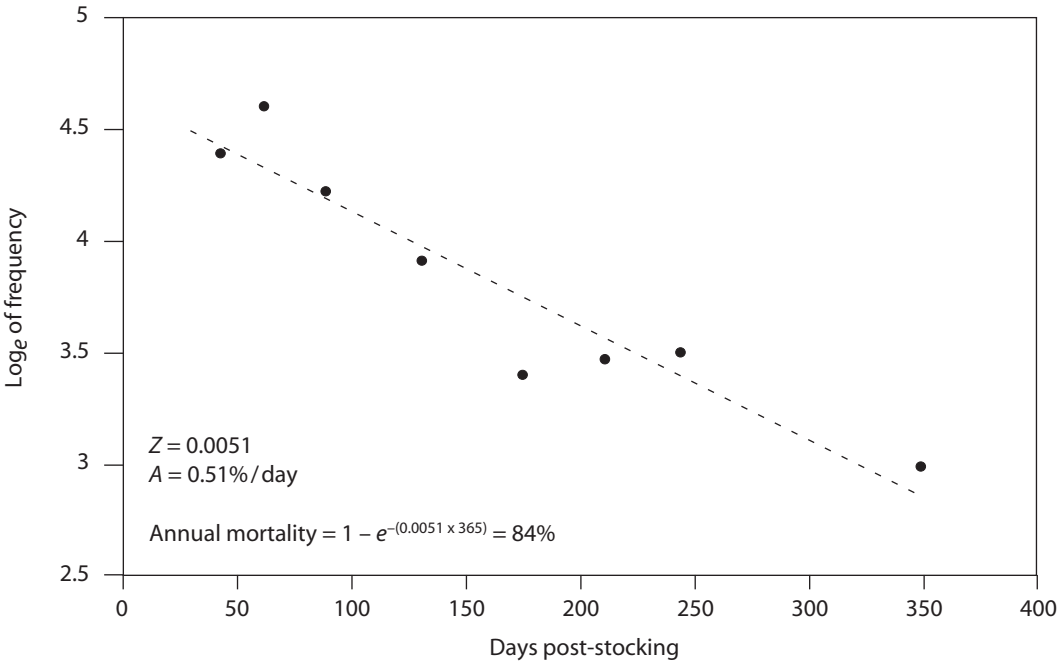


Figure 6.6 Catch curve for microtagged brown trout (N = 17,322) stocked into the Watauga River, Tennessee, March 1998, and subsequently sampled using electrofishing gear on eight dates (Bettoli 1999). Note that the daily instantaneous mortality rate (Z) and the daily interval mortality rate (A) are similar because the size of the interval is small. During a 32-week (224 d) creel survey that began when the fish were stocked, it was estimated that 4,612 of these brown trout were harvested during the survey period; thus, the interval exploitation rate (μ) was 4,612/17,322, or 27%. The total mortality rate (A) over that same interval was $1 - e^{-(Z \cdot 224\ d)} = 68\%$. Thus, the interval natural mortality rate (v) was $A - \mu = 41\%$.

5) to convert length to age. Like catch-curve models, assumptions of length-based models include (1) recruitment is constant within the period covered by the length distribution, or at least recruitment has varied in a random fashion, (2) mortality is constant over ages, (3) only lengths fully recruited to the gear are included (equivalent to the descending portion of a catch curve), (4) growth is constant and adequately described by the growth model, and (5) the sampling gear adequately represents the standing length distribution. Another assumption made by length-based models is that recruitment into the smallest length considered for analysis is constant through time each year, so that the shape of the length distribution and mean length does not vary seasonally. This last assumption is violated in populations that exhibit seasonal instead of continuous recruitment but may be avoided by taking multiple samples within the year and pooling them before analysis (Ralston 1989) or by limiting analysis to longer (i.e., older) fish for which length at age is generally more variable and recruitment spread out over a year. Given these stringent assumptions, length-based estimates should be used when only a rough approximation will do or there is no better option.

Box 6.7 Cohort Catch Curves

Spring electrofishing samples at 40 sites in Normandy Reservoir indicated that recruitment by spotted bass varied more than twofold among years (Sammons and Bettoli 1998); therefore, analysis of cohort catch curves was employed. The catch from the 1992 cohort in annual samples taken between 1993 and 1998 was as follows.

Sample year	1993	1994	1995	1996	1997	1998
Age (years)	1	2	3	4	5	6
Number	65	66	27	6	4	1

These and other data suggested that fish were not fully recruited to the gear until age 2. The Chapman–Robson estimator (Box 6.3) was used to estimate annual survival, S, and ages 2 through 6 were assigned coded ages of 0 to 4. Thus, $N = 104$; $T = 55$;

$$S = \frac{T}{N + T - 1} = 0.35,$$

$SE_S = 0.046$, and
$CI_{0.95} = 0.35 \pm 1.96(SE_S) = 0.35 \pm 0.090$.

Alternatively, the slope of the catch curve could be calculated to estimate Z. The low catch of age-6 fish restricted the analysis to ages 2 through 5. Natural logarithms were taken of the catch data, and a regression line was fit to the points, yielding a slope of –0.99 and a SE of 0.14. Thus, annual survival for ages 2 through 5 was $S = e^{-0.99} = 0.37$. Although this estimate was similar to the 35% Chapman–Robson estimate, the CIs (calculated with equation [6.7]) were broad (20–68%) because of the small number of age-classes in the regression model. In this example, the Chapman–Robson estimate and variance are clearly superior to the regression estimate.

6.4.1. Estimates from Average Length

The rationale behind these methods is that as mortality increases, the average length of fish in a population is expected to decrease. Various models have been developed to convey this relation (reviewed by Hoenig et al. 1983), but the most common is that attributed to Beverton and Holt (1956):

$$Z = K \frac{L_\infty - L_{\text{mean}}}{L_{\text{mean}} - L_x}, \tag{6.8}$$

where K and L_∞ are von Bertalanffy growth parameters, L_x the length above which all fish are equally vulnerable to capture by the collection gear, and L_{mean} the mean length of fish larger than L_x.

A similar method but based on median length instead of mean length was developed by Hoenig et al. (1983):

$$Z = \frac{0.693\text{K}}{\text{Y}_{\text{median}} - \text{Y}_{\text{x}}}, \tag{6.9}$$

where $Y_{median} = -\log_e(1 - L_{median}/L_\infty)$; $Y_x = -\log_e(1 - L_x/L_\infty)$; and L_{median} is the median length of fish above L_x. Hoenig et al. (1983) indicated that estimates based on the median length were more robust because median length is less sensitive to variability in growth and year-class strength than is mean length. Box 6.8 shows how equations (6.8) and (6.9) are applied to estimate mortality.

Approximate variances for Z in equations (6.8) and (6.9) were derived by Hoenig et al. (1983) but are not reproduced here because of their length. Alternatively, variances may be derived by bootstrapping from the expected distributions of K, L_∞, L_x, L_{mean}, and L_{median}. Bootstrapping (Efron and Tibshirami 1998; Haddon 2001) is a method for estimating variance based on resampling from the statistical distribution of each variable included in the computation of Z.

6.4.2 Estimates from Length-Frequency Distributions

When a length-frequency distribution is available a catch curve may be constructed through conversion of lengths to age relative to L_∞. Pauly (1984) developed a length-converted catch-curve procedure that consists of regressing the logarithm of the number of fish in the ith length interval (N_i, dependent variable) against the relative age t_i' of fish in the interval:

$$\log_e(N_i) = a - bt_i', \tag{6.10}$$

where $t_i' = -\log_e(1 - [L_{mid}/L_\infty])$, and L_{mid} is the midpoint of the ith length interval. The slope of this regression (b) represents $1 - (Z/K)$, and thus $Z = K(1 - b)$. An example is given in Box 6.8. A variance equation for Z has not been derived; however, the variance may be estimated by bootstrapping from the distributions of K and b.

When estimates of L_∞ and K are not available, several methods may be used to derive approximations. First, dividing the mean length of the three largest fish known from the population stock by 0.95 may adequately approximate L_∞ when the population is not too heavily exploited (Pauly 1984). Second, L_∞ may be estimated from the maximum length of fish observed (L_{max}) with an empirical equation derived by Froese and Binohlan (2000; $\log_e L_\infty = 0.044 + 0.984\log_e L_{max}$; length in centimeters). Note that L_∞ is smaller than L_{max} because L_∞ represents a population mean, whereas L_{max} represents the largest fish. Third, L_∞ may be approximated through regression of L_x (L_x = lower limit of each length interval in the length-frequency distribution) on $L_{mean} - L_x$ (L_{mean} = mean length of fish larger than L_x in the length-frequency distribution) as suggested by Wetherall et al. (1987):

$$(L_{mean} - L_x) = a - bL_x, \tag{6.11}$$

where $L_\infty = -a/b$. Once an estimate of L_∞ is obtained by one or more of these methods, K may be estimated by rearranging the growth equation, and if an estimate of length at time t (L_t) is available and t_0 is assumed equal to zero,

$$K = \frac{-\log_e(1 - [L_t/L_\infty])}{t}. \tag{6.12}$$

Box 6.8 Mortality Estimation with Length-Based Models

We use a largemouth bass data set from Columbus Lake, Mississippi, to illustrate mortality computations from length-based models. The von Bertalanffy model parameters (K and L_∞) were available from a parallel study in Columbus Lake and were $K = 0.226$ and $L_\infty = 636$ mm (see Chapter 5 for calculations). All largemouth bass 150 mm or longer were considered equally vulnerable to the collection gear (electrofishing); thus, $L_x = 150$. The mean and median length of fish 150 mm or longer in the data set were $L_{mean} = 260$ and $L_{median} = 255$. Therefore, based on equation (6.8),

$$Z = K(L_\infty - L_{mean})/(L_{mean} - L_x)$$
$$= 0.226(636 - 260)/(260 - 150) = 0.773.$$

Based on equation (6.9),

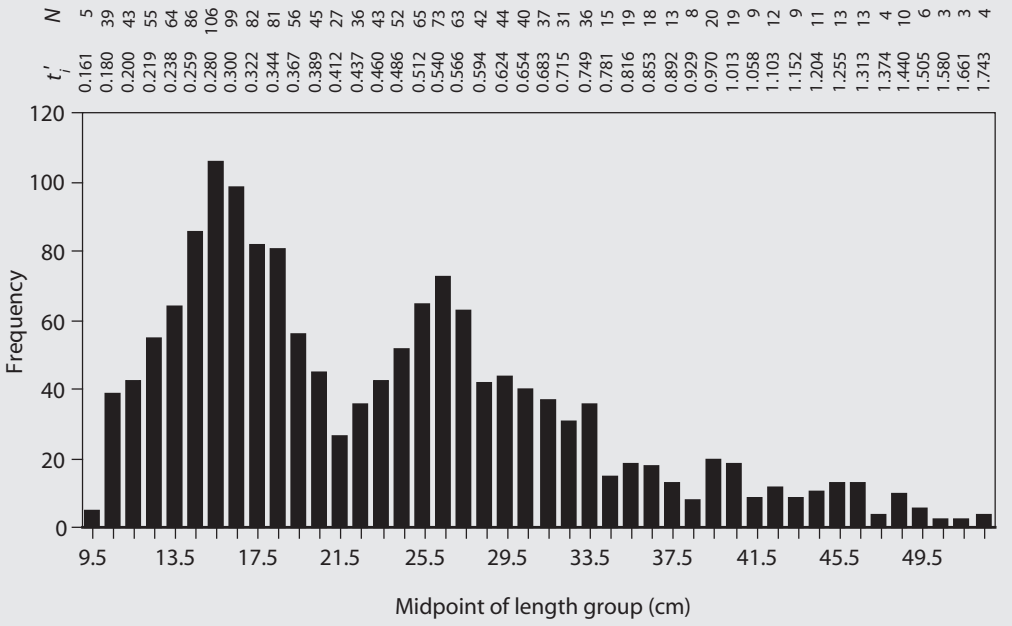
$$Z = 0.693K/([-\log_e\{1 - L_{median}/L_\infty\}] - [-\log_e\{1 - L_x/L_\infty\}])$$
$$= 0.693(0.226)/[-\log_e(1 - 255/636) - (-\log_e(1 - 150/636)] = 0.644.$$

Mortality can also be estimated using a length-converted catch curve (Pauly 1984). The length-groups in the length–frequency distribution of the largemouth bass population in Columbus Lake (see figure in this box) are converted to relative age t_i' and regressed on $\log_e$ of the number of fish in the ith length interval as in equation (6.10):

$$\log_e(N_i) = a - bt_i'$$
$$= 4.97 - 2.27t_i'.$$

Then, $Z = K(1 - b) = 0.226[1 - (-2.27)] = 0.739$. Regression was limited to length-groups 15 cm and greater.

Figure Length–frequency distribution of the largemouth bass population in Columbus Lake. The number of fish in each length-group is represented by N, and t_i' is as defined in equation (6.10).



6.5 MARK–RECAPTURE MODELS

Mortality can be measured directly by marking individual fish. Historically, mark–recapture models have been developed to estimate abundance, which naturally leads to methods for estimating mortality (i.e., reductions in abundance); however, more recently, the focus of mark–recapture models has shifted towards estimation of mortality (Lebreton et al. 1992). Although an extensive literature on mark–recapture models exists (see reviews by Ricker 1975; Seber 1982; Lebreton et al. 1992; Schwarz and Seber 1999), we describe only three approaches for estimating Z. Additional details about the use of mark–recapture models to estimate abundance are given in Chapter 8.

Our presentation is brief because tagging is not extensively used to assess mortality of fish populations, mostly due to the cost and the practical difficulties related to tagging a representative sample of a population and obtaining unbiased recovery data. Major commercially or recreationally exploited fish stocks are usually large and distributed over a wide area. Thus, mark–recapture estimates depend on tagging large numbers of fish and often cooperation from fishers to find and report marked fish. In the past, mark–recapture estimation has sometimes failed because too few fish have been tagged or because fishers and other members of the industry have been reluctant to report recoveries (Hilborn and Walters 1992; Miranda et al. 2002).

Mark–recapture models make many assumptions about the tagged sample and untagged population. Assumptions include (1) the tagged sample is representative of the entire population; (2) there is no tag loss or it can be accounted for (e.g., Seber 1982; Fabrizio et al. 1996); (3) mortality rates are not influenced by tagging; (4) all tagged fish within a tagged cohort have constant mortality and recovery probabilities in a given period; (5) mortality and recovery probabilities do not depend on age and are the same for the tagged sample and untagged population; (6) F and M are additive and independent; (7) M is constant within and between periods; (8) fishing mortality imposed by a user group is constant for the period of the year that the fishery is operating; and (9) tagging takes place over a short period (although there are models that account for continuous tagging; see Ricker 1975 and Seber 1982). These assumptions are not made by all models, and not all models make the same assumptions.

6.5.1 Single Tagging Event

If fish are tagged only once, mortality may be estimated from the decline of tagged individuals. Estimates of losses may be obtained by recapturing tagged fish at various time intervals or by relying on the fishery to catch and report tagged fish. The former approach is applicable if it is possible to tag a large proportion of the population, so that the expectation of collecting tagged fish in subsequent samples is reasonably high. The latter approach is applicable when there is a high likelihood that tags will be recognized and reported by commercial or recreational fishers. Both approaches assume that effort is constant, or at least known, so that catch in a given period can be standardized per unit of effort. Whichever method is used, declines in number of tagged fish can be equated to declines in number

of fish in a cohort and analyzed with the various catch-curve techniques described earlier. For example, the fraction of fish bearing a tag for any two successive periods of recaptures will indicate interval mortality rate:

$$A_i = 1 - \frac{r_{i+1}}{r_i}, \tag{6.13}$$

where r_i = number of fish recaptured during period t_i, and r_{i+1} = number of fish recaptured during period t_{i+1} (Box 6.9). When recaptures are available from a series of periods, a regression of either (1) $\log_e$ of the fraction of fish caught bearing a tag or (2) $\log_e$ of the number caught per unit of effort, as a function of time, would produce a decreasing slope equivalent to *Z*:

$$\log_e(r_i/n_i) = a - bt_i, \text{ or} \tag{6.14}$$

$$\log_e(r_i/f_i) = a - bt_i, \tag{6.15}$$

where n_i = number of fish caught in t_i, f_i = fishing effort in t_i, a = regression parameter, and b = slope parameter representing *Z*. Which approach is used will depend on the data available and the assumptions that are appropriate. The assumption of constant recruitment is no longer relevant because the user is dealing with a single group of fish of known initial abundance.

6.5.2 Multiple Tagging Events

Whereas a single tagging event assumes constant survival to estimate mortality, multiple tagging events allow relaxation of this assumption. Studies based upon two tagging events followed by one recapture event (Ricker's method, Ricker 1975; Seber 1982) can account for variable mortality if recruitment is assumed constant. A triple-catch study is based upon two tagging events with recaptures collected during the second tagging event and during a third sampling event (Bailey 1951; Ricker 1975). A triple-catch study can account for variable recruitment (which includes immigration) and variable mortality (which includes emigration). Multiple mark–recapture data are best handled by a model proposed independently by both Jolly (1965) and Seber (1965) that accounts for variable recruitment and mortality. The Jolly–Seber model is more general and powerful than any of the other methods and can estimate population size and recruitment in addition to mortality using four or more mark–recapture periods; estimates are limited to sizes of fish that were tagged. Example applications for Ricker's method are given by Ricker (1975), triple-catch method by Fairfield and Mizroch (1990) and Evans and Lockwood (1994), and Jolly–Seber method by Hightower and Gilbert (1984), Law (1994), and Fabrizio et al. (1997). Below, we describe the Ricker and Jolly–Seber methods for estimating mortality.

Ricker (1975) and Seber (1982) describe similar methods for determining mortality from tagging in two successive years. With both methods, tagging occurs at the start of two periods (e.g., seasons or years) using tags that distinguish between the two tag groups. With Ricker's method, recaptures are taken during both years from

fishers. If mortality is assumed constant over years, then mortality is estimated by equation (6.13). Thus, only one marking followed by two recapture periods are needed. If mortality cannot be assumed constant over years, Ricker's method estimates mortality in period 1 as

$$A_1 = 1 - \frac{r_{12}\, m_2}{r_{22}\, m_1}, \tag{6.16}$$

with variance

$$V(A_1) = A_1^2 \left(\frac{1}{r_{12}} + \frac{1}{r_{22}} + \frac{1}{m_1} + \frac{1}{m_2} \right), \tag{6.17}$$

where m_1 = fish marked at the start of time 1, m_2 = fish marked at start of time 2, r_{22} = fish marked and recaptured in time 2, and r_{12} = fish marked in time 1 and recaptured in time 2. Seber's method also uses equation (6.16), but r_{12} and r_{22} are measured through samples taken soon after the second release. Both of these methods assume that natural mortality is constant over ages. The equations for both methods are the same because the expectation of the ratio r_{12}/r_{22} is unchanged through time 2. For Ricker's method, it is not essential that all recaptured fish be reported, only that reporting rate is constant over years. Both Seber and Ricker provide equations modified to compensate for small number of recaptures. An example application of Ricker's method is given in Box 6.9.

The Jolly–Seber method estimates mortality by evaluating changes in population size, including increases (recruitment and immigration) and decreases (deaths and emigration), from multiple mark–recapture samplings on an open population. Thus, estimates of mortality represent death only when emigration is zero. Fish are captured and marked during brief collection periods (e.g., days), and in between are longer periods (e.g., months) in which recapturing is not attempted and no tags are released. During the first collection period, fish are marked with numbered tags that distinguish individuals, and during the last period, fish are checked for marks. During intermediate periods, fish are checked for marks, unmarked individuals are tagged, and marked individuals are noted and released. Categories of marked and recaptured are tallied by collection period as shown in Table 6.2. Then, the interval mortality rate between collection period i and collection period $i + 1$ is estimated as

$$A_i = 1 - \frac{\beta_{i+1}}{\beta_i - r_{\cdot i} + m_i}, \tag{6.18}$$

where the number of marked fish in the population at the time of the ith sample, β_i, equals $r_{\cdot i} + m_i k_i / r_{i\cdot}$, and $r_{\cdot i}$, m_i, k_i, and $r_{i\cdot}$ are as defined in Table 6.2. Seber (1965) proposed a modified estimator of β_i, $\beta_i^* = r_{\cdot i} + 1 + (m_i + 1) k_i / (r_{i\cdot} + 1)$ for a small number of recaptures. Variance equations are given by Seber (1982) and by programs listed in Table 6.3. We illustrate application of the Jolly–Seber method in Box 6.9.

Box 6.9 Total Mortality Estimation from Marked Recaptures

Single tagging event

In late winter 1995, before intense fishing began, 1,596 crappies were tagged in Sardis Reservoir, Mississippi. Of these, 655 were recaptured and reported by anglers during the first year after tagging, 225 in year 2, 89 in year 3, and 34 in year 4 (in this example, recaptures have been preadjusted for tag loss and nonreporting; Miranda et al. 2002). Following equation (6.13), where

r_i = number of fish recaptured during period t_i,
$A_i = 1 - (r_{i+1}/r_i)$,
$A_1 = 1 - (225/655) = 0.66$,
$A_2 = 1 - (89/225) = 0.60$, and
$A_3 = 1 - (34/89) = 0.62$.

Alternatively, regression of $\log_e(r_i/f_i)$ as a function of t_i (equation [6.15] assuming constant $f_i = 1$, where f_i = fishing effort in t_i,) yields

$$\log_e(r_i/f_i) = a - bt_i$$
$$= 7.4 - 0.98t_i,$$

which indicates $Z = 0.98$ and thus $A = 0.62$.

Multiple tagging events

Crappies were tagged at the beginning of two consecutive years in Lake Sham. In all, 1,700 crappies were marked in year 1 (m_1) and 1,500 in year 2 (m_2). In year 1, 430 crappies were recaptured (r_{11}); in year 2, 360 of the crappies tagged earlier that same year were recaptured (r_{22}), and 249 tagged the previous year (r_{12}). If annual mortality can be assumed constant, then A for year 1 (and year 2) may be estimated with equation (6.13) as $A_1 = 1 - (249/430) = 0.42$. However, if mortality is suspected to vary over years, then A_1 and $V(A_1)$ may be estimated with Ricker's method (equations [6.16] and [6.17]) as

$$A_1 = 1 - \frac{r_{12}m_2}{r_{22}m_1} = 1 - \frac{249 \times 1{,}500}{360 \times 1{,}700} = 0.39,$$

$$\text{and } V(A_1) = A_1^2\left(\frac{1}{r_{12}} + \frac{1}{r_{22}} + \frac{1}{m_1} + \frac{1}{m_2}\right) = 0.39^2\left(\frac{1}{249} + \frac{1}{360} + \frac{1}{1{,}700} + \frac{1}{1{,}500}\right) = 0.0012.$$

Estimation of mortality for year 2 would require a third year of marking and recaptures.

6.5.3 Other Mark–Recapture Methods

Many Jolly–Seber-type models have been developed in recent years (Buckland 1982; White 1983; Burnham et al. 1987; Cormack 1989; Lebreton et al. 1992; Pradel et al. 1997; Schwarz and Seber 1999). In particular, Pollock and Mann (1983) extended the Jolly–Seber model to enhance application to fisheries by accounting for differential mortality over age-groups. Advances in computer technology have facilitated development of these models and allowed a number of extensions such as constraining of the model parameters (e.g., fixing mortality

A 5-year tagging program was completed to monitor mortality (as well as population size and recruitment, which are not shown in this example) of largemouth bass in Lake Travesty. Fish were marked and recaptured annually during a 2-week collection period in spring each year, and results were analyzed with the Jolly–Seber method.

Table Five-year history of marking and recaptures (recaptures were preadjusted for tag loss) in a largemouth bass fishery. See Table 6.2 for explanation of symbols.

		r_{ii}					
t_i	m_i	t_{1i}	t_{2i}	t_{3i}	t_{4i}	$r_{\cdot i}$	k_i
1	643						
2	489	43				43	43
3	712	28	31			59	40
4	630	12	16	48		76	31
5		3	9	19	37	68	
$r_{i\cdot}$		86	56	67	37		

Computations of annual interval mortalities are made with equation (6.18) as shown below. As an example,

$$A_i = 1 - \frac{\beta_{2+1}}{\beta_2 - r_{\cdot 2} + m_2} = 1 - \frac{484}{418 - 43 + 489} = 0.44 .$$

Table Computations of annual interval mortalities. The number of marked fish in the population at the time of the ith sample is given by β_i .

t_i	m_i	$r_{i\cdot}$	$r_{\cdot i}$	k_i	β_i	A_i
1	643	86	0	0	0	0.35
2	489	56	43	43	418	0.44
3	712	67	59	40	484	0.47
4	630	37	76	31	604	
5			68			

between 0 and 1) and setting selected parameters constant (e.g., mortality over time). Parameters can also be modeled as functions of ordinary variables, with a regression equation built into the recapture model; thus, mortality can be made dependent on environmental conditions or capture rates dependent on measures of effort. Maximum likelihood estimation of model parameters and associated probabilities is facilitated by computer power, superseding traditional deterministic estimates. Treatment of these computer models is beyond the scope of this chapter; however, we list many in Table 6.3.

Table 6.2 Classification of marked and recaptured fish in a Jolly–Seber-type model with five mark–recapture periods. Only marking takes place in the first period, mark and recapture in periods two through four, and only recaptures in the last period. Time period is t_i; m_i represents the number of fish marked in t_i; r_{ii} the recaptures in time period t_i of fish marked at an earlier t_i ; r_i. the total number of recaptures that were originally tagged in t_i ; $r_{\cdot i}$ the total number of recaptures in t_i regardless of when they were tagged; and k_i the total number of recaptures made after t_i of fish marked before t_i. Application is illustrated in Box 6.9.

		r_{ii}					
t_i	m_i	t_{1i}	t_{2i}	t_{3i}	t_{4i}	$r_{\cdot i}$	k_i
1	m_1						
2	m_2	r_{12}				$r_{\cdot 2}$	$r_{13}+r_{14}+r_{15}$
3	m_3	r_{13}	r_{23}			$r_{\cdot 3}$	$r_{14}+r_{15}+r_{24}+r_{25}$
4	m_4	r_{14}	r_{24}	r_{34}		$r_{\cdot 4}$	$r_{15}+r_{25}+r_{35}$
5		r_{15}	r_{25}	r_{35}	r_{45}	$r_{\cdot 5}$	
r_i.		$r_{1\cdot}$	$r_{2\cdot}$	$r_{3\cdot}$	$r_{4\cdot}$		

6.6 SEPARATION OF FISHING FROM NATURAL MORTALITY

Fisheries scientists may need to know the proportionate effects of several components of mortality. Most commonly, we wish to isolate the effect of fishing from the effect of all other influences on mortality, a group often lumped together as natural mortality. It is possible to obtain estimates of fishing and natural mortality independently. Most commonly, and given $Z = F + M$ (Table 6.1), Z and F are measured and M estimated as the difference. However, this approach results in estimates of M that are not independent of F. At least six different approaches may be used to estimate M, F, or both, including (1) regression of Z as a function of fishing effort to estimate M, (2) catch-curve analysis to estimate M, (3) mark–recapture to estimate F, (4) direct census to estimate F, (5) production modeling to estimate M, and (6) meta-analysis to estimate M, F, and Z.

6.6.1 Regression of *Z* as a Function of Fishing Effort to Estimate *M*

Natural mortality *(M)* is commonly estimated as the difference between Z and F. For unfished populations or segments of populations, M equals Z and may be estimated using methods described earlier. As unfished populations are rare, other approaches must be used. Changes in fishing effort can lead to changes in Z, and the relation between fishing effort and Z can be used to achieve the separation of F and M (Paloheimo 1958). Thus, with Z as the dependent variable and fishing effort as the independent variable, the slope of the line becomes a catchability coefficient and the intercept (i.e., when effort is zero) becomes M (Figure 6.7). This method requires a minimum of two x–y pairs, but more is better.

There are at least three drawbacks for this method. First, because the independent variable (i.e., fishing effort) is estimated with considerable error, a basic

Table 6.3 Selected computer programs for analyzing mark–recapture data from multiple tagging events in open populations. Many of these programs include the ability to fit customized log-linear and constrained maximum log-likelihood models and impose arbitrary temporal, group, and covariate constraints to select the best model.

Program	Description	References
BAND2	Estimates number of animals that must be marked to achieve a specified level of precision for mortality estimates.	Wilson et al. (1989)
BROWNIE	Estimates mortality and recovery rates for two age-classes (e.g., juvenile and adult) in open populations.	Brownie et al. (1985)
CAPQUOTA	Estimates expected coefficients of variation of mortality and capture probability.	Pollock (1981)
CONTRAST	Compares estimates of mortality when variances and covariances are available (analogous to means comparisons in ANOVA).	Hines and Sauer (1989)
ESTIMATE	Estimates mortality and recovery rates for one-age-class (e.g., juvenile or adult) open populations. More flexible than BROWNIE but not as flexible as MARK.	Brownie et al. (1985)
JOLLY	Estimates mortality and capture probability for one-age-class open populations. Widely used but not as flexible as MARK.	Pollock et al. (1990)
JOLLYAGE	Similar to program JOLLY, it estimates mortality and capture probability for two-age-class open populations.	Pollock et al. (1990)
MARK	Estimates mortality and capture probability for open populations. Allows a wider class of encounter histories and constraints than do other programs and was developed primarily for mortality estimation. It will handle various mark–recapture models, the joint live-recapture and dead-recovery models, robust-design models, and multi-strata models. It is the newest and potentially most complete package.	White and Burnham (1999)
MULT	Estimates mortality and recovery rates with additional models for estimating reporting rate.	Conroy et al. (1989)
POPAN	Estimates mortality and capture probability for open populations.	Arnason and Schwarz (1999)
RELEASE	Estimates mortality and goodness-of-fit tests for a large class of mortality models for open populations. Originally developed to estimate survival for a large suite of fish mark–release experiments.	Burnham et al. (1987)
SURGE	Estimates mortality and allows easy implementation of linear models.	Pradel and Lebreton (1993); Cooch et al. (1996)
SURPH	Estimates mortality using mark–recapture data as a function of environmental and experimental effects. These effects may apply to a population (such as ambient temperature) or an individual (such as body length).	Smith et al. (1994)
SURVIV	Estimates mortality with multinomially distributed data. SURVIV is very flexible and used heavily as a research tool. However, one needs a FORTRAN compiler to run program and a healthy appetite for programming to get things to work. Not recommended for novices.	White (1992)
TMSURVIV	Estimates mortality and capture probability and the proportion of "transients" in open populations.	Pradel et al. (1997)

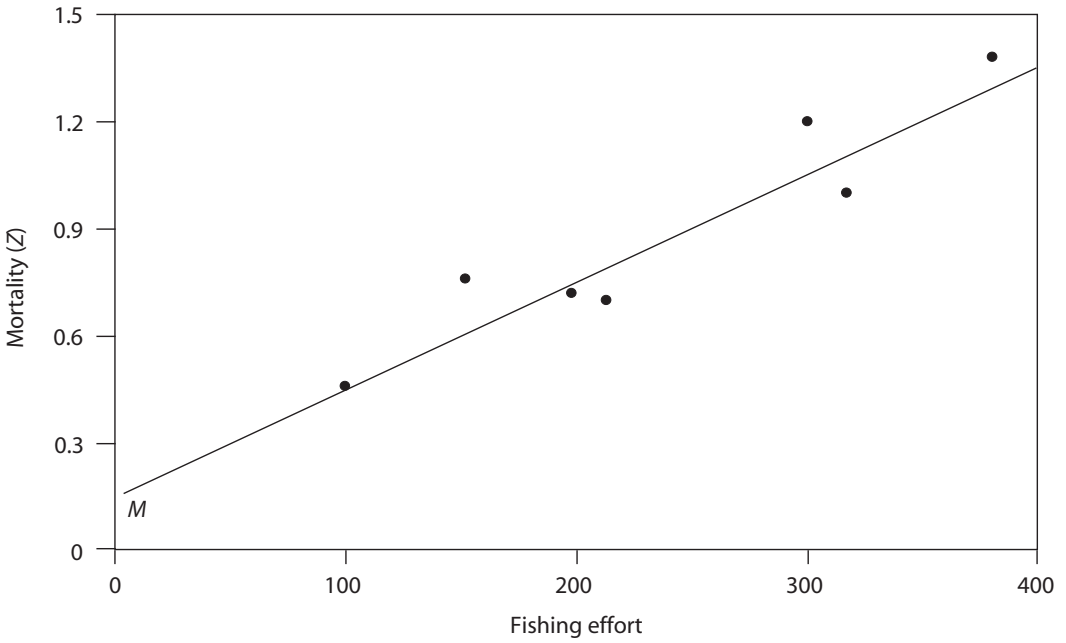


Figure 6.7 Separation of fishing *(F)* and natural *(M)* mortality by regression of *Z* as a function of fishing effort. The intercept of regression represents *M*.

assumption of regression analysis is violated. The effect is to flatten the slope because, as the measurement error in the independent variable increases, any relationship between the dependent and independent variables becomes indistinguishable, driving the slope toward zero. Flattening of the slope can produce an overestimate of the intercept and thus *M*. Second, this method is applicable only when the relation between effort and *Z* is linear (i.e., catchability is constant). Conceivably, the same fishing effort may not encounter the same catchability in different years because of changes in population density or gear efficiency. Third, an unreliable estimate of the *y*-intercept (i.e., *M*) will result if the fishing effort does not vary greatly; ideally, estimates of *Z* would be available over a wide range of fishing effort, including very low levels.

6.6.2 Catch-Curve Analysis to Estimate *M*

Under limited conditions, the linearized catch-curve analysis described in section 6.3.2 may be used to estimate *M*. Conceivably, some of the age-groups available for analysis may not be available to the fishery. The slope of a line fitted through these points may be interpreted as *M*. For instance, in situations where a length-limit regulation exists and catch-and-release mortality and illegal harvest are virtually zero, fishing mortality for protected fish is in effect zero. Hence, any estimates of *Z* will constitute estimates of *M* for fish in those protected lengths.

6.6.3 Mark–Recapture to Estimate *F* and *M*

Estimates of F can be derived from tagged fish recaptured by fishers if concurrent estimates of Z are available. Estimation procedures depend on whether one or multiple release periods are employed. Various methods are available (reviewed by Seber 1982), but we limit our presentation to methods counterpart to those identified for estimating Z in section 6.5. An additional assumption is that fishers report tagged fish; violation results in an underestimate of F. Various methods have been designed to adjust for underreporting (Zale and Bain 1994; Hearn et al. 1999), but none of the underreporting adjustments are fully satisfactory (Miranda et al. 2002).

If fish are tagged in only one marking period, fishing mortality may be estimated from the proportion of tagged individuals captured in the fishery. This approach is applicable when there is a high likelihood that tags will be recognized and reported by commercial or recreational fishers. Equation (6.13) estimated A as the fraction of fish bearing a tag in two successive periods of recaptures. If the number of fish bearing tags in the first period (m_1) and the number of tagged fish captured by fishers in this period (f_1) are known, the interval fishing mortality, μ, can be estimated as

$$\mu_1 = \frac{f_1}{m_1}, \tag{6.19}$$

and $F = \mu Z/A$. A variance equation for (6.19) was given by Ricker (1975) and Jagielo (1991). If μ is assumed constant over several recapture periods, a weighted estimate of mean exploitation is obtained as

$$\mu = \frac{f_1 + f_2 + \ldots + f_{n-1}}{m_1(1 + S_1^1 + S_2^2 + \ldots + S_{n-1}^{n-1})}, \tag{6.20}$$

where S_i is the survival in each period. These computations are illustrated in Box 6.10.

If fishing mortality cannot be assumed constant, and mark–recapture is conducted over two or more periods, estimates of μ_i for each period i can be obtained by making successive estimates with equation (6.19). If mark–recapture is continued for three or more periods, estimates of μ_i for each period i can be estimated as (Ricker 1975)

$$\mu_i = \frac{f_{i.} f_{.i}}{m_i k_i}, \tag{6.21}$$

where m_i is the number of fish marked at the start of period i, $f_{i.}$ is the number of fish marked in year i caught by fishers over all years, $f_{.i}$ is the number of marked

Box 6.10 Fishing Mortality Estimation from Marked Recaptures

Single tagging event

Consider the 1,596 crappies tagged in Sardis Reservoir and used in Box 6.9 to illustrate computation of total mortality. Recall that 655 were recaptured and reported by anglers during the first year after tagging, 225 in year 2, 89 in year 3, and 34 in year 4 (in this example, recaptures have been preadjusted for tag loss and nonreporting; Miranda et al. 2002). Following equation (6.19),

$$\mu_1 = \frac{f_1}{m_1} = \frac{655}{1{,}596} = 41\%,$$

where m_1 = number of fish bearing tags in the first period and f_1 = number of tagged fish captured by fishers in this period. If μ can be assumed constant over the 4 years of tag returns, a weighted estimate of mean exploitation can be obtained with equation (6.20) as

$$\mu = \frac{f_1 + f_2 + \ldots + f_{n-1}}{m_1(1 + S_1^1 + S_2^2 + \ldots + S_{n-1}^{n-1})} = \frac{655 + 225 + 89 + 34}{1{,}596\,(1 + 0.34 + 0.40^2 + 0.38^3)} = \frac{1{,}003}{2{,}482} = 40\%.$$

Multiple tagging events

The 5-year tagging program for largemouth bass in Lake Travesty described in Box 6.9 provided data to estimate exploitation. Fish were marked and recaptured annually during a 2-week collection period in spring each year, and anglers were asked to report tagged fish they harvested. The 4-year history of tag reports is summarized below (recaptures were preadjusted for tag loss and nonreporting).

fish caught each year i, regardless of when they were marked, and k_i is the number of marked fish caught after year i of fish marked before year i. When mark–recapture occurs over three or more periods, equation (6.21) is preferred over successive estimates with equation (6.19) because equation (6.21) incorporates more recapture information.

Some relatively new approaches integrate changes in fishing effort with traditional multiperiod mark–recapture data to estimate F and M, and possibly tag-reporting rate, from a data matrix like the one illustrated at the bottom of Box 6.10. Hoenig et al. (1998) describe two approaches, one that estimates F and M from the pattern of effort over the course of a year or other period and another that estimates them from the pattern of effort over years. Brooks et al. (1998) develop a method to separate F from M in situations where two user groups (e.g., commercial and recreational fisheries) are exploiting a fish population. Separation of M and F is made possible by differences in recapture rates and seasonal

Table Tag reports for the largemouth bass fishery in Lake Travesty. The symbol m_i represents the number of fish tagged in year i; $f_{i.}$ is the number of fish marked in year i caught by fishers over all years; $f_{.i}$ is the number of marked fish caught each year i, regardless of when they were marked; k_i the total number of tagged fish caught after year i of fish tagged before year i.

		Recaptures by anglers of fish marked in year$_i$					
Year and $f_{i.}$	m_i	1	2	3	4	$f_{.i}$	k_i
1	643	89				89	173
2	489	60	63			123	202
3	712	19	35	92		146	212
4	630	5	20	41	75	141	141
$f_{i.}$		173	118	133	75		

Computations of annual exploitation estimated with equation (6.21) are as follows.

$$\mu_1 = \frac{f_{1.}\, f_{.1}}{m_1 k_1} = \frac{173(89)}{643(173)} = 14\%,$$

$$\mu_2 = \frac{118(123)}{489(202)} = 15\%,$$

$$\mu_3 = \frac{133(146)}{712(212)} = 13\%, \text{ and}$$

$$\mu_4 = \frac{75(141)}{630(141)} = 12\%.$$

effort between users. Hearn et al. (1998) present a method for estimating F and M from twice-a-period tagging (e.g., twice per year over several years). Tagging takes place before a heavy fishing episode and once again at the end of this episode; M and F are sorted out by comparing rates of returns from the two markings, over years.

Radio tags may also be used to estimate Z, M, and F in large-bodied species. For instance, Hightower et al. (2000) applied telemetry to estimate natural mortality of striped bass. The general approach was to locate repeatedly live and dead radio-tagged fish at fixed time intervals. The rate of decline in the number of live fish located over time provided information to estimate Z, whereas locations of dead fish provided information to estimate M. Fishing mortality may be estimated indirectly by subtraction or directly if the circumstances allow for inventorying harvest of radio-tagged fish. A key advantage of this approach is the information gained about the timing and causes of mortality. Telemetry studies are labor

intensive but may be pertinent to estimating mortality in closed populations, particularly where the effort can fulfill other information requirements (e.g., movement pattern or habitat use).

6.6.4 Direct Census to Estimate *F*

Fishing mortality can be derived from estimates of μ and Z (Table 6.1). Values of μ may be obtained through mark–recapture (section 6.6.3) or through direct census of harvest and population size. Direct census of harvest (H) involves estimating the total number of fish taken by the fishery during a time period (reviewed by Malvestuto 1996; and Fabrizio and Richards 1996), and direct census of the population (N) involves estimating the average population size during the same period (reviewed by Seber 1982; and Schwarz and Seber 1999). These two censuses estimate exploitation as $\mu = H/N$. Instantaneous fishing mortality is then estimated as $F = \mu Z/A$.

The phenomenon of catch-and-release mortality in recreational fisheries has received much attention in recent decades. When catch-and-release mortality is low or negligible, conventional estimates of F and M will not be grossly affected. For instance, catch-and-release mortality was 3% for cutthroat trout in the Yellowstone River (Schill et al. 1986) and 2% for common snook in southern Florida waters (Taylor et al. 2001). However, high levels of catch-and-release mortality will confound what otherwise might be a straightforward measurement of F and M. For instance, 67% of striped bass died after being caught and released in a Tennessee reservoir during summer (Bettoli and Osborne 1998), and reef fishes such as red snapper often experience high (>30%) catch-and-release mortality (Gitschlag and Renaud 1994). Catch-and-release mortality represents unaccounted fishing mortality, and high levels of catch-and-release mortality will inflate estimates of natural mortality. When M is high, the success of harvest regulations depends on the level of catch-and-release mortality (Waters and Huntsman 1986).

If estimates of release rates (P_r) and catch-and-release mortality (P_m) are available, these can be used to adjust exploitation rate (μ') as

$$\mu' = \mu + P_r P_m \frac{\mu}{1 - P_r}. \tag{6.22}$$

For example, suppose that a reward-tag study estimated $\mu = 40\%$. If a concurrent creel survey indicated that 50% of fish caught are released, and an independent study indicated that 10% of the fish released do not survive, the adjusted exploitation rate would be 44%.

6.6.5 Production Modeling to Estimate *M*

Csirke and Caddy (1983) estimated M from the relation between yield and Z. This method assumes a parabolic relation between Z and yield, and represents an extension of the traditional Graham-Schaefer production model (Ricker 1975). If

total catch (yield) and Z are known for a series of years (at least three, more is better), yield (Y_i) can be modeled in terms of Z_i with a quadratic equation as

$$Y_i = b_0 + b_1 Z_i - b_2 Z_i^2. \tag{6.23}$$

This equation corresponds to a parabola with a convex-downward curvature. When $F = 0$, yield becomes zero and Z becomes M, so that equation (6.23) becomes

$$Y = b_0 + b_1 M - b_2 M^2 = 0 \text{ (when } F = 0). \tag{6.24}$$

Solving for M,

$$M = (-b_1 + [b_1^2 - 4b_0 b_2]^{0.5})/2b_2. \tag{6.25}$$

Figure 6.8 provides an example application. A limitation of this method is the assumption of a parabolic relation between Y and Z. The model may be made more realistic by using new formulations of the basic parabola model (several models are reviewed by Quinn and Deriso 1999). Another limitation is that to produce reliable regression coefficients there must be enough contrast in the values of Z.

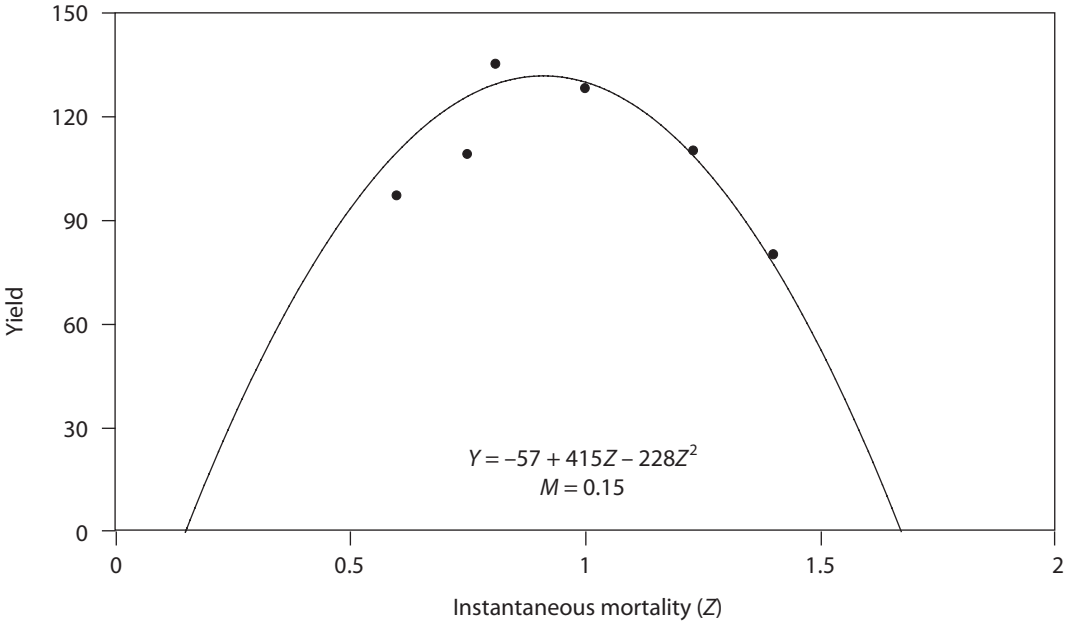


Figure 6.8 Estimating M from the relation between Z and yield (Y), assuming a traditional Graham–Schaefer-type curve. When $F = 0$, yield becomes zero and Z becomes M.

6.6.6 Meta-Analyses to Estimate *M*, *F*, or *Z*

Meta-analysis is a method for objectively synthesizing information from the literature and subjecting that information to statistical analysis (Wolf 1986). Unlike traditional literature reviews, the methodology of meta-analyses is clearly presented so that others can see how decisions were made. Meta-analysis has clearly defined procedural steps to translate the findings of different research to a common parameter defined statistically.

Meta-analyses can be used to develop empirical regression equations predictive of mortality. Natural mortality is consistently related to factors such as growth rate, ultimate body size, fecundity, age at sexual maturity, and temperature. For example, fish populations with slow growth tend to have low *M* values; a slow-growing species or population simply cannot bear high natural mortality without becoming extirpated. Likewise, fishing mortality is related to factors such as fishing effort and diversity of target species. Meta-analyses make use of these natural associations between mortality and allied variables to develop regional or wide-ranging, single or multispecies, predictive models. Selected examples are listed in Table 6.4.

Given that sources of mortality are difficult to sort out, empirical models derived through meta-analyses are sometimes used to estimate mortality components. For example, Campana (1987) used Pauly's meta-analysis (Table 6.4) to

Table 6.4 Selected meta-analyses that use associations between mortality and allied variables to develop regional or wide-ranging predictive models.

Meta-analysis	Reference
In high-latitude stocks, there was a close association between *M* (annual) and the age (years) when 50% of the population was sexually matured (Tm_{50}). The equation was $M = 1.52Tm_{50}^{-0.72} - 0.155$.	Rikhter and Efanov (1976)
With data on 10 species the relation between *M* (annual) and gonadosomatic index (GSI = gonad weight/total weight) was estimated as $M = 4.64\text{GSI} - 0.37$.	Gunderson (1980)
Annual natural mortality (*M*) was analyzed for 175 stocks, including 84 freshwater and marine species of tropical to polar distribution. A predictive equation was derived for M based on the von Bertalanffy growth parameters *K* (annual), L_∞ (cm), and *T* (mean annual surface temperature, °C). The equation was $M = -e^{0.0152 + 0.654\log_e K - 0.279\log_e L_\infty + 0.463\log_e T}$.	Pauly (1980)
Instantaneous total mortality (*Z*) was modeled relative to longevity (Y_{max} = mean age of "the oldest specimens" in a sample). Unfortunately, longevity can be as difficult to estimate as mortality; thus, the value of such relations is limited. The equation was $Z = e^{-1.01\log_e Y_{max} + 1.46}$.	Hoenig (1983)
For 40 largemouth bass populations in North America, *M* (annual) was related to the average number of degree-days (DD) above 10°C in a year as $M = 0.000159\text{DD} + 0.197$.	Beamesderfer and North (1995)
Wilde determined that the fraction of fishing mortality due to tournaments (TM; %) was related to water temperature (*T;* °C) in 45 events: $\text{TM} = 0.1042T^{1.683}$.	Wilde (1998)

estimate *M* for haddock, and Ebbers (1987) used the model to estimate *M* for largemouth bass. The user of models derived through meta-analyses should recognize the limitations of the models. Because mortalities and their predictors are often difficult to estimate, and come from a variety of studies using different techniques, quality of the models derived through meta-analyses may be questionable. Moreover, the models predict only average mortality for a given population characteristic, or a set of characteristics in the case of multivariate models, whereas the study population may fall above or below the mean. Pascual and Iribarne (1993) evaluated the predictive power of several empirical models and found that error around mortality predictions was high. Thus, estimates from models derived through meta-analysis can be unreliable and should be used only as rough approximations in preliminary analyses or exploratory modeling that seek only relative solutions.

6.7 REFERENCE POINTS

Managing a fishery requires adjusting input and outputs to obtain a desired outcome. Reference points are targets or limits that help guide such adjustments. Target reference points represent a desirable condition toward which a population may be guided to obtain a desired outcome; limit reference points represent a danger zone to be avoided.

6.7.1 Reference Points Based on *F*

The relationship between yield-per-recruit (*y*-variable) and *F* (*x*-variable) is generally depicted as a dome-shaped curve. The peak of the curve has a slope of zero and identifies the *F* that produces the maximum yield-per-recruit (F_{max}; Quinn and Deriso 1999). This target reference point is often difficult to estimate because of the flat-topped shape of the yield-per-recruit curve. An easier target reference point to estimate is $F_{0.1}$, which estimates the fishing mortality at which the slope of the dome-shaped yield-per-recruit curve is 10% of its value at the origin. This value is always less than F_{max} and therefore more conservative.

6.7.2 Reference Points Based on *M*

In unfished or lightly fished populations, mortality limit reference points may be established based on *M*. For surplus-production models, Gulland and Boerema (1973) proposed a simple empirical formula (i.e., MSY = $0.5MB_0$) to establish maximum sustainable yield (MSY) in terms of the unfished standing stock (B_0) and the natural mortality at which the slope of a dome-shaped yield curve is zero. Their assumption relies on the symmetrical Schaefer yield model to assume that MSY will occur at 0.5 the unfished standing stock and that $F_{MSY} = M$. Because there is little evidence that $F_{MSY} = M$, this equation has been generalized to MSY = pMB_0, with *p* equal to 0.5 or other fraction. In general, *p* should be higher for long-lived species (low *M*) than for short-lived ones (high *M*). Patterson (1992) suggested

that for small pelagic species, a p near 0.5 (i.e., $F_{MSY} = 0.5M$) should be sustainable. Caddy (1998) suggested that p should decrease as M increases, so that $p = 0.8 - 0.9$ for long-lived ($M = 0.1 - 0.2$) terminal predators and $p = 0.4–0.5$ for short-lived ($M = 1.1 - 1.4$) small prey species.

6.7.3 Reference Points Based on *Z*

Because partitioning mortality into F and M is often difficult, there are advantages in expressing mortality limit reference points in terms of Z. For surplus-production models, Caddy and Defeo (1996) used time series of paired annual Z and catch to approximate the Z values that resulted in MSY. For age-structured models applied to recreational fisheries management, Miranda (2002) derived limits on Z based on size objectives for the fishery stated in terms of mean length or a size structure index. The relation between Z and mean length of fish is described by a decaying exponential curve. Thus, to preserve fisheries with large fish requires maintaining a low Z, although exact levels depended on growth rate. The relation between Z and proportional stock density (PSD; Anderson and Neumann 1996) was described with the model (Miranda 2002)

$$Z = -\frac{\log_e(PSD/100)}{t_Q - t_S}, \tag{6.26}$$

where t_S = number of years it takes fish to grow to stock size and t_Q = number of years to quality size (size is defined according to species by Anderson and Neumann 1996). Thus, fast-growing populations can withstand higher mortality to maintain a target PSD. Equations (6.8) and (6.26) can be used to establish reference points based on threshold size objectives for the fishery (Box 6.11).

6.8 COMPENSATORY AND ADDITIVE MORTALITY

Additive mortality assumes that an increment in F or M results in an equal increment in Z. When increments in F or M lead to disproportionate or no increment in Z, mortality is compensatory (Figure 6.9). Populations near carrying capacity are more likely regulated by compensatory processes and populations at low density by additive processes (Bartmann et al. 1992). Hence, a population may exhibit additive mortality at low density and compensatory mortality at high density, but a continuum of escalating partial compensation between completely additive and completely compensatory mortality is possible (Nichols et al. 1984; Conroy and Krementz 1990).

Adult fishes probably experience lower levels of compensatory mortality than do higher vertebrates because fish are better able to adjust their growth rate to food availability, lengthening the period they can survive with limited food (Weatherley and Gill 1987; Shuter 1990). Nevertheless, compensatory mortality may result from cumulative effects. During periods of reduced growth through

Box 6.11 Establishing Target Mortality Caps in Length-Based Fisheries Management

Consider, for instance, that in the Columbus Lake example (Box 6.8) a fishery management objective is for largemouth bass in the population (and thus perhaps the angler's creel) to average 275 mm total length or better (average length estimate includes only fish fully vulnerable to the collection method). If Z is excessive, whether due to F or M, few fish will live to old age (= large size), and thus the management objective cannot be met. Given the existing growth conditions described by the von Bertalanffy growth model ($K = 0.226$ and $L_\infty = 636$ mm), the length above which all largemouth bass are considered equally vulnerable to electrofishing ($L_x = 150$ mm), and the target mean length ($L_{mean} = 275$ mm), the limit Z may be estimated with equation (6.8) as

$$Z = 0.226\ \frac{636 - 275}{275 - 150} = 0.65\ .$$

Alternatively, if the management objective for the largemouth bass population is expressed in terms of PSD instead of mean length, a limit on Z can be estimated with equation (6.26). Assume that the target PSD is 50 and that it takes 1.1 year for the average largemouth bass in the population to grow from stock to quality size (i.e., $t_Q - t_S = 1.1$); then

$$Z = -\frac{\log_e(50/100)}{1.1} = 0.63.$$

These Z values represent limit reference points above which the management objective cannot be achieved. The limit is intended to prevent overfishing that renders the size distribution of a population undesirable from a fishery perspective. The limit is not a target for management, but instead it helps managers define the upper cap of mortality. If the cap is approached, additional emphasis must be placed on monitoring the fishery. If the cap is exceeded, Z must be immediately reduced through cuts in F that are equal to or larger than the excess Z.

intra- or interspecific competition for resources, other stressors (e.g., disease, parasitism, or predation) may act synergistically to cause density-dependent mortality. Cushing (1981) suggested that predation acts in a density-dependent manner in some pelagic marine fish stocks. Allen et al. (1998) found that mortality was additive in largemouth bass populations but could be compensatory in crappies and northern pike. For crappies, empirical data showed no relation between μ and A at low levels of μ but a positive slope at mid- to high levels of μ; for northern pike, there was no relation between μ and A.

The existence of compensatory mortality can be examined by plotting independent estimates of A and μ to identify potential trends like those shown in Figure 6.9. Such estimates may be obtained from existing data or by experimentally manipulating μ through harvest restrictions. A plot of F on Z would be problematic because computation of F involves Z (i.e., $F = \mu Z/A$), and thus Z would be included in both axes. The relation between A and μ is expected to be direct and linear if mortalities were completely additive. Conversely, the plot is expected to be slopeless or nearly so if mortalities were completely compensatory. Burnham

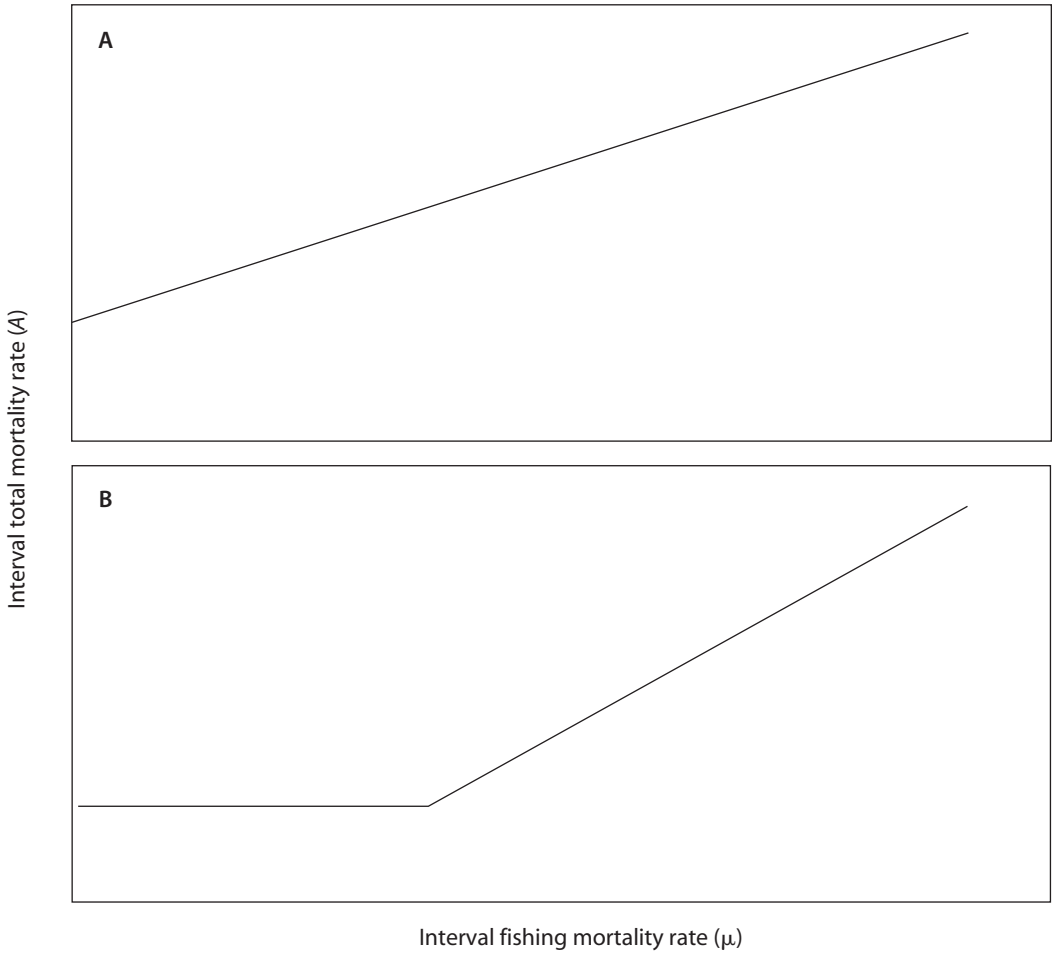


Figure 6.9 Conceptual models of the relation between μ and *A* under (A) additive, and (B) compensatory mortality. The flat portion of the curve in (B) implies natural mortality is changing and compensating for increased fishing mortality.

and Anderson (1984) present statistical models to test whether mortality conforms to either of these extremes or some intermediary level of compensation.

6.9 PERSPECTIVES ON BIASED AND IMPRECISE MORTALITY ESTIMATES

It should be apparent from the preceding sections that estimating mortality with sufficient accuracy and precision is not easy. Even the simplest estimation models require data that are difficult and expensive to collect, apart from having to rely on collection methods that have numerous biases. The models make various assumptions, which are often disregarded by the environment and ignored by fisheries scientists. The saving attribute is that mortality has well-defined lower and upper limits, 0 and 100%, that conveniently bound the estimates. Given these

difficulties, it is appropriate to end this chapter with our views on how to deal with the uncertainties associated with mortality estimates.

Uncertainties result from inaccurate and imprecise estimates (i.e., estimates that have error). Accuracy refers to how close an estimate of mortality matches the true value, whereas precision refers to how close repeated estimates of mortality would agree with each other. Error encompasses both the imprecision and inaccuracies of estimates. Uncertainty about accuracy of mortality estimates is created by error in estimating variables that affect the computation of mortality, such as fish age, size structure, growth rate, harvest rate, tag retention, and tag reporting. Error arises from imperfect representation of the population by the sampling process, inability to meet the assumptions of the estimating model, and lack of complete knowledge about the functioning of populations (e.g., additive versus compensatory mortality). Uncertainty due to poor precision results from the high variability associated with population variables that include sampling and natural components. Sampling error is introduced by the sampling gear, timing, and procedures; this error can be reduced through improved collection methods, proper sampling design, and increased sample sizes. Natural variability results from normal population fluctuations; although this variability does not constitute error, measures of error normally include natural variability. Francis and Shotton (1997) and Charles (1998) provide good reviews with more refined classifications of uncertainties.

Uncertainty in mortality estimates can be reduced by confronting the questions of accuracy and precision. Accuracy of estimates may be verified by comparing multiple estimates made with different methods (e.g., mark–recapture, length-based, and catch-curve models) or by evaluating estimates relative to covarying population or environmental parameters to examine if they follow expected trends (e.g., high Z values are unlikely when fishing effort is low, unless habitat is of poor quality). If two estimates are similar, the fisheries scientist may become increasingly confident about the quality of the estimates and use the average of the two values. Commonly, the estimates are not so similar, and the fisheries scientists ignores the least certain one, takes the average of the two, or develops two recommendations based on each of the estimates. If the estimates were highly different, averaging should be avoided because there is a good possibility that one of them is wrong and averaging would lead to an undesirable estimate (Schnute and Hilborn 1993). When only a single estimate is available and its accuracy is not confirmed by covarying variables, a second estimate should be sought.

Collecting ample, good data with proven protocols under acceptable sampling designs can increase precision. For example, to perform catch-curve analyses a reasonable sample size of aged fish could be about 200 for a heavily exploited, short-lived freshwater species with few age-classes or 500 or more for a species with 10 or more age-classes in the population (Sampson and Yin 1998; Ciepielewski 1999). However, except in cases where every death can be counted, estimates will still contain error. The variability inherent in every estimate should not be ignored by working exclusively with a point estimate. Instead, confidence intervals should be estimated and further application of the mortality esti-

mate must involve the range of values within the confidence band. Confidence limits mix estimation error and natural variability, which is pertinent given that management will be applied to naturally stochastic populations occupying unpredictable environments.

Further analysis may involve appraisal of the effect of uncertainty on possible outcomes and decision making. This step may take the form of an informal qualitative evaluation or a quantitative assessment using simple or complex models. Qualitative evaluations often involve making conservative allowances for uncertainties through arbitrary safety factors. Much attention has been given in the literature to establishment of precautionary reference points for F (e.g., $F_{0.1}$ and other F limits; Caddy 1998). Brown and Patil (1986) provide an example of a qualitative evaluation of uncertainty to establish levels of F. Quantitative evaluations may evaluate the outcome of models relative to the statistical distribution of mortality and other (if any) variables in the model (i.e., sensitivity analysis; Saltelli et al. 2000). These evaluations help identify the range of possible outcomes given the uncertainty of the variables included in the model; however, models can introduce additional uncertainty because they are unlikely to simulate accurately a population's dynamics.

6.10 CONCLUSIONS

We have presented numerous conceptual and mathematical models of mortality in the preceding sections; however, mortality in fish populations should be more than an abstract concept. Knowledge of mortality rates is fundamental to understanding the dynamics of exploited fish populations, and when compared to rates of recruitment and growth, mortality rates are often the easiest to manage using harvest regulations. Size limits, slot limits, creel limits, closed seasons, and gear restrictions are all examples of regulations typically used to modify fishing mortality (Noble and Jones 1999). When you consider that promulgating regulations and evaluating the subsequent response of freshwater fish populations to new regulations is commonplace, it is surprising that mortality rates are not estimated more routinely or scrutinized more intensely by fisheries scientists.

The most common methods used by inland fisheries scientists to calculate mortality are linearized catch curves or Chapman–Robson's catch curves. Although some of the methods presented here have seen little use outside the marine literature, numerous freshwater sport fish and commercial fish populations are exploited in the same manner as marine stocks. For instance, crappie fisheries are almost exclusively catch-and-harvest fisheries (i.e., catch and release of legal-sized fish is unusual), as are most fisheries for paddlefish, catfishes, bluegill, walleye, and sauger. Instead of large fishing fleets operating in marine systems, inland fishers operate singly but with no less determination and zeal. The populations that are exploited by marine and freshwater fishers are also identical with respect to what is important to know about them and what fisheries scientists can do to conserve or enhance these populations. It should also be apparent that many

datasets lend themselves to several different analytical procedures that can produce different estimates of varying accuracy and precision. When estimates agree, confidence in them will be high; however, when they disagree, they provide direction and justification for future efforts.

The greatest difficulty in estimating mortality is partitioning total mortality into fishing and natural mortality. Whereas estimates of total mortality are abundant, rates of exploitation are difficult to obtain and are known for comparatively few populations; natural mortality rates are available for even fewer populations. The need for accurate estimates of all three rates will grow more acute as more fisheries scientists take advantage of recent advances in population models, whose outputs are critically sensitive to mortality rates.

6.11 REFERENCES

Agostinho, A. A., L. E. Miranda, L. M. Bini, L. C. Gomes, S. M. Thomaz, and H. I. Susuki. 1999. Patterns of colonization in neotropical reservoirs, and prognoses on aging. Pages 227–265 *in* J. G. Tundisi and M. Straškraba, editors. Theoretical reservoir ecology and its applications. Backhuys Publishers, Leiden, The Netherlands.

Allen, M. S. 1999. Effects of variable recruitment on catch-curve analysis for crappie populations. North American Journal of Fisheries Management 17:202–205.

Allen, M. S., L. E. Miranda, and R. E. Brock. 1998. Implications of compensatory and additive mortality to the management of selected sportfish populations. Lakes and Reservoirs: Research and Management 3:67–79.

Anderson, R. O., and R. M. Neumann. 1996. Length, weight, and associated structural indices. Pages 447–512 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Arnason, A. N., and C. J. Schwarz. 1999. Using POPAN-5 to analyze banding data. Bird Study 46 (Supplement l):127–168.

Bailey, N. T. J. 1951. On estimating the size of mobile populations from recapture data. Biometrika 38:293–306.

Bartmann, R. M., G. C. White, and L. H. Carpenter. 1992. Compensatory mortality in a Colorado mule deer population. Wildlife Monographs 121:1–39.

Beamesderfer, R. C. P., and J. A. North. 1995. Growth, natural mortality, and predicted response to fishing for largemouth bass and smallmouth bass populations in North America. North American Journal of Fisheries Management 15:688–704.

Bettoli, P. W. 1999. Creel survey and population dynamics of salmonids stocked into the Watauga River below Wilbur Dam. Tennessee Wildlife Resources Agency, Fisheries Report 99–41, Nashville.

Bettoli, P. W., and L. E. Miranda. 2001. Cautionary note about estimating mean length at age with subsampled data. North American Journal of Fisheries Management 21: 425–428.

Bettoli, P. W., and R. S. Osborne. 1998. Hooking mortality and behavior of striped bass following catch and release angling. North American Journal of Fisheries Management 18:609–615.

Beverton, R. J. H., and S. J. Holt. 1956. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. Rapports et Proces-Verbaux des Reunions, Conseil International pour l'Exploration de la Mer 140:67–83.

Brooks, E. N., K. H. Pollock, J. M. Hoenig, and W. S. Hearn. 1998. Estimation of fishing and natural mortality from tagging studies in fisheries with two user groups. Canadian Journal of Fisheries and Aquatic Sciences 55:2001–2010.

Brown, B. E., and G. P. Patil. 1986. Risk analysis in the Georges Bank haddock fishery—a pragmatic example of dealing with uncertainty. North American Journal of Fisheries Management 6:183–191.

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1985. Statistical inference from band recovery data—a handbook. 2nd edition. U.S. Fish and Wildlife Service, Resource Publication 156, Washington, D.C.

Buckland, S. T. 1982. A mark–recapture survival analysis. Journal of Animal Ecology 51: 833–847.

Burnham, K. P., and D. R. Anderson. 1984. Test of compensatory vs. additive hypotheses of mortality in mallards. Ecology 65:105–112.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society, Monograph 5, Bethesda, Maryland.

Caddy, J. F. 1998. A short review of precautionary reference points and some proposals for their use in data-poor situations. FAO (Food and Agriculture Organization of the United Nations) Fisheries Technical Paper 379, Rome.

Caddy, J. F., and O. Defeo. 1996. Fitting the exponential and logistic surplus yield models with mortality data: some explorations and new perspectives. Fisheries Research 25:39–62.

Campana, S. E. 1987. Comparison of two length-based index of abundance in adjacent haddock stocks *(Melanogrammus aeglefinus)* on the Scotian shelf. Journal du Conseil, Conseil International pour l'Exploration de la Mer 39:195–212.

Charles, A. T. 1998. Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian groundfishery experience. Fisheries Research 37:37–50.

Ciepielewski, W. 1999. Sample size in the studies of population age structure of some fish species. Archives of Polish Fisheries 7:35–52.

Conroy, M. J., J. E. Hines, and B. K. Williams. 1989. Procedures for the analysis of band-recovery data and user instructions for Program MULT. U.S. Fish and Wildlife Service, Resource Publication 175, Washington, D.C.

Conroy, M. J., and D. G. Krementz. 1990. A review of the evidence for the effects of hunting on American black duck populations. Transactions of the North American Wildlife and Natural Resources Conference 55:501–517.

Cooch, E. G., R. Pradel, and N. Nur. 1996. Practical guide to mark–recapture analysis using SURGE. Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier, France.

Cormack, R. M. 1989. Loglinear models for capture–recapture. Biometrics 45:395–413.

Csirke, J., and J. F. Caddy. 1983. Production modeling using mortality estimates. Canadian Journal of Fisheries and Aquatic Sciences 40:43–51.

Cushing, D. H. 1981. Fisheries biology: a study in population dynamics, 2nd edition. University of Wisconsin Press, Madison.

Ebbers, M. A. 1987. Vital statistics of a largemouth bass population in Minnesota from electrofishing and angler-supplied data. North American Journal of Fisheries Management 7:252–259.

Efron, B., and R. J. Tibshirami. 1998. An introduction to the bootstrap. CRC Press, Boca Raton, Florida.

Evans, C. R., and A. P. M. Lockwood. 1994. Population field studies of the Guinea chick lobster (*Panulirus guttatus* Latreille) at Bermuda: abundance, catchability, and behavior. Journal of Shellfish Research 13:393–415.

Fabrizio, M. C., M. E. Holey, P. C. McKee, and M. L. Toneys. 1997. Survival rates of adult lake trout in northwestern Lake Michigan, 1983–1993. North American Journal of Fisheries Management 17:413–428.

Fabrizio, M. C., and R. A. Richards. 1996. Commercial fisheries surveys. Pages 625–650 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Fabrizio, M. C., B. L. Swanson, S. T. Schram, and M. H. Hoff. 1996. Comparison of three nonlinear models to describe long-term tag shedding by lake trout. Transactions of the American Fisheries Society 125:261–273.

Fairfield, C. P., and S. A. Mizroch. 1990. Comparison of abundance estimation techniques for the western North Atlantic right whale *(Eubalaena glacialis)*. Pages 119–126 *in* P. S. Hammond, S. A. Mizroch, and G. P. Donovan, editors. Proceedings of the symposium and workshop on individual recognition and the estimation of cetacean population parameters. Report of the International Whaling Commission, Special Issue 12.

Francis, R. I. C. C., and R. Shotton. 1997. "Risk" in fisheries management: a review. Canadian Journal of Fisheries and Aquatic Sciences 54:1699–1715.

Froese, R., and C. Binohlan. 2000. Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. Journal of Fish Biology 56:758–773.

Gitschlag, G. R., and M. L. Renaud. 1994. Field experiments on survival rates of caged and released red snapper. North American Journal of Fisheries Management 14:131–136.

Gulland, J. A., and L. K. Boerema. 1973. Scientific advice on catch levels. U.S. National Marine Fisheries Service Fishery Bulletin 71:321–325.

Gunderson, D. R. 1980. Using r-K selection theory to predict natural mortality. Canadian Journal of Fisheries and Aquatic Sciences 37:2266–2271.

Haddon, M. 2001. Modeling and quantitative methods in fisheries. Chapman and Hall, CRC, Boca Raton, Florida.

Hearn, W. S., T. Polacheck, K. H. Pollock, and W. Whitelaw. 1999. Estimation of tag-reporting rates in age-stuctured multicomponent fisheries where one component has observers. Canadian Journal of Fisheries and Aquatic Sciences 56:1255–1265.

Hearn, W. S., K. H. Pollock, and E. N. Brooks. 1998. Pre- and post-season tagging models: estimation of reporting rate and fishing and natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences 55:199–205.

Heincke, F. 1913. Investigations on the plaice. General report I: the plaice fishery and protective measures. Preliminary brief summary of the most important points on the report. Rapports et Proces-Verbaux des Reunions, Conseil International pour l'Exploration de la Mer 16.

Hightower, J. E., and R. J. Gilbert. 1984. Using the Jolly–Seber model to estimate population size, mortality, and recruitment for a reservoir fish population. Transactions of the American Fisheries Society 113:633–641.

Hightower, J. E., J. R. Jackson, and K. H. Pollock. 2000. Using telemetry methods to estimate natural and fishing mortality of striped bass in Lake Gaston, North Carolina. Transactions of the American Fisheries Society 130:557–567.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.

Hines, J. E., and J. R. Sauer. 1989. Program CONTRAST—a general program for the analysis of several survival or recovery rate estimates. U.S. Fish and Wildlife Service, Fish and Wildlife Technical Report 24, Washington, D.C.

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. U.S. National Marine Fisheries Service Fishery Bulletin 81:898–903.

Hoenig, J. M., N. J. Barrowman, W. S. Hearn, and K. H. Pollock. 1998. Multiyear tagging studies incorporating fishing effort data. Canadian Journal of Fisheries and Aquatic Sciences 55:1466–1476.

Hoenig, J. M., W. D. Lawing, and N. A. Hoenig. 1983. Using mean age, mean length and median length data to estimate the total mortality rate. International Council for the Exploration of the Sea, ICES-CM-1983/D:23, Copenhagen, Denmark.

Jagielo, T. 1991. Synthesis of mark–recapture and fishery data to estimate open-population parameters. Pages 491–505 *in* D. Guthrie, J. M. Hoenig, M. Holliday, C. M. Jones, M. J. Mills, S. A. Moberly, K. H. Pollock, and D. R. Talhelm. Creel and angler surveys in fisheries management. American Fisheries Society, Symposium 12, Bethesda, Maryland.

Jensen, A. L. 1985. Comparison of catch-curve methods for estimation of mortality. Transactions of the American Fisheries Society 114:743–747.

Johnson, B. M., and J. P. Goettl, Jr. 1999. Food web changes over fourteen years following introduction of rainbow smelt into a Colorado Reservoir. North American Journal of Fisheries Management 19:629–642.

Jolly, G. M. 1965. Explicit estimates from capture–recapture data with both death and immigration-stochastic model. Biometrika 52:225–247.

Law, P. M. W. 1994. Simulation study of salmon carcass survey capture–recapture methods. California Fish and Game 80:14–28.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62:67–118.

Maceina, M. J. 1997. Simple application of using residuals from catch-curve regressions to assess year-class strength in fish. Fisheries Research 32:115–121.

Maceina, M. J., and P. W. Bettoli. 1998. Variation in largemouth bass recruitment in four mainstream impoundments on the Tennessee River. North American Journal of Fisheries Management 18:998–1003.

Malvestuto, S. P. 1996. Sampling the recreational creel. Pages 591–623 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Miranda, L. E. 2002. Establishing size-based mortality caps. North American Journal of Fisheries Management 22:433–440.

Miranda, L. E., R. E. Brock, and B. S. Dorr. 2002. Uncertainty of exploitation estimates made from tag returns. North American Journal of Fisheries Management 22:1358–1363.

Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied linear statistical models, 3rd edition. Irwin, Boston.

Nichols, J. D., M. J. Conroy, D. R. Anderson, and K. P. Burnham. 1984. Compensatory mortality in waterfowl populations: a review of the evidence and implications for research and management. Transactions of the North American Wildlife and Natural Resource Conference 49:535–554.

Noble, R. L., and T. W. Jones. 1999. Managing fisheries with regulations. Pages 455–477 *in* C. C. Kohler, and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Paloheimo, J. E. 1958. A method of estimating natural and fishing mortalities. Journal of the Fisheries Research Board of Canada 15:749–758.

Pascual, M. A., and O. O. Iribarne. 1993. How good are empirical predictions of natural mortality? Fisheries Research 16:17–24.

Patterson, K. 1992. Fisheries for small pelagic species: an empirical approach to management targets. Reviews in Fish Biology and Fisheries 2:321–338.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. Journal du Conseil, Conseil International pour l'Exploration de la Mer 39:175–192.

Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM (International Center for Living Aquatic Resources Management) Studies and Reviews 8, Manila, Philippines.

Pollock, K. H. 1981. Capture–recapture models allowing for age-dependent survival and capture rates. Biometrics 37:52l–529.

Pollock, K. H., and R. H. K. Mann. 1983. Use of an age-dependent mark–recapture model in fisheries research. Canadian Journal of Fisheries and Aquatic Sciences 40:1449–1455.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture–recapture experiments. Wildlife Monographs 107.

Pradel, R., J. E. Hines, J. D. Lebreton, and J. D. Nichols. 1997. Capture–recapture survival models taking account of transients. Biometrics 53:60–72.

Pradel, R., and J. D. Lebreton. 1993. User's manual for program SURGE. Version 4.2. Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier, France.

Quinn, T. J., and R. B. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Ralston, S. 1989. Effect of seasonal recruitment on bias of the Beverton–Holt length-based mortality estimator. Pages 190–197 *in* E. F. Edwards and B. A. Megrey, editors. Mathematical analysis of fish stock dynamics. American Fisheries Society, Symposium 6, Bethesda, Maryland.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Rikhter, V. A., and V. N. Efanov. 1976. On one of the approaches to estimating natural mortality of fish populations. International Convention for the Northwest Atlantic Fisheries Research Document 76/VI/8.

Robson, D. S., and D. G. Chapman. 1961. Catch curves and mortality rates. Transactions of the American Fisheries Society 90:181–189.

Saltelli, A., K. Chan, and E. M. Scott, editors. 2000. Sensitivity analysis. Wiley, New York.

Sammons, S. M., and P. W. Bettoli. 1998. Influence of water levels and habitat manipulations on fish recruitment in Normandy Reservoir. Tennessee Wildlife Resources Agency, Fisheries Report 98–42, Nashville.

Sampson, D. B., and Y. Yin. 1998. A Monte Carlo evaluation of the stock synthesis assessment program. Pages 315–338 *in* F. Funk and seven coeditors. Fishery stock assessment models. University of Alaska, Alaska Sea Grant Publications, Fairbanks.

SAS Institute. 1998. SAS/Stat user's guide, version 6: volumes 1 and 2. SAS Institute, Cary, North Carolina.

Schill, D. J., J. S. Griffith, and R. E. Gresswell. 1986. Hooking mortality of cutthroat tout in a catch-and-release segment of the Yellowstone River, Yellowstone National Park. North American Journal of Fisheries Management 6:226–232.

Schnute, J. T., and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences 50:1916–1923.

Seber, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52:249–259.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Griffin, London.

Schwarz, C. J., and G. A. F. Seber. 1999. A review of estimating animal abundance III. Statistical Science 14:427–456.

Shuter, B. J. 1990. Population-level indicators of stress. Pages 145–166 *in* S. M. Adams, editor. Biological indicators of fitness in fish. American Fisheries Society, Symposium 8, Bethesda, Maryland.

Smith, S. G., J. R. Skalski, J. W. Schlechte, A. Hoffmann, and V. Cassen. 1994. SURPH. 1 manual: statistical survival analysis of fish and wildlife tagging studies. University of Washington, School of Fisheries, Center for Quantitative Science, Seattle.

Steel R. G. D., and J. H. Torrie. 1980. Principles and procedures of statistics, 2nd edition. McGraw-Hill, New York.

Taylor, R. G., J. A. Whittington, and D. E. Haymans. 2001. Catch-and-release mortality rates of common snook in Florida. North American Journal of Fisheries Management 21:70–75.

Timmons, T. L., W. L. Shelton, and W. D. Shelton. 1980. Sampling of reservoir sport fish populations with rotenone in littoral areas. Proceedings of the Southeastern Association of Fish and Wildlife Agencies 32(1978):474–484.

Van Den Avyle, M. J., and R. S. Hayward. 1999. Dynamics of exploited fish populations. Pages 127–166 *in* C. C. Kohler, and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Vandergoot, C. J., and P. W. Bettoli. 2001. Evaluation of current management practices and assessment of recruitment, growth, and condition of walleyes in Tennessee reservoirs. Tennessee Wildlife Resources Agency, Fisheries Report 01–44, Nashville.

Waters, J. R., and G. R. Huntsman. 1986. Incorporating mortality from catch and release into yield-per-recruit analyses of minimum size limits. North American Journal of Fisheries Management 6:463–471.

Weatherley, A. H., and H. S. Gill. 1987. The biology of fish growth. Academic Press, New York.

Wetherall, J. A., J. J. Polovina, and S. Ralston. 1987. Estimating growth and mortality in steady-state fish stocks from length-frequency data. ICLARM (International Center for Living Aquatic Resources Management) Conference Proceedings 13, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat.

White, G. C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management 47:716–728.

White, G. C. 1992. PC SURVIV user's manual. Colorado State University, Department of Fishery and Wildlife Biology, Fort Collins.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46 (Supplement):120–138.

Wilde, G. R. 1998. Tournament-associated mortality in black bass. Fisheries 23(10):12–22.

Wilson, K. R., J. D. Nichols, and J. E. Hines. 1989. A computer program for sample size computations for banding studies. U.S. Fish and Wildlife Service, Fish and Wildlife Technical Report 23, Washington, D.C.

Wolf, F. M. 1986. Meta-analysis: quantitative methods for research synthesis, volume 59. Sage, Thousand Oaks, California.

Zale, A. V., and M. B. Bain. 1994. Estimating tag-reporting rates with postcards as tag surrogates. North American Journal of Fisheries Management 14:208–211.

7 Relative Abundance and Catch per Unit Effort

Wayne A. Hubert and Mary C. Fabrizio

7.1 INTRODUCTION

Knowledge of the abundance of fish in a stock is a component of the information used in management of fisheries (Ney 1999). Abundance estimates are used along with data on age and length composition and weight–length relations to make judgments regarding the status of fish stocks. Many methods have been developed to estimate the numerical abundance of fish in a stock including counts within isolated segments of a water body, mark and recapture, and removal methods (Chapter 8; Van Den Avyle and Hayward 1999). However, in many freshwater fisheries these methods require more time and money than can be allocated to the assessment. In these cases, fisheries managers use indices of abundance to estimate relative abundance of fishes (Fabrizio and Richards 1996; Hubert 1996; Ney 1999).

The most common indices of relative abundance are computed from catch per unit effort (C/f) data for samples from a fish stock (Fabrizio and Richards 1996; Hubert 1996; Ney 1999). A C/f index is defined mathematically as

$$C/f = qN, \tag{7.1}$$

where C is the number of fish caught, f is the unit of effort expended, q is the catchability coefficient or probability of catching an individual fish in one unit of effort, and N is the absolute abundance of fish in the stock. When numerical abundance cannot be estimated, fisheries scientists often use C/f to make judgments about the abundance of fish in a stock.

Effort (f) is computed in many ways depending on the sampling gear and habitat in which the target species resides. Units of effort may include individual sets or hauls with a gear, the volume or area of habitat sampled, or the temporal duration of sampling. With passive gears (Hubert 1996), such as gill nets and trap nets, effort is generally expressed in terms of the standard "set" with a specific piece of

gear. For example, a gill-net set might involve placing the net on the bottom overnight for 12 h, and the net may be 100 m long and 2 m high, constructed of 2.54-cm-square-mesh monofilament netting, and have a float line and a lead line. With active gears (Hayes et al. 1996), such as trawls, effort is often described in terms of the duration or length of the haul at a given boat speed. Similarly, effort with seines is often quantified relative to the area or length of shoreline over which the seine is pulled. With small larval fish trawls or push nets (Kelso and Rutherford 1996), the volume of water filtered is often computed, and *C/f* is expressed as the numbers captured per unit of water filtered. The *C/f* of electrofishing samples (Reynolds 1996) is often described in terms of the number of fish caught in a given amount of time (minutes or hours) or length of shoreline sampled.

7.1.1 General Applications in Freshwater Fisheries

Applications of *C/f* to assessment of stocks of freshwater fish occur in both sport and commercial fisheries. A stock is a group of fish or other aquatic animals that can be treated as a single unit for management purposes (Lackey and Hubert 1976). A stock is generally considered to be a self-contained and self-perpetuating population of a single species with no mixing from the outside and within which biological characteristics and impact of fishing are uniform. This definition is accurate when applied to populations in small lakes and impoundments. However, the geographic boundaries of many freshwater fish stocks are vague and unknown, as in streams, rivers, large reservoirs, or large lakes. Consequently, defined areas and not biological populations are sometimes used as the management units.

Catch per unit effort data are commonly used to monitor or assess stocks when the boundaries of the populations are unknown. Sport fisheries are often assessed by sampling with active or passive gears (Hayes et al. 1996; Hubert 1996) or by surveying recreational anglers and sampling creeled fish (Malvestuto 1996). Commercial fisheries are often assessed using onboard or port-side sampling of the catch (Fabrizio and Richards 1996), but sampling protocols with active or passive gears are also used. Commercial fishery sampling programs are often used to estimate the catch of species in a fishery and the amount of fishing effort (Gillis and Peterman 1998). All of these sampling approaches can generate *C/f* data that can be used to assess temporal and spatial trends of fish stocks.

7.1.1.1 Monitoring of Stock Abundance over Time

One of the earliest applications of *C/f* data in inland waters was a description of annual changes in relative abundance of sport fish in Clear Lake, Iowa from 1947 to 1968 based on gill-net data (Carlander 1953; Bulkley 1970). Similarly, the temporal patterns in relative abundance of prey fishes in Lake Michigan from 1973 to 1993 have been described using *C/f* data from trawl sampling (Fabrizio et al. 2000). Also, cyclic patterns in abundance of yellow perch in an oligotrophic lake have been described using *C/f* data (Sanderson et al. 1999). Many similar monitoring programs have been conducted by state, provincial, and federal management agencies.

Most commonly, time series of *C/f* data are used to assess the efficacy of fisheries management actions, such as the response of largemouth bass and bluegill populations to the removal of excess vegetation in lakes (Pothoven et al. 1999). Monitoring of *C/f* is also conducted to determine declines or increases in abundance of rare species, such as Atlantic sturgeon in the Hudson River (Peterson et al. 2000). Other *C/f* monitoring efforts may be used to evaluate restoration efforts, such as those for lake trout in Lake Superior (Hansen et al. 1995). Similarly, the response of fisheries to introductions of exotic species can be assessed using *C/f*, as has been done for Lake Erie fishes relative to the appearance of zebra mussels in the lake (Trometer and Busch 1999).

Several measures of annual variation in *C/f* have been developed to predict the future abundance of fish or the quality of a fishery. For example, *C/f* of small yellow perch in trawls in the southern portion of Lake Michigan has been used to predict the future abundance of fish acceptable to anglers (Shroyer and McComish 1998). Similarly, *C/f* of walleye in gill nets during the fall can be a predictor of angler catch rates the following summer (Isbell and Rawson 1989). Also, *C/f* of age-0 fish has been used as a predictor of recruitment of age-1 fish of some species in reservoirs (Willis 1987; Sammons and Bettoli 1999).

7.1.1.2 *Evaluation of Spatial Distribution Patterns within Stocks*

Another common application of *C/f* data is the evaluation of spatial distribution patterns or patchiness of fish within a stock. For example, *C/f* data have been used to describe spatial distributions of fishes in large (Ward et al. 2000) and small (Hi and Lodge 1990) lakes, as well as reservoirs of various sizes (Hubert and O'Shea 1992; Van Den Avyle et al. 1995; Michaletz and Gale 1999). Habitat associations of fishes may be identified using *C/f* data obtained from different habitats in both lentic (Irwin et al. 1997; Sammons and Bettoli 1999) and lotic (Jackson 1995; Johnson and Jennings 1998) waters. Seasonal patterns in fish distributions have also been described using *C/f* data. For example, seasonal abundance of fishes in tributaries to the Missouri River has been described in this manner (Braaten and Guy 1999). Fisheries scientists also use *C/f* data to ascertain the effects of habitat mitigation efforts on the spatial distribution of fishes (e.g., Moyer et al. 1995; Chipps et al. 1997).

7.1.1.3 *Assessment of Stocks Relative to Other Stocks*

Comparison of fish stocks in two or more water bodies based on *C/f* data obtained by standard fish sampling protocols has also been applied by freshwater fishery managers. For example, among biologists managing small impoundments there is general consensus that electrofishing *C/f* is a good measure of largemouth bass abundance (Flickinger et al. 1999).

7.1.1.4 *Surveys*

Surveys are sometimes conducted in which *C/f* data are used to describe the fish assemblage in a water body. However, the catchability coefficient (q) with a particular gear differs among species, so the actual composition of a fish assemblage

is generally not well represented by *C/f* data. Nevertheless, researchers have attempted to consider the effects of differential encounter probabilities, fish size, fish swimming speed, and retention probabilities of a specific gear to provide a better indicator of actual assemblage composition. For example, Spangler and Collins (1992) made such adjustments to *C/f* data from gill nets to describe fish assemblages in different portions of Lake Huron. Parsley et al. (1989) computed capture efficiencies for small fishes sampled with beach seines to achieve a better estimate of assemblage structure in a reservoir.

7.1.2 Underlying Assumptions

An underlying assumption of using *C/f* as an index of abundance is that the number of fish captured is proportional to the amount of effort expended. When a population is closed, one unit of sampling effort removes a fixed proportion of the total population (Seber 1982). As the population declines in abundance, the number of animals captured by one unit of effort declines. This simple linear relation between *C/f* and abundance has been extended to research and monitoring surveys such that *C/f* data are typically treated as a measure of abundance. However, when the assumption of a linear relation fails, *C/f* can be a misleading indicator of stock abundance.

7.1.2.1 Density As an Index of Abundance

The classic catch equation expresses catch as a proportion of abundance, and this proportion varies with the amount of effort:

$$C = fq\,(N/A), \tag{7.2}$$

where C is catch, f is fishing effort, q is (constant) catchability, N is abundance, and A is the area in which the stock occurs (Gulland 1969). This equation can be re-arranged to $C/f = q(N/A)$, so if catchability is known, C/f is a measure of fish density (N/A).

Assumptions of this model are (1) the population is in equilibrium (i.e., birth, recruitment, and immigration rates are balanced by death and emigration rates); (2) units of effort (such as individual trap or net sets) operate independently (one unit of fishing gear does not interfere with other units); (3) q is constant throughout the sampling period; and (4) every individual in the stock has the same probability of capture (Seber 1982). The fourth assumption concerns the spatial distribution of fish and is met when fish are uniformly distributed within the boundaries of the stock. Additionally, when sampling of fish within a stock is without replacement (i.e., live fish are not returned to the water), it is assumed that the effects of such removals are negligible.

7.1.2.2 Constant Catchability

Technically, the constancy of the catchability coefficient (q) determines how well *C/f* serves as an index of abundance (Gulland 1969). The assumption of equal

capture probability for each fish in the population implies that fish are uniformly distributed in space and that all occupied areas are accessible to the gear and are randomly sampled. However, neither fishing effort nor fish are uniformly distributed (Paloheimo and Dickie 1964). Even when effort is uniform, such as in research studies using standardized sampling methods, variation in catchability arises when changes occur in the spatial distribution of fish. It has long been recognized that C/f data reflect changes in animal distributions as often as they reflect changes in abundance (Paloheimo and Dickie 1964). Changes in fish distribution (and availability to the gear) may occur vertically (e.g., changes in the thermocline affecting the vertical distribution of fish) or horizontally (e.g., different habitats are occupied such that the proportion of a population occurring outside the survey area changes). Catch-per-unit-effort data are further confounded when changes in spatial distribution occur concurrently with changes in abundance. For example, at low abundance a relatively greater proportion of Atlantic cod were found in shallow regions outside a survey area and were unavailable to the trawl, thereby reducing catchability during times of low abundance (Swain et al. 1994).

Care must be taken to restrict interpretations of C/f estimates to the portion of a stock actually sampled. For example, when fish in a stock are spatially distributed among exploited and unexploited regions, and individuals move from an unexploited to an exploited region, C/f estimates from the exploited segment are not a good measure of total stock abundance (Sampson 1991). Catch-per-unit-effort estimates from the exploited region are representative of the entire stock only when the rates of movement between the two regions are random. Effects of shifts in distribution on C/f estimates have been recognized for some time and have been incorporated into equilibrium models of production for exploited fisheries (Die et al. 1990).

Variations in catchability decrease the accuracy of C/f estimates as indices of abundance. Catchability can vary with size, sex, or other intrinsic characteristics of fish. Catchability can also vary with time of day, season, sampling site, water temperature, dissolved oxygen levels, or other environmental features that may affect the ability of the gear to capture fish or the distribution of fish relative to the gear (i.e., availability).

There are several ways to address departures from the constant catchability assumption. One approach is to stratify sampling of a stock according to the feature of interest. For example, when catchability varies with fish length, then an estimate of catchability for a stock is really the average q for all fish in the stock. As long as the length structure of the fish in the stock does not change, the average q will be a reasonable estimate for the entire stock. However, the length structure of fish in a stock is generally not constant, and catchability may best be estimated separately for individual length-classes in the sample (Seber 1982). Another approach is to adjust C/f data to account for changes in extrinsic factors such as changes in fishing power (Kimura 1981). For example, trawlers with larger engines consistently catch more fish and have greater fishing power than do trawlers with smaller engines, all other things being equal (Gulland 1977). Still another approach is to estimate catchability independently, for example, from a tagging

study that yields estimates of stock abundance through time (e.g., Paloheimo 1963). Whichever approach is taken, the factors affecting catchability should be measured and used to adjust *C/f* data.

Numerous relationships between catchability and environmental factors have been found by examining correlations between *C/f* and these factors. For instance, swimming speed of fish generally decreases at lower temperatures and fish become more vulnerable to capture by a trawl. At the same time, the spatial distribution of fish in a stock may change as temperature decreases, thereby changing their availability to the gear. However, although *C/f* may be significantly correlated with environmental conditions, catchability may not be affected by those factors (Swain et al. 2000). Careful examination and appropriate experimental designs are needed to understand the nature of the relationship between catchability and environmental factors.

7.1.2.3 Validation of Assumptions

The distinction between density and abundance is often overlooked, but the key to understanding their difference is the validity of the constant area assumption. Typically, the area occupied by a stock is assumed to remain constant. As abundance changes, the expectation is that density will also change, and *C/f* estimates will remain proportional to abundance. Although the proportional relation between *C/f* estimates and abundance is convenient, it is not universal (e.g., Crecco and Savoy 1985). In some cases, as abundance increases, fish may increase their spatial distribution and spread into adjacent nonsampled areas.

In other cases, *C/f* may exhibit "hyperdepletion" in relation to abundance (Hilborn and Walters 1992). In this situation, the rate of change for *C/f* is higher than it is for abundance. This relation is observed when *C/f* decreases faster than abundance because the most vulnerable animals are captured first, leaving behind less vulnerable individuals (Ricker 1975; Miller 1990; Hilborn and Walters 1992).

An opposite effect is "hyperstability," which occurs when *C/f* remains high even as abundance decreases (Hilborn and Walters 1992). This relationship occurs when the search for fish is highly efficient, effort is concentrated in areas of high densities, and the fish remain concentrated as abundance declines (Hilborn and Walters 1992). This has been observed among commercial (Rose and Kulka 1999) and recreational (Peterman and Steer 1981) fisheries. Aggregation of fish in a small portion of the stock's boundaries during a period of declining abundance is termed hyperaggregation (Rose and Kulka 1999). For example, anglers experienced high catchabilities of Chinook salmon during periods of low riverine abundance because both fish and anglers were concentrated in small areas of the river (Peterman and Steer 1981).

The assumption of constant catchability has been investigated for commercial fisheries because of known changes in fishing efficiency associated with vessel power, learning by crews, and technological improvements in commercial fleets through time (Fabrizio and Richards 1996). These factors increase catchability and introduce systematic error in *C/f* data. Thus, long-term *C/f* data from commercial fisheries must be adjusted prior to computation of *C/f* estimates. Without

such adjustments, increased catchability leads to overestimation of stock abundance from commercial fishery statistics. Variations in fishing power may also characterize research surveys when more than one crew, vessel, or unit of gear is used (Munro 1998). Often, such variations must be explored experimentally to derive conversion coefficients (e.g., Pelletier 1998). When C/f data are adjusted for differences in catchability there is a tendency to overestimate the variance of C/f; thus, Munro (1998) developed a method to determine when adjustments are warranted.

The assumption of independence of fish-sampling units (i.e., no interference of one unit of gear with another) has been considered in a few studies. Interference among highly aggregated gill nets has been documented (Rose and Leggett 1989). Among anglers, interference is commonly observed when total effort is high (e.g., during holidays when crowding can lead to lower catchabilities; Ricker 1975). In general, data are insufficient to determine how frequently interference may occur (Gillis 1999). Simulation studies show that when stock abundance is low, C/f estimates can fail to reflect abundance even under low levels of interference (Gillis and Peterman 1998).

In some instances, mark–recapture experiments can be conducted to estimate stock abundance (N) and relate these estimates to C/f data to obtain an estimate of the catchability: $q = (C/f)/N$. For example, electrofishing catchability has been related to largemouth bass abundance estimated by mark–recapture methods in small impoundments (Hall 1986) and lakes (Coble 1992).

7.2 SAMPLING DESIGN

The importance of sampling design cannot be overemphasized when using C/f estimates as an index of stock abundance. Catch per unit effort can vary widely because fish distributions are patchy and fish exhibit spatial and temporal variation in their distribution and activity patterns. Mean C/f estimates often have high variance (Peterman and Bradford 1987; Allen et al.1999), thereby introducing uncertainty when using C/f to assess differences in stock abundance. Thus, sampling designs that minimize variation in C/f should be used. For example, in an effort to reduce the variation in C/f, fisheries scientists often sample with the same gear, in the same locations, and at the same time each year when assessing annual changes in abundance of fishes (e.g., Fabrizio et al. 2000).

Generally, sampling designs are developed to minimize variation in C/f that is due to factors other than the true abundance of fish. Sampling locations and times are selected based on knowledge of the life history, movement, and habitat associations of a species (Pope and Willis 1996). Substantial literature is dedicated to identifying where and when to sample different species. For example, Mero and Willis (1992) assessed seasonal variation in gill-net catches of walleye and sauger from Lake Sakakawea, North Dakota, to determine when C/f was highest and the coefficient of variation of the C/f data was lowest. Similarly, variation in C/f data for largemouth bass sampled by electrofishing is minimized when sampling only shoreline areas (McInerny and Cross 2000).

A study is dependent on the objectives of the fisheries scientist, and objectives must be clearly identified. For example, objectives may be to (1) define annual trends in abundance of walleye in a prairie lake, (2) evaluate changes in relative abundance of channel catfish in a river in response to implementation of a minimum-length limit, or (3) determine effects of shoreline restoration efforts on the relative abundance of largemouth bass in a small impoundment. Each objective may require a different sampling design dependent not only on the question but also on the species and type of water body.

The experimental designs described in Chapter 3 have the potential of being used in studies in which *C/f* estimates are the response variable. Simple random sampling is generally not appropriate when *C/f* estimates are used to assess fish stocks because low precision generates *C/f* data that are too variable to detect trends or differences that may occur. Within the fisheries literature, we have found no examples of simple random sampling where *C/f* estimates were the response variable, but it is possible that situations may occur for which such a design may be applicable, particularly in small water bodies with homogeneous habitat features.

There is a strong tendency among fisheries scientists to use stratified random sampling designs, especially when assessing temporal trends in *C/f*. Four general reasons for using stratified random sampling to assess fish stocks are (see Cochran 1977) (1) calculation of *C/f* statistics may be required for different portions of a stock, such as different bays within a large lake; (2) sampling constraints may necessitate using different sampling methods in different areas, such as trawling in offshore areas and beach seining in nearshore areas of a large lake; (3) stratification may result in a gain in precision of *C/f* estimates for the whole stock; and (4) administrative convenience may require stratification in different areas, such as different states or provinces around one of the Great Lakes. Michaletz and Gale (1999) provide an example of the application of stratified random sampling where *C/f* estimates were used to assess both spatial and temporal patterns of abundance (also, see example in section 7.5.2).

A systematic sampling design is another approach that may be considered for studies based on *C/f* data. In this approach, sampling begins at a randomly selected site or time and continues at equally spaced locations or time intervals. Systematic sampling may be used effectively in rivers to gather information on the relative abundance of organisms along a gradient of environmental conditions (Karr 1999). Although several estimators exist for the variance of the mean from a systematic sample, all estimators require data from replicated systematic samples (Cochran 1977). The variance of the mean from a single systematic sample may be estimated, but estimators studied to date are biased and inconsistent (Skalski et al. 1993). For example, when mean abundance estimates are obtained from hydroacoustic surveys, systematic designs may (Simmonds and Fryer 1996) or may not (Jessop and Harvie 1990; Skalski et al. 1993) yield highly precise estimates of abundance. In general, systematic sampling provides less precise estimates of the mean than does stratified random sampling (Cochran 1977) and should be considered only when the objectives of the study are not compromised by the lower precision of systematic sampling estimators or when preliminary analyses indicate

that lower sampling costs associated with systematic designs outweigh the need for precision. Systematic sampling designs are probably most useful when combined with stratified random sampling in a two-stage approach (Schweigert et al. 1985; Chapter 3). Additional sampling designs may be applicable to assessment of *C/f* (see Chapters 2 and 3).

7.3 ASPECTS OF SAMPLING EFFORT

Design and construction of a sampling gear and factors associated with its operation contribute to variations in gear efficiency. For instance, catch rates for traps are affected by soak time and the type of bait used to attract animals (Miller 1990). In addition, gear efficiency may be affected by the interaction of captured animals and the gear itself (the saturation effect) and may vary according to life stage of the target species (Miller 1990). When conducting research or monitoring, care must be taken to standardize gear, not just in terms of the design and construction, but also in terms of the operation. Standardization also pertains to techniques used by operators of the gear. It is widely recognized that even though using the same gear, some operators can obtain a higher catch than others. By standardizing gear design, construction, and operation, fisheries scientists minimize variation in catchability and *C/f* data.

Often, preliminary sampling is necessary to identify factors associated with variation in catchability of a target species. A good example of an informative preliminary analysis is described in Bernard et al. (1991). They examined diurnal changes in catchability, optimal baiting strategies, optimal soak duration, and hoop-net size effects (among other factors) on the efficacy of hoop nets for capturing burbot in Alaskan lakes. These results were used to design surveys of stocks of burbot in 15 Alaskan lakes (Bernard et al. 1993).

7.3.1 Selectivity and Saturation

Gear performance is species and habitat specific (Choat et al. 1993). For example, light-trap selectivity for larval fish sampling depends on the attraction of different species to light, and not all taxa are equally phototaxic (Choat et al. 1993). Encounter rates of some species or sizes can be increased by deploying the gear in appropriate habitats and exploiting behavioral differences among species or life stages. However, encounters with gear do not necessarily result in capture. Fish are captured when they encounter the gear and are also retained. The probability of retention is termed selectivity. With some gear, retention of organisms will vary with mesh size and the likelihood of extrusion through the mesh. Body size, body shape, and pressure exerted by fish across the net mesh are three factors that determine the likelihood of extrusion.

Gear saturation is another factor affecting gear efficiency and catchability. Saturation occurs when the present catch reduces the potential for additional catch by reducing the number of new captures, increasing escapement, or both (Miller 1990). As a gear becomes saturated, the likelihood of capturing additional animals

decreases. Good examples of saturation effects can be found for gill nets (the presence of entangled fish may scare other fish away), longlines (as more fish are captured, the number of vacant hooks decreases, and eventually no additional fish are caught), and baited pots or traps (captured animals deplete the bait or discourage other animals from entering the trap). In traps, reduced entry is thought to be due to intimidation by trapped organisms using odor, posture, or sound to prevent entry of additional animals into the trap (Miller 1990). Longlines are notoriously prone to the effects of saturation and interspecific competition for hooks. Under these conditions, time fished is not a good indicator of true effort. An extreme example occurs when all hooks bear fish at time t but the longline is retained in place until $t + i$; in this case, C/f is biased low because true effort was (over) estimated by $t + i$. New methods have been developed to more accurately estimate effort associated with longlines based on time to capture as measured by fish-activated timing devices placed on each hook (Somerton and Kikkawa 1995).

7.3.2 Sampling Issues Specific to Gear Types

7.3.2.1 Passive Gears

Passive gears rely on the movement of organisms, either schooling or more directed migrations such as spawning, to bring organisms in contact with the gear (Hubert 1996). Schooling behavior creates density differences that affect the estimation of relative abundance. About one-fourth of teleosts are obligate schoolers and exhibit schooling behavior throughout their life, and about half of all teleost species school as juveniles (Shaw 1978). Schooling increases the vulnerability of fish to capture by fishing gear. Increased vulnerability of individuals in schools leads to less time expended in capturing fish, thus leading to biased C/f estimates. For example, catch rates with passive gears may be higher when environmental factors cause fish movements and increase their vulnerability to capture (Rose and Leggett 1989).

Because the effective area fished by passive gear is impossible to measure, effort is measured in terms of soak time. Although it may seem that longer soak times should produce greater catches, in fact, as soak time increases, the gear may become saturated and catch per unit of time will decrease. At this point, C/f does not provide an index of relative abundance (Hansen et al. 1998). For traps and pots, as soak time increases, the catch actually may decrease as more organisms escape than enter (Miller 1990). For baited longlines, saturation begins to occur as the odor concentration from baited hooks decreases (Sigler 2000).

The relation between catch and soak time is specific for particular gear types (Miller 1990) and must be determined experimentally. When designing experiments or surveys involving traps, Miller (1990) suggests the following: (1) determine the relation between catch and soak time; (2) ensure that catch rates are uniform throughout the study area; (3) standardize bait quantity and quality; (4) standardize time of setting and hauling; (5) standardize trap spacing; (6) maintain

traps in good repair; and, if following an experimental protocol, (7) randomize sampling spatially and temporally within strata.

With other passive gears such as gill nets and drift nets, catchability may be related to the visual acuity of fish, which, in turn, is affected by turbidity, light intensity, or other environmental conditions (e.g., Cui et al. 1991). Light intensity varies over a daily cycle, but lunar phase also affects light intensity. In addition, the visibility of the net depends on the color of the mesh.

Catchability of passive gears may be affected by changes in activity of fish associated with light levels. For example, most decapods are more active during dawn, dusk, or generally at night, and catchability increases during these times (Miller 1990). Some fish are also more active at night, increasing their vulnerability to passive gear. However, this relation of increased activity and light intensity may change seasonally. For example, burbot are nocturnal in spring and summer but diurnal in the fall (Bernard et al. 1993).

7.3.2.2 Active Gears

Active gears often have different catchabilities depending on light intensity. This may be due to diel vertical movements of fish or reduced visibility (Walsh 1991; Casey and Myers 1998; Korsbrekke and Nakken 1999). Catchability may also be affected by the ability of fish to escape, and that ability depends on the behavior of individuals during herding and capture (Godø et al. 1999).

Electrofishing is a highly effective active sampling gear for fish in streams and littoral zones of lakes (Reynolds 1996). Electrofishing tends to be more effective for larger fish and for species that float at the surface when stunned. Some species exhibit relatively low catchabilities to electrofishing gear. For instance, benthic fishes exhibit low catchabilities because the likelihood of seeing immobilized individuals is low, whereas other pelagic species avoid the electric field (Bohlin et al. 1989), thereby reducing their catchability. Additionally, increasing water levels can reduce electrofishing catchabilities in rivers (Bohlin et al. 1989). Standardization of electrofishing techniques is important when using C/f as an index of abundance.

7.3.3. Standardization of Effort

The appropriate units for measuring effort for a given gear can vary depending on the target species and habitat sampled. For example, electrofishing C/f is usually reported as catch per minute (e.g., Tillma et al. 1998), especially for highly abundant species or life stages such as age-0 bluegills in Midwestern lakes. At times, electrofishing C/f may be reported as catch per hour (e.g., Paragamian 1989; Miranda et al. 1996), but this usually occurs when the species of interest is rarely captured. In some cases the shoreline of a lake or reservoir may serve as the sampling unit in an electrofishing survey and C/f is measured as catch per area (e.g., fish per 100 m^2 of littoral zone) or catch per length of shoreline (e.g., fish per 100 m).

Fishing power should be standardized to maintain constant catchability. For example, *C/f* estimates from trawl surveys can change with vessel speed, so vessel speed must be constant. With all types of sampling, fisheries scientists must consider how procedural changes may affect fishing power. For example, electrofishing equipment is typically standardized by using constant voltage or constant amperage. Variations in electrical power (wattage) have caused 12–15% increases in variation in electrofishing catch rates (Burkhardt and Gutreuter 1995).

7.3.3.1 *Multiple Gears*

In some instances it is desirable to use two or more gear types to sample organisms and to combine *C/f* estimates from the different gears (Seber 1982). However, care must be taken to ensure that catchability of each gear remains constant. Although catchability may be constant through time for one gear, it may not be for another (e.g., gill nets and trap nets for Atlantic cod, Rose and Leggett 1989). At times it may be possible to calibrate *C/f* from multiple gears, but patterns of variation in *C/f* may be due to behavioral characteristics of the species under study (Methven and Schneider 1998).

7.3.3.2 *Effects of Seasonal and Daily Variation*

Catch per unit effort can change seasonally due to variations in recruitment, growth, and mortality, but such changes may not be the same for all species or for a given species in all habitats (Pope and Willis 1996; Richards et al. 1996). For example, seasonal variation in *C/f* for the virile crayfish was observed in Minnesota lakes but not for the northern clearwater crayfish in streams (Richards et al. 1996). For some species, catchability may increase temporarily during a particular season as fish increase activity levels in response to environmental factors such as temperature and photoperiod (e.g., Bernard et al. 1993; Braaten and Guy 1999; Gordoa et al. 2000). Also, seasonal patterns in *C/f* may not be observed every year due to climate variability (Gordoa et al. 2000).

When collecting *C/f* data over time to examine trends, care must be exercised to sample at appropriate times if seasonal variation in density exists. In the case of sampling during a seasonal spawning migration, annual *C/f* data will reflect relative abundance only if the seasonal timing of the migration remains the same from year to year (Fréon and Misund 1999). Thus, when sampling migratory species, the timing within the run is critical. For example, two-thirds of the annual emigration of Chinook salmon smolts occurred during a new or waning moon (Roper and Scarnecchia 1999). If using *C/f* data to compare abundance of a species from various areas, then care must be exercised to sample the areas of interest during the same time. For example, a comparative survey of bluegill abundance in Minnesota lakes found that data collected at different times of the year should not be compared because about 40% of the variation in *C/f* was explained by day of the year (Cross et al. 1995).

Daily or circadian variations in *C/f* are well known (e.g., Walsh 1991). Such variations may be related to the visual acuity of fish or diel vertical movements in lentic systems (Stoner 1991). Similarly, electrofishing catchability may (Paragamian

1989; Dumont and Dennis 1997) or may not (Maceina et al. 1995; Van Zee et al. 1996; Dumont and Dennis 1997) increase at night depending on the target species or the habitat in which the species occurs (Kessler 1999).

7.3.3.3 *Consideration of Life History and Behavior*

Catchability may be affected by life history or physiological stage of the target species. This is illustrated among decapod crustaceans. Typically, the increased activity levels of decapods during warm temperatures increase their vulnerability to capture in baited traps, but their vulnerability ceases during molting (e.g., Somers and Green 1993; Richards et al. 1996). Similarly, catchability of the American lobster in traps decreases to near zero during molting, and because males and females may molt at different times of the year, sex-specific catchability varies (Miller 1990). In Minnesota streams, catchability of northern clearwater crayfish in baited traps was highest between molting periods when animals were actively feeding (Richards et al. 1996). Thus, behavioral changes associated with life history events should be taken into account when interpreting C/f estimates.

Some of the most effective fishing gears use the behavioral responses of organisms to maximize encounter rates and retention. A pertinent example is how olfactory cues can be used to elicit behavioral responses of fish to enhance encounter probabilities. For example, when Gerhardt and Hubert (1989) baited hoop nets, the C/f of channel catfish was doubled during the postspawning period.

Some species or life stages are photopositive, so gear catchability can be increased by using light lures at night. While lighted traps and other nets may increase nocturnal catches of certain fishes or life stages, the phase of the moon may interact with catchability if fish activity varies with lunar phase. For example, Rooker et al. (1996) found that nocturnal catches of larval fishes increased significantly when lighted lift nets were used during the new moon.

The presence of predators or competitors may influence catchability. For example, in Ontario lakes, crayfish catchability in baited traps declined in the presence of rock bass and smallmouth bass and with increasing numbers of cooccurring crayfish species (Collins et al. 1983; Somers and Green 1993). These affects were noted only in lakes with relatively high abundance of predatory fishes (Collins et al. 1983).

Habitat preferences and behavior of organisms contribute to variation in C/f (Fréon and Misund 1999). Juvenile and adult fish may be distributed in areas varying in depth (Hubert and Sandheinrich 1983; Bernard et al. 1993). Individuals of some species may segregate spatially on the basis of sex (Miller 1990). If a substantial portion of a stock occupies a habitat that is inaccessible to the sampling gear, then the proportion available to the gear is likely to vary through time depending on environmental factors that alter habitat selection (Fréon and Misund 1999). For example, tidal currents in the Barents Sea have been shown to influence the vertical distribution of cod and haddock such that they are available to bottom trawls only during periods of low or decreasing tidal currents (Michalsen et al. 1996).

Habitat preferences of fish are sometimes exploited to enhance catchability. For example, some species prefer areas with cover and fisheries scientists may

purposefully sample in these areas. Often, nets are set or gear is towed in areas likely to contain fish, and the sampling locations are not truly standardized or random. This type of selection mimics the manner in which commercial fisheries operate by locating areas with potentially high densities of fish and fishing only in these areas. If the objective is to compare changes through time in an impoundment or lake, then such judgment sampling may be appropriate (Hubbard and Miranda 1988). When sampling for largemouth bass, electrofishing in areas near weed beds, stump fields, or flooded timber in the littoral zone would constitute judgment sampling (Hubbard and Miranda 1988). As long as the judgment sampling sites are constant over time (i.e., permanent sampling sites), this approach can yield an efficient means to assess temporal trends in relative abundance within a given water body.

7.3.3.4 Consideration of Gear Efficiency in Different Habitats

The efficiency of a given gear can vary substantially among habitat types. For example, electrofishing efficiency can vary widely among habitat types. In habitats with low water clarity (transparencies less than 1 m) and depths greater than a few meters, electrofishing is not very efficient (Bohlin et al. 1989). For example, Dewey (1992) reported that in turbid, highly vegetated waters, electrofishing was less efficient than were other gears because low visibility and entanglement of fish in the vegetation reduced capture efficiency.

7.3.4 The Need to Minimize Variance and Bias

One of the most common approaches to increasing the precision of *C/f* estimates is to increase the number of samples. Assuming the sampling design is appropriate (see section 7.2; Chapters 2 and 3) and catchability is constant, increasing sample size will likely increase precision. However, factors affecting catchability must remain constant during the sampling period. For instance, if catchability varies greatly with light intensity and samples are collected throughout a 24-h period without regard to this factor, then an increase in the number of samples may not improve precision. We recommend that variation in catchability be studied with respect to factors that influence the magnitude of *C/f* estimates including those that influence availability of animals to the gear and vulnerability to capture. Once these factors are known, then the value of increasing the number of samples can be determined.

In stratified random sampling designs the optimal sampling plan may not involve equal sampling among all strata, but rather optimal sampling intensity may vary according to stratum size. Minimizing the variation of *C/f* data is particularly important when these data are used to evaluate changes in relative abundance. Trends in abundance may be difficult to discern or detect when the data are highly variable (see Box 7.1).

The considerations we discussed to maximize precision of *C/f* estimates are not exhaustive, and additional considerations should be made. For example, only fully recruited age-classes should be considered in deriving *C/f* estimates;

Box 7.1 Detection of Changes in Relative Abundance with Highly Variable Catch per Unit Effort (*C/f*) Data

The time series data below illustrate the effects of highly variable *C/f* data on the ability to detect relative abundance changes for a hypothetical fish population that is declining through time. The first column is the year; the second column, population abundance (*N*), shows a decline of 5% each year; the third column gives *C/f* as 0.001*N;* the forth column shows *C/f* varying randomly by 5% or 10% above or below 0.001*N;* and the fifth column is *C/f* varying randomly by 20% or 40% above or below 0.001*N*.

Table Times series data for a hypothetical fish population.

Year	Population abundance (*N*)	*C/f*	*C/f* ± 5% or 10%	*C/f* ± 20% or 40%
1	10,000	10.0	10.5	6.0
2	9,500	9.5	8.6	13.3
3	9,025	9.0	8.1	5.4
4	8,573	8.6	7.7	10.3
5	8,145	8.1	7.7	4.9
6	7,739	7.7	6.9	4.6
7	7,351	7.4	8.1	5.9
8	6,983	7.0	7.7	5.6
9	6,634	6.6	7.2	9.2
10	6,302	6.3	6.0	8.8
11	5,987	6.0	6.3	8.4
12	5,688	5.7	5.4	3.4
13	5,404	5.4	5.7	6.5
14	5,133	5.1	4.8	3.1
15	4,877	4.9	5.4	6.9

A significant correlation ($r = 0.64$; $P < 0.001$) is observed between *N* and the *C/f* ± 5 or 10% measurement error, but the correlation ($r = 0.33$; $P = 0.23$) between *N* and the *C/f* ± 20% or 40% measurement error is not significant. Note that with the *C/f* ± 20% or 40% measurement error the *C/f* in year 15 exceeds the *C/f* in year 1.

Both levels of measurement error used in this example are within the range of what may be encountered in the field when sampling fish populations and obtaining *C/f* data. This illustrates how *C/f* measurement error can mask changes in actual abundance of fish populations.

otherwise, recruitment variability will induce variation in catchability (Seber 1982). A single gear will not capture all components of a stock in proportion to their abundance, so a key piece of information is the selectivity of the gear for various life stages or length-classes of the target species. In addition, all habitats inhabited by a species will not be equally sampled. Sampling should be conducted during a time when factors affecting catchability are similar if *C/f* data are used to compare across time or space (Richards and Schnute 1986; Miller 1990). A good approach is to focus the unit of study and define it properly, and then consider gains in precision through replication.

7.3.5 Assessment of Sample Sizes Prior to Initiation of Sampling

A power analysis allows the researcher to determine the level of effort (i.e., sample size) necessary to detect a change of a predetermined magnitude given a measure of the variability in the factor of interest (see Box 7.2). For instance, the number of trap-nights necessary to detect a 25% change in relative abundance of bluegill can be determined using an estimate of the variance of the mean *C/f*. In general, power analysis requires the assumption that *C/f* data follow a normal distribution. If the *C/f* data are not normally distributed, it becomes important to identify a transformation that yields an approximately normal distribution (Gryska et al. 1997). An example of statistical power analysis applied to *C/f* data from electrofishing samples is given in Paller (1995). In general, many samples will be necessary to detect small (<20%) differences among means, but when *C/f* is low, an even greater number of samples is necessary (Paller 1995).

7.4 STATISTICAL ANALYSIS

A common approach to analysis of *C/f* data has been to compute means and assume normal distributions of the data. However, the frequency distributions of *C/f* data are seldom normal. This is not surprising because *C/f* is a ratio estimator having catch and effort as random variables (Cochran 1977). Testing hypotheses regarding *C/f* generally involve application of statistical tests that assume the variables have a continuous scale of measure, the data exhibit a normal frequency distribution, and standard deviations are independent of the mean. Statistical analyses that require these assumptions can lead to reductions in power and misleading results when the assumptions are not met. Nonparametric statistical procedures have less restrictive assumptions regarding distributions, but it is difficult to assess the magnitude of difference between treatments or change over time based on nonparametric procedures.

7.4.1 Normalization of *C/f* Distributions

The shapes of *C/f* sample distributions can vary widely and may include normal frequency distributions and negative binomial distributions. It is common for *C/f* distributions to have standard deviations that are about equal to the mean, to be positively skewed (Moyle and Lound 1960), and to have standard deviations that increase proportionally with the mean—indications of distributions that are not normal. Among 703 published studies on larval fish abundances estimated from replicated sampling, Cyr et al. (1992) found many positive relationships between the variance and the mean, indicating that *C/f* frequency distributions were not normal in many studies. It has been suggested that the shape of *C/f* sample distribution changes with fish abundance (see Hubert 1996). At very high fish densities, *C/f* data may be normally distributed, but as fish densities decline, the mode shifts to the left and the distribution becomes skewed to the right. At relatively

Box 7.2 Power Analysis Assessment of Sampling Effort

Preliminary sampling of channel catfish in two hypothetical small impoundments is conducted with traps in early summer to obtain *C/f* data as the first step in establishing an annual monitoring program to assess temporal variation in mean *C/f*. In each reservoir, 20 traps are set at randomly selected locations, left overnight, and retrieved the following day. The following *C/f* data (i.e., fish/trap-night) and statistics are obtained for each impoundment. Each C/f value is transformed as $\log_{10}(C/f + 1)$ to assess the effects of data transformation on *C/f* statistics and estimates of needed sampling effort.

Table Catch per unit effort data and summary statistics for channel catfish in two hypothetical impoundments.

Net set and summary statistic	Impoundment A		Impoundment B	
	C/f	$\log_{10}(C/f + 1)$	*C/f*	$\log_{10}(C/f + 1)$
1	0	0	1	0.301
2	0	0	1	0.301
3	0	0	2	0.477
4	0	0	2	0.477
5	0	0	3	0.602
6	0	0	3	0.602
7	0	0	3	0.602
8	0	0	4	0.699
9	1	0.301	4	0.699
10	1	0.301	4	0.699
11	1	0.301	4	0.699
12	1	0.301	5	0.778
13	2	0.477	5	0.788
14	2	0.477	5	0.788
15	3	0.602	6	0.845
16	3	0.602	6	0.845
17	7	0.903	8	0.954
18	9	1.000	10	1.041
19	12	1.114	12	1.114
20	20	1.322	20	1.322
Mean	3.1	0.385	5.4	0.731
SD	5.2	0.422	4.4	0.254

Observation of the data and the summary statistics suggests that the *C/f* data for Impoundment A are highly skewed and substantially depart from a normal distribution; furthermore, the logarithmic transformation did little to affect the shape of the distribution. For both forms of *C/f* data the standard deviation exceeds the mean. The *C/f* data for Impoundment B are less severely skewed, and a logarithmic transformation reduces the standard deviation and creates a frequency distribution that more closely resembles a normal distribution.

Power analysis allows definition of the required sampling effort to determine specified changes in mean *C/f* at predetermined levels of significance (α) and power ($1 - \beta$) to guard against type I

(Box continues)

Box 7.2 *(continued)*

(rejecting the null hypothesis of no difference when it is true) and type II (failing to reject the null hypothesis of no difference between the means when it is false) errors (Brown and Austen 1996; Gryska et al. 1997). Power analysis is conducted by computing needed sampling effort at various levels of significance, power, and detectable effect sizes. Means and variances of *C/f* data from previous or preliminary sampling periods are used in the computations. The detectable effect size is the specified difference in two means when the null hypothesis is rejected at a specified α and β (Cohen 1969). For example, if the mean *C/f* is 3.1 fish/trap-night (as was observed for Impoundment A during the preliminary sampling) and the fisheries scientist specifies that the desire is to detect a change in *C/f* in either direction of 10% or more, the detectable effect size is 0.31 fish/trap-night. The desire is that the null hypothesis would be rejected if future sampling means differ from the preliminary sampling mean by 0.31 fish/trap night or more. How much sampling effort is needed to detect such a difference at various probabilities of type I and type II error?

Calculations are performed using the formula of Snedecor and Cochran (1989):

$$n = 2\,(z_\alpha + z_\beta)^2\,(s^2/d^2);$$

n = number of samples needed;
z_α = standard normal deviation for the probability of a type I error at a given level of probability (significance);
z_β = standard normal deviation for the probability of a type II error at a given level of probability (power = $1 - \beta$);
s = standard deviation of the preliminary *C/f* data; and
d = the detectable effect size as an absolute number.

Standard normal deviations, or z scores, are easily obtained from tables in reference books or programs in various statistical software packages.

An example computation is conducted using the logarithmic transformation of *C/f* data from Impoundment *B* because it most closely resembles a normal distribution and yields the smallest estimates of needed sampling effort. The mean $\log_{10}(C/f + 1)$ is 0.731 and SD (s) is 0.254. If it is specified that the detectable effect size is 10% of the mean, or 0.0731, α is 0.05, and β is 0.10, then $z_\alpha = 1.65$, $z_\beta = 1.28$, and

$$n = 2\,(1.65 + 1.28)^2\,(0.254^2/0.0731^2) = 207 \text{ trap nights.}$$

It is unlikely that the needed sampling effort could be achieved by practicing fisheries scientists as part of a routine monitoring program.

If the specified criteria are relaxed, lesser amounts of sampling effort are needed. For example, if it is specified that the detectable effect size is 20% of the mean, or 0.146, α is 0.10, and β is 0.20, then

$$n = 2\,(1.28 + 0.84)^2\,(0.254^2/0.146^2) = 27 \text{ trap nights.}$$

Although substantially less sampling effort is needed, the magnitude of change in *C/f* that would occur before that change is detected is doubled, and the probabilities of both type I and type II errors are doubled.

low fish densities, the most frequent catch is no fish and the distribution is likely to approximate a negative binomial probability (Power and Moser 1999). Because most fish stocks occur in relatively low densities and have patchy spatial distributions, *C/f* sample distributions that resemble negative binomial probability distributions are fairly common.

The negative binomial distribution is widely recognized as a descriptor of animal distribution patterns, and it has been argued that the negative binomial distribution is a reasonable probability distribution for the overall description of *C/f* data (Moyle and Lound 1960; Power and Moser 1999). Often, *C/f* data are characterized by a high frequency of zeroes (Bannerot and Austin 1983; Power and Moser 1999), and occasionally one or more *C/f* values are excessively large, thereby exerting excessive influence on the arithmetic mean (Pennington 1996; Kappenman 1999). The variance of the mean *C/f* is often large; thus, it is difficult to discern if mean *C/f* estimates differ among groups or over time using parametric statistical testing (e.g., a *t*-test or analysis of variance [ANOVA]; Bannerot and Austin 1983). In addition, if the mean *C/f* is small and the variance is large, the probability of observing zero catches will be high (Power and Moser 1999) if the fisheries scientist assumes that data are from a normal probability distribution. For these reasons, mean *C/f* calculated from data distributed as a negative binomial distribution may not provide a reasonable statistic for comparison of samples.

Because the occurrence of negative binomial distributions of *C/f* data have been recognized, logarithmic transformations ($y = \log e\,[x + 0.001]$ or $y = \log_{10}\,[x + 0.001]$) have been applied frequently in an attempt to normalize distributions (Bulkley 1970; Bagenel 1972) but with quite variable success. It has become common practice to apply logarithmic transformations to *C/f* data prior to conducting statistical tests and to assume that the transformation sufficiently normalizes the distribution so that test assumptions are not grossly violated. Fisheries scientists who follow this practice seldom carry out statistical tests to determine if normal distributions are achieved by the transformation. It is our experience that logarithmic transformations of *C/f* data seldom yield a normal distribution but can reduce the variance relative to the mean (see Box 7.2). Other transformations of *C/f* data have been applied in attempts to normalize the distributions (Shroyer and McComish 1998), but none was found to have wide success.

7.4.2 Appropriate Sample Statistics

Fisheries scientists occasionally use statistics other than the arithmetic mean to describe *C/f* distributions, primarily the geometric mean, median, and the frequency of occurrences of the target species among samples.

The back-transformed mean of the logarithmically transformed *C/f* data is called the geometric mean (Sokal and Rohlf 1981) and is used by fisheries scientists as a measure of central tendency for *C/f* data (Craig and Fletcher 1982; Hamley and Howley 1985; Hansen et al. 1995). It seems to be a logical expression of *C/f* when the data are logarithmically transformed for analysis. However, because the scale

is not familiar to many, it is difficult to grasp the magnitude of difference or change using the geometric mean.

The median of the *C/f* data distribution has an equal number of observations on either side of it (Sokal and Rohlf 1981) and also has been used by fisheries scientists as a measure of central tendency (Moyle and Lound 1960; Moyer et al. 1995). Moyle and Lound (1960) provided a method for computing confidence limits around median *C/f* estimates.

Another statistic applied to *C/f* data is based on enumeration of the frequency of occurrence of the target species among individual units of effort (Bannerot and Austin 1983; Counihan et al. 1999). If the frequency distribution of *C/f* data resembles a negative binomial, Bannerot and Austin (1983) suggested comparing the frequency of zero catches, which they found was a less biased index of abundance than mean *C/f*. The frequency of zero catches was more responsive to changes in stock abundance than mean *C/f* for a marine fishery (Bannerot and Austin 1983). Similarly, Counihan et al. (1999) stated that an index based on the proportion of individual units of effort when a target species is captured may have advantages over mean *C/f* because it is robust to biases and errors in sampling and insensitive to extremely high *C/f* values. Presence–absence indices generate proportional data that can be analyzed for differences among groups or over time (Sokal and Rohlf 1981). This approach circumvents issues of normal distributions associated with using the mean *C/f* and statistical tests requiring the assumption of normal *C/f* frequency distributions.

7.4.3 Bootstrap and Jackknife Techniques

Bootstrap and jackknife techniques are used to answer the same question: how precise is a particular statistic? These techniques can provide estimates of precision of *C/f* statistics (Dixon 1993). These techniques release fisheries scientists from the restrictive assumption that *C/f* data conform to a normal frequency distribution (Krebs 1989). Because both techniques compute a standard error for a statistic, they allow us to compute *t*-tests.

Bootstrap and jackknife techniques can be applied to statistics computed from *C/f* data, including the arithmetic mean, geometric mean, and median. They provide measures of the precision of the statistics and enable statistical comparisons of two samples. For statistics that are bounded in range (such as any of the three *C/f* statistics mentioned above, which are always greater than or equal to zero), these techniques may work more satisfactorily if the data are subjected to logarithmic transformation. Programs for bootstrap and jackknife routines are available, but some computational shortcuts may yield erroneous results. We recommend that fisheries scientists who want to apply these techniques work with statisticians to develop programs appropriate for their applications. It should be noted that the bootstrap and jackknife techniques will not usually yield the same answer (Dixon 1993), and there is no agreement on which technique is "better" for analysis of *C/f* data (Krebs 1989).

Some examples of using bootstrapping to obtain estimates of precision from *C/f* data can be found in the fisheries literature. In one example, Kimura and Balsiger (1985) applied bootstrapping to estimate the precision of *C/f* data for sablefish captured in pot gear off the Pacific Coast. Estimates of precision (i.e., coefficient of variation) of *C/f* data were obtained from different sampling areas and water depths and were then used to develop recommendations for the number of locations that should be sampled. Kimura and Balsiger (1985) also used the bootstrap technique to compute *Z*-statistics to estimate the statistical significance of observed differences in *C/f* between years within specified locations and depths. Similarly, Stanley (1992) used bootstrapping to estimate the variance and confidence limits for *C/f* data from four trawl fisheries along the Pacific Coast, and this information was then used to estimate the number of hauls needed to estimate mean *C/f* at ±25% of the actual rate 80% of the time ($\alpha = 0.2$). Bernard et al. (1993) applied bootstrapping to an assessment of burbot in Alaskan lakes. The bootstrap procedure was used to generate an empirical sampling distribution from which the variance was estimated for individual lakes. More recently, Smith (1997) developed bootstrap confidence limits for groundfish trawl survey estimates of mean *C/f*.

7.4.4 Comparison of Two Samples

In a classical statistical approach, a comparison of two samples is undertaken by testing the equality of means. Assuming the observations are selected randomly from a normal frequency distribution, the arithmetic mean provides a measure of central tendency. Because the mean is computed from a sample (and not the entire population), the uncertainty of the mean can be measured by the variance, another characteristic of the distribution. Under these conditions, a comparison of two samples is fairly simple: estimates of the variance are used to calculate standard errors, and confidence intervals around the means are constructed. However, as mentioned earlier, *C/f* data generally violate many of the assumptions of classical statistical approaches. For example, using traditional statistical approaches (*t*-tests and analysis of covariance) and logarithmically transformed data, only order-of-magnitude differences in larval fish abundance could be detected among sampling areas or time periods (Cyr et al. 1992).

To compare mean *C/f* values from different sampling locations or across time, an estimate of the variance (hence, standard error) is needed. Although some investigators advocate estimating the variance with regression methods (e.g., regressing catch on effort, Smith 1980) or jackknifing the variance of the ratio (Smith 1980), these approaches assume a linear relation between catch and effort. Because this assumption is not often met, we recommend seeking alternate approaches. One such approach uses maximum-likelihood methods to estimate *C/f* and its variance from the bivariate distribution of catch and effort (Richards and Schnute 1992).

Recently, Power and Moser (1999) applied an approach based on the assumption that the distribution of the catch data follows a negative binomial and variances

need not be homogeneous. Their generalized linear model permits comparison of catch rates among two or more samples and allows catch rates to vary as linear (or nonlinear) functions of exogenous variables. Generalized linear models share the same structure as general linear models (GLMs), but unlike general linear models, generalized linear models are not constrained by the assumption of normality. Using bootstrap simulations, Power and Moser (1999) demonstrated that the linear model with negative binomial errors performed better than the *t*-test in detecting differences between the means of two samples; furthermore, this was true when the *t*-test was applied to either raw data or logarithmically transformed data.

7.4.5 Analysis of Variance

7.4.5.1 Comparisons Based on Blocked Designs

When testing hypotheses based on ecological experiments, particularly field experiments, experimental units are sometimes grouped together in blocks. The purpose of this is to identify groups of similar experimental units so there is more similarity within blocks than among blocks. Blocks are not randomly assigned (as are treatments) but are either intrinsic characteristics of the experimental units (e.g., year-classes of a particular species, where the species is the experimental unit) or arbitrary segments of the experimental unit (e.g., 0.5-ha area of sandy bottom, where sections of sandy areas are the experimental units in a lake; Newman et al. 1997). In fisheries fieldwork, the blocking factors most likely to be encountered are areas fished (e.g., station or individual lakes), length-classes, or age-classes. The purpose of blocking is to improve the precision of the estimator. The example in Box 7.3 demonstrates how blocking can improve the ability to detect effects associated with factors of interest to fisheries scientists.

7.4.5.2 Other Analysis of Variance Models

Catch-per-unit-effort data are frequently collected from surveys and are intended to measure changes in abundance of fish in a stock. Depending on the design, ANOVA can be used to analyze these data under certain assumptions and constraints. Because surveys typically consist of repeated measures (e.g., C/f is estimated from a set of fixed or random stations through time), a repeated-measures ANOVA could be used to analyze such data (see Maceina et al. 1994). This approach accommodates temporal autocorrelation among observations—that is, it explicitly accounts for the fact that two observations taken closely apart in time will likely be correlated, and the correlation is likely to decrease as observations further apart in time are considered. The repeated-measures approach is recommended when (1) the study includes only fixed effects, (2) the data are balanced, and (3) the variance–covariance structure of the data conforms to a restrictive form (i.e., compound symmetry; Neter et al. 1996). Some adjustments exist for incorporating random effects into a repeated-measures ANOVA and for relaxing the assumption of compound symmetry (Neter et al. 1996). However, in many fishery surveys, C/f measures are repeated not just across time but also through space. In these cases, a different approach must be taken to accommodate correlations in

Box 7.3 Illustration of Blocked Design in One-Way Analysis of Variance (ANOVA)

An investigator wishes to determine if the abundance of bluegill differs among vegetated and nonvegetated areas of lakes. Bluegills are sampled using trap nets set within a vegetated and a nonvegetated area in each of eight lakes. In this study, the vegetation (presence or absence) is a fixed effect, the blocking factor is lake, and the response is the *C/f* of bluegill (fish/trap-night). Using the following hypothetical data set, we show how ignoring the blocking factor (lake) can lead to erroneous conclusions about the effect of vegetation on the relative abundance of bluegill.

Table Catch per unit effort data for bluegills.

	Lake							
Vegetation	A	B	C	D	E	F	G	H
Absent	63	45	30	50	80	67	48	55
Present	70	56	42	68	87	75	49	62

Program

The following SAS program is employed.

```
data bluegill;
input lake $ vegetation $ cpue;
lines;
[input data]
proc glm data=bluegill;
class lake vegetation;
model cpue=lake vegetation;
title 'One-Way ANOVA, Block Design';
prog glm data=bluegill;
class vegetation;
model cpue=vegetation;
title 'ANOVA without Blocking';
run;
```

Results

Table Results of one-way ANOVA with block design. The number of observations in the data set is 16. Abbreviations are as follows: mean square error (MSE); coefficient of variation (CV); and sum of squares (SS).

Class Level Information

Class	Levels	Values
Lake	8	A B C D E F G H
Vegetation	2	absent present

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	8	3339.000	417.375	34.20	0.0001
Error	7	85.438	12.205		
Corrected	15	3424.438			
R^2	0.975	Root MSE	3.494		
CV	5.903	*C/f* mean	59.188		

(Box continues)

Box 7.3 *(continued)*

Analysis of Variance *(continued)*					
Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Model	8	3339.000	417.375	34.20	0.0001
Lake	7	3023.938	431.991	35.39	0.0001
Vegetation	1	315.063	315.063	25.81	0.0014

Table Results of ANOVA without blocking. The number of observations in the data set is 16.

Class Level Information		
Class	Levels	Values
Vegetation	2	absent present

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model (vegetation)	1	315.063	315.063	1.42	0.253
Error	14	3109.375	222.098		
Corrected total	15	3424.438			
R^2	0.092	Root MSE	14.903		
CV	25.179	C/f mean	59.188		

Interpretation

A cursory examination of the data reveals that within a lake, bluegill *C/f* is higher in vegetated areas than in nonvegetated areas. However, on closer inspection, we see that bluegill *C/f* in nonvegetated areas of lakes A and F was just as high as *C/f* in vegetated areas of lakes H and D. Results from the blocked ANOVA indicate that vegetation significantly affects the mean relative abundance of bluegills (vegetation is a significant factor in the ANOVA). In addition, the relative abundance of bluegills varied significantly among lakes (lake is a significant factor in the ANOVA). Thus, by blocking the design and accounting for lake to lake differences in the relative abundance of bluegills, the investigator was able to examine the effect of vegetation on bluegill abundance within lakes. Incidentally, if the investigator had randomly sampled one set of lakes (say lakes L, M, N, O, P) to estimate bluegill *C/f* in vegetated areas, and another set of lakes (say lakes Q, R, S, T, U) to estimate bluegill *C/f* in nonvegetated areas, then lake would not be a blocking factor because the treatment (vegetated versus nonvegetated) was randomly applied across lakes. In our blocked design, the treatment (vegetated versus nonvegetated) occurred in the same lake; because of this, the response to the two treatment levels was measured from the same lake. In this respect, data from a blocked design could also be analyzed as a paired *t*-test (see Sokal and Rohlf [1981] for further discussion).

In the second analysis, blocking is ignored, and it is not possible to detect the effect of vegetation on bluegill abundance. This occurs because in the unblocked design, the variation associated with lakes is considered part of the error term (compare the sums of square error terms from both models). By ignoring the blocking effect, we would interpret the results of this ANOVA in the following manner: on average across all lakes sampled, variation in relative abundance of bluegills is not affected by the presence of vegetation.

space in addition to serial correlations in time. Fabrizio et al. (2000) demonstrated how the mixed-model procedure (MIXED) in SAS (SAS Institute 1998) can be used to study changes in fish abundance from a complex repeated-measures design. Procedure (PROC) MIXED was used to fit a linear model with correlated errors to a 20-year time series of catch data from Lake Michigan, but the approach is applicable to other fixed-station fishery surveys. The linear model used by Fabrizio et al. (2000), $y = X\beta + e$, is similar to a GLM (one fit with SAS' PROC GLM) except that in the GLM, the vector e is a vector of independent random variables, and in the linear model with correlated errors, e is a vector of possibly correlated random errors with covariance matrix R (Littell et al. 1996). In this notation, y is a vector of observations, X is a matrix of fixed effects values, and β is a vector of fixed effects coefficients. Another difference between the two models is that instead of simply modeling the mean and a single variance of y (the GLM), the mean, variance, and covariance of y are modeled in the linear model with correlated errors.

Procedure MIXED is a flexible approach that works well for unbalanced data. It can be used to fit a variety of models, including mixed models that contain both fixed effects and random effects: $y = X\beta + Z\mu + e$, where y, X, β, and e are defined as before, Z is the design matrix (usually a matrix of 0s and 1s), and μ is a vector of random effects parameters (Littell et al. 1996). To use PROC MIXED, the investigator must identify and specify the type of the variance–covariance structure that defines the error term of the model. Version 6.12 of SAS offers about 20 options for the structure of the variance–covariance matrix, and Wolfinger (1996) discusses useful variance–covariance structures for models fit to repeated-measures data. A new procedure, PROC NLMIXED, available with SAS version 7 and higher, can fit nonlinear models using likelihood-based methods (Wolfinger 1999).

7.4.6 Nonparametric Alternatives to Analysis of Variance

Because nonparametric tests do not require the assumption that the data come from a normal distribution, these tests have been recommended when assumptions of parametric statistics cannot be met. Standard parametric tests such as the *t*-test and ANOVA have nonparametric counterparts that can be easily implemented once the raw data are rank transformed. For example, the counterpart to the *t*-test is the Mann–Whitney *U*-test or Wilcoxon's rank-sum test. However, it should be noted that the null hypothesis of a nonparametric test is not equivalent to that tested by the analogous parametric method. In the case of the Mann–Whitney or Wilcoxon test, the null hypothesis is that two distributions are identical. Nonparametric tests will detect differences not just in central tendency but also differences in the spread or shape of distributions (Johnson 1995). When performing a nonparametric comparison of two samples, a significant test result provides no information on whether the difference is due to the mean, variance, shape, or some other characteristic of the distribution (Johnson 1995).

Although nonparametric techniques obviate the need for normally distributed data and appear to be well suited for analysis of C/f data, relations between C/f and other variables must still be considered. For example, Richards and Schnute (1986) found that C/f data had to be standardized prior to applying a Kruskal–

Wallis test when evaluating the effects of sea surface condition, time of day, and tidal phase on *C/f*. These data were standardized by working with observations from a restricted time (only when sea surface conditions were calm) when catches were thought to be most reliable.

With nonparametric tests investigators still must apply a priori significance levels to tests and consider the trade-offs between type I and type II errors. Because some fisheries scientists perceive the power of nonparametric tests to be low, larger alpha levels (e.g., $\alpha = 0.10$ or 0.15) are sometimes used in significance testing. However, nonparametric tests often have as much power as their parametric counterparts.

7.4.7 Time Series Analysis

Fish abundance measures that are estimated repeatedly through time are typically examined for patterns of change through time by means of regression analysis. The relation between observations close in time may be similar—that is, values in a given year may be influenced by values in the previous year. This relationship generally decreases with increasing time intervals. Such data are said to exhibit positive autocorrelation. The presence of autocorrelation (or serial dependence) in fish abundance data compromises statistical interpretation of correlation and regression analyses that may be undertaken to relate changes in fish abundance to environmental or biological variables (Pyper and Peterman 1998). The reason is that most parametric statistical tests assume independence (correlation equals 0) of observations. Hypothesis tests on autocorrelated data require adjustments to the degrees of freedom to reflect the lack of independence among observations (Pyper and Peterman 1998).

For predictive modeling or exploratory analyses, autocorrelated data must be transformed. Several transformations have been used with autocorrelated data, including smoothing, first-differencing, and prewhitening. A smoothed data series results from the computation of a series of weighted averages from nearby points. A simple smoothing technique is the running average (moving average). Smoothing is an effective transformation for removing high-frequency variation, which appears as rapid changes over short time scales. An example of high-frequency variation is measurement error. Sometimes it may be desirable to remove the signal associated with slow, long-term changes from a time series of *C/f* data. These changes are typical of low-frequency variation and first-differencing or prewhitening may be used to transform the data series (Pyper and Peterman 1998). In first-differencing, the observation at time $t - 1$ is subtracted from the observation at time t. Prewhitening is typically applied when the analyst wishes to relate *C/f* data to one or more environmental variables. For example, to determine if the pattern of variation in *C/f* is associated with the pattern of variation in temperature data, the time series of temperature data are modeled with an appropriate time series model. That model is then applied to the series of *C/f* data to prewhiten the *C/f* series. Model identification, parameter estimation, and diagnostic checking procedures are beyond the scope of this chapter but are

well described in Box and Jenkins (1976). Additional information on smoothing, first-differencing, and prewhitening is available in Pyper and Peterman (1998), along with an effective method for adjusting the degrees of freedom for statistical testing of autocorrelated data.

It should be noted, however, that the decision to employ any of the transformations should be taken with extreme caution. Pyper and Peterman (1998) point out that if low-frequency variation is removed from a time series of data, the effect of a slowly changing variable on the dynamics of the population will be difficult to detect, but the effect of a quickly changing variable (high-frequency variability) will be well detected if it is the dominant source of covariation. When the dominant source of covariation is low frequency, Pyper and Peterman (1998) recommend adjusting the degrees of freedom because this approach has greater statistical power.

7.4.8 Assessment of Relationships between *C/f* and Other Variables

Regression analysis is commonly applied to *C/f* data to make predictions. For example, Isbell and Rawson (1989) found that *C/f* of walleye captured in experimental gill nets was a predictor of angler catch rates in western Lake Erie. Mean *C/f* in gill nets was used as the predictor variable, and mean *C/f* among anglers was used as the response variable. Similarly, Shroyer and McComish (1998) predicted the future *C/f* of quality-length (>200 mm total length) yellow perch based on *C/f* of stock-length (>130 mm total length) yellow perch in trawl samples in Indiana waters of Lake Michigan.

Regression analysis has also been used to predict *C/f* of fish from various habitats. For example, Johnson and Jennings (1998) assessed the habitat associations of small fishes around islands in the upper Mississippi River based on *C/f* as an index of abundance. They predicted *C/f* from measures of habitat. Similarly, Irwin et al. (1997) assessed the habitat associations of age-0 largemouth bass along the shoreline of a large reservoir. Regression analysis was used to determine if measured habitat features accounted for variation in *C/f* of age-0 largemouth bass among 43 discrete shoreline sections.

These approaches are fairly straightforward when habitat variables are static characteristics of the environment. However, when habitats or environmental conditions are dynamic (such as salinity and water temperature), it is advisable to remove the high frequency (<24 h) variation of these dynamic physical variables prior to using such data in a regression (Rose and Leggett 1989).

Fishery scientists have long recognized the problem of using *C/f* as a predictor variable in regression analysis (see Ricker 1975) because the predictor variable is assumed to be known accurately. This is an untenable assumption with *C/f* data that are fraught with measurement errors. Measurement errors in *C/f* data are a good example of what statisticians call the "errors in variables" problem. In general, errors in variables tend to flatten the probability density function and increase dispersion; such changes lead to upwardly biased variance estimates and downwardly

biased estimates of the mean (Chesher 1991). Ricker (1975) demonstrated how to use functional regression analysis to estimate regression parameters from *C/f* data, but Ricker's approach is now considered ad hoc (Hilborn and Walters 1992). A good review of the fisheries-related work on the errors in variables problem is provided in Hilborn and Walters (1992). Although errors in variables are unavoidable in fisheries modeling, the magnitude of the bias associated with the errors in variables problem can be investigated using Monte Carlo simulation techniques (Hilborn and Walters 1992). When modeling the stock–recruitment relation, even small measurement error (mean = 0 and SD = 0.2) can lead to erroneous conclusions about the nature of the relationship (Walters and Ludwig 1981). Several other approaches have been proposed to work with *C/f* data subject to the errors in variables problem, including techniques for data containing measurement error in both the dependent and independent variables (e.g., Richards and Schnute 1986; Kimura 2000).

Most linear models, including regression analysis, do not address extra-Poisson or extra-binomial variation, and, because of this, such models may not provide reliable confidence intervals or significance tests for parameters of interest (Casey and Myers 1998). As discussed in section 7.4.3, the jackknife or bootstrap approach may be useful in estimating precision of regression parameters. Another approach is to use simulation modeling. In addition to these techniques, the randomization approach may be used to estimate confidence intervals for a regression parameter, particularly if the standard significance levels or standard errors of the parameter estimates are not reliable (Casey and Myers 1998).

Detection of trends in *C/f* data has been recently pursued with regression tree methods (Watters and Deriso 2000). Observed trends were ascribed to changes in catchability and to actual changes in abundance. This regression tree application required estimation of 139 parameters for 30 years of monthly data on bigeye tuna from the Pacific Ocean, so it is not likely to be appropriate for short time series. Regression trees are useful in examining the interaction of factors such as area (e.g., latitude–longitude grids) and time (e.g., specific months) and may be more parsimonious (fewer parameters) than GLMs or spatially explicit models that account for variations in environmental conditions (Watters and Deriso 2000).

7.5 INTERPRETATION AND APPLICATION OF *C/f* STATISTICS

Monitoring changes in fish stock abundance through time is a costly activity undertaken typically by federal and state agencies. The goal of these surveys is to provide long-term information on the status of species so that changes in abundance can be detected. These surveys require an investment in gear and personnel as well as an institutional commitment to multiyear support. Analysis of data collected from monitoring surveys is often difficult due to the nature of the *C/f* data, which may also reflect the vagaries of the weather and the reliability of the equipment and gear.

7.5.1 Example of a Temporal Monitoring Program

In this section, we illustrate how to assess patterns of change in *C/f* over time. We use mean *C/f* of bay anchovy from two regions of a mid-Atlantic estuary as an example (see Box 7.4). The *C/f* data were collected from fixed stations along a salinity gradient in the estuary and are considered repeated measures. To begin analysis, components that represent the treatment structure and those representing the design structure must be designated so that the appropriate statistical model can be identified. In general, the treatment structure refers to the components of the experimental design whose effects are of interest. In a temporal monitoring program, this typically includes effects of time and may also include effects of other factors such as region (in this example), habitat type, or habitat manipulation. Design structure components are elements necessary to conduct or construct the experiment and assist in addressing the components of the treatment structure. Randomization and blocking are two examples of design structure elements. Statistical testing is focused on components of the treatment structure and generally not on the design structure. In this example, we use PROC MIXED to examine changes in mean *C/f* between two regions in the estuary and through time. The treatment structure consists of two fixed effects: region (bay versus river) and time. The design structure incorporates a random component (i.e., stations, which are nested within regions).

In a typical repeated-measures design, the response is measured from the same subject multiple times. In this temporal monitoring example, the "subjects" are stations from which *C/f* data were sampled; the repeated measures are *C/f*. To designate stations as the experimental units, we must restrict our inferences to the two regions sampled and assume that the *C/f* data from the two regions are uncorrelated. (Technically, because our treatment [region] cannot be applied to each station, the stations are not independent experimental units, and further investigation of the dependency among stations may be pursued prior to modeling. However, this line of inquiry is beyond the scope of this example and will not be illustrated here.)

Although PROC GLM provides several approaches for analyzing repeated-measures data (using split-plot or multivariate approaches), we advise against using this SAS procedure to analyze fisheries survey data such as these. The split-plot and multivariate approaches to analysis of repeated measures are useful for analysis of some ecological data, but in those cases, the experimental units are typically groups of organisms or samples that can be randomly assigned a treatment. In addition, the split-plot repeated-measures design has been used to analyze fisheries data obtained at various points in time before and after an experimental manipulation, where the manipulation affects all possible sample sites within a water body (e.g., to test the effects of vegetation removal on length structure of fish in a small lake; Maceina et al. 1994). The split-plot approach is well suited to surveys of individual water bodies from which replicates are taken and treatments can be applied to each experimental unit (station and replicate). The split-plot approach

Box 7.4 Analysis of *C/f* Data from a Temporal Monitoring Program

Bay anchovies were sampled with an otter trawl every 6 months from November 1996 through May 2000 for a total of eight sampling periods. Trawl tows were taken at 13 randomly selected sites in the river and 11 randomly selected sites in the bay. A single 5-min tow was completed at each station during each sampling period, the area swept was calculated, and *C/f* was computed as the number captured per unit of area swept. The analysis was designed to address two questions. Did mean *C/f* change through time? Did mean *C/f* differ between the river and bay?

The data set contained 192 otter trawl samples, but 134 samples contained no bay anchovies. Because the *C/f* data contained many zeros, had a large coefficient of variation, and had a variance that exceeded the mean, the *C/f* observations were transformed as $\log_e(x + 0.0001)$. River sites were coded as 0 and bay sites as 1 in a variable termed region. Sampling periods were coded 1 through 8 in a variable termed time. Sampling sites were coded 1 to 13 in the river and 1 to 11 in the bay in a variable termed station.

We first identified the type of covariance matrix that best described the random component (i.e., stations sampled repeatedly) in this study. The keyword TYPE identifies the covariance matrix in the repeated statement of PROC MIXED. Because the *C/f* data are repeated measures, temporal correlation among samples from a site may occur with the correlation decreasing as the time interval increases. If the decrease is exponential, then the covariance structure can be modeled using a first-order autoregressive structure (TYPE = AR[1]). If the correlations are equal across time intervals, the covariance can be modeled using a compound symmetric structure (TYPE = CS). For maximum flexibility in modeling the correlations, an unstructured covariance matrix may be specified (TYPE = UN). We fitted these three covariance structures to the data and compared the model fit using a likelihood-based criterion, Akaike's Information Criterion (AIC; Littell et al. 1996). The best model provides the smallest AIC, which is reported in a SAS output under "Fit Statistics." We found TYPE = AR(1) provided the best description of the covariance structure for these data.

Procedure MIXED was used to test for region and time effects. Before fitting the model to the data we examined the interaction plot to determine the relation between region and time.

Program

The following SAS program was employed.

```
   /* Plot the mean data through time for each region — Interaction Plot */
   proc univariate noprint data=cpue.anchovy;
   var lcpue_a;
   by region t;
   output out=anc_out mean=ybar;
   proc plot data=anc_out;
   plot ybar*t=region/box;
   proc mixed data=cpue.anchovy;
   class region station t;
   model lcpue_a = region t region*t / outp=predict1;
   random station(region) / s;
   /* This is the error term for testing the region effect */
   repeated / subject=station type=ar(1);
   ods listing exclude solutionr;
   ods output solutionr=randsoln;
   title 'Proc Mixed Results for Anchovy Data';

   /* Assess residuals for approximate normality at the whole plot (region)
level. Actually, these are estimated random effects. */
   proc univariate data=randsoln plot normal;
   var tvalue;
   probplot tvalue / normal;
   title 'Residuals — Estimated Random Effects';
```

```
/* Assess residuals for approximate normality at the subplot (time) level. */
proc univariate data=predict1 plot normal;
var resid;
probplot resid / normal;
title 'Residuals — From Predicted'; run;
```

In the PROC MIXED statements, note that region, station, and time are class variables. The model statement includes region, time, and the region × time interaction, which we suspect may be significant. The output option (outp=) in this statement specifies that the output SAS data set is called predict1. All the modeled effects are fixed effects. The single random effect identified in the random statement was station nested within region. The s option in the random statement requests estimation of the solution, which will be used to evaluate normality of random effects. The repeated statement defines the subjects of this repeated-measures analysis, which were stations. Two statements are included to control the output delivery system (ods). The first (listing exclude solutionr) suppresses the listing of the model estimates for each of the 24 random effects. The ods output statement places the random effects estimates in a SAS data set (randsoln). Portions of the output from the PROC MIXED analysis are given below.

Results

Table Portion of results for mixed-model ANOVA of bay anchovy *C/f.* Abbreviations are first-order autoregressive structure (AR[1]); time (*t*); numerator (Num); and denominator (Den).
Covariance Parameter Estimates

Covariance Parameter Estimates		
Covariance parameter	Subject	Estimate
Station (region)		0.5511
AR(1)	Station	0.2283
Residual		4.4531

Type 3 Tests of FIxed Effects				
Source	Num *df*	Den *df*	*F*-value	*P* > *F*
Region	1	22	2.81	0.1077
t	7	154	7.01	<0.0001
Region*t	7	154	2.79	0.0093

Interpretation

The interaction plot indicated that mean *C/f* may be changing differently in the two regions, hinting that an interaction of time and region may occur (see Quinn and Keough 2002). The output reveals the estimate of the variance of the mean *C/f* among stations nested within regions (0.5511). The correlation coefficient (0.2283) indicates a relation in *C/f* between adjacent 6-month sampling times. The *F*-test identifies a significant interaction of region and time (region*t, *F* value = 2.79, *P* > *F* = 0.0093), as suggested by the interaction plot, making interpretation of region and time effects difficult to assess. The interaction indicates that processes contributing to changes in mean *C/f* differ in the river and bay. Options for further analysis may be to test for the presence of trends in *C/f* for each region separately or omit the distinction between regions and assess the collective data set for temporal trends.

cannot be applied to our example data set because each whole plot (station in bay or river) would need to contain both treatments (regions). Furthermore, for some moderately complex designs, the GLM procedure is known to compute incorrect standard errors. When analyzing repeated-measures data with PROC GLM, the analyst should be aware of potential problems that occur with missing data, especially when the random statement is used or when modeling multivariate contrasts (Littell et al. 1996). We note that for most field studies, missing data are common due to experimental failures, weather-related loss of sampling opportunity, and other unplanned problems. Procedure MIXED was developed to address some of the limitations of PROC GLM for modeling data from experiments that incorporate random components either in the design structure, treatment structure, or both.

7.5.2 Example of an Assessment of Spatial Patterns

Fisheries scientists are frequently faced with the need to identify spatial patterns in fish distributions. Given the target species and type of water, an approach may entail the use of *C/f*. A hypothetical example of an assessment of spatial distribution patterns of a fish stock based on *C/f* data may be illustrated by means of yellow perch in a stratified Midwestern lake (Box 7.5). The fisheries scientists wanted to know if the midday depth distribution patterns of yellow perch differed between June and August. A stratified random sampling design was used. The strata were the two sampling months and five sampling depths (2.5, 5, 10, 15, and 20 m). Within each depth stratum, three redundant locations were randomly selected for sampling during each sampling month. Gill nets were set perpendicular to the shore 1 h before midday and retrieved 2 h later. Fish/net/h was used as an index of yellow perch relative abundance. Two-way ANOVA (i.e., PROC GLM) was used to assess variation in *C/f* among months and sampling depths (see Box 7.5). No significant difference in mean *C/f* was found between June and August, but mean *C/f* differed significantly among sampling depths. Sampling month and depth exhibited no significant interaction, indicating that patterns in the depth distribution of yellow perch were similar in June and August. The mean *C/f* was greatest at the 10-m sampling depth during both months. The data suggest that yellow perch are most abundant between 5 and 15 m with lower numbers near shore (2.5 m) and at 20 m.

7.5.3 Example of the Use of a Regression Estimator

A common problem encountered by fisheries scientists is the need to identify relations between fish abundance and habitat features. Such relationships help define habitat features needed by a species, determine habitat quality, or define the likely responses of aquatic organisms to improvement or degradation of habitat (Orth and White 1999; Summerfelt 1999). Because experiments involving manipulation of habitat are difficult, time-consuming, and expensive to conduct,

Box 7.5 Assessment of Depth Distribution Patterns of Yellow Perch Based on *C/f* Data

Data on *C/f* (fish//net/h) of yellow perch captured with gill nets in a Midwestern lake were obtained during midday at five depths during 2 months with three randomly selected sites sampled at each depth during each month. A two-way ANOVA was used to assess effects of sampling depth and month as well as the interaction between the two.

Table Data on *C/f* of yellow perch in a Midwestern lake. Three sites were sampled at each of five depths during two dates (June = 1 and August = 2).

Month	Depth (*m*)	*C/f*	Month	Depth(*m*)	*C/f*
1	2.5	2	2	2.5	0
1	2.5	4	2	2.5	2
1	2.5	7	2	2.5	0
1	5.0	6	2	5.0	10
1	5.0	10	2	5.0	10
1	5.0	12	2	5.0	9
1	10.0	8	2	10.0	13
1	10.0	12	2	10.0	40
1	10.0	33	2	10.0	46
1	15.0	10	2	15.0	6
1	15.0	10	2	15.0	5
1	15.0	17	2	15.0	12
1	20.0	0	2	20.0	0
1	20.0	2	2	20.0	0
1	20.0	3	2	20.0	0

Program
The following SAS program was employed.

```
data yelperch;
input month depth catch;
cards;
[input data]
proc glm;
class month depth;
model catch=month depth month*depth;
proc sort;
by month depth;
proc means mean stderr;
by month depth;
var catch;
```

Results
The ANOVA indicated that *C/f* varied significantly among sampling depths, but no significant difference occurred among sampling months and no interaction occurred.

(Box continues)

Box 7.5 *(continued)*

Table Two-way ANOVA of *C/f* data for yellow perch. There were 30 observations.

Class Level Information		
Class	Levels	Values
Month	2	1 2
Depth	5	2.5 5 10 15 20

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	9	2659.633333	295.514815	5.48	0.0008
Error	20	1079.333333	53.966667		
Corrected total	29	3738.966667			
R^2	0.711328	Root MSE	7.346201		
CV	76.25814	Catch mean	9.633333		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Month	1	9.633333	9.633333	0.18	0.6772
Depth	4	2249.800000	562.450000	10.42	<0.0001
Month*depth	4	400.200000	100.050000	1.85	0.1582

The mean *C/f* and SE for each sampling month and depth are given below so that comparisons can be made between months.

Table Mean *C/f* and SE for yellow perch data.

	June		August	
Depth (*m*)	Mean *C/f*	SE	Mean *C/f*	SE
2.5	4.33	1.45	0.66	0.66
5.0	9.33	1.76	9.66	0.33
10.0	17.67	7.75	33.00	10.15
15.0	12.33	2.33	7.67	2.19
20.0	1.66	0.88	0	0

fisheries scientists often rely on a regression analysis to make inferences on the relationships between fish abundance and habitat features. Many examples exist in the literature in which a regression analysis was used to identify habitat features that might be related to fish abundance, particularly as measured by *C/f* (Irwin et al. 1997; Tillma et al. 1998; Braaten and Guy 1999). Cause and effect relations between measured habitat features and *C/f* cannot be proven by means of a

regression analysis, but substantial insight and predictive capabilities can be generated if studies are designed properly and analyses are conducted carefully.

We provide a hypothetical example to illustrate application of a regression design to examine habitat quality when *C/f* data are used to index fish abundance. In this example, fisheries scientists wanted to identify habitat features that may affect the abundance of age-0 smallmouth bass in shoreline areas and to develop the ability to predict abundance from measured habitat features. Box 7.6 contains a hypothetical data set for age-0 smallmouth bass and demonstrates how habitat features along the shoreline of a small natural lake may be associated with *C/f* of age-0 fish. Twenty sites representing the range of shoreline habitats were selected from the periphery of the lake. At each site, a 50-m segment was sampled between the shoreline and the 1-m-depth contour in late July. The mean bottom slope, proportion of the bottom composed of gravel–cobble substrate, and proportion of the bottom covered by aquatic macrophytes were measured at each site. Over each 50-m segment, one pass was made at night with a boat-mounted electrofishing unit, and all age-0 smallmouth bass captured during the pass were counted. The *C/f* (number/50 m of shoreline) was used as an index of age-0 smallmouth bass abundance at each site.

Pearson's correlation coefficients were computed to assess relations among the three habitat features. A significant correlation was found between the proportion of gravel and the bottom slope indicating that these two independent variables may be redundant measures of the same ecological feature, which may or may not be important to age-0 smallmouth bass. Linear regression analysis was next used to evaluate relationships between *C/f* and each habitat feature. A $\log_{10}(x + 1)$ transformation of the *C/f* data was made to improve the linear relation. Gravel accounted for significant variation in *C/f* (see Box 7.6). When *C/f* was transformed, the coefficient of determination (r^2) increased and the probability (P) that the relation was due to chance declined, indicating a more linear relationship. Vegetation, which was not correlated with gravel, did not account for additional variation in *C/f* when included in a multiple-regression model with gravel. Based on the high coefficient of determination ($r^2 = 0.84$), the relation between gravel and $\log_{10}(C/f + 1)$ would be judged as a good predictor of age-0 smallmouth bass abundance in shoreline areas. However, fisheries scientists using this model should note that cause and effect relations were not defined by the regression model. In this case, it is likely that small gravel is a suitable spawning substrate for smallmouth bass and that age-0 fish are abundant where spawning was concentrated not necessarily because gravel is a needed habitat feature for age-0 fish. The model should be tested with several independent data sets before it is used for management decisions.

7.6 SUMMARY

When assessing temporal or spatial trends in fish stocks, freshwater fisheries scientists often use *C/f* as an index of relative abundance. Underlying assumptions associated with the relationship between *C/f* data and actual population abundance

Box 7.6 Regression Analysis to Assessment of Habitat Features when *C/f* Data Are Used As the Response Variable.

This hypothetical problem focuses on defining the habitat features affecting the densities of age-0 smallmouth bass around the shoreline of a natural lake in the Midwestern United States.

Table Data: for 20 sites sampled along 50-m segments of shoreline of a Midwestern natural lake in late July. The mean bottom slope, proportion of the bottom composed of gravel–cobble substrate, and proportion of the bottom covered by aquatic macrophytes were measured at each site, and one pass was made at night with a boat-mounted electrofishing unit for age-0 smallmouth bass.

C/f (fish/50 m)	Gravel (%)	Vegetation (%)	Slope (%)
0	0	0	7.3
5	5.2	10.1	1.5
1	0.3	1.1	2.2
10	5.5	13.6	1.2
12	6.7	7.3	1.7
0	0.8	10.9	5.3
1	0.1	10.0	8.3
3	1.9	0	3.0
25	8.3	0	1.5
0	0.5	5.0	4.3
98	11.1	1.8	1.5
2	0.9	4.0	4.3
15	6.6	1.3	1.1
60	10.0	11.0	1.4
0	3.0	0	5.9
1	1.0	7.7	8.0
7	5.9	1.0	1.3
1	1.0	4.0	4.4
0	4.0	0	8.7
5	4.3	7.6	1.4

The data were entered into a spreadsheet, and the *C/f* data were transformed [$\log_{10}$ (*C/f* + 1)] to create a second response variable. Correlations were assessed among the habitat features to avoid inclusion of redundant variables in regression models. Simple linear regressions were computed between each of the three habitat features and each of the two measures of relative abundance.

Program

The following SAS program was employed.:

```
data bass;
input cpue logcpue gravel vegetation slope;
gravveg=gravel*vegetation;
cards;
[...input .data...]
proc corr;
var gravel slope vegetation;
proc reg;
model cpue=gravel;
model cpue=slope;
model cpue=vegetation;
model logcpue=gravel;
model logcpue=slope;
model logcpue=vegetation;
```

Results

The correlation analysis indicated that the proportion of gravel and the shoreline slope were negatively correlated ($r = -0.649, P = 0.002$).

Table Pearson's correlation coefficients (*r*), ($n = 20$) for habitat variables and the probability of a greater $|r|$ under the null hypothesis that rho = 0.

	Gravel	Slope	Vegetation
Gravel	1.00000	0.03204 0.8933	–0.64916 0.0020
Slope	0.03204 0.8933	1.00000	–0.09309 0.6963
Vegetation	–0.64916 0.0020	–0.09309 0.6963	1.00000

The regression analyses indicated that the strongest linear relationship between relative abundance and a measured habitat feature occurred between $\log_{10} C/f$ and the proportion of gravel ($r^2 = 0.837$, $P < 0.0001$):.

Table Regression analysis of $\log_{10}(C/f + 1)$ of smallmouth bass and the proportion of gravel.

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model (gravel)	1	5.93149	5.93149	92.59	<0.0001
Error	18	1.15315	0.06406		
Corrected tTotal	19	7.08464			
R^2	0.8372	Root MSE	0.25311		
Adjusted R^2	0.8282	Dependent mean	0.66490		
CV	38.06690				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	0.04083	0.08608	0.47	0.6410
Gravel	1	0.16189	0.01682	9.62	<0.0001

Interpetation

The model that used the untransformed *C/f* as the response variable and gravel as the dependent variable was significant (<0.0001), but the amount of variability in *C/f* accounted for by gravel was substantially less ($r^2 = 0.636$).

Multiple-regression models were computed with both the proportion of gravel and the proportion of vegetation as habitat variables, as well as with the interaction term (gravveg = gravel × veg). The following SAS program was employed:.

```
Proc reg;
model logcpue=gravel vegetation;
model logcpue=gravel vegetation gravveg;
```

Results

Neither the proportion of vegetation nor the interaction term was significant. Slope was not included in the multiple-regression model because it was significantly correlated with the proportion of gravel. Therefore, the regression analysis suggests that the relative abundance of age-0 smallmouth bass can be predicted from the proportion of gravel along shoreline areas.

must be considered, or *C/f* can be a misleading indicator of abundance. While there are several assumptions to be considered, the assumption of constant catchability may be the most critical and commonly violated. Substantial effort should be made to assure constant catchability in management assessments and research designs. In order to minimize uncontrolled sources of variation (error) in *C/f*, stratified random and systematic sampling designs are commonly used. Such designs incorporate standardization of gear and effort and identification of sampling times and locations. Assessing the extent of variability in *C/f* with a particular design and identifying the sampling effort required to detect changes over time or to detect differences among sampling sites through preliminary sampling are necessary components of management and research efforts.

A major problem in the application of *C/f* sampling data is that the distribution is seldom normally distributed. Negative binomial distributions are common among *C/f* data sets, but they cannot be assumed to occur. A variety of descriptive statistics have been used to characterize the distribution of *C/f* data, but none are universally applicable. The power of classical statistical methods is substantially reduced when *C/f* data are incorrectly assumed to be normally distributed. Furthermore, changes in *C/f* over time or among different locations may not be detectable when, in fact, differences in fish abundance exist. However, recent applications of general linear models and mixed models that incorporate temporally and spatially autocorrelated errors into *C/f* analyses provide substantial promise for more powerful analyses. Similarly, advances in regression analyses beyond classic least-squares regression are providing better descriptors of relations between *C/f* and other variables. The historic and emerging statistical methods described in this chapter have utility in management and research; however, users of these techniques are advised to seek consultation of professional statisticians to assure that the most appropriate analytical methods are used and to avoid misleading results or interpretations.

7.7 REFERENCES

Allen, M. S., M. M. Hale, and W. E. Pine, III. 1999. Comparison of trap net and otter trawls for sampling black crappie in two Florida lakes. North American Journal of Fisheries Management 19:977–983.

Bagenel, T. B. 1972. The variability in the number of perch, *Perca fluviatilis* L., caught in traps. Freshwater Biology 2:27–36.

Bannerot, S. P., and C. B. Austin. 1983. Using frequency distributions of catch per unit effort to measure fish-stock abundance. Transactions of the American Fisheries Society 112:608–617.

Bernard, D. R., J. F. Parker, and R. Lafferty. 1993. Stock assessment of burbot populations in small and moderate-size lakes. North American Journal of Fisheries Management 13:657–675.

Bernard, D. R., G. A. Pearse, and R. H. Conrad. 1991. Hoop traps as a means to capture burbot. North American Journal of Fisheries Management 11:91–104.

Bohlin, T., S. Hamrin, T. G. Heggberget, G. Rasmussen, and S. J. Saltveit. 1989. Electrofishing—theory and practice with special emphasis on salmonids. Hydrobiologia 173:9–43.

Box, G. E. P., and G. M. Jenkins. 1976. Time series analysis, forecasting and control. Holden-Day, Oakland, California.

Braaten, P. J., and C. S. Guy. 1999. Relations between physiochemical factors and abundance of fishes in tributary confluences of the lower channelized Missouri River. Transactions of the American Fisheries Society 128:1213–1221.

Bulkley, R. V. 1970. Fluctuations in abundance and distribution of common Clear Lake fishes as suggested by gillnet catch. Iowa State Journal of Science 44:413–422.

Burkhardt, R. W., and S. Gutreuter. 1995. Improving electrofishing catch consistency by standardizing power. North American Journal of Fisheries Management 15:375–381.

Carlander, K. D. 1953. Use of gill nets in studying fish populations, Clear Lake, Iowa. Proceedings of the Iowa Academy of Science 60:621–625.

Casey, J. M., and R. A. Myers. 1998. Diel variation in trawl catchability: is it as clear as day and night? Canadian Journal of Fisheries and Aquatic Sciences 23:579–586.

Chesher, A. 1991. The effect of measurement error. Biometrika 78:451–462.

Chipps, S. R., D. H. Bennett, and T. J. Dresser, Jr. 1997. Patterns of fish abundance associated with a dredge disposal island: implications for fish habitat enhancement in a large reservoir. North American Journal of Fisheries Management 17:378–386.

Choat, J. H., P. J. Doherty, B. A. Kerrigan, and J. M. Leis. 1993. A comparison of towed nets, purse seine, and light-aggregation devices for sampling larvae and pelagic juveniles of coral reef fishes. U.S. National Marine Fisheries Service Fishery Bulletin 91:195–209.

Coble, D. W. 1992. Predicting population density of largemouth bass from electrofishing catch per effort. North American Journal of Fisheries Management 12:650–652.

Cochran, W. G. 1977. Sampling techniques, 3rd edition. Wiley, New York.

Collins, N. C., H. H. Harvey, A. J. Tierney, and D. W. Dunham. 1983. Influence of predatory fish density on trapability of crayfish in Ontario lakes. Canadian Journal of Fisheries and Aquatic Sciences 40:1820–1828.

Counihan, T. D., A. I. Miller, and M. J. Parsley. 1999. Indexing the relative abundance of age-0 white sturgeons in an impoundment of the lower Columbia River from highly skewed trawling data. North American Journal of Fisheries Management 19:520–529.

Craig, J. F., and J. M. Fletcher. 1982. The variability in the catches of charr, *Salvelinus alpinus* L., and perch, *Perca fluviatilis* L., from multi-mesh gill nets. Journal of Fish Biology 20:517–526.

Crecco, V. A., and T. F. Savoy. 1985. Density-dependent catchability and its potential causes and consequences on Connecticut River American shad, *Alosa sapidissima*. Canadian Journal of Fisheries and Aquatic Sciences 42:1649–1657.

Cross, T. K., M. C. McInerny, and D. H. Schupp. 1995. Seasonal variation in trap-net catches of bluegill in Minnesota lakes. North American Journal of Fisheries Management 15:382–389.

Cui, G., C. S. Wardle, C. W. Glass, A.D. F. Johnstone, and W. R. Mojsiewicz. 1991. Light level thresholds for visual reaction of mackerel, *Scomber scombrus* L., to coloured monofilament nylon gillnet materials. Fisheries Research 10:255–263.

Cyr, H., J. A. Downing, S. Lalonde, S. B. Baines, and M. L. Pace. 1992. Sampling larval fish populations: choice of sample number and size. Transactions of the American Fisheries Society 121:356–368.

Dewey, M. R. 1992. Effectiveness of a drop net, pop net, and an electrofishing frame for collecting quantitative samples of juvenile fishes in vegetation. North American Journal of Fisheries Management 12:808–813.

Die, D. J., V. R. Restrepo, and W. W. Fox, Jr. 1990. Equilibrium production models that incorporate fished area. Transactions of the American Fisheries Society 119:445–454.

Dixon, P. M. 1993. The bootstrap and the jackknife: describing the precision of ecological indices. Pages 290–318 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Dumont, S. C., and J. A. Dennis. 1997. Comparison of day and night electrofishing in Texas reservoirs. North American Journal of Fisheries Management 17:939–946.

Fabrizio, M. C., J. Raz, and R. R. Bandekar. 2000. Using linear models with correlated errors to analyze changes in abundance of Lake Michigan fishes: 1973–1992. Canadian Journal of Fisheries and Aquatic Sciences 57:775–788.

Fabrizio, M. C., and R. A. Richards. 1996. Commercial fisheries surveys. Pages 625–650 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Flickinger, S. A., F. J. Bulow, and D. W. Willis. 1999. Small impoundments. Pages 651–588 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Fréon, P., and O. A. Misund. 1999. Dynamics of pelagic fish distribution and behavior: effects on fisheries and stock assessment. Fishing News Books, Blackwell Scientific Publications, Malden, Massachusetts.

Gerhardt, D. R., and W. A. Hubert. 1989. Effectiveness of cheese bait to capture channel catfish in hoop nets. North American Journal of Fisheries Management 9:343–351.

Gillis, D. M. 1999. Behavioral inferences from regulatory observer data: catch rate variation in the Scotian Shelf silver hake *(Merluccius bilinearis*) fishery. Canadian Journal of Fisheries and Aquatic Sciences 56:288–296.

Gillis, D. M., and R. M. Peterman. 1998. Implications of interference among fishing vessels and the ideal free distribution to the interpretation of CPUE. Canadian Journal of Fisheries and Aquatic Sciences 55:37–46.

Godø, O. R., S. J. Walsh, and A. Englandås. 1999. Investigating density-dependent catchability in bottom-trawl surveys. ICES Journal of Marine Science 56:292–298.

Gordoa, A., M. Maso, and L. Voges. 2000. Monthly variability in the catchability of Namibian hake and its relationship with environmental seasonality. Fisheries Research 48:185–195.

Gryska, A. D., W. A. Hubert, and K. G. Gerow. 1997. Use of power analysis in developing monitoring protocols for the endangered Kendall Warm Springs dace. North American Journal of Fisheries Management 17:1005–1009.

Gulland, J. A. 1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. FAO (Food and Agriculture Organization of the United Nations) Manuals in Fisheries Science 4, Rome.

Gulland, J. A. 1977. Fish population dynamics. John Wiley and Sons, New York.

Hall, T. J. 1986. Electrofishing catch per hour as an indicator of largemouth bass density in Ohio impoundments. North American Journal of Fisheries Management 6:397–400.

Hamley, J. M., and T. P. Howley. 1985. Factors affecting variability of trapnet catches. Canadian Journal of Fisheries and Aquatic Sciences 42:1079–1087.

Hansen, M. J., R. G. Schorfhaar, J. W. Pcek, J. H. Selgegy, and W. W. Taylor. 1995. Abundance indices for determining the status of lake trout restoration in Michigan waters of Lake Superior. North American Journal of Fisheries Management 15:830–837.

Hansen, M. J., R. G. Schorfhaar, and J. H. Selgeby. 1998. Gill-net saturation by lake trout in Michigan waters of Lake Superior. North American Journal of Fisheries Management 18:847–853.

Hayes, D. B., C. P. Ferreri, and W. W. Taylor. 1996. Active fish capture methods. Pages 193–220 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Hi, X., and D. M. Lodge. 1990. Using minnow traps to estimate fish population size: the importance of spatial distribution and relative species abundance. Hydrobiologia 190: 9–14.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hubbard, W. D., and L. E. Miranda. 1988. Competence of non-random electrofishing sampling in assessment of structural indices. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 40 (1986):79–84.

Hubert, W. A. 1996. Passive capture techniques. Pages 157–192 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Hubert, W. A., and D. T. O'Shea. 1992. Use of spatial resources by fishes in Grayrocks Reservoir, Wyoming. Journal of Freshwater Ecology 7:219–225.

Hubert, W. A., and M. B. Sandheinrich. 1983. Patterns of variation in gill-net catch and diet of yellow perch in a stratified Iowa lake. North American Journal of Fisheries Management 3:156–162.

Irwin, E. R., R. L. Noble, and J. R. Jackson. 1997. Distribution of age-0 largemouth bass in relation to shoreline landscape features. North American Journal of Fisheries Management 17:882–893.

Isbell, G. L., and M. R. Rawson. 1989. Relations of gill-net catches of walleyes and angler catch rates in Ohio waters of western Lake Erie. North American Journal of Fisheries Management 9:41–46.

Jackson, D. C. 1995. Distribution and stock structure of blue catfish and channel catfish in macrohabitats along riverine sections of the Tennessee–Tombigbee Waterway. North American Journal of Fisheries Management 15:845–853.

Jessop, B. M., and C. J. Harvie. 1990. Evaluation of designs of periodic count surveys for the estimation of escapement at a fishway. North American Journal of Fisheries Management 10:39–45.

Johnson, B. L., and C. A. Jennings. 1998. Habitat associations of small fishes around islands in the upper Mississippi River. North American Journal of Fisheries Management 18:327–336.

Johnson, D. H. 1995. Statistical sirens: the allure of nonparametrics. Ecology 76:1998–2000.

Karr, J. R. 1999. Defining and measuring river health. Freshwater Biology 41:221–234.

Kappenman, R. F. 1999. Trawl survey based abundance estimation using data sets with unusually large catches. ICES Journal of Marine Science 56:28–35.

Kelso, W. W., and D. A. Rutherford. 1996. Collection, preservation, and identification of fish eggs and larvae. Pages 255–302 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Kessler, R. K. 1999. Potential influence of predator presence on diel movements of small riverine fishes in the Ohio River, Kentucky. Journal of the Kentucky Academy of Science 60:108–112.

Kimura, D. K. 1981. Standardized measures of relative abundance based on modeling log(c.p.u.e.), and their application to Pacific Ocean perch *(Sebastes alutus)*. Journal du Conseil International pour l'Exploration de la Mer 39:211–218.

Kimura, D. K. 2000. Using nonlinear functional relationship regression to fit fisheries models. Canadian Journal of Fisheries and Aquatic Sciences 57:160–170.

Kimura, D. K., and J. W. Balsiger. 1985. Bootstrap methods for evaluating sablefish pot index surveys. North American Journal of Fisheries Management 5:47–56.

Korsbrekke, K., and O. Nakken. 1999. Length and species-dependent diurnal variation of catch rates in the Norwegian Barents Sea bottom-trawl survey. ICES Journal of Marine Science 56:284–291.

Krebs, C. J. 1989. Ecological methodology. Harper and Row, New York.

Lackey, R. T., and W. A. Hubert. 1976. Analysis of exploited fish populations. Virginia Polytechnic Institute and State University, Sea Grant Extension Division, VPI-SG-76-04, Blacksburg.

Lawson, G. L., and G. A. Rose. 1999. The importance of detectability to acoustic surveys of semi-demersal fish. ICES Journal of Marine Science 56:370–380.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, North Carolina.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot ANOVA design for repeated measures fishery data. Fisheries (Bethesda) 19(3):14–20.

Maceina, M. J., W. B. Wrenn, and D. R. Lowery. 1995. Estimating harvestable largemouth bass abundance in a reservoir with an electrofishing catch depletion technique. North American Journal of Fisheries Management 15:103–109.

Malvestuto, S. P. 1996. Sampling the recreational creel. Pages 591–624 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

McInerny, M. C., and T. K. Cross. 2000. Effects of sampling time, intraspecific density, and environmental variables on electrofishing catch per effort of largemouth bass in Minnesota lakes. North American Journal of Fisheries Management 20:328–336.

Mero, S. W., and D. W. Willis. 1992. Seasonal variation in sampling data for walleye and sauger collected with gill nets from Lake Sakakawea, North Dakota. Prairie Naturalist 24:231–240.

Methven, D. A., and D. C. Schneider. 1998. Gear-independent patterns of variation in catch of juvenile Atlantic cod *(Gadus morhua)* in coastal habitats. Canadian Journal of Fisheries and Aquatic Sciences 55:1430–1442.

Michaletz, P. H., and G. M. Gale. 1999. Longitudinal gradients in age-0 gizzard shad density in large Missouri reservoirs. North American Journal of Fisheries Management 19:765–773.

Michalsen, K., O. R. Godø, and A. Fernö. 1996. Diel variation in the catchability of gadoids and its influence on the reliability of abundance indices. ICES Journal of Marine Science 53:389–395.

Miller, R. J. 1990. Effectiveness of crab and lobster traps. Canadian Journal of Fisheries and Aquatic Sciences 47:1228–1251.

Miranda, L. E., W. D. Hubbard, S. Sangare, and T. Holman. 1996. Optimizing electrofishing sample duration for estimating relative abundance of largemouth bass in reservoirs. North American Journal of Fisheries Management 16:324–331.

Moyer, E. J., M. W. Halon, J. J. Sweatman, R. S. Butter, and V. P. Williams. 1995. Fishery responses to habitat restoration in Lake Tohopekaliga, Florida. North American Journal of Fisheries Management 15:591–595.

Moyle, J. B., and R. Lound. 1960. Confidence limits associated with means and medians of series of net catches. Transactions of the American Fisheries Society 89:53–58.

Munro, P. T. 1998. A decision rule based on the mean square error for correcting relative fishing power differences in trawl survey data. U.S. National Marine Fisheries Service Fishery Bulletin 96:538–546.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models, 4th edition. Richard D. Irwin, Chicago.

Newman, J. A., J. Bergelson, and A. Grafen. 1997. Blocking factors and hypothesis tests in ecology: is your statistics text wrong? Ecology 78:1312–1320.

Ney, J. J. 1999. Practical use of biological statistics. Pages 167–192 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Orth, D. J., and R. J. White. 1999. Stream habitat management. Pages 249–284 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Paller, M. H. 1995. Interreplicate variance and statistical power of electrofishing data from low-gradient streams in the southeastern United States. North American Journal of Fisheries Management 15:542–550.

Paloheimo, J. E. 1963. Estimation of catchabilities and population size of lobsters. Journal of the Fisheries Research Board of Canada 20:59–88.

Paloheimo, J. E., and L. M. Dickie. 1964. Abundance and fishing success. Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer 155:152–163.

Paragamian, V. L. 1989. A comparison of day and night electrofishing: size structure and catch per unit effort for smallmouth bass. North American Journal of Fisheries Management 9:500–503.

Parsley, M. J., D. E. Palmer, and R. W. Burkhardt. 1989. Variation in capture efficiency of a beach seine for small fishes. North American Journal of Fisheries Management 9: 239–244.

Pelletier, D. 1998. Intercalibration of research survey vessels in fisheries: a review and an application. Canadian Journal of Fisheries and Aquatic Sciences 55:2672–2690.

Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. U.S. National Marine Fisheries Service Fishery Bulletin 94:498–505.

Peterman, R. M., and M. J. Bradford. 1987. Statistical trends in fish abundance. Canadian Journal of Fisheries and Aquatic Sciences 47:2–15.

Peterman, R. M., and G. J. Steer. 1981. Relation between sport-fishing catchability coefficients and salmon abundance. Transactions of the American Fisheries Society 110:585–593.

Peterson, D. L., M. B. Bain, and N. Haley. 2000. Evidence of declining recruitment of Atlantic sturgeon in the Hudson River. North American Journal of Fisheries Management 20:231–238.

Pope, K. L., and D. W. Willis. 1996. Seasonal influences on freshwater fisheries sampling data. Reviews in Fisheries Science 4:57–73.

Pothoven, S. A., B. Vondracek, and D. L. Pereira. 1999. Effects of vegetation removal on bluegill and largemouth bass in two Minnesota lakes. North American Journal of Fisheries Management 19:748–757.

Power, J. H., and E. B. Moser. 1999. Linear model analysis of net catch data using the negative binomial distribution. Canadian Journal of Fisheries and Aquatic Sciences 56:191–200.

Pyper, B. J., and R. M. Peterman. 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. Canadian Journal of Fisheries and Aquatic Sciences 55:2127–2140. (See also Pyper, B. J., and R. M. Peterman. 1998. Erratum: Comparison of methods to account for autocorrelation in correlation analyses of fish data. Canadian Journal of Fisheries and Aquatic Sciences 55:2710.)

Reynolds, J. B. 1996. Electrofishing. Pages 221–254 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Richards, C., F. J. Kutka, M. E. McDonald, G. W. Merrick, and P. W. Devore. 1996. Life history and temperature effects on catch of northern orconectid crayfish. Hydrobiologia 319:111–118.

Richards, L. J., and J. T. Schnute. 1986. An experimental and statistical approach to the question: is CPUE an index of abundance? Canadian Journal of Fisheries and Aquatic Sciences 43:1214–1227.

Richards, L. J., and J. T. Schnute. 1992. Statistical models for estimating CPUE from catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences 49:1315–1327.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Rooker, J. R., G. D. Dennis, and D. Goulet. 1996. Sampling larval fishes with a nightlight lift-net in tropical inshore waters. Fisheries Research 26:1–15.

Roper, B. B., and D. L. Scarnecchia. 1999. Emigration of age-0 chinook salmon *(Oncorhynchus tshawytscha)* smolts from the upper South Umpqua River basin, Oregon, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 56:939–946.

Rose, G. A., and D. W. Kulka. 1999. Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod *(Gadus morhua)* declined. Canadian Journal of Fisheries and Aquatic Sciences 56 (Supplement 1):118–127.

Rose, G. A., and W. C. Leggett. 1989. Predicting variability in catch-per-effort in Atlantic cod, *Gadus morhua,* trap and gillnet fisheries Journal of Fish Biology 35 (Supplement A):155–161.

Sammons, S. M., and P. W. Bettoli. 1999. Spatial and temporal variation in electrofishing catch rates of three species of black bass (*Micropterus* spp.) from Normandy Reservoir, Tennessee. North American Journal of Fisheries Management 19:454–461.

Sampson, D. B. 1991. Local catch per unit effort as an index of global fish abundance. Pages 275–284 *in* I. G. Cowx, editor. Catch effort sampling strategies: their application in freshwater fisheries management. Fishing News Books, Blackwell Scientific Publications, Cambridge, Massachusetts.

Sanderson, B. L., T. R. Hrabik, J. J. Magnuson, and D. M. Post. 1999. Cyclic dynamics of a yellow perch *(Perca flavescens)* population in an oligotrophic lake: evidence for the role of intraspecific interactions. Canadian Journal of Fisheries and Aquatic Sciences 56: 1534–1542.

SAS Institute. 1998. SAS/Stat user's guide, version 6: volumes 1 and 2. SAS Institute, Cary, North Carolina.

Schweigert, J. F., C. W. Haegele, and M. Stocker. 1985. Optimizing sampling design for herring spawn surveys in the Strait of Georgia, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 42:1806–1814.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Charles Griffin & Co., London.

Shaw, E. 1978. Schooling fishes. American Scientist 66:166–175.

Shroyer, S. M., and T. S. McComish. 1998. Forecasting abundance of quality-size yellow perch in Indiana waters of Lake Michigan. North American Journal of Fisheries Management 18:19–24.

Sigler, M. F. 2000. Abundance estimation and capture of sablefish *(Anoplopoma fimbria)* by longline gear. Canadian Journal of Fisheries and Aquatic Sciences 57:1270–1283.

Simmonds, E. J., and R. J. Fryer. 1996. Which are better, random or systematic acoustic surveys? A simulation study using North Sea herring as an example. ICES Journal of Marine Sciences 53:39–50.

Skalski, J. R., A. Hoffman, B. H. Ransom, and T. W. Steig. 1993. Fixed-location hydroacoustic monitoring designs for estimating fish passage using stratified random and systematic sampling. Canadian Journal of Fisheries and Aquatic Sciences 50:1208–1221.

Smith, S. J. 1980. Comparison of two methods of estimating the variance of the estimate of catch per unit effort. Canadian Journal of Fisheries and Aquatic Sciences 37: 2346–2351.

Smith, S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. Canadian Journal of Fisheries and Aquatic Sciences 54:616–630.

Snedecor, G. W., and W. G. Cochran. 1989. Statistical methods. Iowa State University Press, Ames.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry, 2nd edition. W. H. Freeman, San Francisco.

Somers, K. M., and R. H. Green. 1993. Seasonal patterns in trap catches of the crayfish *Cambarus bartoni* and *Orconectes virilis* in six south-central Ontario lakes. Canadian Journal of Zoology 71:1136–1152.

Somerton, D. A., and B. S. Kikkawa. 1995. A stock survey technique using the time to capture individual fish on longlines. Canadian Journal of Fisheries and Aquatic Sciences 52:260–267.

Spangler, G. R., and J. J. Collins. 1992. Lake Huron fish community structure based on

gill-net catches corrected for selectivity and encounter probabilities. North American Journal of Fisheries Management 12:585–597.

Stanley, R. D. 1992. Bootstrap calculation of catch-per-unit-effort variance from trawl logbooks: do fisheries generate enough observations for stock assessment? North American Journal of Fisheries Management 12:19–27.

Steel, R. G. D., and J. H. Torrie. 1960. Principles and procedures of statistics, with special reference to the biological sciences. McGraw-Hill, New York.

Stewart-Oaten, A. 1995. Rules and judgments in statistics: three examples. Ecology 76:2001–2009.

Stoner, A. W. 1991. Diel variation in the catch of fishes and penaeid shrimps in a tropical estuary. Estuarine, Coastal and Shelf Science 33:57–69.

Sullivan, P. J. 1991. Stock abundance estimation using depth-dependent trends and spatially correlated variation. Canadian Journal of Fisheries and Aquatic Sciences 48:1691–1703.

Summerfelt, R. C. 1999. Lake and reservoir management. Pages 285–320 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Swain, D. P., G. A. Nielsen, A. F. Sinclair, and G. A. Chouinard. 1994. Changes in catchability of Atlantic cod *(Gadus morhua)* to an otter-trawl fishery and research survey in the southern Gulf of St. Lawrence. ICES Journal of Marine Science 51:493–504.

Swain, D. P., G. A. Poirier, and A. F. Sinclair. 2000. Effect of water temperature on catchability of Atlantic cod *(Gadus morhua)* to the bottom-trawl survey in the southern Gulf of St Lawrence. ICES Journal of Marine Science 57:56–68.

Tillma, J. S., C. S. Guy, and C. S. Mammoliti. 1998. Relations among habitat and population characteristics of spotted bass in Kansas streams. North American Journal of Fisheries Management 18:886–893.

Trometer, E. S., and W.-D. N. Busch. 1999. Changes in age-0 fish growth and abundance following the introduction of zebra mussels *Dreissena polymorpha* in the western basin of Lake Erie. North American Journal of Fisheries Management 19:604–609.

Van Den Avyle, M. J., J. Boxrucker, P. Michaletz, B. Vondracek, and G. R. Ploskey. 1995. Comparison of catch rate, length distribution, and precision of six gears used to sample shad populations. North American Journal of Fisheries Management 15:940–955.

Van Den Avyle, M. J., and R. S. Hayward. 1999. Dynamics of exploited fish populations. Pages 127–166 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Van Zee, B. E., D. W. Willis, and C. C. Stone. 1996. Comparison of diel sampling data for sauger collected by electrofishing. Journal of Freshwater Ecology 11:139–143.

Walsh, S. J. 1991. Diel variation in availability and vulnerability of fish to a survey trawl. Journal of Applied Ichthyology 7:147–159.

Walters, C. J., and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences 38:704–710.

Ward, C., R. L. Eshenroder, and J. R. Bence. 2000. Relative abundance of lake trout and burbot in the main basin of Lake Michigan in the early 1930s. Transactions of the American Fisheries Society 129:282–295.

Watters, G., and R. Deriso. 2000. Catches per unit of effort of bigeye tuna: a new analysis with regression trees and simulated annealing. Bulletin of the Inter-American Tropical Tuna Commission 21:531–552.

Willis, D. W. 1987. Use of gill-net data to provide a recruitment index for walleyes. North American Journal of Fisheries Management 7:591–592.

Wolfinger, R. D. 1996. Heterogeneous variance–covariance structures for repeated measures. Journal of Agricultural, Biological, and Environmental Statistics 1:205–230.

Wolfinger, R. D. 1999. Fitting nonlinear mixed models with the NLMIXED procedure. SAS Users Group International Conference Proceedings 24: Paper 287.

8 Abundance, Biomass, and Production

Daniel B. Hayes, James R. Bence, Thomas J. Kwak, and Bradley E. Thompson

8.1 INTRODUCTION

Fisheries scientists face a challenge in that virtually all methods of fish capture or observation are selective. Further, most fish capture methods can be applied to only a fraction of the entire area of interest. Thus, measures such as catch per unit effort (C/f) or catch per area can only be regarded, at best, as being proportional to the true population abundance (see Chapter 7). The methods presented in this chapter are designed to address these problems and provide estimates of absolute abundance or "true" fish density. In general, these methods require additional sampling beyond that required to estimate relative abundance. As such, careful consideration should be given to whether relative measures of abundance are adequate or if the need for estimates of absolute abundance justifies the additional cost.

In many cases, relative abundance is sufficient to answer important research or management questions. One example is when the principal goal is to determine if abundance has changed over time. As long as vulnerability to the gear remains constant over time, trends in C/f can accurately indicate changes in abundance (see Chapter 7). In such cases, the extra effort required to determine absolute abundance is better spent in sampling more sites. In general, estimates of absolute abundance are needed when catchability is likely to vary across time or between sampling sites, confounding comparisons of C/f across space or time. Absolute abundance estimates are also important when harvest quotas are being computed.

Whether relative or absolute measures of abundance are desired, it is critical to define the population of interest carefully. In many cases, some part of the population is excluded from consideration because of limitations of the sampling gear. For example, population estimates of yellow perch in midsummer conducted by means of gill nets would likely not include age-0 fish because they would not be

vulnerable to the gear. Similarly, care must be taken in defining the spatial extent of the target population. Sometimes one is interested in the population in only a particular stream reach, whereas in other situations, the desired scale is an entire watershed, which would likely need to be subsampled.

Another consideration common to both relative and absolute measures of abundance is the precision and accuracy required for the task. Accuracy, bias, and precision are defined in Chapter 3. Applying these concepts to population estimates, it is important to recognize that failures to meet assumptions often reduce both accuracy and precision. Therefore, we emphasize methods for checking assumptions in addition to the methods commonly used to provide point estimates and measures of variability.

8.2 DIRECT OBSERVATION METHODS

In some situations, direct observation of all fishes in a given area (sampling site) is possible, providing a complete census of the area searched. This approach has been applied in small streams (Hankin and Reeves 1988) or in other situations where fish are tightly constrained. Likewise, counts of fish in hydroacoustic surveys are often assumed to represent all individuals within the hydroacoustic beam path. In situations in which counts are assumed to be accurate and complete, the total population is estimated as the product of the mean density in the sites sampled times the total area. The precision of total population estimates depends principally on the variability between sampling sites (Hankin and Reeves 1988) and the sampling design used (e.g., stratified random sampling). Methods of computing the variance for several sampling designs are presented in Chapter 3 and can be applied directly to data collected through complete censuses at selected sites. One specialized design not included in Chapter 3 is hydroacoustic surveys for which counts are collected along the path of the boat (i.e., along a transect). If data are collected along a single transect, specialized statistical methods are necessary to calculate the variance of the population estimate because of the autocorrelation between counts at adjacent points. (Foote and Stefansson 1993; Vondracek and Degan 1995). If two or more randomly placed transects are followed, however, each transect can be treated as a sampling site, and the methods described in Chapter 3 can be applied.

In many situations, visual observation misses some proportion of the population, even in situations where fish are constrained. Because of this, estimates of density for individual sites are imprecise and contribute to the overall imprecision of total density estimates. In order to estimate the proportion observed within a sampling site, additional information needs to be gathered. The most commonly used method is to measure the distance that each animal observed lies off the transect (i.e., the right-angle distance from each animal seen to the transect) or from the center of a fixed point of observation. Depending on the observation technique, this distance can be determined directly, or the distance and angle of departure from the transect can be determined and the right-angle distance calculated by simple geometry. Generally, the proportion of fish present that are

detected (i.e., sightability) declines farther from the point of observation or from the transect surveyed (Figure 8.1). Assuming that fish are randomly distributed with respect to the transect and sightability is 100% at or near the center of the transect, the proportion observed can be estimated as a function of distance from the transect.

Critical assumptions for applying the direct observation approach include (1) fish are randomly and independently distributed, and movement of the observer does not attract or repel fish prior to observation; (2) distances are measured accurately; (3) fish are not counted more than once; (4) fish are detected at their original position with respect to the transect; and (5) sighting of each fish is independent of other fish, meaning that the likelihood of seeing an individual fish

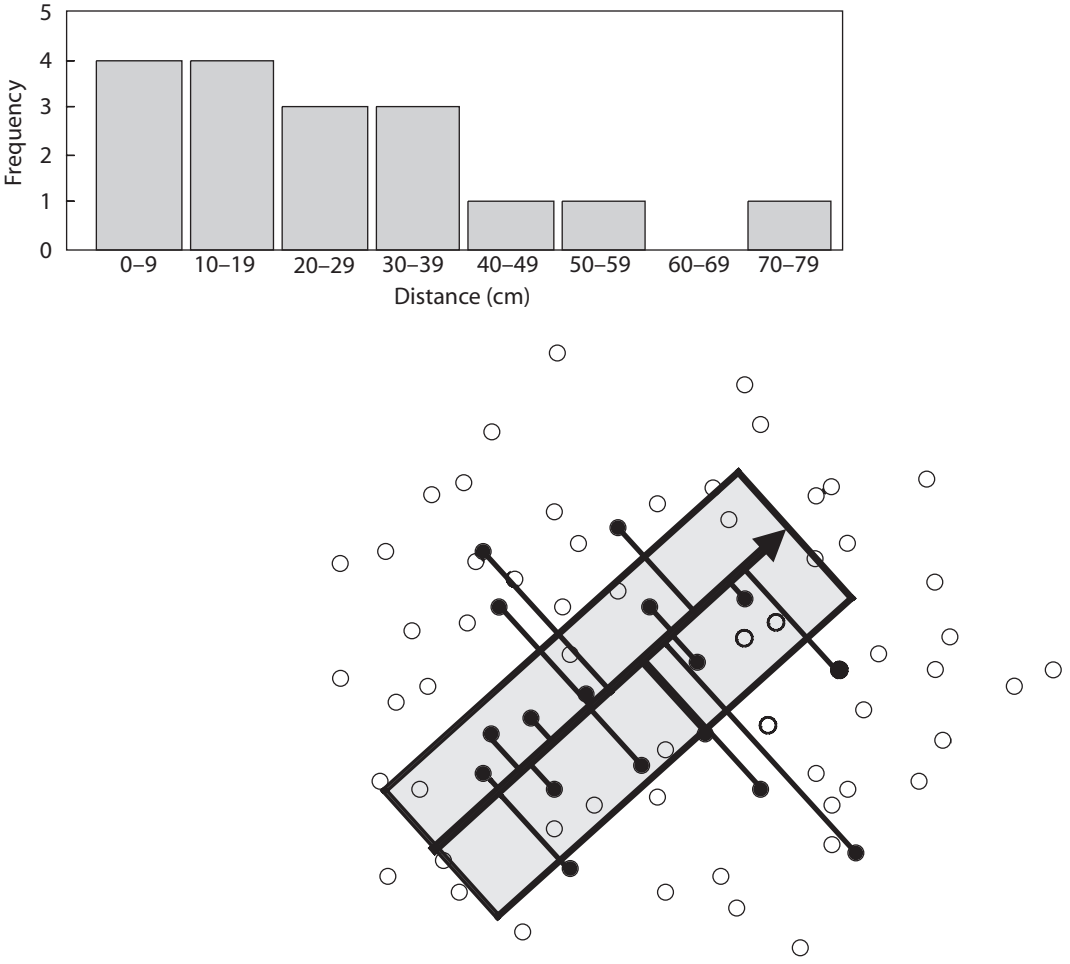


Figure 8.1 Example of animals sighted in a transect survey. The histogram depicts the relative frequency of observations within 0.1-m intervals from the transect. The shaded box depicts the effective width of the transect. Open circles indicate fish that are not sighted and closed circles indicate fish that are sighted. Figure modified from Thompson et al. (1998).

does not depend on the number of other fish in the vicinity (Seber 1982; Buckland et al. 1993; Thompson et al. 1998). Carefully implemented field techniques can help ensure that assumptions 1–4 are met. The assumption of independent sightings, however, depends on the behavior of fish and their schooling behavior and patchiness. When fish are sighted in groups, but the proportion of fish sighted is constant with fish density, the precision of population estimates is generally reduced, but the population estimate is not necessarily biased (Buckland et al. 1993). In cases in which the unit of observation is a school or other aggregation of animals, we refer the reader to Buckland et al. (1993) for methods for appropriately analyzing these data. When sightability varies as a function of density or school size, estimates of fish density are likely to be biased, and the applicability of this approach should be reconsidered.

For a single line-transect survey, the general formula for density is (Buckland et al. 1993)

$$\hat{D} = \frac{n}{2L\hat{w}}, \tag{8.1}$$

where $\hat{D}$ = estimated density; n = number of fish observed; $\hat{w}$ = estimated effective width of transect from center; and L = transect length.

When counts are conducted from a single fixed point (point plot survey), the area surrounding the point is observed, resulting in a circular search area. In this situation, the general formula for density is (Buckland et al. 1993)

$$\hat{D} = \frac{n}{2\pi\hat{w}^2}, \tag{8.2}$$

where $\hat{w}$ = estimated effective search radius.

In applying these formulae, a critical component is estimating w, the effective width of the transect or search radius from a point. Essentially w corresponds to an equivalent transect for which all fish out to w are detected and all fish beyond w are not. In order to estimate this quantity accurately, it is necessary to select a function describing the pattern of sightability with distance. Many functions can be used to describe the sightability function. We apply two of these functions to illustrate that the choice of sightability function matters, and we provide formulae for estimating total population abundance from density and the total area of the study site in Box 8.1. Buckland et al. (1993) provide a thorough discussion of various sightability functions and methods for selecting among these functions.

The variance for the density estimate (and population size) for a single transect within a site can be estimated approximately based on the binomial distribution describing observed and unobserved fish (Box 8.1), assuming that fish are randomly and independently distributed. When multiple transects or points are observed, the variance among transects should be determined based on the overall sampling design, following methods outlined in Chapter 3.

Specialized software packages are available to estimate population size based on distance sampling (for example, the comprehensive package, DISTANCE; Thomas et al. 2001; available at http://www.ruwpa.st-and.ac.uk/distance/).

Box 8.1 Estimation of Abundance and Density Based on Distance Sampling

An investigator snorkels along a 100-m transect that is randomly located in a stream reach containing 500 m^2. Thirty brook trout are observed at the following right-angle distances (m) from the center of the transect: 0.7, 0.1, 0.6, 0.3, 0.4, 0.1, 3.2, 0.4, 0.6, 1.4, 0.2, 0.1, 2.5, 0.4,4.6, 2.2, 0.5, 1.6, 0.4, 0.4, 1.5, 0.8, 0.0, 0.2, 2.1, 0.4, 0.4, 0.1, 1.1, and 0.6. The investigator would like to estimate the density of brook trout in the section and the total population in the reach.

We define the following variables:

- n = number animals observed;
- N = total population in reach;
- A = total area of reach (m^2);
- D = density of fish (number/m^2);
- L = length of transect (m);
- y = right angle distance (m) from transect for each animal;
- w = effective strip width;
- $V(\hat{N})$ = estimated variance of population estimate; and
- CI = confidence interval.

Based on the assumption that sightability drops off exponentially with distance from the transect, and that fish are independently distributed in the reach, we have the following (Seber 1982):

$$\hat{w} = \frac{\sum y}{n-1} = \frac{27.9}{30-1} = 0.962;$$

$$\hat{D} = \frac{n}{2L\hat{w}} = \frac{30}{2 \cdot 100 \cdot 0.962} = 0.156;$$

$$\hat{N} = \frac{nA}{2L\hat{w}} = \frac{30 \cdot 500}{2 \cdot 100 \cdot 0.962} = 78;$$

$$V(\hat{N}) = \frac{n}{\left(\frac{n}{\hat{N}}\right)^2}\left(1 - \frac{n}{\hat{N}} + \frac{n}{n-2}\right) = \frac{30}{\left(\frac{30}{78}\right)^2}\left(1 - \frac{30}{78} + \frac{30}{30-2}\right) = 342;$$

$$\text{CI} = \hat{N} \pm Z_{\alpha/2}\sqrt{V(\hat{N})} = 78 \pm 1.96\sqrt{342} = 78 \pm 36 = 42,\ 114.$$

Based on the assumption that the sightability function follows a half-normal distribution, the formula for effective width is (Buckland et al. 1993)

$$\hat{w} = \frac{1}{\sqrt{\frac{2}{\pi\sum(y^2/n)}}} = \frac{1}{\sqrt{\frac{2}{\pi \cdot 1.956}}} = 1.752,$$

and density is calculated as above:

$$\hat{D} = \frac{30}{2 \cdot 100 \cdot 1.752} = 0.086.$$

If sightability drops off exponentially, the estimated population is 78 with an approximate CI of 42 to 114. Note that the density (and hence total abundance) based on a half-normal distribution is approximately half that obtained with an exponential model, highlighting the need to test the assumed sightability function (see Buckland et al. 1993 for these methods).

8.3 POPULATION ESTIMATION: MARK–RECAPTURE METHODS

8.3.1 Closed Population Mark–Recapture Methods

The underlying concepts and assumptions of mark–recapture methods of population estimation have a long history in the fishery literature. Because of the extensive reviews available on this subject (Otis et al. 1978; Seber 1982; Burnham et al. 1987), we will emphasize common applications and methods that provide a base for specialized or particularly complex situations.

8.3.1.1 Single Marking Period and Single Recapture Period

In the simplest case, fish are randomly collected from a closed population, and easily recognizable, permanent marks are applied to captured individuals. These individuals are then released and allowed to mix completely with the remainder of the population. A second sampling is undertaken, and the ratio of marked to unmarked fish can be used to estimate the total population. Assumptions in basic mark–recapture studies include (1) the population is geographically closed, with no immigration or emigration, (2) the population is demographically closed, with no birth or deaths, (3) no marks are lost or missed (4) marking does not change fish behavior or vulnerability to capture, (5) marked fish mix at random with unmarked fish, and (6) all animals have an equal probability of capture that does not change over time (Otis et al. 1978; Seber 1982). A number of formulae have been developed for this basic situation. In practical terms, all give similar results when reasonable numbers of marked fish are recaptured (e.g., at least 2–3, but preferably greater than 10; Chapman 1951; Robson and Regier 1964). Because of its widespread use and theoretical basis, we recommend the use of the Chapman estimator (Seber 1982):

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 , \tag{8.3}$$

where n_1 = number caught and marked in first sampling period; n_2 = number caught in second sampling period; and m_2 = number of marked animals in second sampling period.

The variance of this estimator can be approximated as (Seber 1982)

$$V(\hat{N}) = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2 (m_2 + 2)} - 1 . \tag{8.4}$$

Numerous approaches are available to develop confidence intervals (CIs) for $\hat{N}$. Unfortunately, a diversity of methods occurs because of different distributional assumptions and different approximations for small and modest sample sizes. As such, a single method has not yet been identified as being generally best. For large sample sizes (e.g., $m_2 > 50$), the normal approximation is generally adequate (Seber 1982), and $([100 - \alpha]\%)$ confidence limits can be calculated as

$$\hat{N} \pm Z_{\alpha/2}\sqrt{V(\hat{N})}\,. \tag{8.5}$$

For a 95% CI, $\alpha = 0.05$, and $Z_{\alpha/2} = 1.96$. When there are fewer than 50 recaptures, Chapman (1948; reproduced in Seber 1982 and Appendix) provides a table from which CIs can be calculated based on the number of recaptured fish.

8.3.1.2 *The Schnabel Method*

When multiple marking and recapture samples are collected over a short period (so that the population is closed with no immigration, emigration, recruitment, or mortality), population size can be estimated with the Schnabel method (Schnabel 1938; Seber 1982):

$$\hat{N} = \frac{\sum_{i=2}^{t} n_i M_i}{\sum_{i=2}^{t} m_i + 1}, \tag{8.6}$$

where t = number of sampling occasions; n_i = number of fish caught in ith sample; m_i = number of fish with marks caught in ith sample; and M_i = number of marked fish present in the population for ith sample.

The variance of this estimator can be approximated as (Seber 1982)

$$V(\hat{N}) = \hat{N}^2\left[\frac{\hat{N}}{\sum n_i M_i} + 2\cdot\frac{\hat{N}^2}{(\sum n_i M_i)^2} + 6\cdot\frac{\hat{N}^3}{(\sum n_i M_i)^3}\right]. \tag{8.7}$$

Confidence intervals for $\hat{N}$ with the Schnabel method can be computed following the same recommendations for the Chapman method in a single mark–recapture experiment.

8.3.1.3 *Multiple Recapture Events with Uniquely Marked Individuals*

In many situations, a simple design using a single marking period and single recapture period or a Schnabel-type design is sufficient to estimate population abundance. The effectiveness of such designs, however, rests on adequately meeting the assumptions. Unfortunately, it is generally not possible to test these assumptions using the data collected during a single recapture period or when fish are simply marked as being previously caught. To test the assumptions underlying mark–recapture methods of population estimation, it is generally necessary to sample over multiple periods and to have marks that allow for the capture history of individual fish to be determined (e.g., by using individually numbered tags).

For closed populations with uniquely marked fish, Otis et al. (1978) present a hierarchical suite of models intended to cover a range of situations for which particular assumptions hold (Figure 8.2). The simplest, yet most restrictive, model is that for which all assumptions listed earlier apply (M_o). In the next tier of models, three basic mechanisms causing unequal capture probabilities are addressed.

In model M_t, the probability of capture is allowed to vary among different sample periods (time). Variability in capture over time may occur due to factors such as weather or due to changes in the amount or type of fishing gear deployed. In model M_b, the probability of capture is allowed to vary due to behavioral response to prior capture (i.e., fish become more prone or less prone to capture after being caught, handled, and marked). In surveys of small mammals, for example, investigators find that marked animals may become trap happy or trap shy, thus biasing population estimates if such behavior is not considered (Seber 1982). The final model, M_h, allows for heterogeneity in the capture probability of individual fish. This heterogenity may occur for a variety of reasons, including inherent features of each fish, such as its size, or less obvious factors such as variation in the size of home ranges, resulting in different vulnerabilities to passive gears such as trap nets. Methods have been developed to estimate population size for each of these models and are illustrated below. Because of the complexity of the required analyses, we strongly recommend the use of specialized software when applying these models. The program MARK (White and Burnham 1999; available at http://www.cnr.colostate.edu/~gwhite/mark/mark.htm.) is a very flexible software package designed to analyze data from mark–recapture studies.

In the next tier of models, variations in probability of capture occur through combinations of two of the above factors (Figure 8.2). Thus, model M_{tb} represents the case in which capture probability varies over time, as well as with the prior capture history of an animal (behavior). Estimation methods are also available for each of these models; however, we refer the reader to software, such as MARK, specially designed to handle such situations. Unfortunately, no method has yet been developed to estimate population size and account for these three sources of variation simultaneously (i.e., to estimate the parameters for model M_{tbh}).

A central concept to estimating population size by means of these models is the capture history of an animal. Because the population is assumed to be closed, the number of animals in the population (N) remains constant over all sampling periods. As such, during each of the sampling periods (numbered 1 to t), an animal can either be caught or not. For convenience, the capture history of all animals observed can be recorded in a matrix in which a 1 is used to indicate a capture and 0 to indicate no capture during a particular sampling period.

The second concept central to estimating population size based on these models is the likelihood function. Although this is the foundation for many methods of population estimation (in fact, it is the basis for the Chapman and Schnabel estimators), likelihood functions may be unfamiliar to many readers. We provide a brief synopsis of this topic in the context of population estimators in Box 8.2. Readers should consult texts in mathematical statistics (e.g., Bickel and Doksum 1977; Rice 1995) for a more thorough treatment. In some cases, likelihood methods result in a formula for directly estimating population abundance. In most situations, however, there is no direct formula relating the data to the population estimate. Instead, the likelihood function is repeatedly evaluated at trial values of N (or related parameters that determine N) until a value of N is found that produces the maximum value of the likelihood function. This is chosen as the best or

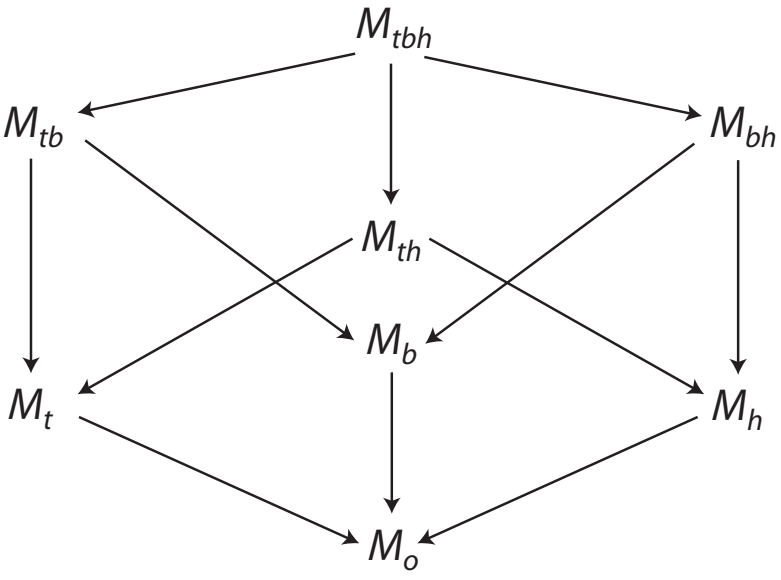


Figure 8.2 Hierarchical organization of models for capture–recapture methods of population estimation as described by Otis et al. (1978). Model M_o allows all assumptions of the mark–recapture model to apply; model M_t allows the probability of capture to vary among sample periods (time); model M_b allows the probability of capture is to vary due to behavioral response to prior capture; model M_h allows for heterogeneity in the capture probability of individual fish; and the remaining models allow variations in probability of capture through combinations of the above factors. Figure modified from Otis et al. (1978).

most likely population estimate, $\hat{N}$. With modern computing power, a "brute force" solution can be found simply by starting with a trial value of N set equal to the total number of distinct individuals caught (i.e., the minimum possible population) and then evaluating the likelihood function for each integer value of N up to an arbitrary maximum. Generally, the logarithm of the likelihood function is evaluated because it is often easier to compute and because it provides a useful basis for comparing between models and for estimating CIs. Further, the likelihood function can often be maximized focusing on one parameter at a time, producing what is called a concentrated likelihood (Seber and Wild 1989). Following precedent in the literature, and for simplicity, we will generally not distinguish between concentrated likelihood functions or likelihood functions that include all parameters.

The log-likelihood function ($\log_e L$) used to estimate $\hat{N}$ for model M_o (Otis et al. 1978) is

$$\log_e L(N \mid X) = \left\{ \log_e \left(\frac{N!}{(N - M_{t+1})!} \right) + [n_\cdot \log_e (n_\cdot)] + [(tN - n_\cdot) \log_e (tN - n_\cdot)] - [tN \log_e (tN)] \right\}, \quad (8.8)$$

where X = capture history matrix; M_{t+1} = total number of distinct fish caught; $n.$ = total number of fish captured; and t = number of sampling periods.

Once the maximum-likelihood estimate, $\hat{N}$, has been calculated by finding the N that maximizes equation (8.8), the corresponding maximum-likelihood estimate for the probability of capture is

$$\hat{p} = \frac{n.}{t\hat{N}}. \tag{8.9}$$

An asymptotic estimate of the variance for $\hat{N}$ is (Otis et al. 1978)

$$V(\hat{N}) = \frac{\hat{N}}{\left[(1-\hat{p})^{-t} - \left(\frac{t}{1-\hat{p}}\right) + (t-1)\right]}. \tag{8.10}$$

Confidence intervals for $\hat{N}$ can be obtained in a number of ways. The first method is to estimate variance of $\hat{N}$ by means of equation (8.10) and calculate upper and lower bounds based on equation (8.5). This approach assumes that $\hat{N}$ has a normal distribution, which should be a reasonable approximation when more than 30 animals are recaptured. An alternate method, discussed in Box 8.2, is to use the likelihood function itself to determine CIs. Trial values of N that produce likelihood values that differ from the maximum likelihood by more than 3.841, which is the critical value for a χ^2 distribution with 1 df and an α of 0.05, define the bounds of the CI. The likelihood method for determining CIs is often preferred because it does not require the assumption of normality, thereby allowing for asymmetric CIs for modest sample sizes.

In model M_t, the capture probability for individual animals varies over time. As such, this model has $t+1$ parameters: N, which is the population abundance, and $p_1, p_2, \ldots, p_t$, which are the time-specific capture probabilities. The log-likelihood function for model M_t is (Otis et al. 1978)

$$\begin{aligned}\log_e L(N|X) = \log_e \left[\frac{N!}{(N-M_{t+1})!}\right] + \left[\sum_{j=1}^{t} n_j \log_e(n_j)\right] \\ + \left[\sum_{j=1}^{t} (N-n_j)\log_e(N-n_j)\right] - [tN\log_e(N)].\end{aligned} \tag{8.11}$$

The N that maximizes equation (8.11) is the maximum-likelihood estimate, $\hat{N}$. The corresponding maximum-likelihood estimates for the probability of capture for each time period can be determined by

$$\hat{p}_j = \frac{n_j}{\hat{N}}. \tag{8.12}$$

Box 8.2 Application of Likelihood Functions in Population Estimation

Here, we illustrate the ideas underlying likelihood functions in the context of estimating population size. For this example, consider the situation in which 60 fish are present in a pool within a stream and we have a 40% chance of catching each fish with one electrofishing pass. In this example, we theoretically could catch between 0 and 60 fish. Assuming that the probability a fish is caught is independent among fish, the probability a specific number of fish will be caught in one pass is given by a binomial probability distribution. For example, the probability of capturing 20 in one pass (i.e., number caught = n =20), assuming catchability is 0.4, is given by the formula

$$P(n = 20|N = 60, q = 0.4) = \frac{N!}{n!(N-n)!}\, q^n (1-q)^{N-n}$$

$$= \frac{60!}{20!(60-20)!}\, 0.4^{20}(1-0.4)^{60-20} = 0.0616.$$

Applying this formula for each possible outcome, we can see that the outcome with the highest probability (i.e., the most likely outcome) is 24 fish captured (Figure 8.3A). Equations of this type are known as probability functions for discrete distributions or probability density functions for continuous distributions.

When estimating population size by maximum likelihood, we reverse the role of parameters and data. We know our data (or, in this case, datum, i.e., n = 20) and ask what is the most likely population size that would have produced our observation. For simplicity in this example, we assume that q = 0.4 and is known. Now, we can write

$$P(N = 60|n = 20, q = 0.4) = \frac{N!}{n!(N-n)!}\, q^n (1-q)^{N-n}$$

$$= \frac{60!}{20!(60-20)!}\, 0.4^{20}(1-0.4)^{60-20} = 0.0616.$$

Note that this is mathematically identical to the previous equation. However, we now refer to $P(N = 60|n = 20, q = 0.4)$ as the likelihood. When using the likelihood, we generally take the view that the parameter we are estimating (N) can be varied to maximize this likelihood. The process of calculating the likelihood for a series of different parameter values over some range is referred to as profiling the likelihood. The fundamental concept of statistical likelihood is that our observations (the data) occur through a stochastic or random process with a defined probability structure. Through this process, we are likely to observe data in proportion to their probabilities as described in the formulae above.

As shown in Figure 8.3B, the likelihood given the data (n = 20) is maximized for N = 50. Our estimate is less than the true value of 60 because we happened to capture somewhat less than one might typically capture. Note that although N = 60 does not maximize the likelihood, it has a likelihood that is reasonably high (Figure 8.3B). Thus, we cannot rule out N = 60, as it could have reasonably generated the observed data. On the other hand, the likelihood is very low for N = 100. If we repeated the process of sampling, sometimes our population estimates would be above and other times they would be below the true value, but our estimates would very rarely be above 100. As such, the likelihood is a measure of how consistent the data are with different population sizes.

(Box continues)

Box 8.2 *(continued)*

In this simple example, we assumed that q was known. If we had not, we could not have computed a unique solution (e.g., our data could have resulted from a combination of smaller q and larger N). As indicated in the introduction to this chapter, we generally need more information than catch per unit effort (C/f) from a single sampling event to estimate true abundance. In our simple example, the additional information we need is the probability of capture with a single pass (i.e., catchability).

To be somewhat more realistic, assume that q is unknown and we apply a depletion sampling experiment (see section 8.4) to the stream and catch 24 fish in the first sampling pass, 17 fish in the second sampling pass, and 8 fish in the third sampling pass. We will also make the usual assumptions that the population is closed and that all fish have equal vulnerability and that this is consistent over time. The details of the likelihood function for the removal method are presented in section 8.4, equation (8.24). Note that the $\log_e$ of the likelihood is often used to make the computations more tractable. Applying the formula to various levels of catchability from 0 to 1.0, we can profile the likelihood for these data as shown in Figure 8.4A.

From Figure 8.4A, it is apparent that it is possible that catchability (q) is equal to 0.6, but it is not very likely relative to other possible values of q. Likewise, q could be 0.01, but that too is not very consistent with our observations. In this example, the value of q that is most consistent with our observations is 0.40. Thus, we term this the maximum-likelihood estimate of q. Because we sampled the population three times, the estimated cumulative proportion of the population removed is $1 - (1 - q)^3 = 1 - (0.6)^3 = 0.784$. Given that we caught a total of 49 fish, the most likely estimate of N is $49/0.784 = 62$ (Figure 8.4B).

There are several ways to estimate variances and CIs associated with maximum-likelihood estimates. One way is to consider how the likelihood changes when the parameters move small distances away from the maximum-likelihood value. The first derivative of the logarithm of the likelihood measures how quickly the likelihood changes relative to a change in the parameter and is equal to zero at the maximum. Variance is estimated by taking the negative of the reciprocal of the second partial derivative of the logarithm of the likelihood with respect to each parameter (Seber 1982). The second partial derivative measures the curvature of the log-likelihood portrayed in Figure 8.4. If the magnitude of the second derivative is large, the likelihood falls off rapidly as we move the parameters away from the maximum-likelihood estimate; the estimated variance would be relatively small because alternative values very far from the estimate are unlikely. Confidence intervals can be constructed from the variance estimated above, assuming a distribution (often normal) for the estimate. The profile likelihood can also be used to construct CIs directly by determining values for the parameters that give a log-likelihood value that is less than the maximum value of the log-likelihood by 3.841. This method is based on the fact that, under the null hypothesis, this difference approximates a χ^2 distribution with 1 df, and 3.841 is the 5% critical value for the χ^2 distribution with 1 df. As shown in Figure 8.4B, the maximum-likelihood estimate of N is 62 with a 95% CI of 51 to approximately 1,650.

Both approaches for computing variances and CIs produce approximations based on asymptotic (i.e., large sample) statistical properties and require relatively large sample sizes to be accurate. The profile likelihood method often performs better because the shape of the likelihood profile is examined and no assumption of normality is made. The better performance of the profile likelihood method comes at the cost of greater computation, however.

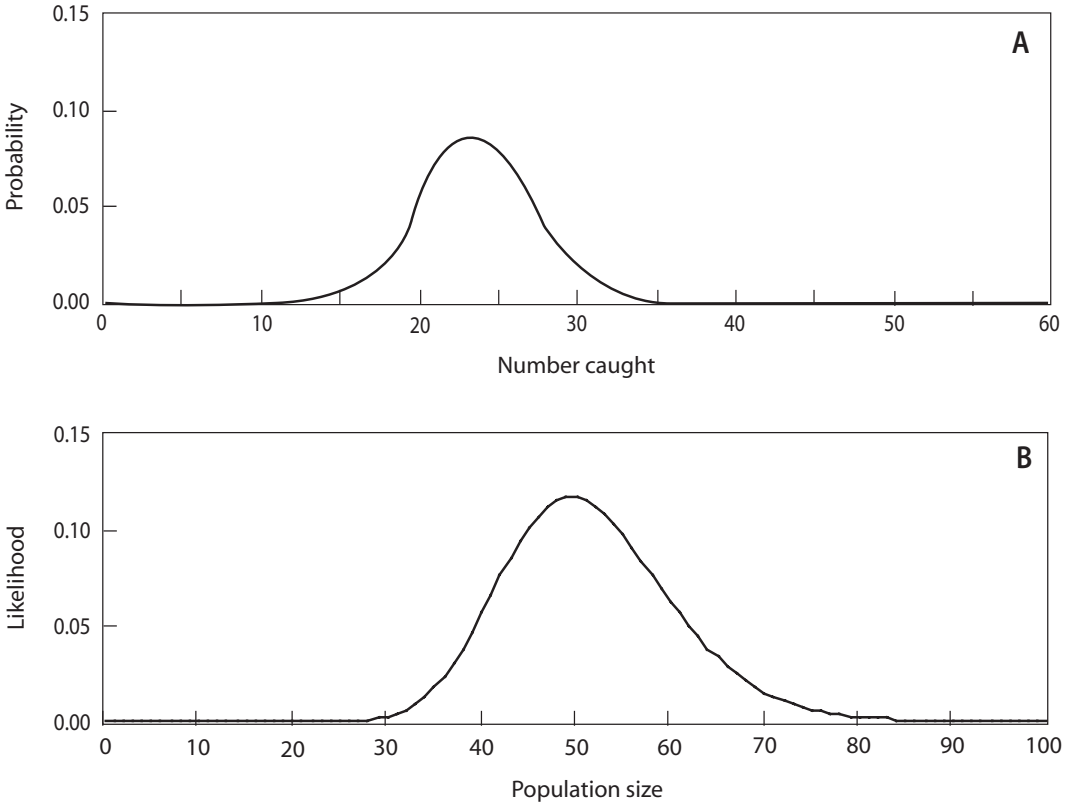


Figure 8.3 (A) Probability of capturing *n* fish from a population with 60 individuals, each with a 40% chance of capture, and (B) log-likelihood of the observation ($n = 20$ fish caught) as a function of population size *(N)*.

The asymptotic variance of $\hat{N}$ under model M_t is (Otis et al. 1978)

$$V(\hat{N}) = \frac{\hat{N}}{\dfrac{1}{\prod_{j=1}^{t}(1-\hat{p}_j)} + (t-1) - \sum \dfrac{1}{1-\hat{p}_j}}. \tag{8.13}$$

As with model M_o, CIs for $\hat{N}$ under model M_t can be estimated using the variance of $\hat{N}$ and an assumption of normality or through a likelihood-based approach as outlined in Box 8.2.

Model M_b (variability in capture probability due to changes in behavior after capture) has three parameters: N, p, the probability of capturing an unmarked animal, and c, the probability of capturing an animal that was previously captured, marked, and released. The parameter c can be estimated separately from estimation of N and p by (Otis et al. 1978)

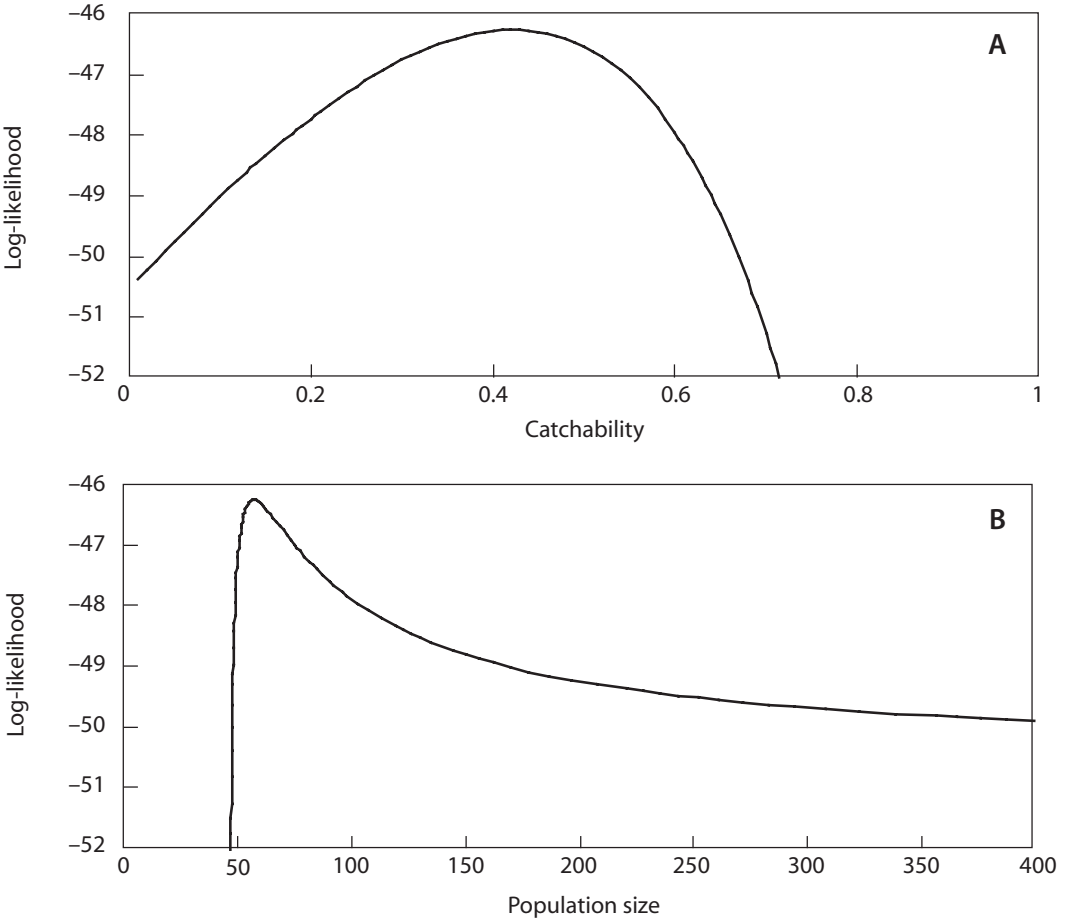


Figure 8.4 Log-likelihood as a function of (A) catchability *(q)* and (B) population size (*N)* given depletion sampling experiment described in Box 8.2.

$$\hat{c} = \frac{m_.}{M_.}, \tag{8.14}$$

where $m_. = \Sigma m_j$; m_j = number of marked animals in jth sample; $M_. = \Sigma M_j$; and M_j = number of marked animals in the population for the jth sample.

The likelihood function for model M_b is (Otis et al. 1978)

$$\begin{aligned} \log_e L(N) = \log_e &\left[\frac{N!}{(N - M_{t+1})!}\right] + \left[M_{t+1}\log_e(M_{t+1})\right] \\ &+ \left[(tN - M_. - M_{t+1})\log_e(tN - M_. - M_{t+1})\right] \\ &- \left[(tN - M_.)\log_e(tN - M_.)\right] + m_.\log_e(\hat{c}) + (M_. - m_.)\log_e(1 - \hat{c}). \end{aligned} \tag{8.15}$$

Once $\hat{N}$ has been found by maximizing equation (8.15), the maximum-likelihood estimate of p is calculated as (Otis et al. 1978)

$$\hat{p} = \frac{M_{t+1}}{t\hat{N} - M.}. \tag{8.16}$$

An asymptotic variance estimate for $\hat{N}$ is (Otis et al. 1978)

$$V(\hat{N}) = \frac{\hat{N}(1-\hat{p})^t\left[1-(1-\hat{p})^t\right]}{\left[1-(1-\hat{p})^t\right]^2 - t^2\hat{p}^2(1-\hat{p})^{t-1}}. \tag{8.17}$$

Estimation of the parameters for model M_h is more problematic than it is for model M_o, M_t, or M_b. The reason for this is that each fish (including unobserved fish) has its own individual catchability. A number of approaches have been taken to solve this problem, generally by making an assumption regarding the statistical distribution of catchabilities. For details of computation for this model, we refer the reader to Otis et al. (1978) and to the program MARK (White and Burnham 1999).

An example applying models M_o, M_t, and M_b is given in Box 8.3. Beyond being able to estimate the parameters for each of these models, an important question is how to choose among them. The most common way of doing this is to compare the maximum-likelihood value for each model and select the model with the highest maximum likelihood. Because the maximum likelihood that can be obtained generally increases as more parameters are added, the likelihood obtained from models with more parameters is typically "penalized" for the additional flexibility offered. The most widely used adjustment to the likelihood function is Akaike's Information Criterion (AIC; Akaike 1973), which is calculated as

$$\text{AIC} = -2\log_e(\text{likelihood}) + 2\,(\text{number of parameters}). \tag{8.18}$$

After computing the AIC, one then selects the model that has the lowest AIC value (Box 8.3).

8.3.2 Open Population Mark–Recapture Methods

Open populations are characterized by having immigration, emigration, mortality, or recruitment occur during the study period. As in closed populations, general models developed to estimate abundance in open populations also make use of the encounter history matrix as the basis for maximum-likelihood estimators and assume that each fish is uniquely marked. Conceptually, the encounter history matrix is important because it defines which animals are observed at particular times. From this, we can also infer which time periods the animal is known to

Box 8.3 Estimation of Population Abundance for a Closed Population Based on Otis et al.'s (1978) Mark–Recapture Models

An investigator conducts a mark–recapture study on a closed population of largemouth bass in a farm pond in order to determine the abundance of adult fish. The sampling consists of four sampling events; fish captured in each event are given a uniquely numbered Floy Tag and released. The capture–recapture data are arranged into a capture matrix in which each cell of the matrix (X_{ij}) is referenced by fish$_i$ in row i and sample period$_j$ in column j. An entry of 1 in the matrix indicates that a fish was caught, and a 0 indicates that the fish was not caught during that sampling period. Fish 1, for example was caught in all four sampling periods, whereas fish 4 was caught in only the first sample period.

Table Data matrix for mark–recapture study of closed population of largemouth bass.

Fish	Sample 1	Sample 2	Sample 3	Sample 4
1	1	1	1	1
2	1	1	0	0
3	1	0	1	0
4	1	0	0	0
5	1	1	0	1
6	1	0	1	1
7	1	0	0	0
8	0	1	1	0
9	0	1	0	0
10	0	1	0	1
11	0	1	0	0
12	0	1	0	0
13	0	1	1	1
14	0	0	1	0
15	0	0	1	0
16	0	0	1	0
17	0	0	1	1
18	0	0	0	1
19	0	0	1	1
20	0	0	0	1

From these data, the investigator explores which of the Otis et al. (1978) suite of capture–recapture models is most appropriate. For this investigation, we obtain the following basic statistics that are used in the estimation of population abundance, for which t = number of sampling occasions; n_i = number of fish caught in ith sample; m_i = number of fish with marks caught in ith sample; and M_i = number of marked fish present in the population for ith sample.

$t = 4$;
$M_{t+1} = 20$;
$n_1 = 7, n_2 = 9, n_3 = 10, n_4 = 9, n. = 35$;
$m_1 = 0, m_2 = 3, m_3 = 5, m_4 = 7, m. = 15$; and
$M_1 = 0, M_2 = 7, M_3 = 13, M_4 = 18, M. = 38$.

Starting with Model M_o (see Figure 8.2), we compute the log-likelihood for trial values for $\hat{N}$ by applying equation (8.8). Two examples for trial values are 30 and 23. Using these values, we obtain

$$\log_e L(\hat{N} = 30|X) = \left\{ \log_e \left(\frac{30!}{(30-20)!} \right) + [35\log_e(35)] + [(4\cdot 30 - 35)\log_e(4\cdot 30 - 35)] - [4\cdot 30\log_e(4\cdot 30)] \right\} = -12.890; \text{ and}$$

$$\log_e L(\hat{N} = 23|X) = \left\{ \log_e \left(\frac{23!}{(23-20)!} \right) + [35\log_e(35)] + [(4\cdot 23 - 35)\log_e(4\cdot 23 - 35)] - [4\cdot 23\log_e(4\cdot 23)] \right\} = -11.299.$$

When the log-likelihood is computed and plotted for trial values of $\hat{N}$ ranging from 21 to 100 for model M_t (equation [8.11]), we find that the maximum of the log-likelihood is

$$\begin{aligned}\log_e L(\hat{N} = 30|\mathrm{X}) = \log_e \left[\frac{30!}{(30-20)!} \right] &+ [7\log_e(7) + 9\log_e(9) + 10\log_e(10) + 9\log_e(9)] \\ &+ \{[(30-7)\log_e(30-7)] + [(30-9)\log_e(30-9)] + [(30-10)\log_e(30-10)] \\ &+ [(30-9)\log_e(30-9)]\} - [4\cdot 30\log_e(30)] = -12.492; \text{ and}\end{aligned}$$

$$\begin{aligned}\log_e L(\hat{N} = 23|\mathrm{X}) = \log_e \left[\frac{23!}{(23-20)!} \right] &+ [7\log_e(7) + 9\log_e(9) + 10\log_e(10) + 9\log_e(9)] \\ &+ \{[(23-7)\log_e(23-7)] + [(23-9)\log_e(23-9)] + [(23-10)\log_e(23-10)] \\ &+ [(23-9)\log_e(23-9)]\} - [4\cdot 23\log_e(23)] = -10.854.\end{aligned}$$

The maximum of the log-likelihood for model M_t is –10.854 at $\hat{N} = 23$ (Figure 8.5).

When model M_b is employed (equation [8.15]), the log-likelihood for the same trial values is

$$\begin{aligned}\log_e L(30|X) = \log_e \left[\frac{30!}{(30-20)!} \right] &+ [20\log_e(20)] + [(4\cdot 30 - 38 - 20)\log_e(4\cdot 30 - 38 - 20)] \\ &- [(4\cdot 30 - 38)\log_e(4\cdot 30 - 38)] + 15\log_e(0.395) + (38-15)\log_e(1-0.395)] = -11.491; \text{ and}\end{aligned}$$

$$\begin{aligned}\log_e L(23|X) = \log_e \left[\frac{23!}{(23-20)!} \right] &+ [20\log_e(20)] + [(4\cdot 23 - 38 - 20)\log_e(4\cdot 23 - 38 - 20)] \\ &- [(4\cdot 23 - 38)\log_e(4\cdot 23 - 38)] + 15\log_e(0.395) + (38-15)\log_e(1-0.395)] = -11.270.\end{aligned}$$

The maximum of the log-likelihood for model M_b is –11.247 at $\hat{N} = 24$ (Figure 8.5). The Akaike's Information Criterion (AIC) for each model is

AIC for $M_o = -2(-11.299) + 2(2) = 26.598$;
AIC for $M_t = -2(-10.854) + 2(5) = 31.708$; and
AIC for $M_b = -2(-11.247) + 2(3) = 28.494$.

Based on the AIC, we would choose model M_o as the best model among those considered. The likelihood for this model is not substantially lower than for M_t and M_b, but it requires fewer parameters, resulting in a more parsimonious model.

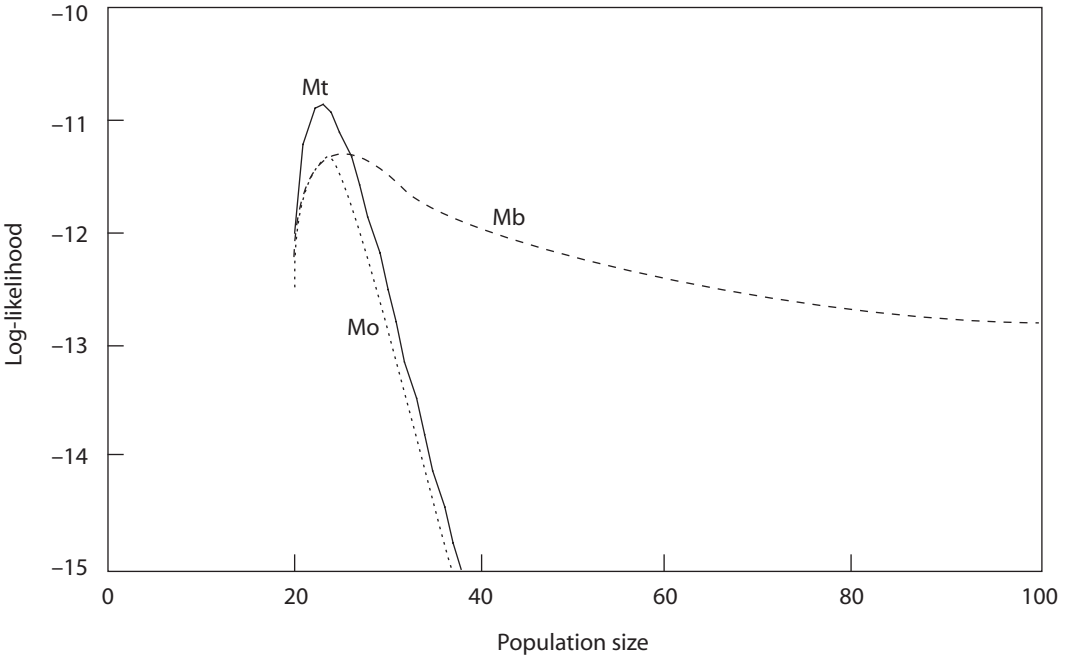


Figure 8.5 Log-likelihood as function of population size *(N)* for Otis et al.'s (1978) hierarchy (see Figure 8.2) based on example in Box 8.3. An approximate confidence interval for *N* under model M_o is where the *x*-axis crosses the log-likelihood curve.

be alive even if it is not observed. For example, a fish marked at the beginning of an experiment, but not observed again until the end, is known to have survived through all intervening sampling events. Open population models are also similar to the Otis et al. (1978) hierarchy of models in that numerous factors and assumptions can be represented in a suite of models applicable to open populations. Compared with closed population models, additional parameters describing losses and additions to the population are necessary for open population models. The additional parameters necessary to describe open populations often lead to a decline in the precision of population estimates. Further, allowing for an open population requires stronger adherence to some of the model assumptions. In particular, heterogenity in capture probability becomes increasingly important and can lead to bias in population estimates.

In Box 8.4, we demonstrate computations for a basic model illustrating the underlying approach applicable to a broader range of possible models for open populations. Lebreton et al. (1992) and Seber (1982) provide in-depth coverage of this broader family of models, commonly referred to as Cormack–Jolly–Seber (CJS) survival models. The software program MARK, referred to earlier, can accommodate CJS models. The program RELEASE (available at http://www.warnercnr.colostate.edu/~gwhite/software.html) was also recommended by Lebreton et al. (1992).

Box 8.4 represents a commonly used open population model in which the abundance of animals changes over time due to births and deaths, survival varies over time, but capture probability is constant over time and across all individuals in the population. As such, this model is analogous to model M_o, with the addition of time-varying population abundance and survival. In the CJS models, four basic sets of parameters are estimated: population abundance (N_i), capture probability (p_i), apparent survival (φ_i), and additions (births and immigrant) to the population (B_i). The term apparent survival is used instead of survival because, in most cases, it is impossible to distinguish any losses due to emigration from mortality. If the population is geographically closed, φ_i is an estimator for actual survival rate. Each of the above parameters are indexed by time, but care must be taken in understanding that φ_i indicates the survival rate from time i to $i + 1$. Further, not all quantities are estimable; for example, abundance at the beginning of the study (N_1) generally cannot be determined. The application of this model is illustrated in Box 8.4. For simple models, closed-form equations exist to estimate population size and other necessary parameters. In more complex situations, an iterative (i.e., starting with an initial guess, and then using a numerical optimization to improve the fit) approach is necessary to solve the likelihood equations.

8.4 POPULATION ESTIMATION: REMOVAL METHODS

8.4.1 Closed Population Removal Methods

Like mark–recapture methods, removal methods rely on sequentially sampling the target population. During each sampling period, the number of fish captured are recorded, and captured fish are temporarily (e.g., during monitoring surveys) or permanently (e.g., in recreational or commercial fisheries) removed from the population. Through the reduction in the population, catch in subsequent sampling periods is reduced. The rate at which catch declines gives a measure of the proportion of the original population that has been removed.

As with mark–recapture methods, removal methods generally rely on the population being closed and individuals in the population having equal vulnerability to the sampling gear. Typically, equal amounts of effort are expended during each sampling period, and it is assumed that the capture probability is equal across all sampling periods. Historically, regressions relating C/f to cumulative catch (Leslie method, Leslie and Davis 1939) or cumulative effort (De Lury method, De Lury 1947) were used to estimate population size in removal experiments. These methods are still commonly used and often result in reasonable population estimates. Currently, there is a shift away from the regression-based methods to likelihood-based methods. The principal advantage of likelihood methods over regression methods is that they provide means for testing some of the assumptions of the removal method and creating models that can accommodate a relaxed set of assumptions. For example, the assumption of equal catchability over all sampling periods can be relaxed if a function can be used to describe how catchability changes over time.

Box 8.4 Estimation of Abundance Based on a Cormack–Jolly–Seber Model for Open Populations

In order to determine the conservation status of desert pupfish, a graduate student performs a 3-year capture–recapture experiment on the population in a desert pool that is closed to immigration and emigration but where recruitment and mortality occur on an annual basis.

Table Capture matrix from capture–recapture experiment with desert pupfish.

	Year		
Fish idenfication	1998	1999	2000
1	1	1	1
2	1	1	1
3	1	1	0
4	1	1	0
5	1	0	1
6	1	0	1
7	1	0	1
8	1	0	1
9	1	0	0
10	1	0	0
11	1	0	0
12	1	0	0
13	1	0	0
14	0	1	1
15	0	1	1
16	0	1	1
17	0	1	1
18	0	1	0
19	0	1	0
20	0	1	0
21	0	1	0
22	0	0	1
23	0	0	1
24	0	0	1
25	0	0	1
26	0	0	1
27	0	0	1
28	0	0	1
29	0	0	1
30	0	0	1

From these data, we define the following components of the Cormack–Jolly–Seber model for an open population:

$$\hat{M}_i = m_i + \frac{R_i z_i}{r_i};$$

$$\hat{N}_i = \frac{n_i \hat{M}_i}{m_i} = \text{estimated population abundance;}$$

$$\hat{p}_i = \frac{m_i}{\hat{M}_i} = \text{estimated capture probability;}$$

$$\hat{\varphi}_i = \frac{M_{i+1}}{\hat{M}_i - m_i + R_i} = \text{estimated apparent survival;}$$

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\varphi}_i(N_i - n_i + R_i) = \text{estimated additions (births and immigration) to the population,}$$

where M_i = number of marked animals in population at start of sample *i;* m_i = number of marked fish caught in *i*th sample; n_i = number of fish caught in *i*th sample; R_i = number of fish caught in *i*th sample that are marked and released; z_i = number of fish caught before *i*th sample that are not captured in *i*th sample but are caught at a later time; and r_i = number of fish released at the *i*th sample that are later recaptured.

From these components, we can estimate the following quantities:

$m_1 = 0, m_2 = 4, m_3 = 10;$

$n_1 = 13, n_2 = 12, n_3 = 19;$

$r_1 = 0, r_2 = 6, r_3 = 0;$

$z_1 = 0, z_2 = 4, z_3 = 0;$

$\hat{M}_1 = 0$ (by definition), $\hat{M}_2 = 4 + \frac{12 \cdot 4}{6} = 12$, $\hat{M}_3$ = not estimable;

$\hat{N}_1$ = not estimable, $\hat{N}_2 = \frac{12 \cdot 12}{4} = 36$, $\hat{N}_3$ = not estimable;

$\hat{p}_1$ = not estimable, $\hat{p}_2 = \frac{4}{12} = 0.33$, $\hat{p}_3$ = not estimable; and

$\hat{\varphi}_1 = \frac{12}{0 - 0 + 13} = 0.92$, $\hat{\varphi}_2$ = not estimable.

As this example illustrates, estimating abundance for open populations is much more difficult than for closed populations, with several important population parameters not being estimable.

In the simplest situation, all of the above assumptions hold, and sampling is done during two periods with equal effort. Catch (i.e., the number removed, n_i) is recorded for each sampling period. Zippin (1956, 1958; Seber 1982) showed that the maximum-likelihood estimator in this situation is

$$\hat{N} = \frac{n_1^2}{n_1 - n_2}. \tag{8.19}$$

Note that n_1 must be greater than n_2 to estimate $\hat{N}$. An estimate of catchability, $\hat{q}$, is given by

$$\hat{q} = \frac{n_2}{n_1}. \tag{8.20}$$

The variance estimator for $\hat{N}$ is

$$V(\hat{N}) = \frac{n_1^2 n_2^2 (n_1 + n_2)}{(n_1 - n_2)^4}. \tag{8.21}$$

Because of the dependence on only two data points, the precision of population estimates in this situation is often poor. Moreover, variability in catches can result in higher C/f in the second sampling period, resulting in no estimate of abundance. Heimbuch et al. (1997) present a method for adjusting population estimates in two-pass sampling when many sites are visited. In their method, the catch data are added across sites to allow an estimate of the population for all sites together. By adding data from many sites, variations in the catch at individual sites tend to cancel, resulting in a better estimate of average catchability. Heimbuch et al. (1997) also present extensions to this method by which variability in individual catchability can be accounted for, analogous to model M_h of Otis et al. (1978).

When fish are removed during three sampling periods, the following maximum-likelihood estimators for catchability and population size (Junge and Libosvárský 1965; cited in Seber 1982) may be applied:

$$\hat{q} = \frac{3X - Y - \sqrt{Y^2 + 6XY - 3X^2}}{2X}, \text{ and} \tag{8.22}$$

$$\hat{N} = \frac{6X^2 - 3XY - Y^2 + Y\sqrt{Y^2 + 6XY - 3X^2}}{18(X - Y)}, \tag{8.23}$$

where $X = 2n_1 + n_2$ and $Y = n_1 + n_2 + n_3$.

Variance estimates and CIs for $\hat{N}$ are covered in the general case (below) of four or more removal passes.

When more than three removal passes are conducted, there is no closed-form equation available for directly estimating population size from the data by means of current maximum-likelihood methods. As in the more complex mark–recapture situations, the relative likelihood of parameter values is calculated, and numerical search methods are used to determine which combination of parameter values is most likely given the observed data. When catchability (q) is assumed to be constant over time, the results of this analysis are easy to portray graphically as a profile likelihood (see Box 8.2). In the maximum-likelihood approach, catchability is generally estimated directly, and $\hat{N}$ is calculated from the cumulative catch and estimated cumulative proportion of the population that this represents, based on the estimated catchability (Box 8.5). The likelihood function (dropping those parts of the function that are constants not affecting estimation) for estimating $\hat{q}$ is (Gould and Pollock 1997)

$$\begin{aligned}\log_e(q\,|\,n_1, n_2, n_3, \ldots n_t) = \log_e &\left(\frac{x_{t+1}!}{n_1!, n_2!, n_3!, \ldots n_t!}\right) \\ &+ n_1\log_e\left(\frac{q_1}{1 - q_1 - p_1p_2 - p_1p_2q_3 - \ldots p_1p_2 \ldots q_t}\right) \\ &+ n_2\log_e\left(\frac{q_2p_1}{1 - q_1 - p_1p_2 - p_1p_2q_3 - \ldots p_1p_2 \ldots q_t}\right) \\ &+ n_3\log_e\left(\frac{q_3p_2p_1}{1 - q_1 - p_1p_2 - p_1p_2q_3 - \ldots p_1p_2 \ldots q_t}\right) + \ldots,\end{aligned} \tag{8.24}$$

where t = number of removal passes; n_i = catch in ith sample; x_i = cumulative catch prior to removal pass i; q_i = probability of capture in ith removal pass; and $p_i = 1 - q_i$.

Once q_i has been estimated, $\hat{N}$ is estimated by

$$\hat{N} = \frac{x_{t+1}}{(1 - \hat{q}^t)}. \tag{8.25}$$

When the cumulative removal is relatively large (e.g., greater than 30), the asymptotic variance of $\hat{q}$ and $\hat{N}$ are (Seber 1982)

$$V(\hat{N}) = \frac{\hat{N}(1 - \hat{q}^t)\hat{q}^t}{(1 - \hat{q}^t)^2 - \{[t(1 - \hat{q})]^2\hat{q}^{t-1}}, \text{ and} \tag{8.26}$$

$$V(\hat{q}) = \frac{[(1 - \hat{q})\hat{q}]^2(1 - \hat{q}^t)}{\hat{N}\langle\hat{q}(1 - \hat{q}^t)^2 - \{[t(1 - \hat{q})]^2\hat{q}^t\}\rangle}. \tag{8.27}$$

Box 8.5 Estimation of Abundance Based on the Removal Method in a Closed Population

In order to estimate the abundance of brown trout in a 50-m section of stream below a culvert, a fishery manager conducts a three-pass removal experiment. Fish cannot move upstream because of the culvert, and the manager places a block net on the lower section of the study reach to insure that the population is geographically closed. All three sampling passes are conducted during the same day by means of a backpack electrofishing unit. During sampling, 24 brown trout are caught in the first sampling pass, 17 in the second sampling pass, and 8 in the third sampling pass.

Although the population size can be estimated applying equation (8.23), we illustrate the application of the more general likelihood equation (8.24). Given a trial value for catchability (q) of 0.2,

$$\log_e(q = 0.2|n_1 = 24, n_2 = 17, n_3 = 8) = \log_e\left(\frac{49!}{24!17!8!}\right) + 24\log_e\left(\frac{0.2}{1 - 0.2 - (0.8)(0.2) - (0.8)(0.8)(0.2) - (0.8)(0.8)(0.8)(0.2)}\right)$$

$$+ 17\log_e\left(\frac{(0.2)(0.8)}{1 - 0.2 - (0.8)(0.2) - (0.8)(0.8)(0.2) - (0.8)(0.8)(0.8)(0.2)}\right)$$

$$+ 8\log_e\left(\frac{(0.2)(0.8)(0.8)}{1 - 0.2 - (0.8)(0.2) - (0.8)(0.8)(0.2) - (0.8)(0.8)(0.8)(0.2)}\right) = 60.409.$$

A search across a range of $\hat{q}$ from 0.01 to 0.99 in steps of 0.01 indicates that the most likely value of $\hat{q}$ is 0.40, with a log-likelihood value of –49.832. From this, $\hat{N}$ is calculated as

$$\hat{N} = \frac{49}{(1 - 0.4^3)} = 63,$$

with an estimated variance of

$$\text{Var}(\hat{N}) = \frac{63(1 - 0.4)^3(0.4^3)}{[(1 - 0.4^3)^2] - \{[3(1 - 0.4)]^2(0.4^2)\}} = 10.55.$$

Confidence intervals are typically obtained from the profile likelihood of $\hat{q}$. From the search across values of $\hat{q}$ ranging from 0.01 to 0.99 (in 0.01 increments), the log-likelihood values for $q = 0.65$ and $q = 0.01$ differed from the log-likelihood at $\hat{q} = 0.4$ by 3.841 or more (which is the critical value for the χ^2 distribution with 1 df). The population sizes corresponding to these values of q are 51 and 1,650 and represent approximate 95% CIs for $\hat{N}$.

Confidence intervals can be obtained by assuming that $\hat{q}$ and $\hat{N}$ are normally distributed (equation [8.5]) or from the profile likelihood of (e.g., Box 8.2). Once $\hat{q}$ and $\hat{N}$ have been estimated, the goodness of fit of the estimates can be assessed by comparing the expected catches with the observed catches. Expected catch for each removal pass is predicted by

$$\tilde{\chi}_1 = \hat{N}\hat{q}; \tag{8.28}$$

$$\tilde{\chi}_2 = \hat{N}\hat{q}(1 - \hat{q}); \tag{8.29}$$

$$\tilde{\chi}_3 = \hat{N}\hat{q}(1-\hat{q})^2 \ldots ; \text{and} \tag{8.30}$$

$$\tilde{\chi}_t = \hat{N}\hat{q}(1-\hat{q})^{t-1}. \tag{8.31}$$

Goodness of fit can then be assessed by a χ^2 test by comparing the observed catches with the expected catches. This provides a useful diagnostic test to determine if the assumption of constant catchability over time is reasonable.

$$\chi^2 = \sum \frac{(\chi_i - \tilde{\chi}_i)^2}{\tilde{\chi}_i}, \tag{8.32}$$

where χ_i = observed catch in pass i, and $\tilde{\chi}_i$ = expected catch in pass i.

One of the more common violations of the assumptions in the removal method is that individual fish often differ in their catchability, analogous to model M_h for mark–recapture studies. Two approaches can be used to estimate population size in this situation. The first approach rests on the observation that removal studies can be viewed as a special case of mark–recapture model M_b, where the "response" to capture is removal from the vulnerable population (this is equivalent to setting c in model M_b equal to 1.0). Heterogeneity in individual catchability can then be accounted for by fitting model M_{bh} in the Otis et al. (1978) hierarchy. The calculations for this model are complex, but program MARK includes this option.

The second approach for handling variations in catchability is to fit a time-varying function to $\hat{q}$. Because fish with higher catchability tend to be captured and removed earlier in the sampling process, the average catchability of the remaining population tends to decline as the population is depleted. Thus, additional parameters describing how $\hat{q}$ declines with each sampling pass can be estimated. We refer the reader to Schnute (1983) for more detailed description of this approach.

Several software packages are available to estimate abundance from removal experiments. White and Burnham's (1999) MARK handles removal data well and has the option of fitting alternate models as described above. Van Deventer and Platts' (1989) MicroFish is a software package available through the American Fisheries Society that is designed for removal studies. Its particular strength is that removal experiments from multiple sites and multiple species can be analyzed from a single data file.

8.4.2 Open Population Removal Methods

The application of removal methods to open populations is much more difficult than it is for closed populations because mortality and recruitment need to be estimated in addition to population size. Furthermore, removal methods are generally applied to open populations only when there is a fishery harvesting a substantial portion of the population. As such, the timing and magnitude of removals are often out of the fisheries scientist's control. Further, there is the potential

problem of under (or over) reporting of catch, resulting in biased estimates of population size. This is not to dissuade readers from pursuing removal methods for open populations—this is often the only feasible approach given the data available. Rather, we emphasize that the particular details of the data and the fishery will determine which model is most appropriate. In this chapter, we present a relatively simple formulation requiring minimal data to illustrate the essence of these methods.

Consider a population that is closed to immigration and emigration but is open to natural mortality (M), fishery harvest (C), and recruitment (R). One representation of the dynamics of the population is (Collie and Sissenwine 1983)

$$N_{t+1} = (N_t - C_t + R_t)\, e^{-M} + E_t. \tag{8.33}$$

In this model, E_t represents random variations in mortality that are not included in either catch or natural mortality (which is assumed to be constant). The parameter E_t reflects what is often called a process error, meaning the unaccounted variation in the underlying dynamical processes. Including this in the population dynamic equation (8.33) is important because process error actually influences system dynamics, and these process errors can accumulate over time. This model implicitly assumes that recruitment and fishery removals occur at the beginning of the year. Natural mortality operates at a constant rate for the remainder of the year, and a proportion (e^{-M}) survive to the beginning of the next year. Alternative formulations can be derived for populations for which the fishery and recruitment occur throughout the year (see Ricker 1977).

For the model described above, information on harvest alone is insufficient to estimate population abundance. Additional information in the form of relative abundance indices (e.g., C/f) for the adult stock (n_t) and recruits (r_t) are also required. Age-structured measures of C/f and population dynamic equations can also be used, leading to methods such as virtual population analysis or statistical catch-at-age. We refer the reader to Ricker (1977), Edwards and Megrey (1989), and Hilborn and Walters (1992) for a detailed discussion of these extensions.

If we assume that the expected C/f for adults and recruits is directly proportional to the true population size (N_t and R_t) and that all members of the population are equally vulnerable to the survey gear, we have

$$n_t = \hat{n}_t \eta_t = qN_t \eta_t, \text{ and} \tag{8.34}$$

$$r_t = \hat{r}_t \delta_t = qR_t \delta_t. \tag{8.35}$$

where q = proportionality constant relating survey C/f to true abundance (i.e., catchability in survey); η_t = measurement error term for adults with a mean of 1.0; and δ_t = measurement error term for recruits with a mean of 1.0.

We have assumed that adults and recruits have equal vulnerability to the survey gear. This assumption or a known ratio of recruit to adult vulnerability is generally required when using these Collie–Sissenwine catch survey models (Mesnil 2003).

A critical concept underlying equations (8.34) and (8.35) is that C/f, which is based on samples from the entire population, is generally estimated with considerable variance. The variance associated with these estimates is often termed measurement error and, in the context of population modeling, implies that C/f should not be treated as an exact measure of relative abundance but rather needs to be treated as being imprecise. Using equations (8.34) and (8.35) leads to the following dynamic equation describing the trajectory of the expected value for adult C/f:

$$\hat{n}_{t+1} = (\hat{n}_t - qC_t + \hat{r}_t)\, e^{-M} + \epsilon_t \,. \tag{8.36}$$

Here, $\epsilon_t = qE_t$ and is the process error as it influences adult C/f. The estimation procedure attempts to minimize these process errors as well as the measurement errors (see Box 8.6).

Generally, M is assumed to be known and constant over time. Under the additional assumption that the measurement errors are negligible (i.e., all are close to 1.0), equation (8.36) can be rewritten in a form by which standard linear regression can be used to estimate q (and thereby N_t and R_t). However, as Collie and Sissenwine (1983) state, n_t and r_t are generally both measured with substantial imprecision. Because of this, we recommend the methods of Collie and Sissenwine (1983; illustrated in Box 8.6) over a regression approach because the assumption of negligible measurement error is rarely credible.

We are not aware of any software program that handles the broad range of situations that are likely to occur when using removal methods in open populations. As such, practitioners must either use specialized software previously developed for special cases similar to theirs or develop the models and associated estimation routines in a general programming environment (e.g., C++, Visual Basic, or SAS), a spreadsheet environment (e.g., Microsoft Excel), or a specialized programming environment designed for statistical parameter estimation (e.g., AdModel Builder [Otter Research, Sidney, British Columbia]). Schnute et al. (1998) discuss some of the trade-offs faced in choosing software for such modeling.

8.5 BIOMASS AND YIELD ESTIMATION: SURPLUS PRODUCTION METHODS

In situations where a geographically closed population is subjected to a significant fishery (e.g., where the population has been substantially reduced by fishing; Hilborn and Walters 1992), it is sometimes possible to estimate biomass from the pattern of yield (biomass of fish removed) and fishing effort over time. Conceptually, surplus production models (also known as biomass dynamic models, Hilborn and Walters 1992) are based on the idea that the biomass in a given year (B_t) depends on the biomass in the previous year (B_{t-1}) plus recruitment and growth minus yield and natural mortality. It is often convenient to group recruitment and growth into a single term representing processes that contribute to biomass. If this production is in excess of natural mortality, the surplus production will increase the biomass from one year to the next. Alternately, the surplus production

Box 8.6 Estimation of Abundance Based on the Removal Method in an Open Population

A population of lake trout subjected to a commercial fishery was studied from 1985 to 2001 with the goal of determining trends in abundance over time. The population is sampled each year by a fishery-independent otter trawl survey. Data collected in the survey provide measures of relative abundance (C/f) for fish large enough to be vulnerable to capture in the commercial fishery (adults) and prerecruits that are not vulnerable to the commercial fishery. The number of fish harvested in the commercial fishery is recorded each year and is assumed to occur at the beginning of the year.

Table Lake trout catch in annual otter trawl survey.

Year	Catch (number of fish)	Adult C/f	Prerecruit C/f
1985	94,500	43.15	11.24
1986	99,154	38.46	7.99
1987	74,201	29.70	14.17
1988	65,827	32.85	19.15
1989	66,569	35.07	10.37
1990	69,000	34.38	17.56
1991	93,633	34.91	9.52
1992	78,069	31.05	14.06
1993	78,614	23.73	21.20
1994	82,258	37.11	12.41
1995	60,351	22.92	17.05
1996	48,212	23.49	13.23
1997	45,449	27.77	3.50
1998	34,020	28.58	21.12
1999	38,488	38.09	6.60
2000	44,865	32.04	8.75
2001	47,680	34.31	9.01

To proceed, we need initial values for the measurement errors (η_t for the adult C/f index and δ_t for the prerecruit C/f). A good initial guess would be to set all values to 1.0, but we show the table below with the final estimates. From these initial guesses of 1.0, we then fill in the columns for the expected indices by dividing the observed C/f values by these multiplicative errors. The final column in the table below is a forecast of the adult survey index at time $t + 1$ from the expected index at time t, ignoring process error (see equation [8.33]). To fill in this column, a value for q (catchability) is required. This is unknown (to be estimated), but an initial guess is needed to get started for this quantity also. One approach is to use past experience to obtain an initial value for exploitation rate in recent years, say 25%, and thus approximate N as four times C. Then, the initial value for q would be N/n, perhaps based on an average of such values over years. We illustrate calculations, however, with the final estimate for this parameter also. Once a value of q is available, forecast values are obtained by application of equation (8.36), dropping the process error and denoting the forecast value as $\tilde{n}_t$. For example, the forecasted value for 1986, assuming $q = 0.00011$ and M (natural mortality) = 0.2, is

$$\tilde{n}_{1986} = (\hat{n}_{1985} - qC_{1985} + \hat{r}_{1985})e^{-M} = (46.42 - 0.000110 \cdot 94{,}500 + 11.46)e^{-0.2} = 38.88$$

Note that no prediction is made for the first year of the time series (1985) because survey indices are not available for the year prior.

Table Forecasted values for the lake trout fishery. Given are measurement errors (η_t for the adult *C/f* index and δ_t for the prerecruit *C/f*); the expected, $\hat{\boldsymbol{n}}_t$, and forecasted, $\tilde{\boldsymbol{n}}_t$, adult survey index; and the expected recruitment survey index, $\hat{\boldsymbol{r}}_t$.

Year	η_t	$\hat{n}_t$	δ_t	$\hat{r}_t$	$\tilde{n}_t$
1985	1.076	46.42	1.020	11.46	
1986	1.032	39.68	1.007	8.04	38.88
1987	1.132	33.63	1.063	15.06	30.14
1988	0.964	31.68	0.979	18.76	33.18
1989	1.111	38.97	1.033	10.71	35.36
1990	0.969	33.32	0.984	17.28	34.68
1991	1.115	38.91	1.031	9.82	35.21
1992	0.810	25.14	0.914	12.85	31.46
1993	1.222	28.99	1.198	25.41	24.07
1994	0.709	26.29	0.902	11.20	37.46
1995	0.879	20.15	0.910	15.52	23.29
1996	1.089	25.59	1.050	13.89	23.76
1997	1.315	36.51	1.040	3.64	27.99
1998	1.017	29.07	1.013	21.39	28.78
1999	0.972	37.01	0.995	6.56	38.25
2000	1.183	37.90	1.050	9.19	32.21
2001	1.000	34.31	1.000	9.01	34.51

Assuming that *M* is known to be 0.2, the above model has three sets of parameters: *q*, which applies to both adults and prerecruits; η_t, which encapsulates measurement errors for the adult index; and δ_t, which encapsulates measurement errors in the recruitment index. Parameter estimation is accomplished by minimizing the following sum (representing the sums of squared deviations, or errors [SSE]):

$$\text{SSE} = \sum_t \log_e(\eta_t)^2 + \sum_t \log_e(\delta_t)^2 + \sum_t \epsilon_t^2$$

The first two components are directly calculated from parameter estimates that are also entries in the above table. The values of ϵ_t, the process error, depend jointly on all the estimated parameters and are calculated as $(\hat{n}_t - \tilde{n}_t)$. Thus, *q*, the ηs, and the δs are adjusted through an iterative search procedure from the initial guesses so as to minimize SSE. Note that underlying this minimization are assumptions that the process errors are normal, the logarithms of the measurement errors are normal (i.e., the original terms are lognormal), and the variances for each of these are equal. The terms in this sum could be weighted to represent different variances for each type of error (Collie and Sissenwine 1983).

Outputs of the model include estimates of the above parameters, as well as the annual population size and number of recruits entering the population. Using an Excel spreadsheet to do the calculations, and the solver function to minimize the SSE by changing the parameters, we obtained the following estimates for population size and recruitment.

(Box continues)

Box 8.6 *(continued)*

Table Estimate of lake trout fishery population size ($\hat{N}$) and recruitment ($\hat{R}$).

Year	$\hat{N}$	$\hat{R}$
1985	404,106	99,766
1986	345,247	70,078
1987	293,215	130,682
1988	275,807	163,343
1989	339,441	93,111
1990	290,035	150,535
1991	338,783	85,313
1992	220,131	110,506
1993	253,001	220,645
1994	231,912	94,364
1995	175,706	134,730
1996	223,029	120,793
1997	318,281	31,410
1998	253,196	186,325
1999	322,324	57,199
2000	330,591	79,522
2001	299,049	78,541

Collie and Sissenwine (1983) provide details on how to calculate the variance of the parameter estimates.

may be harvested and still maintain the population biomass. Generally, surplus production is related to the standing biomass; at low biomass levels, surplus production is low due to limited recruitment. At high biomass levels, surplus production is also generally low due to density-dependent growth, recruitment, or both. Surplus production typically peaks at intermediate levels of biomass.

Because of the relatively simple representation of population dynamics, surplus production methods do not require age-specific data. As such, these methods are often used in the analysis of difficult-to-age marine fish stocks. Moreover, simulation studies have suggested that management advice based on surplus production methods may be as robust as population estimates based on age-structured analyses using only yield and effort data (Ludwig and Walters 1985). Although surplus production models have not been widely used in freshwater fishery analysis, they are likely to be applicable and beneficial in some situations for which data are limited.

We approach the problem by developing a model of the biomass dynamics, using that model to predict fishery C/f over time, and then fine-tuning the parameters of the model so that the predicted C/f best fits the observed time series. Hilborn and Walters (1992) provide a thorough review of the principal approaches for fitting surplus production models to data in order to estimate biomass, recruitment, and

density dependence. Although not the only approach to estimation, they indicate that the time series approach we follow here appears to be the best. Although there are many variations of surplus production models, a common model is (Hilborn and Walters 1992)

$$B_t = B_{t-1} + rB_{t-1}\left(1 - \frac{B_{t-1}}{K}\right) - C_{t-1}\,; \tag{8.37}$$

$$\text{Observed } C/f_t = \frac{C_t}{E_t}\,; \text{ and} \tag{8.38}$$

$$\text{Predicted } C/f_t = \hat{q}\hat{B}_t\,, \tag{8.39}$$

where C_t = yield during year t; E_t = effort during year t; r = intrinsic rate of increase; K = carrying capacity; and $\hat{q}$ = catchability.

This formulation treats the observed yield as an exact measure of removals and C/f as an inexact measure of relative abundance. Although equation (8.37) directly involves only yield information, experience has shown that parameter estimation generally requires additional information on relative abundance over time. Here, we use fishery C/f as this auxiliary information calculated according to equation (8.38) and predicted by equation (8.39). The biomass at the start of the time series is also often estimated as a parameter in the model, allowing the iterative solution of equations (8.37) through (8.39) in order to fit best the observed and predicted time series of C/f. An example of the application of this approach is provided in Box 8.7. A useful software package for surplus production modeling is ASPIC (available at http://sefsc.noaa.gov/mprager/ASPIC.html).

8.6 BIOMASS ESTIMATION

Most of the methods presented in this chapter produce estimates of numerical abundance. In some situations, however, biomass (i.e., weight of the population) may be a better measure of the "size" of a population. Generally, biomass is estimated indirectly by multiplying the numerical abundance by the mean weight or by applying methods such as surplus production models that directly estimate biomass. In this section, we will cover indirect methods for estimating biomass.

In the simplest situation, biomass is estimated as

$$\hat{B} = \hat{N} \cdot \bar{w}\,, \tag{8.40}$$

where $\hat{B}$ = estimated biomass (g); $\hat{N}$ = estimated abundance; and $\bar{w}$ = mean weight of fish in the population (g).

In this equation, $\hat{N}$ can be estimated using any of the methods presented earlier, and mean weight is estimated from a random sample representative of the size- or age-groups contained in $\hat{N}$ (Anderson and Neumann 1996).

Box 8.7 Application of Surplus Production Modeling

The commercial fishery for a population of alewife was monitored from 1985 to 2001. Each year, the total weight of the catch (kg) and the total effort (days fished) were recorded, providing C/f as a measure of relative abundance. These data were analyzed using a surplus production model to estimate carrying capacity (K), initial biomass (B_0), catchability (q), and the intrinsic rate of growth (r) for this fishery population.

Table Catch and effort data for alewife fishery.

Year	Effort (days fished)	Catch (kg)	C/f (kg/d)
1985	825	90,000	109
1986	1,008	113,300	112
1987	1,411	155,860	110
1988	1,828	181,128	99
1989	2,351	198,584	84
1990	2,074	198,395	96
1991	1,877	139,040	74
1992	1,566	109,969	70
1993	1,139	71,896	63
1994	893	59,314	66
1995	1,029	62,300	61
1996	727	65,343	90
1997	658	76,990	117
1998	953	88,606	93
1999	1,012	118,016	117
2000	1,203	108,250	90
2001	1,034	108,674	105

With $B_0 = 800{,}000$ kg, $K = 4{,}000{,}000$, $q = 0.0001$, and $r = 0.17$ as initial guesses for the parameters of equations (8.37) and (8.39), we can predict catch and C/f as follows:

$$\hat{B}_{1986} = \hat{B}_{1985} + r\hat{B}_{1985}\left(1 - \frac{\hat{B}_{1985}}{\hat{K}}\right) - C_{1985}$$

$$= 800{,}000 + 0.17 \cdot 800{,}000\left(1 - \frac{800{,}000}{4{,}000{,}000}\right) - 90{,}000 = 818{,}800$$

$$\widehat{C/f}_{1986} = q\hat{B}_{1986} = 0.0001 \cdot 818{,}800 = 81.88 \approx 82$$

Table Recursive application of equations (8.37) and (8.39) result in time series of predicted values for the alewife fishery.

Year	Predicted biomass (kg)	C/f (kg/d)	Predicted C/f (kg/d)	Squared deviation for C/f
1985	800,000	109	80	841
1986	818,800	112	82	961
1987	816,203	110	82	841
1988	770,784	99	77	484
1989	695,440	84	70	225
1990	594,526	96	59	1369

Year	Predicted biomass (kg)	*C/f* (kg/d)	Predicted *C/f* (kg/d)	Squared deviation for *C/f*
1991	482,178	74	48	676
1992	415,227	70	42	841
1993	368,518	63	37	729
1994	353,498	66	35	961
1995	348,968	61	35	729
1996	340,818	90	34	3,136
1997	328,477	117	33	7,225
1998	302,742	93	30	3,969
1999	261,707	117	26	8,281
2000	185,270	90	19	5,184
2001	107,057	105	11	9,025

Note the discrepancy in the trend between observed *C/f* and predicted *C/f,* indicating that our initial guesses for parameter values were not very good. We used the solver function in Excel to perform a nonlinear search across the parameter values (i.e., B_0, K, q, and r were used as the "change cells" in solver) to find the combination of parameters that minimized the sum of squared deviations between predicted and observed *C/f.* Solver returned estimates of $\hat{B}_0 = 732{,}506$, $\hat{K} = 1{,}160{,}771$, $\hat{q} = 0.0001484$, and $\hat{r} = 0.4049$ with a sum of squared deviations of 1,616.7. (Note that when *C/f* is rounded to the nearest 1.0, the squared deviations sum to 1,433). Based on these parameter values as the best estimates, the predicted biomass and *C/f* over time is shown below.

Table Predicted values for the alewife fishery given parameter values that minimize the sum of squared deviations.

Year	Predicted biomass (kg)	*C/f* (kg/d)	Predicted *C/f* (kg/d)	Squared deviation for *C/f*
1985	732,506	109	109	0
1986	751,925	112	112	0
1987	745,852	110	111	1
1988	697,932	99	104	25
1989	629,475	84	93	81
1990	547,540	96	81	25
1991	466,259	74	69	25
1992	440,166	70	65	25
1993	440,828	63	65	4
1994	479,629	66	71	25
1995	534,265	61	79	324
1996	588,713	90	87	9
1997	640,836	117	95	484
1998	680,061	93	101	64
1999	705,480	117	105	144
2000	699,496	90	104	196
2001	703,787	105	104	1

Note that the trend in predicted *C/f* matches the observed trend in *C/f* closely after obtaining the best estimates for B_0, K, q, and r.

Assuming that the variance of $\hat{N}$ is estimated through methods described earlier, and the variance of $\bar{w}$ is also estimated, the variance of $\hat{B}$ is approximated as

$$V(\hat{B}) = \bar{w}^2 V(\hat{N}) + \hat{N}^2 V(\bar{w}) - V(\hat{N}) V(\bar{w}). \qquad \text{V(8.41)}$$

This approximation (Goodman 1960) is based on the assumption that $\hat{N}$ and $\bar{w}$ are estimated independently, an assumption that is reasonable in most cases.

Although equations (8.40) and (8.41) provide relatively simple means of obtaining point estimates of biomass and the associated variance, developing CIs for $\hat{B}$ is much more difficult because the distribution of $\hat{B}$ must be known or assumed. We are not aware of any general guidance in the literature suggesting a suitable distribution for $\hat{B}$. Since $\hat{B}$ is computed as the product of two random variables, the lognormal distribution is a reasonable choice (Aitchison and Brown 1976). Assuming a lognormal distribution, approximate 95% CI bounds for $\hat{B}$ are

$$e^{\left(\log_e(\hat{B}) \pm 1.96\sqrt{\mathrm{Var}(\hat{B})}\right)}. \qquad (8.42)$$

In many situations, abundance and mean weight are estimated separately for different age- or size-classes. In such situations, biomass can be estimated as

$$\hat{B} = \sum \hat{N}_i \cdot \bar{w}_i, \qquad (8.43)$$

where $\hat{N}_i$ = estimated abundance for class *i,* and $\bar{w}_i$ = mean weight of fish in class *i.*

In this case, the variance of $\hat{B}$ is

$$V(\hat{B}) = \sum [\bar{w}_i^2 V(\hat{N}_i) + \hat{N}_i^2 V(\bar{w}_i)], \qquad (8.44)$$

and the 95% CI can be computed following equation 8.42.

8.7 PRODUCTION ESTIMATION

8.7.1 Concepts and Definitions

Fish abundance parameters, such as density or biomass, are static measures of a population's status. That is, information on the state of the population is provided only for a single point in time. Conversely, dynamic population measures describe parameters as rate functions over time and may be more descriptive and meaningful for applications in fisheries science. Examples of dynamic population parameters are rates of recruitment, growth, and mortality (Chapters 4–6). Production is the integration of static and dynamic population measures over time, wherein biomass, recruitment, growth, and mortality are synthesized into a single dynamic measure. As such, production is an indicator of ecological success and is

especially responsive to environmental change (Mann and Penczak 1986). Thus, production rate of a fish population can be a useful measure and comparative tool, with many valuable applications for fisheries research and management.

Production is defined as the rate of tissue elaboration over time, regardless of whether it survives to the end of a given period (Waters 1977). It is expressed in units of quantity/space/time, usually kilograms/hectare/year for fish populations. Production rate represents the flow of energy through trophic levels and may also be expressed in units of calories/hectare/year.

The methods and terminology for estimating fish production have evolved to a generally accepted convention, and fish production estimates are routinely found in the literature, especially for fishes of small streams and salmonid species. However, many fisheries scientists rarely consider using this assessment tool, even though they may regularly gather the data required to estimate production. Presumably, this occurs because the computations can be complex and cumbersome and are more so if precision of production estimates and related parameters is estimated. The development of computer software and availability of other technical resources to minimize computation effort and reduce calculation error associated with the process of estimating production may increase the utility of this tool in fisheries science (Railsback et al. 1989; Kwak 1992).

8.7.2 Production Estimation Methods

Five methods to estimate production rate of aquatic animal populations have been developed, refined, and accepted by ecologists (Waters 1977; Bagenal 1978; Chapman 1978). Some of these methods were originally intended for estimating aquatic macroinvertebrate production but were readily adapted for use with fish populations. The five methods include two iterative summation methods, (1) the removal summation and (2) increment summation methods; (3) the instantaneous growth rate method and a graphical representation, (4) the Allen curve; and (5) the size-frequency method. Three of the methods (increment summation, instantaneous growth rate, and size-frequency) have been refined for application to fish populations, and variance estimators for all parameters associated with those methods have been derived. All methods except the size-frequency method are cohort based, meaning that information on the age structure of the fish population is required.

Sampling requirements to estimate fish production are a series of absolute density and biomass estimates (sections 8.3, 8.4, and 8.6) for a population within a 1-year period, with the first and last sampling dates occurring approximately 1 year apart to estimate annual production. Cohort-based methods require stratification and separate estimates by cohort; thus, data must be collected on population age structure. In general, production is estimated by individual cohorts for a single time interval; then, those partial estimates are summed for all cohorts to yield a production estimate for the entire population during the specific interval. The production estimate for a 1-year period (annual production) is the sum of the production estimates for intervals within the annual period. When the size-frequency method is

used, individual losses from one size-group to the next are summed from mean values over the annual period to yield a production estimate.

8.7.2.1 *Summation Methods*

Summation methods to estimate production stem from the concept that tissue lost from, or accumulated by, a population over a series of time increments is equivalent to an estimate of production. The removal summation method involves estimating the number of individuals lost, by mortality or other removal, from a cohort over a time interval and coupling those data with biomass information, which results in an estimate of production for that cohort during that interval. These are then summed over intervals and cohorts to estimate annual production. Removal summation is not typically applied to fish populations, and algorithms to estimate associated variance are not readily available. As such, we recommend the use of the more commonly used increment summation method in preference to the removal summation method.

Similar to the removal summation method, but quantifying accumulation rather than loss, the increment summation method sums the growth increments of a cohort over time. The growth increment is quantified as the increase in mean individual weight over a time interval for each cohort, and this increment is multiplied by the density of the cohort to obtain a production estimate for the cohort during that interval. Production for each cohort is summed for the population, and production for each interval is summed for an annual estimate.

Explicit formulae for increment summation production estimation and associated variance estimators for fish populations were developed by Newman and Martin (1983) as

$$\hat{P} = \bar{N}\Delta\bar{w}, \tag{8.45}$$

where P = production for a given cohort within a specified interval; $\bar{N}$ = estimated arithmetic mean cohort density from time t to $t+1$; and $\Delta\bar{w}$ = estimated change in mean weight of individuals in the cohort from time t to $t+1$ (i.e., $\bar{w}_{t+1} - \bar{w}_t$).

Sampling variance of the production estimate can be estimated as (Goodman 1960)

$$V(\hat{P}) = \bar{N}^2 V(\Delta\bar{w}) + (\Delta\bar{w})^2 V(\bar{N}) - V(\bar{N})\,V(\Delta\bar{w}), \tag{8.46}$$

where $V(\Delta\bar{w})$ and $V(\bar{N})$ are the estimated variances of $\Delta\bar{w}$ and $\bar{N}$, respectively.

Algorithms to estimate variance of $\Delta\bar{w}$ and $\bar{N}$ are found in Newman and Martin (1983) and require estimates of variance for each density and mean weight estimate (sections 8.3 and 8.6). The statistical software Pop/Pro (Kwak 1992; available on CDROM) includes a module to estimate fish production by the increment summation method according to cohort and time interval, including related parameters and associated variances.

8.7.2.2 Instantaneous Growth Rate and Allen Curve Methods

The instantaneous growth rate method was initially developed to estimate production of fish populations (Ricker 1946; Allen 1949). By this method, production is estimated as simply the product of the estimated instantaneous growth rate and estimated mean biomass as

$$\hat{P} = \hat{G}\,\overline{B}, \tag{8.47}$$

where $\hat{P}$ = estimated production for a given cohort within a specified interval, $\hat{G}$ = estimated instantaneous growth rate for the cohort from time t to $t+1$ (i.e., $\log_e \overline{w}_{t+1} - \log_e \overline{w}_t$), and $\overline{B}$ = estimated arithmetic mean cohort biomass from time t to $t+1$ (i.e., $\hat{B}_t + \hat{B}_{t+1})/2$).

From Newman and Martin (1983), the variance of the production estimate may be estimated as

$$V(\hat{P}) = V(\overline{B})\,\hat{G}^2 + V(\hat{G})\,\overline{B}^2, \tag{8.48}$$

where $V(\overline{B})$ and $V(\hat{G})$ are variances of the mean biomass and instantaneous growth rate, respectively. The variance of mean biomass is estimated as

$$V(\overline{B}) = [\,V(\hat{B}_t) + \mathrm{V}(\hat{B}_{t+1})]/4, \tag{8.49}$$

where $V(\hat{B}_t)$ and $\mathrm{V}(\hat{B}_{t+1})$ are the variances of biomass at times t and $t+1$, respectively.

The variance of the instantaneous growth rate may be estimated as

$$V(\hat{G}) = V(\log_e \overline{w}_t) + V(\log_e \overline{w}_{t+1}), \tag{8.50}$$

where $V(\log_e \overline{w}_t)$ and $V(\log_e \overline{w}_{t+1})$ are variances of the natural logarithms of estimated mean weights of individuals of the cohort at times t and $t+1$, respectively. By using a Taylor series expansion (delta method; Seber 1982; Cone and Krueger 1988), $V(\log_e \overline{w}_t)$ can be approximated as

$$V(\log_e \overline{w}_t) = V(\overline{w}_t)/\overline{w}_t^2. \tag{8.51}$$

Mean annual density or biomass estimates (and their variances) computed from multiple intervals of different duration must be weighted according to the number of days in each interval (i.e., equations [8.47] and [8.48] must be modified) and should be computed following formulae in Newman and Martin (1983). Fish production, including related parameters and associated variances, can be estimated according to cohort and time interval by the instantaneous growth rate method by means of Pop/Pro statistical software (Kwak 1992). Additional algorithms are available in the software documentation. An example calculation of fish annual production estimated by the instantaneous growth rate method is presented in Box 8.8.

Box 8.8 Production Estimation Based on the Instantaneous Growth Rate Method

Density, mean weight, and biomass (and associated variances) of a brook trout population in Valley Creek, Minnesota, were estimated in a stream reach with an area of 0.181 ha on four dates between March 1974 and March 1975 (Waters 1999). Population statistics for two of these dates are presented below in order to illustrate how to estimate production using the instantaneous growth rate method.

Table Population statistics for brook trout in Valley Creek, Minnesota.

Age-class and total	Density ($\hat{N}$)	$V(\hat{N})$	Mean weight ($\bar{w}$, g)	$V(\bar{w})$	Biomass ($\hat{B}$, g)	$V(\hat{B})$
Sampling date: 8 March 1974						
1	277.85	1,336.05	6.86	0.13	1,905.34	75,455.97
2	157.54	317.71	28.56	0.79	4,499.83	222,126.45
3	36.11	34.00	107.23	19.89	3,872.13	469,764.75
4	11.17	13.16	170.05	42.09	1,898.89	350,364.34
Total	482.67	1,700.92			12,176.19	1,117,711.51
Sampling date: 29 July 1974						
1	276.45	553.56	24.27	0.62	6,709.17	306,074.53
2	68.08	94.58	77.31	2.49	5,262.90	278,582.66
3	9.76	7.64	146.18	100.67	1,427.00	167,558.97
4	8.12	1.11	194.72	1.67	1,582.12	30,259.81
Total	362.41	656.89			14,981.19	782,475.97

To estimate production for the age-1 cohort during this interval, we follow equation (8.47) as

$$\hat{P} = G\bar{B}, \text{ or}$$
$$\hat{P} = (\log_e 24.27 - \log_e 6.86)[(6{,}709.17 + 1{,}905.34)/2] = 5{,}442.36\text{g}.$$

where $\hat{P}$ is the estimated production and G is the estimated instantaneous growth rate for a given cohort within a specified interval. To estimate the variance of $\hat{P}$, we begin by estimating the variance $\bar{B}$ of from equation (8.49) as

$$V(\bar{B}) = [V(B_t) + V(B_{t+1})]/4, \text{ or}$$
$$V(\bar{B}) = (75{,}455.97 + 306{,}074.53/4 = 95{,}383.62\text{g}.$$

Allen (1951) extended the instantaneous growth rate method to a graphical form to estimate production in what has become known as the Allen curve method. The Allen curve is a growth–survivorship curve for a given cohort, wherein the number of surviving individuals is plotted against the mean weight of those individuals (Figure 8.6). Following this configuration, the biomass of the cohort can be estimated at any point in time on the curve as the corresponding product of number of individuals (or density) and their mean weight. Likewise, the area

Then, we estimate the variance of G from equation (8.50) as

$V(G) = V(\log_e \bar{w}_t) + V(\log_e \bar{w}_{t+1})$, expanded by incorporating equation (8.53) as
$V(G) = V(\bar{w}_t)/\bar{w}_t^2 + V(\bar{w}_{t+1})/(\bar{w}_{t+1})^2$, or
$V(G) = 0.13/6.86^2 + 0.62/24.27^2 = 0.003815$.

Now, we may employ equation (8.48) as

$V(\hat{P}) = V(\bar{B})G^2 + \bar{B}^2V(G)$, or
$V(\hat{P}) = (95{,}382.62)(\log_e 24.27 - \log_e 6.86)^2 + [(6{,}709.17 + 1{,}905.34)/2]^2(0.003815) = 223{,}058$g.

The sampling area was 0.181 ha, so to convert our production estimate to a standard area unit (ha), we divide by the area, and to convert to a standard mass unit (kg), we divide by 1,000.

$$\hat{P} = 5{,}442.36\text{g}/0.181 \text{ ha}/1{,}000 = 30.1 \text{ kg/ha}.$$

Whenever you multiply a statistic by a constant, you multiply the variance of that statistic by the constant squared. Thus, to convert the variance of our production estimate to standard units, we divide by the area (0.181 ha) squared and divide by 1,000 squared as

$$V(\hat{P}) = 223{,}058/0.181^2/1{,}000^2 = 6.8 \text{ kg/ha}.$$

Our production estimate ± approximate 95% CIs $[\pm 1.96V(\hat{P})^{0.5}]$ for this cohort during this interval is 30.1 ± 5.1 kg/ha. This procedure is then repeated for the other age-classes to estimate production for the population (rounded to the nearest tenth) during this interval as 62.8 ± 7.4 kg/ha, as below.

Table Production estimate for population of brook trout during first time interval.

Age-class	Production (kg/ha ± 95% CI)
1	30.068 ± 5.111
2	26.856 ± 4.293
3	4.536 ± 2.664
4	1.303 ± 0.859
Total	62.763 ± 7.240

Note that CIs are not additive, and variances should be summed to compute a CI for a total. The entire procedure is then repeated for the other two intervals within the annual period to estimate annual production and its CI.

under the curve may be calculated in corresponding units as production of the cohort during the specific interval plotted. The Allen curve is rarely presented in recent literature, most likely because explicit variance estimators have not been derived for it and the quantitative form of this concept, the instantaneous growth rate method, is more precise and relatively easy to apply using software applications. Nonetheless, examination of Allen curves can be an instructive means to visualize and elucidate production dynamics of a fish population.

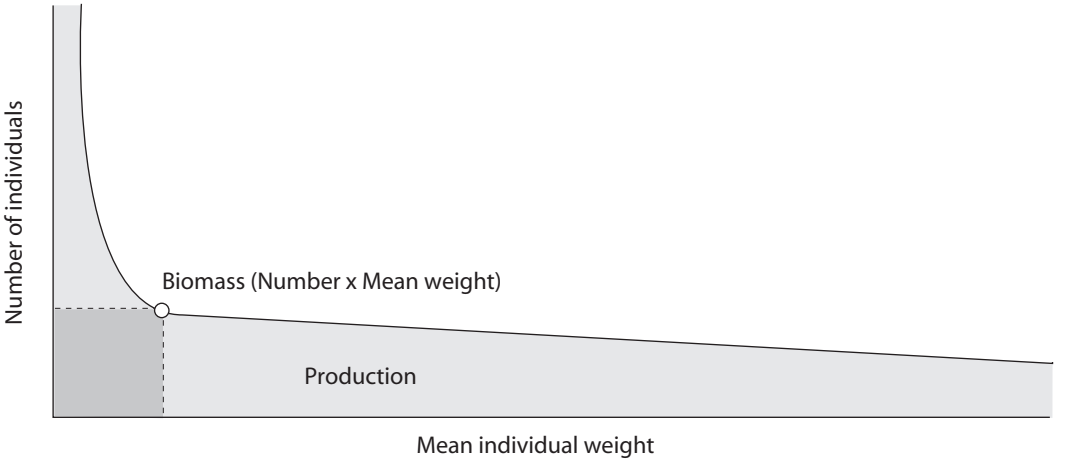


Figure 8.6 A generalized Allen curve depicting growth–survivorship for a fish cohort (after Allen 1951; Waters 1977). Production for the time interval is estimated as the area under the curve (shaded), and biomass at any point in time (dark-shaded area) is the product of the number of individuals (or density) and their mean weight.

8.7.2.3 Size-Frequency Method

The size-frequency method to estimate production was first developed by Hynes (1961) as an alternative means to estimate production when cohort identification was not possible; it was formerly also referred to as the Hynes method. Conceptually, the method is similar to the removal summation method, except that the size-frequency method sums tissue losses between successive size-groups rather than over time intervals. Originally, the method was used to approximate production roughly for multispecies assemblages of aquatic animals, but subsequent criteria and assumptions, established to improve the method, restrict its practical application to single species or closely related species with similar life histories (Waters 1977; Krueger and Martin 1980).

Production is estimated by the size-frequency method for fishes as (Garman and Waters 1983)

$$\hat{P} = 0.5c\left[\overline{w}_1(\overline{N}_1 - \overline{N}_2) + \sum_{k=2}^{c-1}\overline{w}_k(\overline{N}_{k-1} - \overline{N}_{k+1}) + \overline{w}_c(\overline{N}_{c-1} - \overline{N}_c)\right](1/\mathrm{CPI}), \qquad (8.52)$$

where P = production for a given population or multispecies group within a specified interval, $\overline{N}$ = estimated mean density (arithmetic mean of estimates) for a specific length-group, $\overline{w}$ = estimated mean weight (arithmetic mean of estimates) of individuals in a specific length-group, k = index for length-groups, c = number of length-groups, and CPI = the cohort production interval (average maximum age of fish in the population or multispecies group in years).

Estimated variance of the production estimate is computed as

$$V(\hat{P}) = (0.5c)^2 \left\{ \begin{array}{l} (\overline{w}_1 + \overline{w}_2)^2 V(\overline{N}_1) + V(\overline{w}_1)(\overline{N}_1 - \overline{N}_2)^2 \\ + \sum_{k=2}^{c-1} [(\overline{w}_{k-1} - \overline{w}_{k+1})^2 V(\overline{N}_k) + V(\overline{w}_k)(\overline{N}_{k-1} - \overline{N}_{k+1})^2] \\ + (\overline{w}_{c-1} - \overline{w}_c)^2 V(\overline{N}_c) + V(\overline{w}_c)(\overline{N}_{c-1} - \overline{N}_c)^2 \end{array} \right\} (1/\mathrm{CPI})^2, \quad (8.53)$$

where $V(\overline{N})$ and $V(\overline{w})$ are variances of mean density and mean weight, respectively, for length-groups. If the production period spans more than a single time interval (more than two samples), then mean density and mean weight estimates (and associated variances) must be weighted by interval length (days) according to algorithms provided by Garman and Waters (1983). An example calculation of fish annual production estimated by the size-frequency method is presented in Box 8.9.

8.7.3 Production to Mean Biomass ($P/\overline{B}$) Ratio

The annual production to annual mean biomass ($P/\overline{B}$) ratio is of special interest as an ecological index (also referred to as the turnover ratio) and as a simple conversion factor to approximate production (Waters 1977). Based on the premise that $P/\overline{B}$ is relatively constant for an organism or related organisms, it has been suggested that production may be approximated from a biomass estimate using the $P/\overline{B}$ ratio as a multiplier. The ratio of $P/\overline{B}$ for most fish species varies from 0.2 to 4.0. However, $P/\overline{B}$ can be quite variable within and among species (Waters 1977, 1999; Mann and Penczak 1986; Elliott 1994) and may vary with the number of cohorts (or life span) of a population (Waters 1992; Kwak and Waters 1997). Thus, this method should be applied carefully and considered an imprecise approximation of production. The exact $P/\overline{B}$ ratio to employ for such estimates should be species specific and may be refined further if the number of cohorts in a population is known (Waters 1992; Kwak and Waters 1997). Estimates of mean annual biomass should be weighted by interval duration if more than a single interval is included. Newman and Martin (1983) present formulae for estimating mean annual biomass and its variance, as well as a variance estimator for the $P/\overline{B}$ ratio.

8.7.4 Production Estimates in Practice

In general, if age data are available, the instantaneous growth rate method is the preferred approach to estimate fish production and associated parameters; otherwise, the size-frequency method may be used. Estimating production using the $P/\overline{B}$ ratio should be used only when data are lacking for application of more precise methods. Computer software is available for using the increment summation or instantaneous growth rate methods (Kwak 1992; available at http://www4.ncsu.edu/~tkwak), and a spreadsheet application can facilitate calculations by other methods. Estimates of variance (precision or sampling error) should be reported for all estimates of production and related population parameters as approximate 95% CIs.

Box 8.9 Production Estimation Based on the Size-Frequency Method

Density and mean weight (and associated variances) of a rainbow trout population in Valley Creek, Minnesota, were estimated in a stream reach on three dates between April 1977 and April 1978 (Garman and Waters 1983). The catch data were broken into 10 size-groups in order to allow investigators to estimate production using the size-frequency method.

Table Density and weight statistics based on three collection dates for rainbow trout population, Valley Creek, Minnesota.

Length-group	Mean density ($\bar{N}$/ha)	$V(\bar{N})$	Mean weight ($\bar{w}$, g)	$V(\bar{w})$	Mean biomass ($\bar{B}$, g/ha)
1	260.2	2,653.5	3.2	0.2	832.6
2	281.7	1,491.4	6.9	0.1	1,943.7
3	144.9	182.5	12.6	0.1	1,825.7
4	88.8	145.9	27.9	1.4	2,477.5
5	67.7	49.5	49.9	5.1	3,378.2
6	43.1	19.5	75.7	45.0	3,262.7
7	55.9	601.4	109.5	24.8	6,121.0
8	26.9	61.2	158.6	11.8	4,266.3
9	19.8	104.2	196.0	39.9	3,880.8
10	15.0	0.2	260.8	13.1	3,912.0

To estimate annual production for the population, we follow equation (8.52), using 3 years for the cohort production interval (CPI).

$$\begin{aligned}\hat{P} &= 0.5(10)[3.2(260.2 - 281.7) + 6.9(260.2 - 144.9) + \ldots + 260.8(19.8 - 15.0)]/(1/3)\\ &= 5(-68.80 + 795.57 + 2{,}430.54 + 2{,}153.88 + 2{,}280.43 + 893.26 + 1{,}773.90\\ &\quad + 5{,}725.46 + 2{,}332.40 + 1{,}251.84)(0.333)\\ &= 32{,}581.52 \text{ g/ha/year.}\end{aligned}$$

Variance of $\hat{P}$ is then estimated according to equation (8.53) as

$$\begin{aligned}V(P) &= [0.5(10)^2]\,\{(3.2 + 6.9)^2(2{,}653.5) + 0.2(260.2 - 281.7)^2 + [(3.2 - 12.6)^2(1{,}491.4) +\\ &\quad 0.1(260.2 - 144.9)^2] + 7 \text{ other summation terms} + [(196.0 - 260.8)^2(0.2) +\\ &\quad 13.1(19.8 - 15.0)^2]\,\}\,(1/3)^2\\ &= 25(270{,}683.5 + 92.5 + 133{,}109.5 + 84{,}203.5 + 211{,}333.0 + 123{,}750.9 + 75{,}532.9\\ &\quad + 4{,}139{,}575.9 + 473{,}291.6 + 1{,}094{,}002.6 + 1{,}141.6)0.11\\ &= 18{,}168{,}473.1 \text{ g/ha/year.}\end{aligned}$$

To convert our production estimate to a standard mass unit (kg), we divide by 1,000:

$$\hat{P} = 32{,}581.52/1{,}000 = 32.582 \text{ kg/ha/year.}$$

The variance is converted as

$$V(P) = 18{,}168{,}473.1/1{,}000^2 = 18.168 \text{ kg/ha/year.}$$

Thus, our annual production estimate ± approximate 95% intervals $[\pm 1.96V(P)^{0.5}]$ for this population and year is 32.582 ± 8.354 kg/ha/year.

Many assumptions and criteria for applying these methods have been defined and should be considered with application (Waters 1977; Newman and Martin 1983). We suggest that subjective decisions encountered when estimating fish production should be resolved to be conservative, so that the direction of error will be clear, and estimates may be considered minimums. The biomass estimate of age-0 fish, newly recruited into the population, at first sampling should be considered a conservative estimate of production for that cohort during that interval. Negative estimates of production, resulting from negative growth (i.e., weight loss), should be interpreted as no production (zero) for that cohort and interval when using summation methods or the instantaneous growth rate method. However, negative losses (i.e., increase in numbers between size-groups) should be included in the sum when applying the size-frequency method.

Generally, the greater the number of fish population estimates (density and biomass) that are integrated into an annual production estimate, the more accurate that estimate will be. The minimum number of two population estimates will yield a less accurate production estimate than will one based on more estimates within the annual period. A reasonable, general approach to estimating populations over a 1-year period for an annual production estimate is to conduct one estimate prior to spawning (e.g., spring for many temperate fishes), another near the end of the primary growing season (e.g., fall for temperate areas), and a third 1 year after the first estimate (e.g., spring or fall).

Fish production estimates are valuable statistics for understanding population dynamics and elucidating ecological relationships and have great potential for improving fisheries management. Waters (1992) reviewed and proposed the application of annual production, annual $P/\overline{B}$ ratio, and ecotrophic coefficient (annual angler harvest/annual production) to management of stream-dwelling trout fisheries. Incorporating production dynamics into fish assessment and monitoring may provide a broader perspective on the dynamics of harvested fishes. Thus, regulation and assessment of harvest as a proportion of fish tissue produced on an annual basis provides an alternative to the standard approach, based solely on fish density or biomass.

8.8 FUTURE DIRECTIONS

In many studies of fish populations, information is often available beyond that needed to apply the methods outlined in this chapter. In particular, information on the age structure of the population is often collected. When the abundance of a population is estimated on an annual basis, knowledge of the prior age composition is helpful in constraining estimates. As a simple example, the abundance of a cohort cannot be larger than the abundance in the prior year (in a closed population). The constraints imposed by age structure relationships can help improve the precision and accuracy of population estimates. Application of auxiliary information to population estimation opens up a diversity of models. Powerful statistical catch-at-age models (Hilborn and Walters 1992) are an example of a framework that incorporates the extensive information that is often available on

intensively studied fish populations. Because of the complexity of such models and their intensive data needs, such methods are generally applied to marine fish stocks and some stocks in large inland waters (e.g., Great Lakes) where the cost of data collection and analysis is commensurate with the value of the fishery.

Even within the scope of the methods presented in this chapter, there are potential gains to be made by combining data from different sources. In particular, the combination of removal methods with marking fish holds promise for improving population estimates. The methods illustrated here for analyzing mark–recapture data do not make direct use of measures of sampling effort. Removal methods, on the other hand, explicitly assume effort is constant or accommodate changes in effort by standardizing catch to C/f.

The estimation of the variance and CIs for population estimates is an area where substantial improvements need to be made. Although the likelihood methods presented here have a long history of use, and provide a strong statistical basis for estimation, many of the formulae are strictly valid only for large sample sizes or are approximations to the "full" formulae. In many applications, the target population itself may be small (e.g., the number of fish in a 100-m stretch of stream) or the number of marked or recaptured fish is small to moderate (i.e., less than 30). In situations like these, variance estimates and CIs should be interpreted with caution.

A trend we see emerging is the incorporation of a Bayesian approach to data analysis. In many situations, researchers and managers have knowledge from prior experience that is pertinent to the population being studied. Incorporating the experience and beliefs of experts can improve population estimates in many cases (Hilborn and Walters 1992). The Bayesian approach, however, presents several practical concerns regarding how best to represent prior information.

8.9 CONCLUSIONS

In this chapter, we illustrate several approaches for estimating fish abundance, biomass, and production. A foundational concept is that additional information beyond C/f is generally needed to provide accurate population estimates. The incorporation of this information inevitably entails making assumptions about the sampling regime and creating models of this idealized process. Therefore, it is important to test assumptions, where possible, and apply models that best represent the data obtained. Some assumptions can be relaxed by applying alternate models, but some are essential to obtain any valid estimate. The key to meeting critical assumptions is often the judicious planning of the sampling program and the careful application of field methods. Given the wide array of sampling challenges facing fisheries scientists, this chapter should be viewed as an entry into some of the more common and basic methods. Every investigation poses its own set of challenges, but often these problems are not unique. By building on the base developed here, we hope to provide readers with the confidence to face the diversity of situations they are likely to encounter in their professional work.

8.10 REFERENCES

Aitchison, J., and J. A. C. Brown. 1976. The lognormal distribution with special reference to its uses in economics. Cambridge University Press, Cambridge, UK.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principal. Pages 267–281 *in* B. N. Petran and F. Csaki, editors. International symposium on information theory, 2nd edition. Akademiai Kiadi, Budapest.

Allen, K. R. 1949. Some aspects of the production and cropping of fresh waters. Transactions of the Royal Society of New Zealand 77(5):222–228.

Allen, K. R. 1951. The Horokiwi Stream. New Zealand Marine Department Fishery Bulletin 10.

Anderson, R. O., and R. M. Neumann. 1996. Length, weight, and associated structural indices. Pages 447–482 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Bagenal, T. 1978. Methods for assessment of fish production in fresh waters. IBP (International Biological Programme) Handbook No. 3, Blackwell Scientific Publications, Oxford, UK.

Bickel, P. J., and K. A. Doksum. 1977. Mathematical statistics: basic ideas and selected topics. Holden-Day, San Francisco.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance sampling: estimating abundance of biological populations. Chapman and Hall, London.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society, Monograph 5, Bethesda, Maryland.

Chapman, D. G. 1948. A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. International Pacific Salmon Fisheries Commission Bulletin 2:69–85.

Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publications in Statistics 1:131–160.

Chapman, D. W. 1978. Production in fish populations. Pages 5–25 *in* S. D. Gerking, editor. Ecology of freshwater fish production. Blackwell Scientific Publications, Oxford, UK.

Collie, J. S., and M. P. Sissenwine. 1983. Estimating population size from relative abundance data measured with error. Canadian Journal of Fisheries and Aquatic Sciences 40:1871–1879.

Cone, R. S., and C. C. Krueger. 1988. Comparison of survival, emigration, habitat use, marking mortality, and growth between two strains of brook trout in Adirondack ponds. North American Journal of Fisheries Management 8:497–504.

De Lury, D. B. 1947. On the estimation of biological populations. Biometrics 3:145–167.

Edwards, E. F., and B. A. Megrey. 1989. Mathematical analysis of fish stock dynamics. American Fisheries Society, Symposium 6, Bethesda, Maryland.

Elliott, J. M. 1994. Quantitative ecology and the brown trout. Oxford University Press, Oxford, UK.

Foote, K. G., and G. Stefansson. 1993. Definition of the problem of estimating fish abundance over an area from acoustic line-transect measurements of density. ICES Journal of Marine Science 50:369–381.

Garman, G. C., and T. F. Waters. 1983. Use of the size-frequency (Hynes) method to estimate annual production of a stream fish population. Canadian Journal of Fisheries and Aquatic Sciences 40:2030–2034.

Goodman, L. A. 1960. On the exact variance of products. Journal of the American Statistical Association 55:708–713.

Gould, W. R., and K. H. Pollock. 1997. Catch-effort maximum likelihood estimation of important population parameters. Canadian Journal of Fisheries and Aquatic Sciences 54:890–897.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Heimbuch, D. G., H. T. Wilson, S. B. Weisberg, and J. H. Volstad. 1997. Estimating fish abundance in stream surveys by using double-pass removal sampling. Transactions of the American Fisheries Society 126:795–803.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hynes, H. B. N. 1961. The invertebrate fauna of a Welsh mountain stream. Archiv fur Hydrobiologia 57:344–388.

Junge, C. O., and J. Libosvárský. 1965. Effects of size selectivity on population estimates based on successive removals with electrical fishing gear. Zoological Listy 14:171–178.

Krueger, C. C., and F. B. Martin. 1980. Computation of confidence intervals for the size-frequency (Hynes) method of estimating secondary production. Limnology and Oceanography 25:773–777.

Kwak, T. J. 1992. Modular microcomputer software to estimate fish population parameters, production rates and associated variance. Ecology of Freshwater Fish 1:73–75. (Software available at http://www4.ncsu.edu/~tkwak.)

Kwak, T. J., and T. F. Waters. 1997. Trout production dynamics and water quality in Minnesota streams. Transactions of the American Fisheries Society 126:35–48.

Lebreton, J-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62:67–118.

Leslie, P. H., and D. H. S. Davis. 1939. An attempt to determine the absolute number of rats on a given area. Journal of Animal Ecology 8:94–113.

Ludwig, D., and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences 42:1066–1072.

Mann, R. H. K., and T. Penczak. 1986. Fish production in rivers: a review. Polskie Archiwum Hydrobiologii 33:233–247.

Mesnil, B. 2003. The catch-survey analysis (CSA) method of fish stock assessment: an evaluation using simulated data. Fisheries Research 63:193–212.

Newman, R. M., and F. B. Martin. 1983. Estimation of fish production rates and associated variances. Canadian Journal of Fisheries and Aquatic Sciences 40:1729–1736.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference for capture data from closed populations. Wildlife Monographs 62.

Railsback, S. F., B. D. Holcomb, and M. G. Ryon. 1989. A computer program for estimating fish population sizes and annual production rates. Oak Ridge National Laboratory, Environmental Sciences Division Publication 3245, Oak Ridge, Tennessee.

Rice, J. A. 1995. Mathematical statistics and data analysis, 2nd edition. Duxbury Press, Belmont, California.

Ricker, W. E. 1946. Production and utilization of fish populations. Ecological Monographs 16:373–391.

Ricker, W. E. 1977. Computation and interpretation of biological statistics of fish populations. Department of the Environment, Fisheries and Marine Service, Bulletin 191, Ottawa.

Robson, D. S., and H. A. Regier. 1964. Sample size in Petersen mark–recapture experiments. Transactions of the American Fisheries Society 93:215–226.

Schnabel, Z. E. 1938. The estimation of the total fish population of a lake. American Mathematical Monographs 45:348–352.

Schnute, J. 1983. A new approach to estimating populations by the removal method. Canadian Journal of Fisheries and Aquatic Sciences 40:2153–2169.

Schnute, J. T., L. J. Richards, and N. Olsen. 1998. Statistics, software, and fish stock assessment. Pages 171–184 *in* F. Funk and seven coeditors. Fishery stock assessment models. University of Alaska, Sea Grant College Program AK-SG-98–01, Anchorage.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Edward Arnold, London.

Seber, G. A. F., and C. J. Wild. 1989. Nonlinear regression. Wiley, New York.

Thomas, L., and nine coauthors. 2001. Distance 4.0. Beta 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California.

Van Deventer, J. S., and W. S. Platts. 1989. Microcomputer software system for generating population statistics from electrofishing data—user's guide for MicroFish 3.0. U.S. Forest Service, Intermountain Research Station, General Technical Report INT-254, Ogden, Utah.

Vondracek, B., and D. J. Degan. 1995. Among- and within-transect variability in estimates of shad abundance made with hydroacoustics. North American Journal of Fisheries Management 15:933–939.

Waters, T. F. 1977. Secondary production in inland waters. Advances in Ecological Research 10:91–164.

Waters, T. F. 1992. Annual production, production/biomass ratio, and the ecotrophic coefficient for management of trout in streams. North American Journal of Fisheries Management 12:34–39.

Waters, T. F. 1999. Long-term trout production dynamics in Valley Creek, Minnesota. Transactions of the American Fisheries Society 128:1151–1162.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46 (Supplement):120–138.

Zippin, C. 1956. An evaluation of the removal method of estimating animal population. Biometrics 12:163–169.

Zippin, C. 1958. The removal method of population estimation. Journal of Wildlife Management 22:82–90.

Appendix Shortest 95% confidence interval (CI) for the population estimate *N* (Reproduced from Chapman 1948).

Confidence intervals for sample sizes of 50 or less are based on the Poisson distribution. The number of recaptures is denoted by *m*, and estimates of the CI for *N* are obtained by multiplying the table values by the product of the number of fish caught in the first and second sample (i.e., $n_1 \cdot n_2$).

	CI			CI	
m	Lower limit	Upper limit	*m*	Lower limit	Upper limit
0	0.0885		26	0.02478	0.0563
1	0.0720	19.489	27	0.02408	0.0539
2	0.0767	2.821	28	0.02342	0.0516
3	0.0736	1.230	29	0.02279	0.0495
4	0.0690	0.738	30	0.02221	0.0475
5	0.0644	0.513	31	0.02165	0.0457
6	0.0600	0.388	32	0.02112	0.0440
7	0.0561	0.309	33	0.02061	0.0425
8	0.0526	0.256	34	0.02014	0.0410
9	0.0495	0.217	35	0.01968	0.0396
10	0.0468	0.188	36	0.01925	0.0384
11	0.0443	0.165	37	0.01883	0.0372
12	0.0420	0.147	38	0.01843	0.0360
13	0.0400	0.133	39	0.01805	0.0350
14	0.0382	0.121	40	0.01769	0.03396
15	0.0365	0.111	41	0.01733	0.03300
16	0.0350	0.1020	42	0.01700	0.03210
17	0.03362	0.0945	43	0.01668	0.03124
18	0.03233	0.0880	44	0.01636	0.03043
19	0.03114	0.0823	45	0.01606	0.02966
20	0.03004	0.0773	46	0.01578	0.02892
21	0.02901	0.0729	47	0.01550	0.02822
22	0.02806	0.0689	48	0.01523	0.02755
23	0.02716	0.0653	49	0.01498	0.02691
24	0.02632	0.0620	50	0.01475	0.02625
25	0.02552	0.0591			

9 Size Structure

Robert M. Neumann and Micheal S. Allen

9.1 INTRODUCTION

Size structure analysis is one of the most commonly used fisheries assessment tools. The size structure of a fish population at any point in time can be considered a snapshot that reflects the interactions of the dynamic rates of recruitment, growth, and mortality. Thus, length-frequency data provide valuable insight into the dynamics of fish populations and help identify problems such as inconsistent year-class strength, slow growth, or excessive mortality (Anderson and Neumann 1996). In most cases, a thorough interpretation of size structure data is complemented by other population assessment tools, such as catch per unit effort (C/f), age-and-growth analysis, recruitment analysis, mortality, and body condition.

Proper analysis and interpretation of size structure data should begin with a clear understanding of how, when, and where data were collected. Specifically, a fisheries scientist should know how size structure data are influenced by the sampling gear, time of the year, and location where fish were sampled. The fisheries scientist should also consider whether an appropriate sample size was obtained to estimate size structure reliably.

Fisheries scientists use several techniques to analyze size structure data. In the simplest case, a length-frequency histogram (see section 9.2) is constructed or a size structure index is calculated. Oftentimes, the primary objective is to compare size structure among samples. In these cases, a fisheries scientist may be interested in answering several questions. For example, does the size structure of white crappie populations differ among water bodies? Did the size structure of a rainbow trout population change over time in response to a management action? Are the size structures obtained from a channel catfish population different between two or more sampling gears? What factors influence the size structure of walleye populations?

9.2 PRESENTATION OF SIZE STRUCTURE DATA

Three common measures of fish length include total, fork, and standard length. Total length is measured from the anterior-most part of the fish to the tip of the

longest caudal fin ray when the caudal fin is compressed. In this chapter, all lengths are reported as total length. Fork length is measured from the anterior-most part of the fish to the median caudal fin rays, which typically make up the concave portion in a forked caudal fin. Standard length is measured from the anterior-most part of the fish to the end of the caudal peduncle. Anderson and Neumann (1996) described measurements of fish length in detail.

Size structure data are most commonly reported using length-frequency histograms and stock density indices (Anderson and Neumann 1996). Length-frequency histograms show the number or proportion of fish collected in various length categories. The most commonly used length-frequency histogram is the absolute length frequency, which shows the number of fish collected in various length categories (Figure 9.1A). A relative-frequency distribution shows the proportion of all fish that are represented in each length category (Figure 9.1B). For example, in Figure 9.1A, 28 fish are in the 7–9-cm length-group (labeled 8 cm), which represent about 10% of the total number of fish collected (Figure 9.1B). Relative-frequency distributions are useful for comparing length categories that contain different sample sizes, which may result from variable sampling effort or population abundance. An alternative length-frequency distribution is based on C/f (Figure 9.1C), which is used to indicate relative abundance of fish in each length category (see Chapter 7 for treatment of C/f data).

Selection of interval widths is important for interpretation of length-frequency histograms. Anderson and Neumann (1996) suggested using 1-cm intervals for species that reach 30 cm, 2-cm intervals for 60-cm species, and 5-cm intervals for 150-cm species. Effects of interval width on the characteristics of a length-frequency histogram are demonstrated in Figure 9.2. In Figure 9.2, a 1-cm length interval shows more detail with a clear mode at 10 cm, which likely represents age-0 fish collected during fall. The 2-cm interval width shows the mode of young fish less clearly, and 4-cm interval widths mask the first mode completely.

Cumulative-frequency distributions provide an alternate view of length-frequency histograms and are used in some statistical tests comparing two or more distributions. In Figure 9.3, length-frequency histograms of age-0 walleye from three populations are presented. The respective cumulative-frequency distributions are shown in Figure 9.4. Differences in the size structure of walleye among populations are apparent in the length-frequency histograms and in the cumulative-frequency distributions. In Island Lake, the cumulative-frequency line approaches 100% at a shorter length than does the cumulative-frequency line in Lake Thompson because the Lake Thompson sample contains larger walleye (>200 mm) than does Island Lake. The cumulative-frequency lines clearly show that the sample for Lake Mitchell contains longer walleye overall than do the other lakes, but the maximum length of walleye is the same between Lake Mitchell and Lake Thompson (i.e., the cumulative-frequency lines both approach 100% at 210 mm). Another interpretation is that approximately 50% of the walleye in Island Lake and Lake Thompson are shorter than 160 mm, whereas 50% of the walleye in Lake Mitchell are shorter than 190 mm.

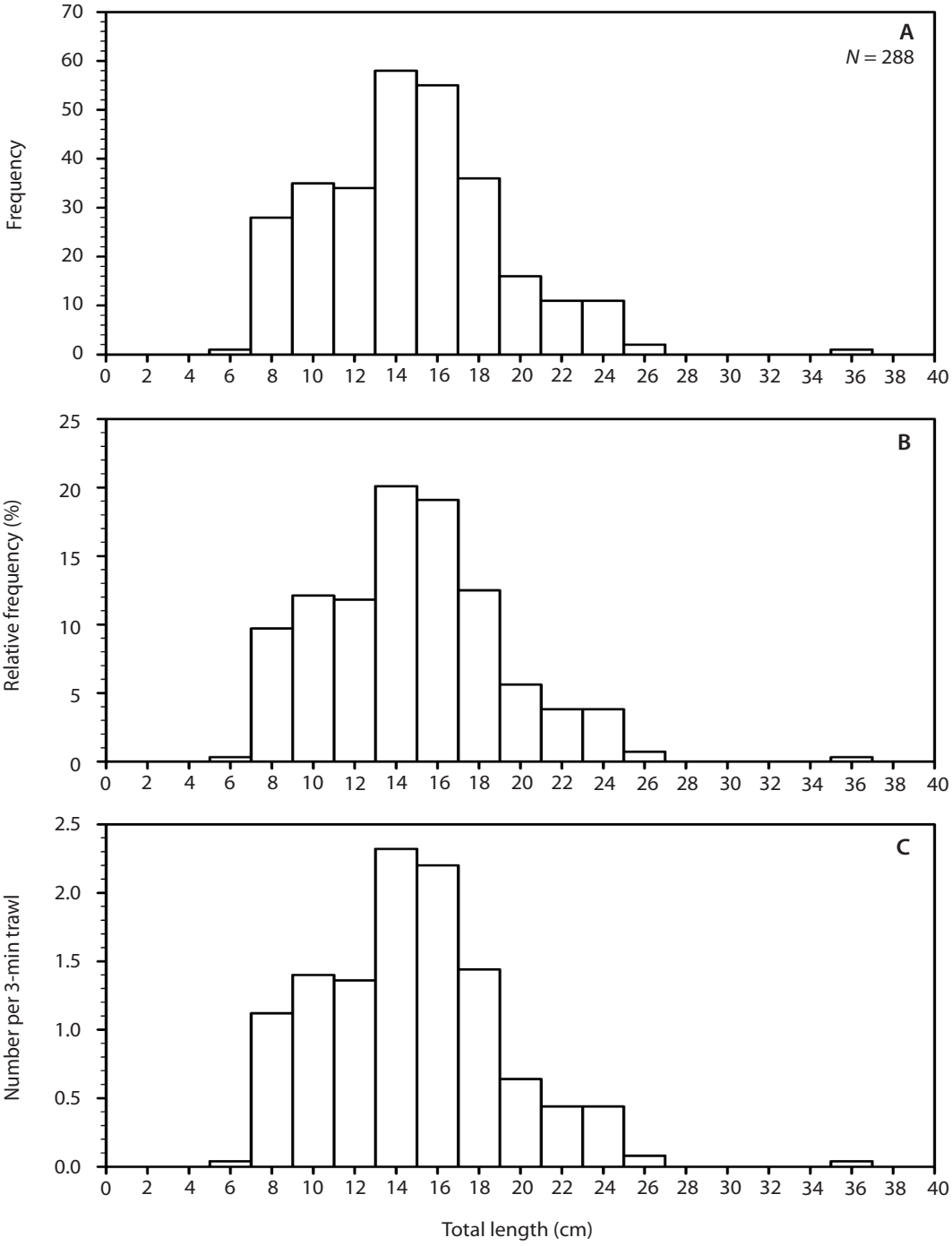


Figure 9.1 Length-frequency histograms for black crappie collected from Lake Jeffords, Florida. Data are displayed using (A) absolute length frequency, (B) relative length frequency, and (C) catch per unit effort.

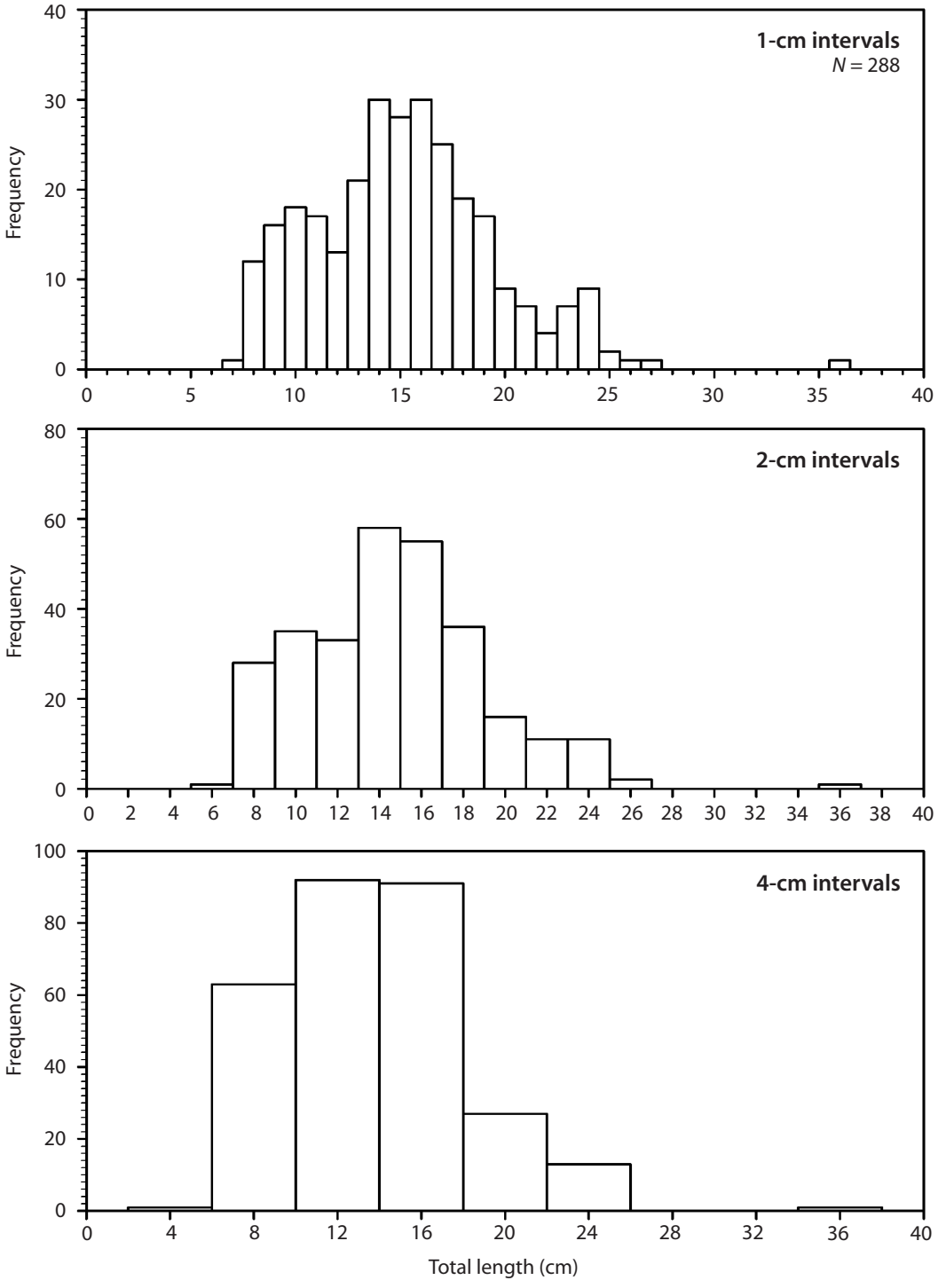


Figure 9.2 Absolute- length-frequency histograms constructed with length interval widths of 1, 2, and 4 cm for black crappie from Lake Jeffords, Florida.

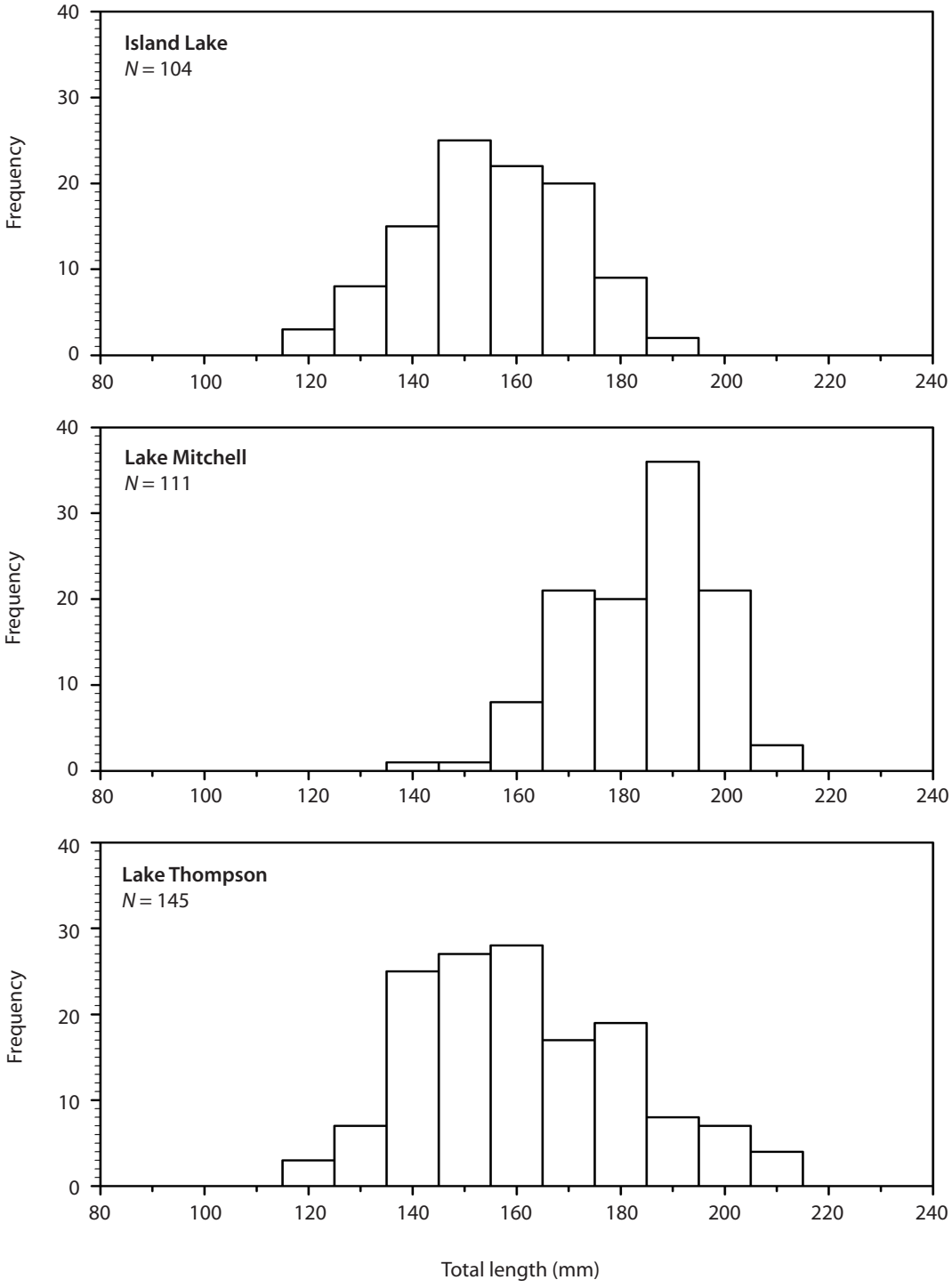


Figure 9.3 Absolute-length-frequency histograms for age-0 walleye collected from three South Dakota lakes (data courtesy of the South Dakota Department of Game, Fish and Parks).

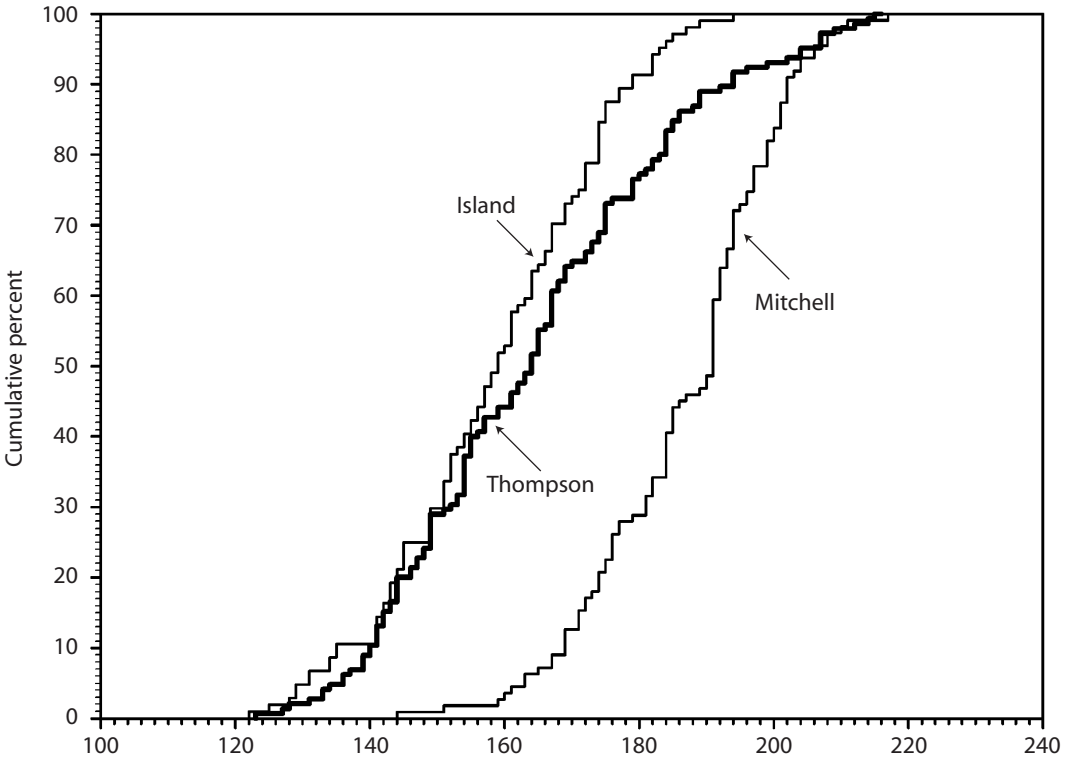


Figure 9.4 Cumulative-frequency distributions for age-0 walleye collected from three South Dakota lakes (data courtesy of the South Dakota Department of Game, Fish and Parks). The respective length-frequency distributions are shown in Figure 9.3.

Stock density indices are used to describe size structure. A detailed review of stock density indices and their calculation can be found in Anderson and Neumann (1996). Proportional stock density (PSD) is calculated as

$$\text{PSD} = \frac{\text{Number of fish} \geq \text{quality length}}{\text{Number of fish} \geq \text{stock length}} \times 100. \tag{9.1}$$

Relative stock density (RSD) is expressed as

$$\text{RSD} = \frac{\text{Number of fish} \geq \text{specified length}}{\text{Number of fish} \geq \text{stock length}} \times 100, \tag{9.2}$$

where the specified length often refers to preferred, memorable, or trophy length. Relative stock densities of preferred-, memorable-, and trophy-length fish are reported as RSD-P, RSD-M, and RSD-T, respectively. The standard convention is to report stock density index values to the nearest whole number without a percent

symbol. Minimum stock, quality, preferred, memorable, and trophy lengths for many species are provided in Anderson and Neumann (1996) and Bister et al. (2000). In traditional stock density index calculations, it is important to emphasize that stock, quality, preferred, memorable, and trophy lengths are minimum lengths. For example, stock and quality lengths for largemouth bass are 20 and 30 cm, respectively. Thus, in a sample of largemouth bass, all fish greater than 20 cm are stock length, and all fish greater than 30 cm are quality length. Length-frequency data can also be indexed using incremental stock density indices (Anderson and Neumann 1996).

The use of PSD alone to index size structure can often lead to loss of data sensitivity. For example, two largemouth bass populations can have PSD values of 60, even though one population may be quite different from the other when the length-frequency histograms are inspected. This is because quality length includes all fish greater than or equal to quality, preferred, memorable, and trophy length. Consider two populations that both have 30 quality-length fish. In one population, all 30 quality-length fish may be between quality and preferred length, whereas in the other, 20 may be between quality and preferred length, and 10 may be between preferred and memorable length. This example illustrates the importance of calculating other stock density indices (e.g., RSD-P) to index size structure precisely. Fisheries scientists should calculate the stock density index for the largest length category of interest, given an appropriate sample size.

Gustafson (1988) provided a formula and easy-to-use tables for determining 80% and 95% confidence intervals around stock density index values (Tables 9.1, 9.2). Confidence interval widths depend on sample size and the magnitude of the stock density index value. Although confidence intervals provide a measure of variation around stock density index values, they should not be used as a test for determining statistically significant differences between two or more values, primarily because confidence intervals for index values with unequal sample sizes were derived from distributions with unequal variances. Trippel and Hubert (1990) cautioned against the use of confidence interval overlap as a test for differences between means unless variance is pooled. Statistical treatment of stock density index values is presented in section 9.4.

9.3 COLLECTION OF SIZE STRUCTURE DATA

In an ideal situation, the size structure of a fish population determined from samples would be the same as the true size structure of the fish population. However, when fisheries scientists collect a sample of fish, the size structure obtained from that sample is often different from the true size structure of the fish population. Size structure from samples can be misrepresentative of the true population because the lengths of fish collected may depend on the type of sampling gear used, the season in which the fish were collected, and the location chosen to collect the fish. To overcome these effects, fisheries scientists use standardized sampling so that changes in size structure over time and comparisons of size structure among water bodies can be adequately assessed.

Table 9.1 Approximate confidence intervals (plus or minus) for proportional stock densities (PSD) as a function of sample size (*N*) of stock-length fish at the 80% confidence interval. Values have been omitted when sample sizes are insufficient for a normal approximation to the binomial distribution (from Gustafson 1988).

	Estimated PSD																		
N	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
10										30									
15							22	22	22	22	22	22	22						
20					16	17	18	18	18	19	18	18	18	17	16				
25				13	14	15	15	16	16	16	16	16	15	15	14	13			
30				12	13	13	14	14	14	14	14	14	14	13	13	12			
35			10	11	12	12	13	13	13	13	13	13	13	12	12	11	10		
40			9	10	11	11	12	12	12	12	12	12	12	11	11	10	9		
45			8	9	10	10	11	11	11	11	11	11	11	10	10	9	8		
50		7	8	9	9	10	10	10	11	11	11	10	10	10	9	9	8	7	
55		6	7	8	9	9	10	10	10	10	10	10	10	9	9	8	7	6	
60		6	7	8	8	9	9	9	9	10	9	9	9	9	8	8	7	6	
65		6	7	7	8	8	9	9	9	9	9	9	9	8	8	7	7	6	
70		6	6	7	8	8	8	9	9	9	9	9	8	8	8	7	6	6	
75		5	6	7	7	8	8	8	8	8	8	8	8	8	7	7	6	5	
80		5	6	7	7	7	8	8	8	8	8	8	8	7	7	7	6	5	
85		5	6	6	7	7	7	8	8	8	8	8	7	7	7	6	6	5	
90		5	6	6	7	7	7	7	8	8	8	7	7	7	7	6	6	5	
95		5	5	6	6	7	7	7	7	7	7	7	7	7	6	6	5	5	
100	3	5	5	6	6	7	7	7	7	7	7	7	7	7	6	6	5	5	3
120	3	4	5	5	6	6	6	6	6	6	6	6	6	6	6	5	5	4	3

9.3.1 Standardized Sampling for Size Structure Data

Willis and Murphy (1996) emphasized the importance of standardized sampling methods because of the numerous gear-, season-, and location-related effects on sampling data for many fishes. They recommended that standardized sampling should consider the use of an effective gear for the fish species being sampled, that the gear be used during an effective time of the year, and that gears be set in standard locations from year to year. Thus, long-term data sets can be established, and trends in sample variables can be monitored over time.

Many fishery management activities focus on the adult portion of a population, whether the goal is to increase abundance of adult fish, increase size structure, or manipulate the adult stock to influence predator–prey dynamics. In cases in which changes in the adult portion of a population are being investigated, the use of a gear that effectively samples fishes that are stock length (see section 9.2) and greater is recommended. Rarely does one gear type effectively sample all lengths of fish in a population. Thus, investigations focusing on recruitment and year-class strength, for which the capture of juvenile fishes is necessary, may require a gear different than that used to capture adult fish. Because each gear

Table 9.2 Approximate confidence intervals (plus or minus) for PSD as a function of sample size (*N*) of stock-length fish at the 95% confidence interval. Values have been omitted when sample sizes are insufficient for a normal approximation to the binomial distribution (from Gustafson 1988).

	Estimated PSD																		
N	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
10										48									
15							34	35	35	36	35	35	34						
20					26	27	28	29	29	29	29	29	28	27	26				
25				21	22	23	24	25	25	25	25	25	24	23	22	21			
30				19	20	21	21	22	22	22	22	22	21	21	20	19			
35			15	17	18	19	20	20	20	20	20	20	20	19	18	17	15		
40			14	15	17	17	18	18	19	19	19	18	18	17	17	15	14		
45			13	14	15	16	17	17	17	17	17	17	17	16	15	14	13		
50		11	12	13	14	15	16	16	16	16	16	16	16	15	14	13	12	11	
55		10	12	13	14	14	15	15	15	15	15	15	15	14	14	13	12	10	
60		9	11	12	13	14	14	14	15	15	15	14	14	14	13	12	11	9	
65		9	10	12	12	13	13	14	14	14	14	14	13	13	12	12	10	9	
70		9	10	11	12	12	13	13	13	13	13	13	13	12	12	11	10	9	
75		8	10	11	11	12	12	13	13	13	13	13	12	12	11	11	10	8	
80		8	9	10	11	12	12	12	12	12	12	12	12	12	11	10	9	8	
85		8	9	10	11	11	12	12	12	12	12	12	12	11	11	10	9	8	
90		7	9	10	10	11	11	11	12	12	12	11	11	11	10	10	9	7	
95		7	8	9	10	10	11	11	11	11	11	11	11	10	10	9	8	7	
100	5	7	8	9	10	10	11	11	11	11	11	11	11	10	10	9	8	7	5
120	5	6	7	8	9	9	9	10	10	10	10	10	9	9	9	8	7	6	5

type, and configurations within a specific gear type, may select for different sizes of fish, combining data from several gears to determine size structure is not recommended. Rather, standard gears should be used that will allow for comparisons of size structure over time or among water bodies.

9.3.2 Effects of Gear Type on Size Structure

Size selectivity of a particular gear can be related to the physical dimensions of a mesh, reaction of fish to a gear, and the location in a water body where the gear is used. Much research has been conducted to achieve a better understanding of the size selectivity of various sampling gears. Hubert (1996) provided information on the size and seasonal biases of many passive gears used in fisheries research.

With gill nets, only a limited range of lengths are sampled within a given mesh size, with fish of a particular length being held most securely. Fisheries scientists should be aware of mesh-size selectivity and mesh-size efficiency to interpret size structure data collected with gill nets properly. Hubert (1996) provided a detailed explanation of mesh-size efficiency and selectivity and referenced methods to correct size structure data from gill nets. Experimental gill nets, which include

several mesh sizes, are often used to sample a broad length range of the species under consideration. The use of experimental gill nets does not ensure that the size structure of the collected fish will be representative of the true size structure of the fish population because mesh-size selectivity and efficiency may still influence the sample size structure. Size selectivity may be reduced when the mesh size and twine diameter complement is carefully chosen and evaluated.

In gears such as gill nets, hoop nets, trap nets, and trawls, the mesh size will determine the minimum length of fish captured (Hubert 1996). Other aspects of net construction, such as mesh material, frame dimensions, and mouth size, also influence size selectivity. Laarman and Ryckman (1982) found that trap nets were selective for larger sizes of some fish species but not others. Holland and Peters (1992) compared length distributions of channel catfish captured from the Platte River, Nebraska, in hoop nets with three different mesh sizes and found that both the minimum and maximum lengths of fish increased as mesh size increased. The effects of mesh size on size selectivity for catfishes has been well documented, and, in general, samples collected with larger mesh sizes produce larger mean lengths (Vokoun and Rabeni 1999).

Electrofishing has also been shown to have size-selective properties. Reynolds and Simpson (1978) demonstrated that in Midwestern ponds, electrofishing efficiency increased as a function of total length for largemouth bass. For bluegill, electrofishing efficiency was higher for 8–15-cm bluegills compared with bluegills less than 8 cm or greater than 15 cm (Reynolds and Simpson 1978). Milewski and Willis (1991) found that compared with trap nets, electrofishing resulted in smaller size structure for smallmouth bass. Robinson (1994) found that large flathead catfish ($\geq$75 mm) were rarely captured when pulsed DC electrofishing was used. Santucci et al. (1999) determined that for channel catfish in a small impoundment, AC electrofishing selected for smaller fish in the population.

Size structure data collected by underwater observation (i.e., snorkel or scuba) were shown to be overestimated because of underwater magnification (Griffith 1981). Mullner et al. (1998) found that length frequencies of three trout species and their hybrids were significantly different between snorkeling and electrofishing samples, and they used an underwater magnification factor of 1.25 to adjust length frequencies. In contrast, Wildman and Neumann (2003) found that when broad length categories were used, size structure estimated by snorkeling and electrofishing were not substantially different for brook trout and brown trout in Connecticut streams.

Size structure is underestimated for most species of fish captured in cove rotenone samples (Hayne et al. 1967). Bayley and Austen (1990) tested the sampling efficiency of rotenone in ponds and coves and found that efficiency was high for large fish in warm water and low for small fish in cool water. Typically, cove rotenone sampling is conducted during mid to late summer, when large individuals of many species move offshore (Willis et al. 1993; Bettoli and Maceina 1996). Thus, summer cove rotenone samples may be selective for small fish.

The use of size structure data obtained from competitive fishing events and angler diaries is becoming more common. As with any collection technique, caution

must be given to how angler data are interpreted because angler data may be selective for larger fish compared with data from traditional sampling gears (Willis et al. 1993). Gabelhouse and Willis (1986) found that tournament anglers in Kansas selected for larger sizes of largemouth bass than did electrofishing, and stock density indices (see section 9.2) calculated from angler data were higher than those based on electrofishing samples. Jacobs et al. (1995) found that in Connecticut lakes, the proportion of largemouth bass greater than 38 cm was usually greater for electrofishing samples compared with tournament samples. In contrast, smallmouth bass greater than 38 cm tended to be underestimated in electrofishing samples compared with tournament samples. Green et al. (1986) found differences in size structure of largemouth bass and smallmouth bass based on data collected by anglers using diaries versus by electrofishing. They provided empirical adjustment factors to predict the size structure of largemouth bass and smallmouth bass in electrofishing samples from the length distribution of the angler catch. In contrast, Ebbers (1987) found that largemouth bass size structure estimated from angler diaries and electrofishing samples were similar in Minnesota. Thus, angler behavior may vary by geographic location and affect data used to determine size structure.

9.3.3 Effects of Sampling Time on Size Structure

Size structure of samples can differ among seasons of the year even when a standard gear is being used. Seasonal changes in size structure occur because of size-dependent changes in fish behavior and physiology throughout the year (Pope and Willis 1996). For example, Carline et al. (1984) found that for largemouth bass sampled by electrofishing in an Ohio impoundment, samples contained larger fish in spring and fall compared to summer. Largemouth bass greater than 30 cm apparently moved offshore after spawning and were not as vulnerable to capture during summer; as water temperature cooled during the fall, large fish returned to inshore areas. Gilliland (1987) and Bettross and Willis (1988) have reported similar seasonal changes in size structure for largemouth bass.

Pope and Willis (1996) provided a review of several studies that documented seasonal changes in size structure. Spring and fall peaks in size structure have been observed for several species, including bluegill captured in trap nets (Bettross and Willis 1988) and yellow perch (Lott and Willis 1991), walleye, and sauger (Mero and Willis 1992) captured in experimental gill nets. Boxrucker and Ploskey (1989) found that greater proportions of larger and older white crappies were captured in trap nets during spring than fall in Oklahoma impoundments. Seasonal patterns in size structure other than spring and fall peaks have also been observed. In a South Dakota lake, size structure of northern pike sampled with experimental gill nets was highest during winter and declined into the summer; significant inverse correlations between size structure and water temperature were observed (Neumann and Willis 1995).

Size structure has also been shown to differ even within a single season and between day and night samples. Across a 1-month period during spring, size structure of largemouth bass captured by electrofishing increased substantially,

apparently due to a greater proportion of largemouth bass greater than 30 cm moving to inshore areas in preparation for spawning (Carline et al. 1984). Paragamian (1989) found that the size structure determined from samples of smallmouth bass was higher at night than that during the day in an Iowa river. In Oklahoma reservoirs, largemouth bass size structure was similar between day and night electrofishing samples in spring, but during fall, day samples produced a narrower range of fish lengths and contained mostly smaller individuals (Gilliland 1987). Size structure of sauger captured by electrofishing during the day in a turbid main-stem reservoir was consistently higher than at night, and sauger greater than 51 cm were collected only during the day (Van Zee et al. 1996).

9.3.4 Effects of Sample Location on Size Structure

Biologists often choose subjective sampling sites based on the likelihood of capturing a large sample size of the target species (Willis et al. 1993). Hubbard and Miranda (1988) found that the size structure of largemouth bass collected by electrofishing from subjective sites was greater than was the size structure obtained from random sites. King et al. (1981) compared sample parameters for several fish species collected by electrofishing from fixed and random sites. They found few statistical differences in population parameters between the two types of sampling sites. However, the fixed sites they sampled over time were initially chosen at random.

Sampling fish from fixed or random sites should depend on the experimental design being used. Sampling at fixed sites is often used to track changes in population characteristics within a single water body, whereas sampling at random sites is more suitable for comparing population characteristics among water bodies. The use of fixed or random sites may also depend on the need to continue standard sampling designs previously developed.

9.3.5 Sample Size Considerations

The sample size necessary to describe the size structure of a fish population adequately is quite large. Anderson and Neumann (1996) recommended that for general stock assessment purposes, at least 100 fish greater than stock length (see section 9.2) should be sampled. Gilliland (1987) compared length frequencies based on various sample sizes of largemouth bass that were sampled by electrofishing in Oklahoma reservoirs and concluded that a sample size of 150 largemouth bass was adequate to estimate size structure, whereas a sample of 50 was not adequate. More recently, Vokoun et al. (2001) estimated the sample size necessary to construct a length-frequency distribution with a given accuracy and precision for bluegill and channel catfish. They compared the length frequency histogram from a known sample to computer generated length frequency histograms by means of bootstrapping methods. Their results demonstrated the importance of using at least 300–400 individuals whenever possible.

Weithman et al. (1980) developed a sequential sampling method that allows a biologist to monitor continuously how many stock-length and quality-length fish are necessary to obtain a reliable estimate of PSD while sampling is being conducted. Miranda (1993) developed a method by which biologists can approximate the sample size required for estimating PSD before collection begins. These sampling methods are further described in Anderson and Neumann (1996). Sample size requirements discussed in this section are recommendations based on existing information. Clearly, the sample size necessary to describe size structure reliably will depend on the species, population structure, sampling constraints, and study objectives.

9.4 STATISTICAL ANALYSES FOR SIZE STRUCTURE DATA

Analysis of size structure data should begin with an exploratory analysis by constructing length-frequency histograms or calculating stock density index values. There are also many statistical tests available to analyze size structure data. In this section, we review several statistical techniques commonly applied to size structure data. Experimental design considerations and statistical assumptions are reviewed in Chapters 2 and 3.

Fisheries scientists are often interested in comparing size structure between two or more samples. For example, comparisons of size structure are often made between different gear types, water bodies, or time periods. Consider the comparison of two hypothetical length-frequency data sets. Many commonly applied statistical tests, such as *t*-tests and analysis of variance (ANOVA), assume that data are normally distributed and, as such, are typically not appropriate for tests of length-frequency data (Brown and Austen 1996). When a broad length range of fish is sampled, length-frequency data are often multimodal, highly skewed, and contain extreme observations. In these cases, nonparametric tests may be more appropriate for comparing length-frequency distributions. Conditions favorable for nonparametric statistics and cautions about their use are described in Brown and Austen (1996) and Chapter 1. Given sufficient sample sizes, and when data approximate a normal distribution, most commonly employed parametric tests are sufficiently robust and can perform well (Zar 1996). Methods to evaluate normality of data and considerations for data transformations are provided in Chapter 3.

9.4.1 Parametric Tests

Assuming a normal distribution, size structure data are commonly compared between two samples using a *t*-test or by an ANOVA in the case of comparing more than two samples. These tests are used to compare the estimated means (e.g., means of length) to determine whether or not the samples come from the same population (Koopmans 1987). The fisheries scientist may use these tests to determine if mean lengths of samples are significantly different. An ANOVA is typically followed by a multiple-comparison test to determine which means are significantly different

from one another. An example of using an ANOVA to compare mean length among three length-frequency samples is provided in Box 9.1.

9.4.2 Nonparametric Tests for Comparing Size Structure

Several nonparametric statistical tests are useful for comparing size structures from two or more samples. Nonparametric tests are usually applied to length-frequency data, primarily because of concerns regarding the distribution of the data. Nonparametric tests commonly applied to length-frequency data include the Kolmogorov–Smirnov two sample, Wilcoxon's rank sum, Kruskal–Wallis, and the chi-square. The Kolmogorov–Smirnov two-sample test is used to determine whether the distribution of a variable (e.g., length) is the same across different groups (e.g., lakes). The test statistic is calculated as the largest absolute distance between the distribution functions (cumulative frequency distributions) associated with the samples (Zar 1996; SAS 1999). This test is often used to determine whether length-frequency distributions are different between samples (Box 9.2). Examples of the application of the Kolmogorov–Smirnov test to examine differences among length-frequency distributions can be found in Cornelius and Margenau (1999), Underwood (2000), Unmuth et al. (2001), Isermann et al. (2002), and Tate et al. (2003).

When applying two-sample tests (such as the Kolmogorov–Smirnov), pairwise tests are performed rather than multiple comparisons. Under these circumstances, the significance level for comparisons should be adjusted using the Bonferroni correction in order to maintain the predetermined experimentwise error rate (Koopmans 1987). This can be achieved by setting the significance level for each subtest equal to the experimentwise error rate divided by the number of subtests. For example, if the experimentwise error rate was $\alpha = 0.05$ and there were three subtests performed, then the significance level for each subtest would be $\alpha = 0.05/3 = 0.017$.

Wilcoxon's rank-sum test for two samples and the Kruskal–Wallis test for several samples are rank-testing procedures and sometimes are considered nonparametric counterparts to the *t*-test and ANOVA, respectively. In fact, the Kruskal–Wallis test is often called "ANOVA by ranks" (Zar 1996). For these tests, the observations from all samples are combined, ordered, and assigned a rank value, and the test statistic is calculated based on rank scores. These tests are used to test for differences in location and scale based on rank scores. Fisheries scientists often use these tests to determine whether length-frequency distributions are different among samples (i.e., does one population tend to yield larger or smaller values than the other). An example of the Kruskal–Wallis test applied to length-frequency data is provided in Box 9.3; additional applications of the Kruskal–Wallis test to length-frequency data can be found in Neumann et al. (1995) and Neal et al. (1999). Several nonparametric multiple-comparison tests are available for use with tests such as the Kruskal–Wallis test (Conover 1980; Zar 1996). Examples of multiple-comparison testing procedures are provided in Box 9.4.

According to Conover (1980), an advantage of the Kolmogorov–Smirnov two-sample test over rank tests (e.g., Wilcoxon-s rank sum and Kruskal–Wallis) is that

the Kolmogorov–Smirnov test is sensitive to detecting differences in location (magnitude of observations) and shape (variance) between distribution functions. Methods based on ranks are sensitive to differences in the magnitude of ranked data among samples, but they may not detect differences in variances or shape of the distributions. Thus, fisheries scientists should visually inspect length-frequency histograms and use statistical tests cautiously when analyzing length-frequency data.

Chi-square tests are commonly used to test for differences in length frequency among samples. Examples of the application of the chi-square test to length-frequency data can be found in Michaletz et al. (1995), Van Den Avyle et al. (1995), Roni and Fayram (2000), and Wildman and Neumann (2003). The chi-square test is used to test that the frequencies of observations among length-groups is independent of the treatment (e.g., water body, gear type, or time period). Chi-square tests are often applied, but not limited to, length-frequency data for which the length-groups are rather large. For example, length data are often categorized using stock density index length categories rather than by more detailed length intervals (Box 9.5)

When size structure is indexed using stock density indices, a fisheries scientist may be interested in statistically comparing stock density index values between two or more samples. Because stock density index values are frequently calculated from a more detailed length-frequency histogram, statistical procedures (as described above) can be applied to the raw length-frequency data, and the outcome of those tests along with stock density index values can be reported. An alternate approach may be to use a chi-square test (Box 9.5) in which stock density index length categories are used as length intervals. Fisheries scientists are often involved in studies in which a treatment (e.g., an experimental harvest regulation) is applied to several water bodies, and additional water bodies are used as a control group. In this case, stock density indices can be calculated for each water body, and the fisheries scientist can test for differences in the mean stock density index values (e.g., mean PSD) between treatments. For example, Margenau and AveLallemant (2000) used two-sample *t*-tests to compare mean stock density index values of muskellunge populations before and after a special harvest regulation was implemented. Proportions (such as PSD) form a binomial distribution rather than a normal distribution (Zar 1996). Thus, PSD values may require a data transformation (e.g., arcsine-root) before analyses (see Chapter 3 for discussion of data transformations).

9.4.3 The Experimental Unit

In each of the examples presented in Boxes 9.1–9.5, catches of fish in each unit of effort were pooled into a single sample, and statistical tests were performed on pooled length-frequency data. By far, this is the most commonly used approach to treating and testing size structure data. In some instances, performing tests on pooled length-frequency data can result in tests with inflated power, resulting in significant differences in length-frequency distributions even though there may be only slight differences between distributions. This is especially the case when large sample sizes are created by pooling length-frequency data. For example, in

Box 9.1 Testing for Differences in Mean Length By Means of Analysis of Variance (ANOVA)

Age-0 walleye were sampled from three eastern South Dakota lakes (Island Lake, Lake Mitchell, and Lake Thompson) by biologists from the South Dakota Department of Game, Fish and Parks in September 2001. In each lake, six 20-min-standardized sites were sampled at night with an electrofishing boat. Because the distributions of lengths in each sample were considered normal, ANOVA was chosen to analyze these data. The analysis was performed using the general linear model procedure (PROC GLM) in SAS (SAS 1999). The purpose of this analysis was to compare mean length of age-0 walleyes among the three lakes. Differences in mean length of age-0 walleye among lakes in fall should indicate differences in growth achieved during the first year of life. The null hypothesis is that there is no difference in mean length among lakes.

Data

The length-frequency histograms for each population are presented in Figure 9.3. All walleye were measured to the nearest millimeter.

Program

```
DATA ONE;
INPUT LAKE $ LENGTH;
CARDS;
ISLAND          122
ISLAND          126
ISLAND          129
[Data input continued]
MITCHELL        145
MITCHELL        152
MITCHELL        160
[Data input continued]
THOMPSON        123
THOMPSON        128
THOMPSON        129
[data input continued]
;
PROC SORT;
BY LAKE LENGTH;
PROC GLM;
CLASS LAKE;
MODEL LENGTH=LAKE;
RUN;
```

Output

Table General linear model (GLM) procedure for length of age-0 walleyes (dependent variable) compared among three South Dakota lakes. The data included 360 observations.

Class Level Information		
Class	Levels	Values
Lake	3	Island Mitchell Thompson

GLM Procedure					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	2	49324.1151	24662.0575	81.73	<0.0001
Error	357	107724.5072	301.7493		
Corrected total	359	157048.6222			

Results

Results of the ANOVA indicated that there was a significant ($F = 81.73$; $P < 0.0001$) difference in mean length among lakes, leading to the rejection of the null hypothesis.

Next, a multiple-comparison test was performed to determine which mean lengths (lakes) were different from one another. In this example, the Tukey's multiple-comparison test was used; it can be invoked using the following code. The program also calls for calculation of mean length for each lake.

Program

```
PROC SORT;
BY LAKE LENGTH;
PROC GLM;
CLASS LAKE;
MODEL LENGTH=LAKE;
MEANS LAKE/TUKEY;
PROC MEANS;
BY LAKE;
VAR LENGTH;
RUN;
```

Output

Table The GLM procedure for Tukey's studentized range (HSD) test for length. This test controls the type I experimentwise error rate. Comparisons significant at the 0.05 level are indicated by ***.

Test Statistics	
Alpha	0.05
Error *df*	357
Error mean square	301.7493
Critical value of studentized range	3.32840

Means Comparisons			
Lake comparison	Difference between means	Simultaneous 95% confidence limits	
Mitchell–Thompson	22.210	17.054	27.366***
Mitchell–Island	28.315	22.736	33.894***
Thompson–Mitchell	–22.210	–27.366	–17.054***
Thompson–Island	6.105	0.852	11.359***
Island–Mitchell	–28.315	–33.894	–22.736***
Island–Thompson	–6.105	–11.359	–0.852***

The MEANS Procedure					
Lake	*N*	Mean	SD	Minimum	Maximum
Island	104	159.4326923	15.8853687	122.0000000	195.0000000
Mitchell	111	187.7477477	13.7136087	145.0000000	218.0000000
Thompson	145	165.5379310	20.5895809	123.0000000	216.0000000

Results

According to this test, all mean lengths are significantly ($P \leq 0.05$) different from one another. Mean length is greatest in Lake Mitchell (188 mm) followed by Lake Thompson (166 mm) and Island Lake (159 mm).

Box 9.2 Testing for Differences among Length-Frequency Distributions by Means of the Kolmogorov–Smirnov Two-Sample Test

The same walleye data analyzed in Box 9.1 (and shown in Figure 9.3) are used in this example. The purpose of this analysis is to compare length-frequency distributions of walleyes among the three lakes by means of a Kolmogorov–Smirnov two-sample test. The analysis was performed using the NPAR1WAY procedure in SAS (SAS 1999). The null hypothesis is that there are no differences in length-frequency distributions (i.e., distribution functions) among lakes. This is a popular nonparametric method to determine differences in length frequencies, as length-frequency data oftentimes deviate substantially from normal. Because this is a two-sample test, only two lakes can be compared simultaneously. Thus, a total of three comparisons (between Mitchell and Thompson, between Island and Thompson, and between Island and Mitchell) were made. In the SAS code shown, Island Lake was deleted from the analysis for the comparison between Lake Mitchell and Lake Thompson. To maintain an experimentwise error rate of $\alpha = 0.05$, the significance level for each comparison ($P = 0.017$) was established by dividing α (0.05) by the number of comparisons (3).

Data

See Box 9.1 and Figure 9.3.

Program

```
DATA ONE;
INPUT LAKE $ LENGTH;
CARDS;
ISLAND          122
ISLAND          126
ISLAND          129
[Data input continued]
MITCHELL        145
MITCHELL        152
MITCHELL        160
[Data input continued]
THOMPSON        123
THOMPSON        128
THOMPSON        129
[Data input continued]
;
DATA TWO; SET ONE;
IF LAKE = "ISLAND" THEN DELETE;
PROC SORT;
BY LAKE LENGTH;
RUN;
PROC NPAR1WAY;
CLASS LAKE;
VAR LENGTH;
RUN;
```

Output

Table Comparison of Lake Mitchell and Lake Thompson. Kolmogorov–Smirnov test for variable LENGTH classified by variable LAKE. The EDF is the empirical distribution function; KS represents the Kolmogorov–Smirnov statistic and KS_a the asymptotic KS; D is the two-sample KS statistic; and $P > KS_a$ is the asymptotic P-value of KS_a, which equals $P > D$.

		Kolmogorov–Smirnov Test	
Lake	*N*	EDF at maximum	Deviation from mean at maximum
Mitchell	111	0.090090	–3.166332
Thompson	145	0.620690	2.770345
Total	256	0.390625	

Maximum deviation occurred at observation	201
Value of LENGTH at maximum	169.0

Kolmogorov–Smirnov Two-Sample Test (Asymptotic)			
KS	0.262950	*D*	0.530600
KSa	4.207193	$P > KS_a$	<0.0001

Table Comparison of Island Lake and Lake Thompson. Kolmogorov–Smirnov test for variable LENGTH classified by variable LAKE.

Kolmogorov–Smirnov Test			
Lake	*N*	EDF at maximum	Deviation from mean at maximum
Island	104	0.846154	0.929386
Thompson	145	0.689655	–0.787098
Total	249	0.755020	

Maximum deviation occurred at observation	204
Value of LENGTH at maximum	175.0

Kolmogorov–Smirnov Two-Sample Test (Asymptotic)			
KS	0.077181	*D*	0.156499
KSa	1.217900	$P > KS_a$	0.1029

Table Comparison of Island Lake and Lake Mitchell. Kolmogorov–Smirnov test for variable LENGTH classified by variable LAKE.

Kolmogorov–Smirnov Test			
Lake	*N*	EDF at maximum	Deviation from mean at maximum
Island	104	0.875000	3.421086
Mitchell	111	0.225225	–3.311458
Total	215	0.539535	

Maximum deviation occurred at observation	129
Value of LENGTH at maximum	176.0

Kolmogorov–Smirnov Two-Sample Test (Asymptotic)			
KS	0.324715	*D*	0.649775
KSa	4.761259	$P > KS_a$	<0.0001

Results

Results of these tests indicate that differences in the length-frequency distributions (i.e., distribution functions) were found among the three lakes, leading to the rejection of the null hypothesis. The length-frequency distribution of age-0 walleye in Lake Mitchell was significantly ($P < 0.0001$) greater than that in Island Lake and Lake Thompson. No difference was observed between Island Lake and Lake Thompson ($P = 0.1037$). Thus, the fisheries scientist may conclude that growth of age-0 walleye was fastest in Lake Mitchell.

Box 9.3 Testing for Differences among Length-Frequency Distributions by Means of the Kruskal–Wallis test

A Kruskal–Wallis test was applied to the same walleye data used in Boxes 9.1 and 9.2. The objective of this analysis was to test whether length-frequency distributions were different among samples (i.e., does one population tend to yield larger or smaller values than the other) based on rank scores. The null hypothesis was that there was no difference among the length-frequency distributions. The Kruskal–Wallis test is an extension of Wilcoxon's rank-sum test for two samples. Results of the Kruskal–Wallis and Wilcoxon's rank-sum tests are provided in the output through execution of the NPAR1WAY procedure in SAS (SAS 1999). The input data are the same as used in Box 9.1 and presented in Figure 9.3.

Program

```
DATA ONE;
INPUT LAKE $ LENGTH;
CARDS;
[See data input in Box 9.1]
;
DATA TWO; SET ONE;
PROC SORT;
BY LAKE LENGTH;
RUN;
PROC NPAR1WAY;
CLASS LAKE;
VAR LENGTH;
RUN;
```

Output

Table Wilcoxon scores (rank sums) for the variable length classified by the variable lake.

Wilcoxon Scores					
Lake	*N*	Sum of scores	Expected under H_0	SD under H_0	Mean score
Island	104	12969.00	18772.00	894.755226	124.701923
Mitchell	111	29767.50	20035.50	911.651297	268.175676
Thompson	145	22243.50	26172.50	968.212806	153.403448

Kruskal–Wallis Test	
Chi-square	118.5671
df	2
P > chi-square	<0.0001

Results

The output indicates that there is a significant ($P < 0.0001$) difference among the three length-frequency distributions, and thus, the null hypothesis is rejected. The mean ranks for Island Lake (124.7), Lake Mitchell (268.2), and Lake Thompson (153.4) are provided in the output under the mean score column.

By default, the NPAR1WAY procedure in SAS provides approximated *P*-values based on asymptotic methods (SAS 1999). Exact *P*-values can be calculated by using the EXACT statement in the NPAR1WAY procedure. Asymptotic methods may not be valid when sample sizes are very small and when data are sparse, skewed, or heavily tied (SAS 1999). When sample sizes are large, asymptotic *P*-values approach exact *P*-values. The EXACT statement in SAS can be computationally time-consuming depending on the sample size and the number of groups. Exact *P*-values for this example can be obtained by using the following code.

```
PROC NPAR1WAY;
CLASS LAKE;
VAR LENGTH;
EXACT;
RUN;
```

Program

In SAS, the Kruskal–Wallis test can also be performed by using a combination of the RANK and GLM procedures (SAS 1990). The overall *F*-test is asymptotically equivalent to the Kruskal–Wallis test in SAS. The program below will perform an ANOVA based on ranked data.

```
PROC RANK OUT=RANKS;
RANKS RLENGTH;
VAR LENGTH;
RUN;
PROC GLM DATA=RANKS;
CLASS LAKE;
MODEL RLENGTH=LAKE;
RUN;
```

Output

Table The GLM procedure for the dependent variable RLENGTH, the rank for the variable length. Abbreviations are as follows: mean square error (MSE); coefficient of variation (CV); and sum of squares (SS).

Analysis of Variance					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	2	1283518.268	641759.134	88.03	<0.0001
Error	357	2602744.232	7290.600		
Corrected total	359	3886262.500			
R^2	0.330271	Root MSE	85.38501		
CV	47.30472	RLENGTH mean	180.5000		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Lake	2	1283518.268	641759.134	88.03	<0.0001

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Lake	2	1283518.268	641759.134	88.03	<0.0001

In this ANOVA, note that $(n - 1)R^2 = 118.57$ and is the same as the chi-square statistic provided for the Kruskal–Wallis test (SAS 1999). There is a significant ($P < 0.0001$) difference in mean ranked length among the three lakes, leading to the rejection of the null hypothesis.

Box 9.4 Performing Multiple Comparisons of Length-Frequency Data

In Box 9.3, three length-frequency distributions were compared using the Kruskal–Wallis test. The distributions were found to be significantly different. Once the null hypothesis is rejected, the fisheries scientist usually will want to determine between which of the samples the significant differences exist. An example of a nonparametric multiple-comparison test (Zar 1996) based on the walleye data presented in Box 9.3 is illustrated below. This particular multiple-comparison test is appropriate in the case of several tied ranks and unequal sample sizes, which are typical characteristics of length-frequency data, especially when fish are measured and reported to the nearest length-group (e.g., centimeters). However, several other multiple-comparison tests are available, depending on the characteristics of the data being analyzed (Conover 1980; Zar 1996). Information in the summary table below can be found in the SAS output for the Kruskal–Wallis test in Box 9.3.

Parameter	Island	Thompson	Mitchell
Mean rank ($\bar{R}$)	124.70	153.40	268.18
Sample size (n)	104	145	111

After the entire data set was rank ordered, the number of groups (lengths) with tied ranks (m) was determined to be 69. Next, calculate T, the tied-rank statistic,

$$T = \sum_{i=1}^{m} (t_i^3 - t_i) ,$$

where t = the frequency of observations with tied ranks in the ith group (length). For example, if in a data set there were two groups (lengths) with tied ranks (three 247-mm fish and two 248-mm fish), T would equal $(3^3 - 3) + (2^3 - 2) = 30$. For the walleye example used in this box, there were many ties, and $T = 20{,}490$.

Next, SEs are calculated for each comparison. The SE for the comparison of Lake Mitchell with Island Lake is calculated as

$$\mathrm{SE} = \sqrt{\left(\frac{N(N+1)}{12} - \frac{T}{12(N-1)}\right)\left(\frac{1}{n_{\text{Mitchell}}} + \frac{1}{n_{\text{Island}}}\right)}$$

$$= \sqrt{\left(\frac{360(361)}{12} - \frac{20{,}490}{12(359)}\right)\left(\frac{1}{111} + \frac{1}{104}\right)} = 14.20$$

For the comparison of Lake Mitchell and Lake Thompson, the SE = 13.12.
For the comparison of Lake Thompson and Lake Island Lake, the SE = 13.37.

The test statistic (Q) for each comparison is calculated as the difference in mean ranks divided by the associated SE. Critical values for Q can be obtained from tables for nonparametric multiple comparisons (e.g., Zar 1996). In this example, the critical value of Q at $\alpha = 0.05$ for three samples is 2.394. For each comparison, the null hypothesis of no difference between length-frequency distributions is rejected if the calculated Q exceeds the critical value of Q.

Table Nonparametric multiple comparison among three lakes of length-frequency distributions of walleye.

Comparison	$\bar{R}_x - \bar{R}_y$	SE	Q	$Q_{0.05,3}$	Conclusion
Mitchell and Island	268.18 – 124.70 = 143.48	14.20	10.10	2.394	Reject H_0
Mitchell and Thompson	268.18 – 153.40 = 114.78	13.12	8.75	2.394	Reject H_0
Thompson and Island	153.40 – 124.70 = 28.70	13.37	2.14	2.394	Accept H_0

The fisheries scientist can conclude that the length-frequency distribution from Lake Mitchell is significantly greater than that of Island Lake and Lake Thompson. Length-frequency distributions were not significantly different between Lake Thompson and Island Lake.

Program

Currently, nonparametric multiple-comparison procedures are not available in SAS. However, some of the required calculations such as a table of ranked lengths, a table of the frequency of observations with tied ranks in the *i*th group, and *T* can be obtained by invoking the following SAS program.

```
PROC RANK OUT=RANKS;
RANKS RLENGTH;
VAR LENGTH;
RUN;
PROC FREQ; TABLES RLENGTH/OUT=FRANK;
RUN;
PROC PRINT DATA=FRANK;
RUN;
DATA CALCT; SET FRANK;
IF COUNT=1 THEN DELETE;
T=((COUNT*COUNT*COUNT)-COUNT);
PROC PRINT;
RUN;
PROC MEANS SUM;
VAR T;
RUN;
```

Multiple comparisons can also be accomplished on the ranked data in the GLM procedure in SAS. The following SAS program performs an ANOVA on the ranked data (see Box 9.3) and uses a Tukey's multiple-range test to determine differences among the mean ranks.

```
PROC RANK OUT=RANKS;
RANKS RLENGTH;
VAR LENGTH;
RUN;
PROC GLM DATA=RANKS;
CLASS LAKE;
MODEL RLENGTH=LAKE;
MEANS LAKE/TUKEY;
RUN;
```

(Box continues)

Box 9.4 *(continued)*

Output

The ANOVA output for this analysis is shown in Box 9.3. The results of the Tukey's multiple-comparison test are shown below.

Table Tukey's studentized range (HSD) test for RLENGTH (the rank for variable length). This test controls the type I experimentwise error rate. Comparisons significant at the 0.05 level are indicated by ***.

Test Statistics	
Alpha	0.05
Error df	357
Error mean square	7290.6
Critical value of studentized range	3.32840

Means Comparisons			
Lake comparison	Difference between means	Simultaneous 95% confidence limits	
Mitchell–Thompson	114.77	89.43	140.12***
Mitchell–Island	143.47	116.05	170.90***
Thompson–Mitchell	–114.77	–140.12	–89.43***
Thompson–Island	28.70	2.88	54.52***
Island–Mitchell	–143.47	–170.90	–116.05***
Island–Thompson	–28.70	–54.52	–2.88***

Results

These results show that there is a significant difference in mean ranked length among each of the three lakes.

In these examples, the multiple-comparison-testing methods had different results. The nonparametric multiple-range test was more conservative that the Tukey test. This clearly demonstrates that different multiple-comparison tests can provide different results. The choice of a multiple-comparison test should be made before the analysis is conducted rather than by searching for significance by performing multiple tests.

Box 9.5 Using Contingency Tables to Test for Differences in Length-Frequency Distributions

The chi-square test is commonly used to test for differences in length-frequency distributions. In this example, DC electrofishing at night was used to collect bluegill in 1996, 1998, and 2000 from a private pond in Connecticut. Bluegills were classified into two length-groups: stock to quality length (80–149 mm) and quality length (≥150 mm). Proportional stock density (see section 9.3) was also calculated for each year. The objective of this analysis was to determine whether length-frequency distributions (summarized by PSD values) were different among years. The chi-square analysis was performed using the frequency procedure (FREQ) in SAS (SAS 1999). The null hypothesis is that the frequency of observations among length-groups (stock to quality length and quality length) is independent of year.

Data

Table The number of bluegill collected in each length-group and proportional stock density (PSD).

	Year		
Size category and length index	1996	1998	2000
Stock to quality length (80–149 mm)	77	124	251
Quality length (≥150 mm)	85	44	34
Total stock length (≥80 mm)	162	168	285
PSD	52	26	12

Program

In the following SAS program, LCAT is the length category (S–Q = stock to quality length and Q = greater than or equal to quality length) and NUM is the number of fish.

```
DATA ONE;
INPUT YEAR LCAT $ NUM;
CARDS;
1996 S-Q   77
1996   Q   85
1998 S-Q  124
1998   Q   44
2000 S-Q  251
2000   Q   34
;
PROC SORT;
BY YEAR LCAT NUM;
RUN;
DATA TWO;
SET ONE;
BY YEAR LCAT NUM;
IF FIRST.LCAT THEN DO;
DO I = 1 TO NUM;
LCAT = LCAT;
YEAR = YEAR;
OUTPUT;
END;
END;
RUN;
PROC FREQ;
TABLES YEAR*LCAT / CHISQ;
RUN;
```

(Box continues)

Box 9.5 *(continued)*

Output

Table Summary statistics for chi-square analysis of length category (LCAT) by year. Sample size is 615.

	Length category		
Year and measure	Q	S–Q	Total
1996			
Frequency	85	77	162
Percent	13.82	12.52	26.34
Row %	52.47	47.53	
Column %	52.15	17.04	
1998			
Frequency	44	124	168
Percent	7.15	20.16	27.32
Row %	26.19	73.81	
Column %	26.99	27.43	
2000			
Frequency	34	251	285
Percent	5.53	40.81	46.34
Row %	11.93	88.07	
Column %	20.86	55.53	
Total			
Frequency	163	452	615
Percent	26.50	73.50	100.00

Box 9.1 an *F*-test with 357 error degrees of freedom results in a high level of power to detect differences among length-frequency distributions. Similarly, Kolmogorov–Smirnov two-sample tests are often highly significant when sample sizes are large, even though the distributions can appear similar. Caution should be applied when using individual fish as experimental units, resulting in very high sample sizes.

An alternative approach to comparing length frequencies would be to treat each group of fish caught in a unit of effort (e.g., trap net or electrofishing station) as a sample. In other words, each unit of effort would be considered a sample or "collection event," and individual fish would be considered subsamples. Consider sampling black crappies with 20 trap nets during a single sampling period in a reservoir. If the 20 nets are set according to a particular sampling design, then each net (location) may be adequate to use as an independent experimental unit. Examples of using units of effort as samples to compare size structure are provided in Boxes 9.6 and 9.7.

Table Chi-square statistics of length category by year.

Statistic	*df*	Value	*P*
Chi-square	2	87.1540	<0.0001
Likelihood ratio chi-square	2	85.5173	<0.0001
Mantel–Haenszel chi-square	1	84.8020	<0.0001
Phi coefficient		0.3764	
Contingency coefficient		0.3523	
Cramer's V		0.3764	

Results

According to the chi-square test, there is a significant ($\chi^2 = 87.15$, $P < 0.0001$) difference in the frequency of observations between length-groups, leading to the rejection of the null hypothesis. To test which years were significantly different from each other, a chi-square test was performed for each combination of years (i.e., 1996 and 1998, 1998 and 2000, and 1996 and 2000) based on 2×2 contingency tables. Although the results for each comparison are not shown, all pairwise tests showed significant differences ($P < 0.0001$) between years. To maintain an experimentwise error rate of $\alpha = 0.05$, the significance level for each comparison ($P = 0.017$) was established by dividing α by the number of comparisons (3). Size structure declined from 1996 (PSD = 52) to 1998 (PSD = 26) to 2000 (PSD = 12).

In this pond, the decrease in PSD of bluegills over the 3 years was probably due to the reduction in density of chain pickerel in the pond. Mean *C/f* (number/h electrofishing) of chain pickerel declined from 82 in 1996 to 39 in 2000. Declines in chain pickerel abundance probably lead to reduced predation on bluegills, resulting in higher abundance and reduced growth of bluegills.

To test for differences in length-frequency data that are summarized using stock density indices other than PSD (e.g., relative stock density preferred length [RSD-P] or quality-to-preferred length [RSD-Q-P]), simply change the length categories in the analysis. For example, to test for differences in length frequency summarized as RSD-P, test for differences in the frequency of occurrence of stock-to-preferred-length fish and preferred-length fish among treatments.

9.4.4 Analysis of Repeated Measures

Fisheries scientists frequently assess changes in size structure on one population through time (e.g., across years). One consideration is that many of the statistical procedures mentioned above (e.g., chi-square test and Kruskal–Wallis) assume the samples are independent. For example, in Box 9.5 bluegill PSD was tested using samples collected at 2-year intervals, and the chi-square test assumes that those samples are independent. Because samples were at 2-year intervals, this assumption may be realistic. However, samples collected over a number of consecutive years are likely not independent (Maceina et al. 1994) because catch rates or size structure in 1 year may influence the size structure in subsequent years (i.e., the same year-classes are sampled over time).

Repeated-measures ANOVA provides a series of models that incorporate time dependency of the data into the analysis (see also Chapter 7 for discussion of

Box 9.6 Testing for Differences in Size Structure by Treating Groups of Fish Caught in Each Unit of Effort as Samples

Fisheries scientists oftentimes evaluate the effectiveness of alternative sampling methods. However, before alternative sampling methods are implemented into standard sampling programs, the fisheries scientist should understand how data (e.g., size structure) collected by the new sampling method compares to the method currently used. For example, the use of angler-collected data in research and monitoring is becoming more popular due to reliability and reduced costs and effort associated with data collection compared with more traditional methods such as electrofishing.

In this example, size structure of largemouth bass obtained from two sampling methods is compared. Largemouth bass were sampled from Mansfield Hollow Reservoir, Connecticut, in spring 2002. Twelve stations along the lake perimeter were sampled at night by means of DC electrofishing, and size structure data were collected at 12 bass fishing tournaments over the same time period. Individual fish were measured to the nearest centimeter total length at the end of each electrofishing station and fishing tournament.

Catches from each electrofishing station and fishing tournament were considered independent samples. Electrofishing stations did not overlap, and catches in one tournament were considered independent of the others. The null hypothesis tested is that the ratio of the number of preferred-length (i.e., ≥38 cm) fish to the number of quality-length (i.e., ≥30 cm) fish was not different between the two sampling methods.

Data

Table Largemouth bass data from Mansfield Hollow Reservoir, Connecticut, in spring 2002. Catches from each electrofishing station and fishing tournament were considered independent samples.

	Number of fish			Number of fish	
Fishing tournament	≥30 cm	≥38 cm	Electrofishing station	≥30 cm	≥38 cm
1	3	2	1	23	12
2	28	4	2	22	2
3	13	1	3	35	8
4	8	0	4	6	2
5	61	16	5	11	1
6	76	12	6	15	7
7	38	10	7	12	3
8	49	12	8	9	6
9	62	24	9	25	5
10	43	10	10	25	8
11	59	18	11	9	1
12	24	5	12	7	1

Program

Electrofishing (ELEC) and fishing tournaments (TOURN) are the sampling methods used, and QUAL and PREF are the number of fish collected in each length category for each electrofishing station and fishing tournament. The variable LOGIT was created, which is the ratio of the number of preferred-length (≥38 cm) fish to the number of quality-length (≥30 cm) fish in each sample, after a value of 0.5 was added to QUAL and PREF to remove zeros prior to log transformation. From a parametric statistics standpoint, using LOGIT has an advantage over using a proportion because it can exceed one and is more likely to be normally distributed.

The GLM procedure (SAS 1999) was used to conduct a *t*-test to determine whether there was a significant difference in mean LOGIT between the two sampling methods. The WEIGHT statement weights each sample based on the number of fish collected in each sample.

```
DATA BASS;
INPUT METHOD $ QUAL PREF;
CARDS;
ELEC    23 12
ELEC    22  2
ELEC    35  8
ELEC     6  2
[Data input continued]
TOURN   28  4
TOURN   13  1
TOURN    8  0
TOURN   61 16
[Data input continued]
;
DATA BASS2; SET BASS;
LOGIT=LOG((PREF+0.5)/(QUAL+0.5));
PROC PRINT;
PROC SORT; BY METHOD;
PROC MEANS; BY METHOD; VAR LOGIT;
WEIGHT QUAL;
PROC GLM;
CLASS METHOD;
MODEL LOGIT=METHOD;
WEIGHT QUAL;
RUN;
```

Output

Table The number of fish collected in each length category for each sampling method. The variable LOGIT is the ratio of the number of preferred-length (≥38 cm) fish to the number of quality-length (≥30 cm) fish in each sample, after a value of 0.5 was added to QUAL and PREF.

	Number of fish		
Method and observation	QUAL	PREF	LOGIT
TOURN			
1	3	2	–0.33647
2	28	4	–1.84583
3	13	1	–2.19722
4	8	0	–2.83321
5	61	16	–1.31568
6	76	12	–1.81156
7	38	10	–1.29928
8	49	12	–1.37624
9	62	24	–0.93649
10	43	10	–1.42139
11	59	18	–1.16821
12	24	5	–1.49393

(Box continues)

Box 9.6 *(continued)*

Method and observation	Number of fish		
	QUAL	PREF	LOGIT
ELEC			
13	23	12	–0.63127
14	22	2	–2.19722
15	35	8	–1.42947
16	6	2	–0.95551
17	11	1	–2.03688
18	15	7	–0.72594
19	12	3	–1.27297
20	9	6	–0.37949
21	25	5	–1.53393
22	25	8	–1.09861
23	9	1	–1.84583
24	7	1	–1.60944

Table Summary statistics (MEANS procedure) based on LOGIT values for two sampling methods.

Method	*N*	Mean	SD	Minimum	Maximum
ELEC	12	–1.3281428	2.2037466	–2.1972246	–0.3794896
TOURN	12	–1.4280742	2.4069863	–2.8332133	–0.3364722

Table Result of GLM procedure to compare mean LOGIT (*t*-test) between the two sampling methods. The WEIGHT statement weights each sample based on the total number of fish (QUAL) collected in each sample. Analysis is based on 24 observations.

Class Level Information		
Class	Levels	Values
METHOD	2	ELEC TOURN

GLM Procedure					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	1	1.3907895	1.3907895	0.26	0.6144
Error	22	117.1509050	5.3250411		
Corrected total	23	118.5416945			
R^2	0.011732	Root MSE	2.307605		
CV	–165.0553	LOGIT mean	–1.398080		

Results

At the top of the output, LOGIT values for each sample are provided by the PROC PRINT statement. Sample LOGIT values are followed by the mean LOGIT for electrofishing (mean LOGIT = –1.33) and fishing tournaments (mean LOGIT = –1.43) weighted by METHOD based on the number of fish in each sample (QUAL). By calculating the inverse $\log_e$ of these values, the mean ratio of the number of preferred-length (i.e., ≥38 cm) fish to the number of quality-length (i.e., ≥30 cm) fish was 0.26 for electrofishing and 0.24 for fishing tournaments. The output for the GLM procedure indicates that there was not a significant difference ($F = 0.26$; $P = 0.6144$) in mean LOGIT between electrofishing and fishing tournaments. The fisheries scientist fails to reject the null hypothesis that the ratio of the number of preferred-length (i.e., ≥38 cm) fish to the number of quality-length (i.e., ≥30 cm) fish was the same between the two sampling methods.

Box 9.7 Using Repeated-Measures ANOVA to Test for Size Structure Differences with Time-Dependent Data

Repeated-measures ANOVA is commonly used to test for differences in a population response when samples are not independent, often because they are collected through time. In this example, we assessed differences in the size structure of a largemouth bass population at Lake Jackson, Florida, after implementation of a 330–431-mm protected-slot-length limit. Data were collected by Florida Fish and Wildlife Conservation Commission biologists. The population had no size limit prior to 1991, and the slot-length limit was enacted in July of 1990. Daytime electrofishing samples were collected at 12 fixed sites during April from 1988 to 1996. Fixed sites in the analysis were treated as subjects sampled through time. The size distribution of largemouth bass is likely to be dependent through time (i.e., size structure of fish present in previous sampling influences size structure at later time intervals). Thus, the analysis should consider that the size structure of fish at a given site is not independent through time. In this example, we assessed whether the size structure of largemouth bass differed before the slot-length limit ($N = 3$ years of data) compared with after the slot limit ($N = 6$ years of data). Largemouth bass were classified into three groups (based on total length): below 200 mm, 200–329 mm, and 330 mm and larger. The objective of this analysis was to test whether the ratio of fish 330 mm and larger to fish between 200 and 329 mm differed before and after the slot-length limit was enacted. Fish below 200 mm were removed from the analysis because the slot limit was not expected to influence abundance of fish below 200 mm. The null hypothesis is that the ratio of fish 330 mm and larger to subslot-size fish (i.e., fish > 200 mm but less than 330 mm) was not different before and after the slot limit was enacted. The mixed-models procedure (PROC MIXED; SAS 1999) was used to conduct the test.

Data: Part I

In the data table below, COUNT is the number of fish in each size-group, and size-groups are given as UND (200–329 mm) and SLOT (330 mm and longer). A SITE was included if a fish was collected in at least one size-group. However, sites that did not contain fish in either size-group were removed from the analysis because collection of no fish provides no information about size structure (i.e., if both the UND and SLOT size groups had a COUNT of zero, the site was not included in the analysis).

Table Size-group data for largemouth bass fishery in Lake Jackson, Florida, before and after slot-length limit implementation.

Year and site	Size-group	Count
1988		
1	UND	4
1	SLOT	1
2	UND	9
2	SLOT	2
[Data continued]		
1996		
11	UND	4
11	SLOT	4
12	UND	9
12	SLOT	4
[Data continued]		

Program: Part I

In the following SAS program, the data were rearranged using PROC TRANSPOSE prior to creating the dependent variable for the test. This procedure changes columns to rows or rows to columns. In this example the column COUNT was changed to rows for both size-groups.

(Box continues)

Box 9.7 *(continued)*

```
DATA A;
INPUT YEAR SITE    SIZEGRP $ COUNT;
CARDS;
1988        1        bund       4
1988        1        slot       1
1988        2        bund       9
1988        2        slot       2
[Data input continued]
;
DATA B; SET A;
IF YEAR LE 1990 THEN PERIOD = 'APRE';
IF YEAR GT 1990 THEN PERIOD = 'BPOST';
RUN;

PROC SORT;
BY PERIOD YEAR SITE;
PROC TRANSPOSE OUT=C;
BY PERIOD YEAR SITE;
VAR COUNT;

DATA D; SET C;
RENAME COL2 = UNDER;
RENAME COL3 = SLOT;
DATA E; SET D;
UNDERT=UNDER+0.5;
SLOTT = SLOT+0.5;
TOTAL =UNDERT+SLOTT;
LOGIT=LOG(SLOTT/(UNDERT));
```

In step DATA E, 0.5 was added to each count to remove zeros prior to the log transformation.

The variable TOTAL is used to weight each transect in PROC MIXED below. Remember that fish shorter than 200 mm were removed, and the ratio of the number of fish greater than or equal to 330 mm to the number of fish between 200 and 329 mm was used.

The variable LOGIT is the log of the ratio of fish greater than or equal to 330 mm relative to fish between 200 and 329 mm. It was predicted that after the slot limit is in place (i.e., anglers cannot keep fish between 330 and 431 mm) the ratio, and thus, the LOGIT, would increase.

```
PROC PRINT;
VAR YEAR PERIOD SITE UNDER UNDERT SLOT SLOTT LOGIT;
RUN;
```

Data: Part II
From the above program, the final data set was created prior to analysis.

Table Final data set for analysis of largemouth bass fishery before and after slot-length limit implementation. Size-groups are UNDER (200–329 mm), SLOT (330 mm and longer), and those two categories transformed (UNDERT and SLOTT).

Period, year, and observation	Site	UNDER	UNDERT	SLOT	SLOTT	LOGIT
Pre-slot limit						
1988						
1	1	4	4.5	1	1.5	–1.09861
2	2	9	9.5	2	2.5	–1.33500
3	3	11	11.5	1	1.5	–2.03688

[Data continued]

Period, year, and observation	Site	UNDER	UNDERT	SLOT	SLOTT	LOGIT
Post-slot limit						
1996						
1	10	4	4.5	2	2.5	–0.58779
2	11	4	4.5	4	4.5	0.00000
3	12	9	9.5	4	4.5	–0.74721
[Data continued]						

Program: Part II

```
PROC UNIVARIATE PLOT NORMAL;
BY PERIOD;
VAR LOGIT;
```

The Wilk's lambda in PROC UNIVARIATE, specified with the NORMAL option, was used to assess whether the dependent variable (LOGIT) was normally distributed for each PERIOD (pre- versus post-slot limit years). In this case the assumptions of normality were met ($P > 0.05$ for both PERIODS).

```
PROC MIXED;
CLASS PERIOD YEAR SITE;
MODEL LOGIT=PERIOD;
WEIGHT TOTAL;
RANDOM YEAR(PERIOD);
REPEATED YEAR/SUBJECT=SITE TYPE=AR(1);
LSMEANS PERIOD/PDIFF;
RUN;
```

The model tests whether the mean LOGIT differs significantly between periods (pre- versus post-slot limit regulation). The WEIGHT statement weights each site based on the number of fish collected (i.e., sites with large catch influence the test proportionally more). The RANDOM statement assumes that among-year variation within each PERIOD was random. The REPEATED statement indicates the consecutive years of sampling (i.e., time variable), and the SUBJECT statement assigns each site as an individual station sampled through time. In this case, sites were not chosen at random so sites are treated as a fixed effect in the model. In SAS, the TYPE statement allows the researcher to investigate various covariance structures to model the time-dependence of the data (discussed below). The LSMEANS statement is a means separation option that will give the overall least-squares means for each period and their significance level (PDIFF).

Output

Table The following output is from the PROC MIXED statement presented above that tests whether the mean LOGIT differs significantly between periods (PRE = pre-slot limit and POS = post-slot limit).

Model Information	
Data Set	WORK.E
Dependent variable	LOGIT
Weight variable	TOTAL
Covariance structures	Variance components, Autoregressive
Subject effect	SITE
Estimation method	REML
Residual variance method	Profile

(Box continues)

Box 9.7 *(continued)*

Fixed effects SE method	Model-based
Degrees of freedom method	Containment

Class Level Information

Class	Levels	Values
PERIOD	2	PRE POS
YEAR	9	1988 1989 1990 1991 1992 1993 1994 1995 1996
SITE	12	1 2 3 4 5 6 7 8 9 10 11 12

Dimensions

Covariance parameters	3
Columns in X	3
Columns in Z	9
Subjects	1
Maximum observations per subject	101
Observations used	101
Observations not used	0
Total observations	101

Iteration History

Iteration	Evaluations	–2Residual log likelihood	Criterion[a]
0	1	269.51966315	
1	2	268.00863114	0.00000795
2	1	268.00828432	0.00000000

Covariance Parameter Estimates

Covariance parameter	Subject	Estimate
Year (period)		0.04188
AR(1)	Transect	0.09518
Residual		7.7798

Fit Statistics

–2Residual log likelihood	268.0
AIC (smaller is better)	274.0
AICc (smaller is better)	274.3
BIC (smaller is better)	274.6

Type III Test of Fixed Effects

Effect	Numerator df	Denominator df	*F*-value	*P* > *F*
PERIOD	1	7	4.60	0.0691

Least Squares Means

Effect	Period	Estimate	Standard error	*df*	*t*-value	*P* > |*t*|

PERIOD	PRE	–0.8506	0.1829	7	-4.65	0.0023
PERIOD	POS	–0.3602	0.1396	7	-2.58	0.0365

Differences of Least Squares Means

Effect	Period	Period	Estimate	SE	*df*	*t*-value	$P > \|t\|$
PERIOD	PRE	POS	–0.4905	0.2286	7	–2.15	0.0691

[a] The convergence criteria were met.

Results

The "Model Information" and "Class Level Information" output show the model configuration and levels of each class going into the model. The "Iteration History" reveals whether the model converged on a solution. The series of "Fit Statistics" allows one to compare various covariance matrix structures to one's data. The Akaike's Information Criteria (AIC), small sample corrected AIC (AICc), and Bayesian Information Criteria (BIC) are model fit statistics commonly used for many ecological modeling applications (see Guthery et al. [2005] for a review and critique). In this example, the AIC statistic is used to assess how well the time-dependent structure of one's data fit the chosen TYPE covariance structure specified in the model (SAS 1999). In this case, TYPE = AR(1) was used, which is the first-order autoregressive structure. The AR(1) structure models correlations between time periods that are linear and decline with the distance in time that observations are made (e.g., assumes years 1 and 2 are more closely related than years 1 and 4; Littell et al. 1996). Littell et al. (1996) described various options for covariance structures in PROC MIXED, and the investigator can choose the structure type with the lowest AIC score (i.e., lowest deviance between the data and the specified structure type). The AR(1) structure is one option for data sampled at regular time intervals, which in this example was appropriate because electrofishing occurred in April of each year. The AR(1) model also obtained the lowest AIC score of several covariance structures considered.

However, we note that the time dependency of the data were not strong based on covariance parameter estimates of 0.095 for the AR(1) variable relative to a high residual value (7.78). Analyses showing strong time dependence typically exhibit covariance parameter estimates of equal or greater magnitude compared with the residual values (authors, personal observation). The lack of a relationship between size structure data in successive years is not atypical given variation around electrofishing data, and would be grounds to ignore time dependency and use a regular one-way ANOVA to test for differences in LOGIT between PERIODS. Thus, samples collected in successive years do not automatically require repeated-measures analyses! Here we'll continue with the output interpretation as an example of the analysis.

Results of this analysis showed that the LOGIT approached significance between pre- and post-size-limit time periods ($P = 0.069$) at an $\alpha = 0.05$. The LSMEANS procedure output the least-squares means of the LOGIT (–0.85 and –0.36). By taking the inverse $\log_e$ of these values, we find that the ratio of fish 330 mm and larger to fish between 200 and 329 mm averaged 0.427 before the slot limit and 0.698 after the slot limit was enacted. Thus, at an α level of 0.10, the ratio increased after the slot was enacted, suggesting that the size structure increased. It is important to note that although there is a significant difference, variables other than the slot limit (e.g., strong year-classes or changes in large fish catchability) could have also influenced the result. This example shows how time dependency in the data can be included in assessment of fish size structure.

repeated-measures data). Maceina et al. (1994) described how a split-plot ANOVA could be used to conduct repeated-measures tests. More recently, mixed-model ANOVA provides multiple options to handle repeated-measures data. The advantage of mixed models over split-plot analyses is that the split-plot ANOVA assumes compound symmetry (Littell et al. 1996). Compound symmetry is defined as constant dependence; in other words, each time period is assumed to be equally related to all other time periods. Mixed-model ANOVA allows the investigator to specify covariance matrix structures other than compound symmetry (Littell et al. 1996). For example, you might expect samples collected in consecutive years to be more highly related than are samples collected 5 years apart. Box 9.7 provides an example of using a repeated-measures ANOVA to test for size structure differences based on time-dependent data. When assessing population size structure on one population through time, or across multiple populations sampled through time, use of repeated-measures designs is recommended.

9.5 INTERPRETATION OF SIZE STRUCTURE

9.5.1 Length-Frequency Distributions

Length-frequency distributions reflect an interaction of the rates of recruitment, growth, and mortality of a fish population. Length-frequency data can provide insight into the dynamics of fish populations and identify problems such as inconsistent year-class strength, slow growth, and excessive mortality (Anderson and Neumann 1996). In most instances, a thorough assessment of a fish population requires other population assessment tools, such as *C/f,* age and growth, or body condition, in addition to length-frequency data.

Length-frequency data for black crappie collected with a trawl from two Florida lakes are presented in Figure 9.5. Based on the length-frequency distribution for Lake Jackson, a fisheries scientist may conclude that the black crappie population is balanced. A balanced population is one that has moderate rates of recruitment, growth, and mortality compared with what is expected for populations in the same geographic region. A length-frequency histogram from a balanced fish population will have a stable decline from the shorter to longer lengths, reflecting a stable age structure produced by consistent recruitment and consistent, moderate rates of mortality among successive age-classes. In exploited populations, the term balance has also been referred to a population that produces sustainable yields of harvestable-size fish. However, balanced populations can also occur in unexploited water bodies (Anderson and Neumann 1996).

The length-frequency histogram for Alligator Lake (Figure 9.5) may indicate that this population is unbalanced. The most striking difference between Lake Jackson and Alligator Lake is that Alligator Lake does not show a stable decline in the numbers of fish with increasing length. Instead, the length-frequency histogram is "interrupted" by length-groups with many individuals bounded by length-groups with fewer individuals. If the strong and weak interruptions corresponded

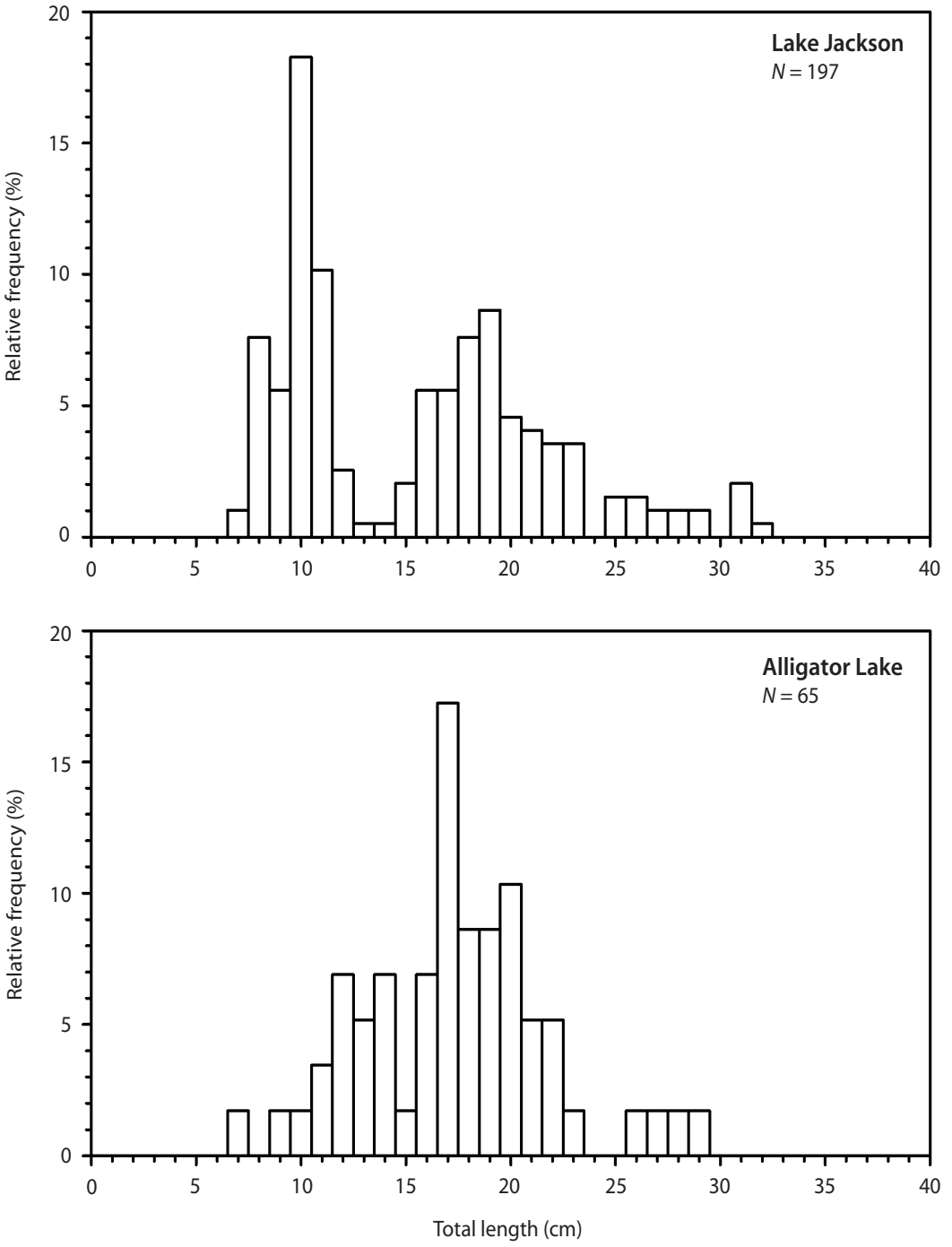


Figure 9.5 Relative-frequency histograms for black crappie from Lake Jackson and Alligator Lake, Florida, collected by means of a trawl. Data were provided by the Florida Fish and Wildlife Conservation Commission.

to age-groups, and if all lengths represented were vulnerable to the sampling gear, then a fisheries scientist might conclude that year-class strength at Alligator Lake is inconsistent compared with Jackson Lake. However, the clearest indication of variable year-class strength would be determined from age-frequency analysis (see Chapter 4).

The length-frequency distribution for largemouth bass collected by means of night electrofishing in a South Dakota pond is presented in Figure 9.6. Note that all largemouth bass sampled were less than quality length; thus PSD = 0. Mortality in this population possibly is high, demonstrated by the lack of largemouth bass greater than quality length. When examining the size structure, a fisheries scientist might arrive at one of several conclusions about the status of this population: (1) low recruitment, slow growth, and moderate to high mortality due to poor habitat; (2) overharvest of largemouth bass greater than quality length; or (3) high density of small, slow-growing largemouth bass due to excessive recruitment. The last condition is often referred to as stunting. In this example, length-frequency information alone could not be interpreted to arrive at the cause for the poor population structure. Other information such as *C/f,* growth, or body condition assessment would be necessary. In Knox Pond, *C/f* was 306 stock-length largemouth

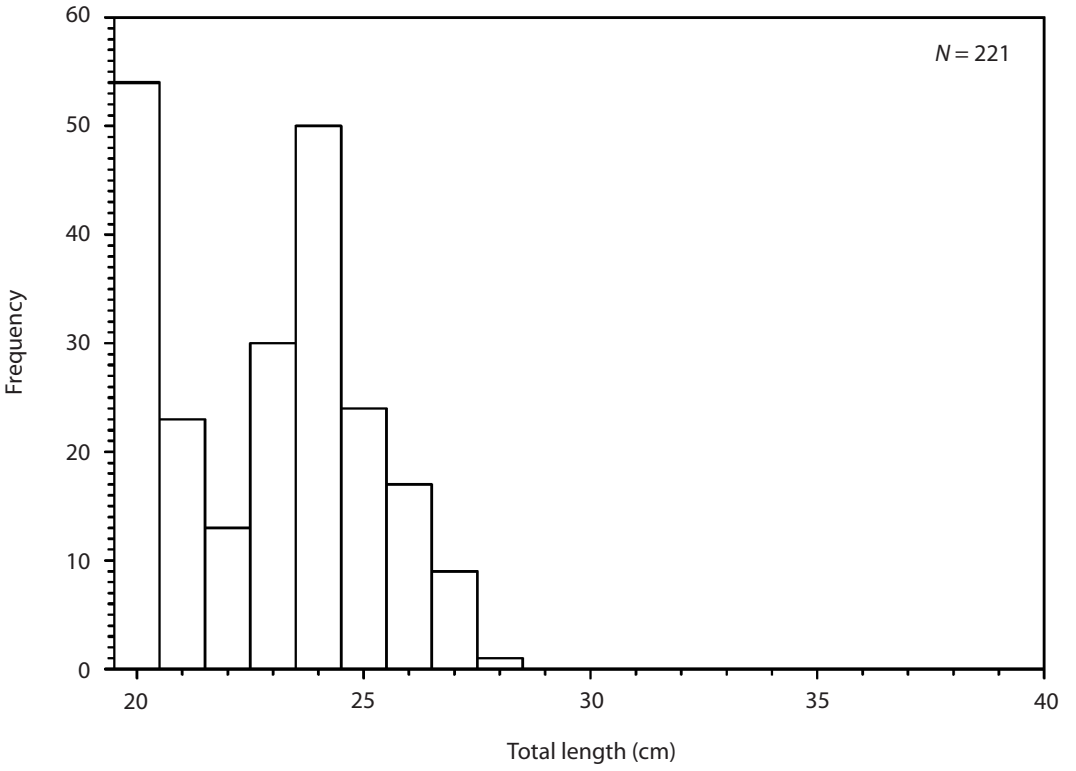


Figure 9.6 Absolute-length-frequency histogram of stock-length (≥20 cm) largemouth bass sampled by electrofishing from Knox Pond, South Dakota.

bass per hour of electrofishing (Neumann et al. 1994), which was high compared with other populations in the state. Mean relative weight (W_r) was 77, and growth rate was well below the state average. Thus, this population represented condition 3 listed above. Condition 1 might be confirmed if C/f was low, growth was slow, and poor habitat was documented. Condition 2 might be confirmed if growth was moderate to fast, habitat conditions were favorable, and creel statistics showed a high harvest of quality-length fish. This example also demonstrates the value of statewide or regional summaries of sampling data for comparative purposes.

9.5.2 Stock Density Indices

The use of stock density indices in size structure assessment should be thought of as a complement, and not a replacement, to other methods of length-frequency analysis. Any size structure assessment should begin with a thorough inspection of length-frequency histograms, as they can provide detail that may be lost when length data are summarized in wide length categories or by an index. A benefit of calculating stock density indices is that the index values can be used to test correlations between size structure and other factors. An appropriate question concerning the use of stock density indices is whether the index value (i.e., size structure) reflects density and dynamics of fish populations (Willis et al. 1993). As the density of a population increases, PSD tends to decrease; declines in size structure can be attributed to slowing of growth and increased mortality as resources become scarce. However, a low PSD value may also occur at low population densities due to overharvest or poor habitat. Negative correlations between PSD and density, C/f, or biomass have been observed for many species, including largemouth bass (Reynolds and Babb 1978; Gabelhouse 1984a; Boxrucker 1987; Guy and Willis 1990; Saffel et al. 1990; Hill and Willis 1993), black crappie (RSD-P; Guy and Willis 1995), black bullhead (Brown et al. 1999), and brook trout (Johnson et al. 1992). Such negative correlations are more likely in small water bodies with simple fish communities.

As growth increases, there is a tendency for PSD to increase. Low density may result in fast growth, whereas high density may result in slow growth. Correlations between stock density indices and growth have been observed for largemouth bass (Miranda 1983; Jacobs and O'Donnell 1996), smallmouth bass (Jacobs and O'Donnell 1996), bluegill (Novinger and Legler 1978; Paukert and Willis 2000), northern pike (Willis and Scalet 1989), yellow perch (Willis et al. 1991; Paukert and Willis 2000; Paukert et al. 2002), and black crappie (Guy and Willis 1995; Paukert and Willis 2000; Paukert et al. 2002).

Several studies have demonstrated that body condition is positively correlated to growth rate (see Chapter 10). Individuals from low-density populations in which PSD is high tend to have high body condition values, and individuals from high-density populations in which PSD is low tend to have low body condition values. Positive correlations between PSD and W_r for species such as largemouth bass (Wege and Anderson 1978), white crappie and black crappie (Gabelhouse 1984a), northern pike (Willis and Scalet 1989), walleye (Murphy et al. 1990), sauger (Guy

et al. 1990), yellow perch (Willis et al. 1991), and brook trout (Johnson et al. 1992) have been observed. However, body condition is an instantaneous measure, and slow-growing fish may exhibit high body condition at times of the year when food is abundant or when gonads are mature during the spawning period.

As total annual mortality increases, there is a tendency for PSD to decrease. In situations in which recruitment is high, as in the Knox Pond example (Figure 9.6), mortality tends to be high and PSD tends to be low. High mortality due to overharvest and poor habitat also results in low PSD values. Negative correlations between PSD and mortality have been observed in largemouth bass (Reynolds and Babb 1978; Miranda 1983; Jacobs and O'Donnell 1996) and smallmouth bass (Jacobs and O'Donnell 1996).

Correlations between stock density indices and density or dynamic rate functions are often moderate in strength, and there is a wide variability in the strength of correlations observed among studies. One reason for this may be that stock density indices may lack sensitivity in some cases; two populations can have the same stock density index value and actually have different length-frequency distributions. Variations in factors such as productivity and growing season can affect establishment of a clear relationship between stock density indices and population parameters (Willis et al. 1993). Additionally, variability in PSD may be related to water body size. For example, largemouth bass in small impoundments may be more recruitment driven than recruitment limited. Jakes (1987) found that size structure of largemouth bass increased in three impoundments ranging in size from 9 to 1,100 ha. Stock density indices also provide more interpretive information when populations are relatively steady state, (i.e., when recruitment, growth, and mortality remain somewhat constant) (Willis et al. 1993). For example, PSD will provide little interpretive information for populations with highly variable recruitment. Willis et al. (1993) provided an example in which the PSD of a black crappie population increased from 3 in spring to 100 in fall. This was the result of a single cohort of black crappies growing over the course of one season. Allen and Pine (2000) found that PSD would often not change significantly in response to minimum length limits if recruitment was highly variable (e.g., coefficients of variation in recruits to age 1 that are greater than 70–90%).

Correlations between predator and prey stock density index values are listed in Table 9.3. Because largemouth bass is a common predator in ponds and small impoundments, most examples listed deal with largemouth bass as the predator, although several examples of prey are listed. In ponds and small impoundments, predator PSD tends to decline as predator density increases. As predator density increases, prey fish density decreases. Thus, prey PSD tends to increase as predator density increases, resulting in an inverse correlation between predator PSD and prey PSD (Willis et al. 1993).

The likelihood of an inverse relationship between predator PSD and prey PSD tends to decline in large water bodies. Carline et al. (1984) suggested that in Ohio impoundments, inverse relationships between size structure of largemouth bass and bluegills may not be expected in impoundments greater than 15 ha in size. In some instances, inverse relationships have been observed in impoundments larger

Table 9.3 Summary of correlation coefficients (*r*) between stock density indices of predator and prey species and other parameters. Parameters compared are proportional stock density (PSD); relative stock density of preferred-length fish (RSD-P); and catch-per-unit-effort (*C/f*).

Predator	Parameter	Prey	Parameter	*r*	Reference
Largemouth bass	PSD	Black bullhead	Mean length	–0.81	Saffel et al. (1990)
	C/f	Bluegill	PSD	0.71	Guy and Willis (1990)
	PSD		PSD	–0.83	Guy and Willis (1990)
	RSD-P		Growth	–0.64	Guy and Willis (1990)
	PSD		PSD	–0.49	Paukert and Willis (2000)
	C/f		PSD	0.52	Paukert and Willis (2000)
	PSD	Crappie[a]	PSD	–0.85	Gabelhouse (1984a)
	RSD-P		PSD	–0.84	Gabelhouse (1984a)
	PSD		*C/f*	0.73	Boxrucker (1987)
	RSD-P		*C/f*	0.88	Boxrucker (1987)
	C/f		PSD	0.56	Boxrucker (1987)
	PSD		PSD	–0.56	Boxrucker (1987)
	C/f	Yellow perch	PSD	0.81	Guy and Willis (1991)
	PSD		PSD	–0.82	Guy and Willis (1991)
	PSD		Growth	–0.95	Guy and Willis (1991)
	C/f		PSD	0.82	Paukert and Willis (2000)
Northern pike	*C/f*	Black bullhead	PSD	–0.54	Brown et al. (1999)

[a] Includes white crappie and black crappie.

than 15 ha (Gabelhouse 1984b; Boxrucker 1987; Guy and Willis 1991; Paukert and Willis 2000; Paukert et al. 2002).

Stock density indices are useful tools not only to report size structure but also to reflect density and population dynamics in certain situations. However, because of the variability in correlations and confounding factors, stock density indices should be used in association with other assessment tools to evaluate fish populations properly.

9.6 CONCLUSIONS

Factors influencing the accuracy and precision of size structure data such as gear selectivity and time of collection should be considered prior to data analysis and interpretation. Standardized sampling that allows for relative comparisons through time or across water bodies provides the most powerful inferences, and planning the study design prior to data collection is imperative. Traditionally, fisheries scientists pool individual fish data from multiple collection events (e.g., electrofishing runs or nets) to develop and test length-frequency histograms. In this chapter, we provided alternative analysis methods that consider the collection event as the experimental unit rather than individual fish. Using the collection event as the experimental unit has advantages because the analysis considers among-sample variation in size structure rather than among-individual-fish variation. Additionally, using individual fish as the experimental unit often causes the error degrees of freedom to be very high, resulting in significant differences when distributions

appear quite similar (e.g., Kolmogorov–Smirnov two-sample test). Thus, the use of collection events as the experimental unit results in a more conservative test of size structure differences, and we recommend these methods when possible. The fisheries scientist should understand the advantages of various statistical tests and match analysis methods as best as possible to design experiments properly. Size structure data can be analyzed as categorical (e.g., chi-square), proportional (e.g., PSD), or continuous (e.g., LOGIT) data, depending on the study design and sample size. Examples in this chapter provide guidance for comparisons across systems, through time, or both, depending on the study objectives. Experimental design and hypothesis testing methods for analyses of length-frequency data will continue to improve.

9.7 REFERENCES

Allen, M. S., and W. E. Pine, III. 2000. Detecting fish population responses to a minimum length limit: effects of variable recruitment and duration of evaluation. North American Journal of Fisheries Management 20:672–682.

Anderson, R. O., and R. M. Neumann. 1996. Length, weight, and associated structural indices. Pages 447–482 *in* B. R. Murphy and D. W. Willis editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Bayley, P. B., and D. J. Austen. 1990. Modeling the sampling efficiency of rotenone in impoundments and ponds. North American Journal of Fisheries Management 10:202–208.

Bettoli, P. W., and M. J. Maceina. 1996. Sampling with toxicants. Pages 303–333 *in* B. R. Murphy and D. W. Willis editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Bettross, E. A., and D. W. Willis. 1988. Seasonal patterns in sampling data for largemouth bass and bluegills in a northern Great Plains impoundment. Prairie Naturalist 20:193–202.

Bister, T. J., D. W. Willis, M. L. Brown, S. M. Jordan, R. M. Neumann, M. C. Quist, and C. S. Guy. 2000. Development of standard weight (*Ws*) equations and standard length categories for 18 warmwater nongame and riverine fish species. North American Journal of Fisheries Management 20:570–574.

Boxrucker, J. 1987. Largemouth bass influence on size structure of crappie populations in small Oklahoma impoundments. North American Journal of Fisheries Management 7:273–278.

Boxrucker, J., and G. Ploskey. 1989. Gear and seasonal biases associated with sampling crappies in Oklahoma. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 42(1988):89–97.

Brown, M. L., and D. J. Austen. 1996. Data management and statistical techniques. Pages 17–62 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition, American Fisheries Society, Bethesda, Maryland.

Brown, M. L., D. W. Willis, and B. G. Blackwell. 1999. Physicochemical and biological influences on black bullhead populations in eastern South Dakota glacial lakes. Journal of Freshwater Ecology 14:47–60.

Carline, R. F., B. L. Johnson, and T. J. Hall. 1984. Estimation and interpretation of proportional stock density for fish populations in Ohio impoundments. North American Journal of Fisheries Management 4:139–154.

Conover, W. J. 1980. Practical nonparametric statistics, 2nd edition. Wiley, New York.

Cornelius, R. R., and T. L. Margenau. 1999. Effects of length limits on muskellunge in Bone Lake, Wisconsin. North American Journal of Fisheries Management 19:300–308.

Ebbers, M. A. 1987. Vital statistics of a largemouth bass population in Minnesota from electrofishing and angler supplied data. North American Journal of Fisheries Management 7:252–259.

Gabelhouse, D. W., Jr. 1984a. An assessment of crappie stocks in small Midwestern private impoundments. North American Journal of Fisheries Management 4:371–384.

Gabelhouse, D. W., Jr. 1984b. An assessment of largemouth bass slot length limits in five Kansas lakes. Kansas Fish and Game Commission, Comprehensive Planning Option Project FW-9-P-3, Pratt.

Gabelhouse, D. W., Jr., and D. W. Willis. 1986. Biases and utility of angler catch data for assessing size structure and density of largemouth bass. North American Journal of Fisheries Management 6:481–489.

Gilliland, E. 1987. Evaluation of Oklahoma's standardized electrofishing in calculating population structure indices. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 39(1985):277–287.

Green, D. M., B. J. Schonhoff III, and W. D. Youngs. 1986. The New York state bass study 1977–1980: use of angler collected data to determine population dynamics. New York State Department of Environmental Conservation, Albany.

Griffith, J. S. 1981. Estimation of the age-frequency distribution of stream dwelling trout by underwater observation. Progressive Fish-Culturist 43:51–53.

Gustafson, K. A. 1988. Approximating confidence intervals for indices of fish population size structure. North American Journal of Fisheries Management 8:139–141.

Guthery, F. S., L. A. Brennan, M. J. Peterson, and J. L. Lusk. 2005. Information theory in wildlife science: critique and viewpoint. Journal of Wildlife Management 69:457-465.

Guy, C. S., E. A. Bettross, and D. W. Willis. 1990. A proposed standard weight (*Ws*) equation for sauger. Prairie Naturalist 22:41–48.

Guy, C. S., and D. W. Willis. 1990. Structural relations of largemouth bass and bluegill populations in South Dakota ponds. North American Journal of Fisheries Management 10:338–343.

Guy, C. S., and D. W. Willis. 1991. Evaluation of largemouth bass–yellow perch communities in small South Dakota impoundments. North American Journal of Fisheries Management 1143–49.

Guy, C. S., and D. W. Willis. 1995. Population characteristics of black crappies in South Dakota waters: a case for ecosystem specific management. North American Journal of Fisheries Management 15:754–765.

Hayne, D. W., G. E. Hall, and H. M. Nichols. 1967. An evaluation of cove sampling of fish populations in Douglas Reservoir, Tennessee. Pages 244–297 *in* Reservoir Committee, editors. Reservoir fishery resources symposium. American Fisheries Society, Southern Division, Reservoir Committee, Bethesda, Maryland.

Hill, T. D., and D. W. Willis. 1993. Largemouth bass biomass, density, and size structure in small South Dakota impoundments. South Dakota Academy of Science 72:31–39.

Holland, R. S., and E. J. Peters. 1992. Differential catch rates by hoop nets of three different mesh sizes in the lower Platte River. North American Journal of Fisheries Management 12:237–243.

Hubbard, W. D., and L. E. Miranda. 1988. Competence of non-random electrofishing sampling in assessment of structural indices. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 40(1986):79–84.

Hubert, W. A. 1996. Passive capture techniques. Pages 157–192 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Isermann, D. A., P. A. Hanchin, and D. W. Willis. 2002. Comparison of two mesh sizes for collecting larval yellow perch in surface trawls. North American Journal of Fisheries Management 22:585–589.

Jacobs, R. P., and E. B. O'Donnell. 1996. An electrofishing survey of selected Connecticut lakes. Connecticut Department of Environmental Protection, Fisheries Division, Federal Aid in Sport Fish Restoration Project F-57-R-14, Final Report, Hartford.

Jacobs, R. P., E. B. O'Donnell, and A. P. Petrillo. 1995. Assessment of experimental length limits on largemouth bass and Lake Saltonstall fisheries investigations. Connecticut Department of Environmental Protection, Fisheries Division, Federal Aid in Sport Fish Restoration Final Report F-57-R-12, Hartford.

Jakes, D. S. 1987. Population structures and prey exploitation of largemouth bass in three unexploited impoundments. Master's thesis. University of Georgia, Athens.

Johnson, S. L., F. J. Rahel, and W. A. Hubert. 1992. Factors influencing the size structure of brook trout in beaver ponds in Wyoming. North American Journal of Fisheries Management 12:118–124.

King, T. A., J. C. Williams, W. D. Davies, and W. L. Shelton. 1981. Fixed versus random sampling of fishes in a large reservoir. Transactions of the American Fisheries Society 110:563–568.

Koopmans, L. H. 1987. Introduction to contemporary statistical methods. PWS Publishers, Duxbury Press, Boston.

Laarman, P. W., and J. R. Ryckman. 1982. Relative size selectivity of trap nets for eight species of fish. North American Journal of Fisheries Management 2:33–37.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, North Carolina.

Lott, J. P., and D. W. Willis. 1991. Gill net mesh size efficiency for yellow perch. Prairie Naturalist 23:139–144.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot analysis of variance design for repeated-measures fishery data. Fisheries (Bethesda) 19(3):14–20.

Margenau, T. L., and S. P. AveLallemant. 2000. Effects of a 40-inch minimum length limit on muskellunge in Wisconsin. North American Journal of Fisheries Management 20:986–993.

Mero, S. W., and D. W. Willis. 1992. Seasonal variation in sampling data for walleye and sauger collected with gill nets from Lake Sakakawea, North Dakota. Prairie Naturalist 24:231–240.

Michaletz, P., J. Boxrucker, S. Hale, and J. R. Jackson. 1995. Comparison of four trawls for sampling juvenile shad. North American Journal of Fisheries Management 15:918–923.

Milewski, C. L., and D. W. Willis. 1991. Smallmouth bass size structure and catch rates in five South Dakota lakes as determined from two sampling gears. Prairie Naturalist 23:53–60.

Miranda, L. E. 1983. Relation between growth, mortality, and proportional stock density of largemouth bass in Texas reservoirs. Annual Proceedings of the Texas Chapter, American Fisheries Society 5:29–38.

Miranda, L. E. 1993. Sample sizes for estimating and comparing proportion-based indices. North American Journal of Fisheries Management 13:383–386.

Mullner, S. A., W. A. Hubert, and T. A. Wesche. 1998. Snorkeling as an alternative to depletion electrofishing for estimating abundance and length-class frequencies of trout in small streams. North American Journal of Fisheries Management 18:947–953.

Murphy, B. R., M. L. Brown, and T. A. Springer. 1990. Evaluation of the relative weight index with new applications to walleye. North American Journal of Fisheries Management 10:85–97.

Neal, J. W., R. L. Noble, and J. A. Rice. 1999. Fish community response to hybrid striped bass introduction in small warmwater impoundments. North American Journal of Fisheries Management 19:1044–1053.

Neumann, R. M., C. S. Guy, and D. W. Willis. 1995. Precision and size structure of juvenile Percichthyid samples collected with various gears from Lake Texoma. North American Journal of Fisheries Management 15:956–962.

Neumann, R. M., and D. W. Willis. 1995. Seasonal variation in gill-net sample indexes for northern pike collected from a glacial prairie lake. North American Journal of Fisheries Management 15:838–844.

Neumann, R. M., D. W. Willis, and D. D. Mann. 1994. Evaluation of largemouth bass slot length limits in two small South Dakota impoundments. Prairie Naturalist 26:15–32.

Novinger, G. D., and R. E. Legler. 1978. Bluegill population structure and dynamics. Pages 37–49 *in* G. D. Novinger and J. G. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, North Central Division, Special Publication 5, Bethesda, Maryland.

Paragamian, V. L. 1989. A comparison of day and night electrofishing: size structure and catch per unit effort for smallmouth bass. North American Journal of Fisheries Management 9:500–503.

Paukert, C. P., and D. W. Willis. 2000. Factors affecting panfish populations in Sandhill lakes. Nebraska Game and Parks Commission, Federal Aid in Sport Fish Restoration Project F-11-R, Completion Report, Lincoln.

Paukert, C. P., D. W. Willis, and J. A. Klammer. 2002. Effects of predation and environment on quality of yellow perch and bluegill populations in Nebraska Sandhill lakes. North American Journal of Fisheries Management 22:86–95.

Pope, K. L., and D. W. Willis. 1996. Seasonal influences on freshwater fisheries sampling data. Reviews in Fisheries Science 4:57–73.

Reynolds, J. B., and L. R. Babb. 1978. Structure and dynamics of largemouth bass populations. Pages 50–61 *in* G. D. Novinger and J. G. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, North Central Division, Special Publication 5, Bethesda, Maryland.

Reynolds, J. B., and D. E. Simpson. 1978. Evaluation of fish sampling methods and rotenone census. Pages 11–24 *in* G. D. Novinger and J. G. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, North Central Division, Special Publication 5, Bethesda, Maryland.

Robinson, J. W. 1994. Sampling procedures to capture flathead catfish and channel catfish. Missouri Department of Conservation, Federal Aid in Sport Fish Restoration Project F-1-R-43, Final Report, Columbia.

Roni, P., and A. Fayram. 2000. Estimating winter salmonid abundance in small western Washington streams: a comparison of three techniques. North American Journal of Fisheries Management 20:683–692.

Saffel, P. D., C. S. Guy, and D. W. Willis. 1990. Population structure of largemouth bass and black bullheads in South Dakota ponds. Prairie Naturalist 22:113–118.

Santucci, V. J., Jr., D. H. Wahl, and D. F. Clapp. 1999. Efficiency and selectivity of sampling methods used to collect channel catfish in impoundments. Pages 317–328 *in* E. R. Irwin, W. A. Hubert, C. F. Rabeni, H. L. Schramm, Jr., and T. Coon, editors. Catfish 2000: Proceedings of the international ictalurid symposium. American Fisheries Society, Symposium 24, Bethesda, Maryland.

SAS. 1999. SAS statistics user's guide. SAS Institute, Cary, North Carolina.

Tate, W. B., M. S. Allen, R. A. Meyers, and J. R. Estes. 2003. Comparison of electrofishing and rotenone for sampling largemouth bass in vegetated areas of two Florida lakes. North American Journal of Fisheries Management 23:181–188.

Trippel, E. A., and J. J. Hubert. 1990. Common statistical errors in fishery research. Pages 93–102 *in* J. Hunter, editor. Writing for fishery journals. American Fisheries Society, Bethesda, Maryland.

Underwood, T. J. 2000. Abundance, length composition, and migration of spawning inconnu in the Selawik River, Alaska. North American Journal of Fisheries Management 20:386–393.

Unmuth, J. M. L., M. J. Hansen, P. W. Rasmussen, and T. D. Pellett. 2001. Effects of mechanical harvesting of Eurasian water milfoil on angling for bluegills in Fish Lake, Wisconsin. North American Journal of Fisheries Management 21:448–454.

Van Den Avyle, M. J., G. R. Ploskey, and P. W. Bettoli. 1995. Evaluation of gill-net sampling for estimating abundance and length frequency of reservoir shad populations. North American Journal of Fisheries Management 15:898–917.

Van Zee, B. E., D. W. Willis, and C. C. Stone. 1996. Comparison of diel sampling data for sauger collected by electrofishing. Journal of Freshwater Ecology 11:139–143.

Vokoun, J. C., and C. F. Rabeni. 1999. Catfish sampling in rivers and streams: a review of strategies, gears, and methods. Pages 271–286 *in* E. R. Irwin, W. A. Hubert, C. F. Rabeni, H. L. Schramm, Jr., and T. Coon, editors. Catfish 2000: proceedings of the international ictalurid symposium. American Fisheries Society, Symposium 24, Bethesda, Maryland.

Vokoun, J. C., C. F. Rabeni, and J. S. Stanovick. 2001. Sample-size requirements for evaluating population size structure. North American Journal of Fisheries Management 21:660–665.

Wege, G. J., and R. O. Anderson. 1978. Relative weight (W_r): a new index of condition for largemouth bass. Pages 79–91 *in* G. D. Novinger and J. G. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, North Central Division, Special Publication 5, Bethesda, Maryland.

Weithman, A. S., J. B. Reynolds, and D. E. Simpson. 1980. Assessment of structure of largemouth bass stocks by sequential sampling. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 33(1979):415–424.

Wildman, T. L., and R. M. Neumann. 2003. Comparison of snorkeling and electrofishing for estimating abundance and size structure of brook trout and brown trout in two southern New England streams. Fisheries Research 60:131–139.

Willis, D. W., C. S. Guy, and B. R. Murphy. 1991. Development and evaluation of a proposed standard weight (*Ws*) equation for yellow perch. North American Journal of Fisheries Management 11:374–380.

Willis, D. W., and B. R. Murphy. 1996. Planning for sampling. Pages 1–16 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Willis, D. W., B. R. Murphy, and C. S. Guy. 1993. Stock density indices: development, use, and limitations. Reviews in Fisheries Science 1:203–222.

Willis, D. W., and C. G. Scalet. 1989. Relations between proportional stock density and growth and condition of northern pike populations. North American Journal of Fisheries Management 9:488–492.

Zar, J. H. 1996. Biostatistical analysis, 3rd edition. Prentice-Hall, Upper Saddle River, New Jersey.

10 Condition

Kevin L. Pope and Carter G. Kruse

10.1 INTRODUCTION

The analysis of fish condition has become a standard practice in the management of fish populations as a measure of both individual and cohort (e.g., age- or size-group) wellness. Condition has been generically described as the well-being or robustness of an individual fish (Le Cren 1951; Bulow et al. 1981; Blackwell et al. 2000). It has typically been estimated by comparing an individual fish weight to a standard weight for a given length and assuming that larger ratios (condition index) reflect a healthier physiological state (Bolger and Connolly 1989; Murphy et al. 1991) or by directly measuring physiological parameters related to the energy stores, such as tissue lipid content (Craig 1977; Fechhelm et al. 1995). All methods of calculating condition share the common goal of controlling for or removing the confounding effects of absolute body size when comparing body mass or other measures of nutritional state (Jakob et al. 1996). This is particularly important for organisms with indeterminate growth, such as fishes (Reist 1985).

Measures of condition are generally intended to be an indicator of tissue energy reserves, with the expectation that a fish in good condition should demonstrate faster growth rates, greater reproductive potential, and higher survival than will a lesser-conditioned counterpart, given comparable environmental conditions. Subsequently, fish condition is of keen interest to fisheries scientists, and numerous studies have investigated the relationship between measures of fish condition and parameters such as growth, fecundity, population structure, life history adaptations, environmental conditions, or management actions such as stocking (Cone 1989; Brown and Murphy 1991; Gabelhouse 1991; Blackwell et al. 2000). Although measures of condition in fish can be sensitive or related to factors that might logically affect energy storage or fitness in an individual, there is commonly substantial interspecies, seasonal, environmental, and spatial variation that influences our ability to interpret changes in fish condition.

Fisheries scientists often must assess population status, effects of management actions, and anthropogenic influences on the resource they are managing (Brown and Austin 1996). Fish condition, if appropriately interpreted, may characterize components of the environment in which the fish exists (e.g., habitat,

prey availability, and competition) and provide insights into ecological and physiological processes (e.g., overwintering mortality, seasonal storage of lipids, and maturation). Thus, measures or indices of fish condition can be valuable components of a fisheries scientist's assessment over multiple ecological scales. A critical component for interpreting fish condition data in a useful and applicable manner is the correct application of statistical methodologies when collecting and analyzing data. The objective of this chapter is to provide a brief overview of fish condition measures, focusing on condition indices, and illustrate commonly used techniques to analyze, summarize, and interpret condition data.

10.2 WEIGHT–LENGTH RELATIONSHIPS

Anderson and Neumann (1996) noted that length and weight statistics are cornerstones in the foundation of fisheries management and research. Weight–length data have generally been used either to describe mathematically the relationship between weight and length (Keys 1928) for purposes of conversion from one to the other or to measure individual variation from an expected weight at a given length as an indicator of condition (Le Cren 1951; Bolger and Connolly 1989). It is often advantageous to describe the weight–length relationship of a population to discern changes in body form. The power function,

$$W = aL^b, \tag{10.1}$$

generally describes the weight–length relationship of most fishes, where W is weight, L is length, a is a constant, and b is an exponent usually between 2.5 and 4.0 (a fish growing isometrically or maintaining the same shape across length categories has an exponent of 3.0). The functional exponent b, which describes the curve of the relationship, is generally different among species and can be sensitive to biotic and abiotic influences, leading to different values of b between sexes or localities, even within the same species.

10.2.1 Regression of Weight –Length Data

Because body form typically changes with increasing length (i.e., allometric growth; $b \neq 3.0$), untransformed weight–length data are related in a curvilinear fashion (Figure 10.1A). Although a curve can be fitted to the weight–length relationship for estimation of the power function coefficients (nonlinear regression), these types of data are more easily analyzed by linear regression after logarithmically transforming the data (Figure 10.1B). Based on the ordinary least-squares regression model ($y_i = \beta_0 + \beta_1 x_i + \varepsilon$), equation (10.1) becomes

$$\log_{10}(W) = a + b(\log_{10}L), \tag{10.2}$$

where W (corresponding to the response or dependent y_i) and L (independent x_i) are weight and length, respectively, a (β_0) is the y-intercept ($\log_{10}$ scaling), and

b (β_1) the slope of the line. The error (ε) associated with estimating y_i (W) from a regression line is implicit in equation (10.2).

The regression assumptions of linearity, normality, homoscedasticity (equal variance of y at each level of x), and independence (no changes in y at a given x due to an influence such as sampling over time) must be met for meaningful interpretation of the regression coefficients (Neter et al. 1989). If a population (i.e., group or cohort of interest) is randomly sampled over a relatively short period, logarithmically transformed weight–length data generally conform to the basic assumptions and are related in a highly significant linear fashion. Biases can be introduced into weight–length data by, among other things, introducing measurement error, combining temporally or spatially separated samples for which physiological or environmental changes may have affected body form (e.g., pre- and postspawn or lotic and lentic individuals), or by incompletely and nonrandomly sampling the entire size structure of the population (e.g., presence or absence of a resource-limited size category). Suspected transgression of the linearity, variance, and independence assumptions can be initially assessed with residual analyses, where residuals (the difference between the observed weight and the corresponding weight predicted by the regression line) or the error associated with using the regression model are plotted against the independent variable (length) or the predicted value of y. Graphically, residuals should appear as a constant band around zero, with no obvious patterns (Figure 10.1C, D, E, and F). Most statistical packages will provide an option for these analyses. The transformed weight–length data generally approximate a normal distribution and small departures from normality do not create serious problems; however, data normality should not be assumed, especially when using the regression coefficients as indices of population condition or the residuals as an index to individual condition. A normal probability plot is a general test to ensure normality of the data (Figure 10.1G).

A linear relation can be a reasonably good approximation for nonlinear data provided the values of the independent variable do not cover a wide range (Steel and Torrie 1980), such as comparisons of individuals in a relatively narrow subset of all lengths sampled (e.g., a small section of the curve). Furthermore, simple linear regression often statistically provides an adequate fit to untransformed weight–length data when assessing statistics such as r^2; however, better results can be obtained with transformation or nonlinear analysis. Thus, it is unadvisable to fit a linear model to curvilinear data. The logarithmic transformation enhances the relationship by accounting for more of the variation in weight (demonstrated by an increased r^2) and minimizing overall model error, or the distance of individual points from the regression line. The logarithmic transformation enhances our ability to predict weight from length and to interpret the slope and intercept of the relationship. A power function (nonlinear regression or curve fitting) of the untransformed data provides the same explanatory power as linear regression of the transformed variables; however, the exponential nature of the relationship makes interpretation and comparison of weight–length relationships more difficult (Box 10.1).

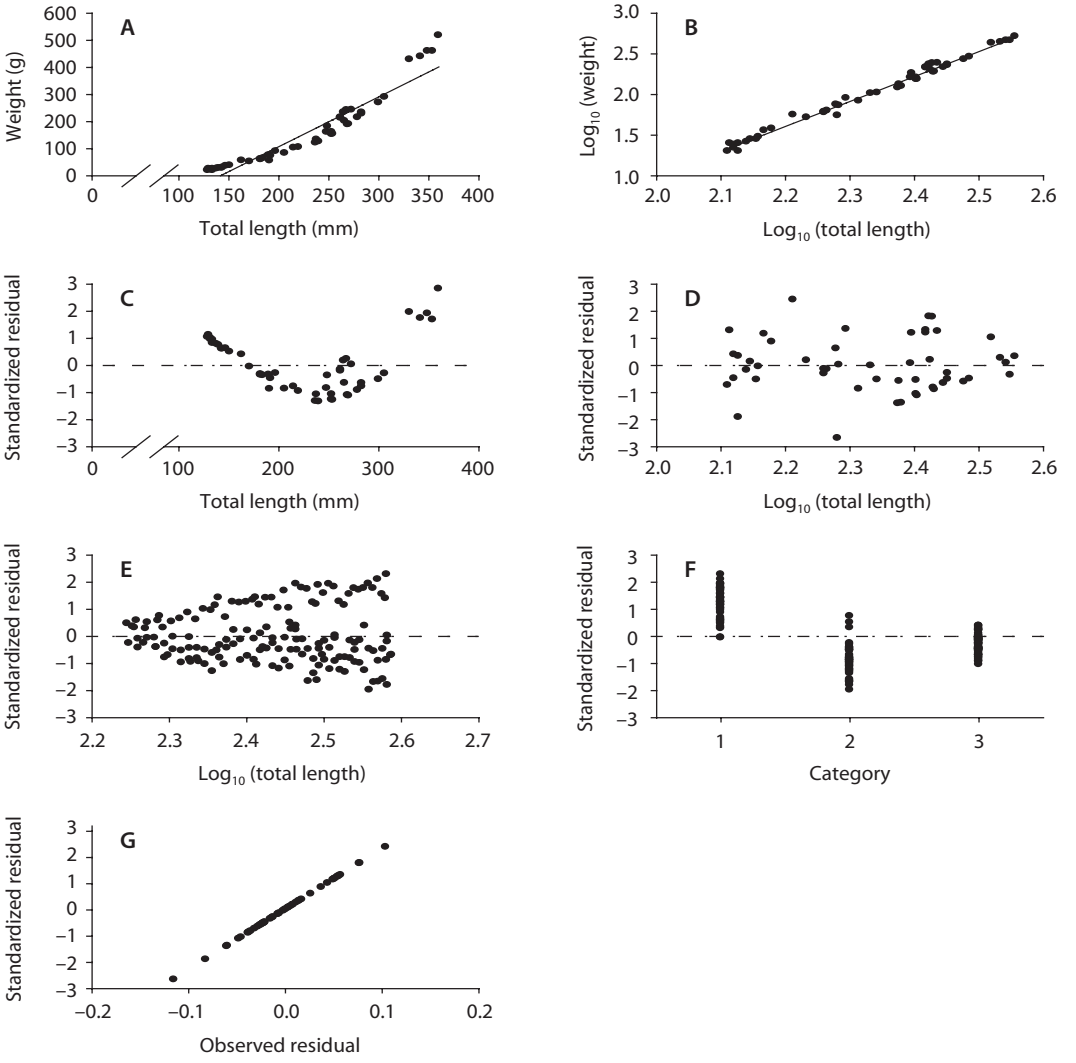


Figure 10.1 Graphical depiction of the curvilinear relationship of (A) untransformed length–weight data from the low-elevation stream Yellowstone cutthroat trout population described in Box 10.1 versus (B) the linear nature of the same data after $\log_{10}$ transformation. (C) A typical diagnostic residual plot clearly illustrates the nonlinearity of the untransformed data, whereas (D) more evenly distributed residuals exist for the transformed data, a pattern that is indicative of linear, homoscedastic, and independent data. (E) The funnel-shaped residual pattern from a separate data set demonstrates unequal variances in the dependent variable (weight), as might be typical when sexually mature fish are collected in pre- and postspawning condition. (F) The up and down pattern of residuals when graphed by sampling time indicate that the data may not be independent but rather influenced by season (1 = prespawn, 2 = postspawn, and 3 = late summer). Normal probability plots can be built or graphed in several ways; here, (G) a normal probability plot of the ranked observed residuals (*x*) versus their paired standardized residual (*y;* calculated assuming a normal distribution) demonstrates the linear relationship indicative of normal weight–length data. A nonlinear relationship would indicate nonnormality or skewness of the data. Other plots, such as a box-plot, can also be used to check data normality.

Box 10.1 Transformation and Regression Analyses of Weight–Length Data—Comparing the Condition of Two Populations

Table Presented are total length (TL; mm), weight (WT; g), and body fat as a percentage of overall wet weight for samples of Yellowstone cutthroat trout collected in midsummer from three locations that could influence individual weight at length: a lower-elevation stream (1,810 m elevation), a lower-elevation lake (1,785 m), and a higher-elevation lake (2,610 m). Fat values were randomly generated for example only. Fish samples were collected via electroshocking, gill nets, and angling.

Low-elevation stream (A)			Low-elevation lake (B)			High-elevation lake (C)		
TL	WT	Fat (%)	TL	WT	Fat (%)	TL	WT	Fat (%)
129	20	5.91	254	181	10.59	180	63	8.32
130	25	12.88	262	186	9.08	191	77	4.78
132	22	7.67	272	136	1.38	198	54	0.6
132	24	11.29	274	191	2.64	203	73	2.63
134	20	2.27	282	245	7.32	231	100	0.99
134	25	11.57	287	236	6.68	234	104	1.21
138	26	6.63	290	168	1.39	236	109	5.91
140	28	11.04	297	263	6.81	239	118	2.55
143	28	9.45	302	290	9.1	239	127	5.37
144	30	9.86	305	290	7.97	241	127	2.87
147	36	8.79	328	327	5.9	244	141	4.97
151	38	13.78	330	354	2.47	244	141	5.51
163	56	11.69	333	363	5.16	244	154	8.21
171	52	6.46	333	390	9.29	246	141	4.37
182	60	8.83	338	372	3.64	246	150	6.27
182	61	8.93	340	417	12.98	246	168	8.44
184	63	8.97	340	417	9.92	249	145	2.34
190	75	11.71	343	399	5.36	249	145	4.51
191	55	4.73	345	408	5.39	249	154	6.44
192	73	7.35	345	445	10.37	249	159	7.88
197	90	12.88	351	463	10.24	251	145	2.99
206	83	4.52	353	390	5.78	251	145	3.04
215	103	9.1	353	390	1.75	251	145	3.57
220	105	6.41	356	467	8.51	254	150	7.98
237	121	3.06	356	472	8.66	254	150	3.64
238	133	6.52	358	408	1.87	254	154	4.63
240	126	3.42	361	545	1.99	254	159	2.95
248	161	9.81	361	472	9.61	254	163	5.94
249	182	12.24	363	481	6.19	257	154	3.73
253	153	7.34	363	508	8.87	257	159	3.11
253	161	11.11	366	481	4.62	257	168	5.98
254	154	4.25	368	476	2.69	257	172	7.38
262	213	12.33	373	544	11.45	257	172	7.83
262	215	12.53	378	526	4.29	259	154	1.23
265	234	12.9	381	535	3.12	259	159	3.93
266	202	10.44	381	535	3.37	259	159	4
268	242	13.04	384	608	9.99	259	172	6.19
269	189	4.27	386	572	5.84	259	181	7.9
270	190	4.29	389	562	7.05	262	181	7.25
273	243	12.66	391	590	4.86	262	181	6.31
279	215	9.63	394	635	3.9	264	159	1.87

(Box continues)

Box 10.1 *(continued)*

Table (*continued*)

Low-elevation stream (A)			Low-elevation lake (B)			High-elevation lake (C)		
TL	WT	Fat (%)	TL	WT	Fat (%)	TL	WT	Fat (%)
283	228	6.83	396	590	2.48	267	191	6.74
283	233	8.62	399	581	2.43	269	136	3.45
300	270	5.45	401	603	2.43	272	163	0.73
306	290	7.62	406	703	7.24	282	195	2.28
331	429	11.87	411	703	6.36	284	231	9.7
342	440	10.71	414	676	2.51	284	245	10.89
349	460	9.82	425	752	4.33	290	231	4.8
354	460	7.63	433	780	3.05	290	231	4.81
360	518	10.77	462	1170	9.99	290	240	5.74

Program

The following SAS program is configured to provide two regression analyses—linear regression of the weight–length data after $\log_{10}$ transformation on sample data from both the low-elevation stream (A) and lake (B) Yellowstone cutthroat trout populations and nonlinear regression of the untransformed data from the stream (A) sample. Only output relevant to the following discussion is provided. Hereinafter, all references to weight–length data transformations refer to a $\log_{10}$ transformation.

```
OPTIONS PS=54 LS=75;
DATA TROUT;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
CARDS;
A   129   20
B   254   181
[Input complete data set];
PROC SORT; BY POP;
PROC REG; BY POP; MODEL LOGWT=LOGTL/CLB;
PROC NLIN; BY POP; PARMS A=0.000001 B=3; MODEL WT=A*(TL**B);
RUN;
```

Output

Table Linear regression of transformed weight–length data for population A. The dependent variable is $\log_{10}$WT (LOGWT). Abbreviations are sum of squares (SS), coefficient of variation (CV), mean square error (MSE), and $\log_{10}$TL (LOGTL).

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	8.54267	8.54267	4545.72	<0.0001
Error	48	0.09021	0.00188		
Corrected total	49	8.63287			
r^2	0.9896	Root MSE	0.04335		
Adjusted r^2	0.9893	Dependent mean	2.01073		
CV	2.15597				

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	$P > \|t\|$
Intercept	1	–5.14432	0.10630	–48.39	<0.0001
LOGTL	1	3.06874	0.04552	67.42	<0.0001

Variable	*df*	95% Confidence limits	
Intercept	1	–5.35805	–4.93059
LOGTL	1	2.97722	3.16025

Table Nonlinear regression of weight–length data for population A. The dependent variable is WT, which is modeled as a constant (A) times TL raised to a power (B). The convergence criterion was met. An intercept was not specified for this model.

Iteration	A	B	SS
0	$1 \cdot 10^{-6}$	3.0000	1656903
1	$1.832 \cdot 10^{-6}$	3.2801	58540.3
2	$2.141 \cdot 10^{-6}$	3.2545	53615.7
3	$2.686 \cdot 10^{-6}$	3.2164	49075.3
4	$3.413 \cdot 10^{-6}$	3.1775	41355.6
5	$4.911 \cdot 10^{-6}$	3.1183	31198.0
6	$5.801 \cdot 10^{-6}$	3.1049	11611.2
7	$5.868 \cdot 10^{-6}$	3.1056	11159.7
8	$5.869 \cdot 10^{-6}$	3.1056	11159.7

Estimation Summary

Method	Gauss-Newton
Iterations	8
Subiterations	7
Average Subiterations	0.875
R	$2.223 \cdot 10^{-7}$
PPC(A)	$6.674 \cdot 10^{-8}$
RPC(A)	0.000045
Object	$5.848 \cdot 10^{-8}$
Objective	11159.7
Observations Read	50
Observations Used	50
Observations Missing	0

Regression Model

Source	*df*	SS	Mean square	*F*-value	Approximate $P > F$
Regression	2	2001114	1000557	4303.59	<0.0001
Residual	48	11159.7	232.5		
Uncorrected total	50	2012274			
Corrected Total	49	838762			

(Box continues)

Box 10.1 *(continued)*

Table (*continued*)

Parameter Estimates				
Variable	*df*	Approximate SE	Approximate 95% confidence limits	
A	$5.869 \cdot 10^{-6}$	$2.215 \cdot 10^{-6}$	$1.416 \cdot 10^{-6}$	0.000010
B	3.1056	0.0659	2.9731	3.2381

Approximate Correlation Matrix		
	A	B
A	1.0000000	–0.9995920
B	–0.9995920	1.0000000

Table Linear regression of transformed weight–length data for population B. The dependent variable is $\log_{10}$WT.

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	$P > F$
Model	1	1.66271	1.66271	914.86	<0.0001
Error	48	0.08724	0.00182		
Corrected total	49	1.74995			
r^2	0.9501	Root MSE	0.04263		
Adjusted r^2	0.9491	Dependent mean	2.63145		
CV	1.62008				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	$P > \|t\|$
Intercept	1	–5.36936	0.26459	–20.29	<0.0001
LOGTL	1	3.14307	0.10391	30.25	<0.0001

Variable	*df*	95% confidence limits	
Intercept	1	–5.90135	–4.83737
LOGTL	1	2.93413	3.35200

Interpretation

Regression of the transformed weight–length data from sample (A) shows a highly significant relationship ($P < 0.0001$) that explains 99% of the variation in weight (r^2). Regression of the transformed data provides a more precise estimation of fish weight than can be obtained by linear regression of the untransformed data and is a useful tool for inferring changes in overall condition (weight) temporally within or spatially across populations. Often a linear equation fitted to the entire range of untransformed data predicts that a fish must be of substantial size before the

weight exceeds zero, overestimates weights for mid-length fish, and underestimates the weight of larger fish—thus, the equation is not biologically relevant. Here, the transformed equation ($\log_{10}WT = -5.144 + 3.069\log_{10}TL$) estimates that individuals incrementally gain mass once they exceed 1 mm in length and demonstrates a strong linear relationship between weight and length.

The nonlinear regression of the untransformed data (equation [10.1]) provides the power function $WT = 0.000005869(TL)^{3.1056}$ and predicts weights very similar to the linear regression (equation [10.2]) based on the transformed data (e.g., a 300-mm cutthroat trout is predicted to weigh 287 g with the transformed equation and 289 g with the power function). In fact, equations (10.1) and (10.2) are the exact same model ($b_{\text{equation [10.1]}} = b_{\text{equation [10.2]}}$ and $a_{\text{equation [10.1]}} = 10^{a\ \text{equation [10.2]}}$). In our example, the coefficients have slightly different values and they provide slightly different predictions because the power function assumed homoscedastic error variances when, in reality, the larger fish had more variance in weight than did the smaller ones. Even with these similarities, interpretation and comparison among populations based on nonlinear regression is intuitively and statistically more difficult, and transformation to a linear equation is the preferred approach.

Often a primary question is whether differences in condition exist between or within specific populations or groups of fish across space and time. Comparisons of the regression coefficients associated with a given set of weight–length data can be used to determine whether a population (or group) of fish is significantly heavier and, by extension, in better condition at a given length. An interesting comparison might be whether Yellowstone cutthroat trout from a lake habitat are better conditioned than are stream-dwelling individuals found at similar elevations because a lake environment could be perceived as energetically favorable (e.g., no current or warmer). The estimated slope and intercept for the sample of transformed weight–length data from the stream population (A) are, respectively, 3.069 and –5.144 compared with 3.143 and –5.369 for the lake population (B). These equations suggest that average fish of 250 mm and 450 mm in length would weigh 164 g and 997 g, respectively, in the stream environment and 147 g and 933 g, respectively, in the lake environment. The regression results seem to indicate that stream fish are heavier at a given length than are their lake counterparts, at least in the sampled locations. However, in order to make meaningful statements regarding this relationship, we need to determine whether these populations are significantly different, given natural variation in weight at length.

Confidence intervals (CIs) around the estimated parameter (slope in this example) can be used as an initial assessment of differences in condition, if any, between populations. Using equation (10.2) (or the values provided by the SAS output) one can calculate the CIs around the parameter estimates. For example, the 95% CI around the estimated slopes (the actual parameter estimate is parenthetically enclosed) are 2.977–(3.069)–3.160 for stream fish and 2.934–(3.143)–3.352 for the lake population. These CIs overlap almost completely, and at least one interval encompasses the slope estimate of the other (in this case both intervals encompass the other slope estimate—the slope of the stream fish falls within the CI for the lake fish, and vice versa), indicating that the slopes are not significantly different, or that weight gain as the fish grows (body form) is similar between these two sites. Similar analyses show that the intercepts of these two populations are not significantly different. Thus we conclude that although the respective transformed equations predict different average weights, neither population is significantly heavier or better conditioned than is the other, contrary to our a prior expectation. If the CIs for the two slope estimates had not overlapped, it would have been an indication that the two values were indeed significantly different.

(Box continues)

Box 10.1 *(continued)*

A clearer comparison that provides a relevant level of precision, and one that must be performed if intervals overlap but neither one encloses the slope estimate of the other, is a CI around the difference between the slopes. For example suppose there are two populations with slopes 3.6 (SE = 0.2) and 3.0 (SE = 0.15). Based on a sample of 62 (60 df) individuals from each and a 95% confidence level, intervals for these slopes overlap but do not encompass the other slope estimate. The SE of the difference between these two slope values is

$$\sqrt{SE_1^2 + SE_2^2} = \sqrt{0.2^2 + 0.15^2} = 0.25,$$

and the CI for the difference is $0.6 \pm 1.98(0.25)$, where 0.6 is the difference between the slopes of the two populations and 1.98 is the *t*-value for an $\alpha = 0.05$ with 120 df (equation [10.4]). This interval does not include zero, which is indicative of significantly different slopes. Completing this calculation for the Yellowstone cutthroat trout example above reveals that the CI around the differences in slopes ($0.074 \pm 2.013[0.113]$, based on 48 df) includes zero, which indicates that the slopes are not different (as was previously concluded).

The least-squares regression coefficients estimated from the log-transformed data can be used to compare relative condition differences among populations or to assess temporal changes in condition within a population (Cone 1989). Bolger and Connolly (1989) indicated that the regression coefficients can suggest significant differences among populations but that estimates of intercept and slope should be considered together to provide a valid interpretation. If the regression slopes of two populations are similar, a larger intercept could indicate a population in better overall condition, or at least heavier fish at a given length. Likewise, a steeper slope would indicate increasingly (with length) better condition if population intercepts were similar. Intersecting regression lines (one population having a greater slope but lesser intercept than another) could indicate general differences in condition among small and large individuals. Carlander (1969) suggested that slopes less than 3.0 might indicate populations in crowded or stunted condition. However, Murphy et al. (1991) cautioned that coefficient analysis should be used to compare only the general form of specific populations because it tends to average out differences in condition between size-classes, an important component of condition analysis if, for example, a fisheries scientist were assessing the effect of prey abundance on different size-classes of fish (e.g., Marwitz and Hubert 1997).

Differences in weight–length regression lines can be cursorily assessed by comparing confidence interval (CI) overlap around the coefficients generated by the regression analysis. However, more precise statistical contrast includes determining the CI around the difference between two like coefficients or conducting analysis of covariance (ANCOVA). A CI, or the range of values within which the

regression coefficient is likely to fall over $1 - \alpha$ percent of all samples from the population of interest, is calculated by

$$g \pm t_{(1-\alpha/2;\, n-2)} \bullet s(g), \tag{10.3}$$

where g is the coefficient estimate, $s(g)$ is the standard error of g, and t is the t-value for a given confidence level (α) and df ($n - 2$). A quick assessment, with no relevant statistical precision, is to calculate interval overlap. If CIs around two linear regression slopes (or some other coefficient) developed from independent samples do not overlap, then they are significantly different. Furthermore, if the CI from one slope encloses the estimated value of the other then the two are not significantly different. A comparison that does provide statistical precision, and one that is required if CIs overlap (but neither encloses the estimated slope of the other), is determination of whether the interval around the difference in slopes contains zero; if so, the difference between the two estimated coefficients is statistically nonsignificant. The interval around the difference in slopes is given by

$$(RS_1 - RS_2) \pm t_{(1-\alpha/2;\, df)} \bullet \sqrt{SE_1^2 + SE_2^2}\,, \tag{10.4}$$

where RS_1 is the first regression slope, RS_2 is the second regression slope, SE_1 is the standard error of the first slope, and SE_2 is the standard error of the second slope. The df is equal to the sum of $(n_1 - 2)$ and $(n_2 - 2)$. Box 10.1 compares CIs of regression coefficients from samples of lake and stream Yellowstone cutthroat trout populations, where one might expect differences in population condition resulting from environmental influences. Interval analyses are relatively simplistic tests of regression line differences but are adequate for contrasting samples where sample size *(n)* and distribution (length categories of individuals captured) are similar, especially if the latter approach of testing the difference between coefficients is employed, or for preliminary analyses for general discussion purposes (interval overlap comparison). When the size-ranges of fish captured become uneven (e.g., larger fish in one sample but not the other), a test such as ANCOVA (section 10.2.2) that controls for size differences across time or habitat is more appropriate.

10.2.2 Analysis of Covariance to Test Differences in Regression Lines

Comparisons of weight-at-length (condition) data across multiple populations are often an important consideration, but the length range of individuals sampled often varies in space and time, and different-sized groups of fish may be in better or worse condition. The ANCOVA can control for the effects of differing size ranges (length as the covariate) and is a more powerful test for homogeneity of regression coefficients (i.e., test for differences in slopes between two or more lines with the null hypothesis that coefficients are equal; Zar 1984) where spatial (e.g., elevation) or temporal (e.g., season) effects might influence inferences regarding population wellness, as modeled by weight. Simply because the length variable is

not statistically significantly different between or among the populations of interest using a means comparison test (*t*-test or analysis of variance [ANOVA]) does not mean length will not confound a comparison of population condition. Rather it is the strength of the covariates' association to both the treatment and response variables together that determines the covariates' influence on our inference regarding condition. On the other hand, ANCOVA should be used with caution when length distributions are completely disparate, as interpretation of the results may become more speculative than meaningful (Agresti and Finlay 1986).

The general assumptions of ANCOVA when applied to weight–length data are (1) that length measurements are fixed, measured without error, and independent of treatments; (2) the regression of weight on length disregarding the treatment is linear (linearity of within-group regressions); (3) there is homogeneity of within-group regressions, and (4) the residuals are normally and independently distributed with zero mean and common variance. The ANCOVA is an inappropriate tool when heterogeneity of regression coefficients and residual variances exists. Assumption two is regularly achieved by some sort of data transformation. Similarly, weight is typically normally distributed and, furthermore, data transformation has a normalizing effect. Assumption three requires that the regression lines associated with the treatment groups have a common slope or parallelism; slope discrepancies will result in a conservative ANCOVA *F*-test, for which the likelihood of type I error (rejecting a true null hypothesis) is actually lower than the nominal alpha. Heterogeneity of error variances is of most concern when sample sizes among groups differ and will result in a conservative *F*-test if the larger and smaller samples sizes are associated with the larger and smaller variances, respectively. If the opposite is true, then the test becomes liberal (i.e., the true alpha is greater than then the nominal alpha) (Vila-Gispert and Moreno-Amich 2001).

We initially want to determine slope similarity. Building on equation (10.2), the complete ANCOVA model contains the response variable (weight, W), an intercept (β_0), two independent variables, the covariate (length, L) and a dummy variable that represents potential effects on weight that are of interest (X; for example, habitat effects are coded 1 for low-elevation stream and 0 for low-elevation lake), and an interaction term (length × habitat code, L_iX_i) in the form

$$W_i = \beta_0 + \beta_1 L_i + \beta_2 X_i + \beta_3 L_i X_i + \varepsilon_i , \tag{10.5}$$

where weight and length are $\log_{10}$ transformed. The relationship can be modeled using a general linear model (GLM) approach or using regression. If the two slopes differ, the interaction term will be significant in the model, indicating that the regression lines intersect at some point (note that point may be outside the range of data collected) and the trend lines are different. This type of result suggests that individual fish in the two populations gain weight at different rates as they increase in length and may indicate, among other things, resource limitations (or availability) for different size categories within (temporal comparisons) or

between (spatial comparisons) populations. If the slopes are statistically different (i.e., we know the lines are different), further testing of intercept differences is difficult to interpret and often of little interest because magnitude of treatment effect varies depending on length and the intercept of a weight–length relationship (length = 0) is generally not relevant.

If fish from two populations maintain similar incremental weight gains with increasing length, then the slopes will not be significantly different; however, one population could be significantly heavier or better conditioned at a given length than another. Thus, we generally want to determine the magnitude of the elevational difference between the lines by assessing the *y*-intercepts. In other words, are the lines truly the same or are they separated in regression space with similar slopes? Here equation (10.5) is reduced to the form

$$W_i = \beta_0 + \beta_1 L_i + \beta_2 X_i + \varepsilon_i \tag{10.6}$$

by removing the interaction term from the analysis. Separate lines or intercept differences are noted by a significant test of the dummy variable (X) in the model.

In its simplest form, ANCOVA is used, as described above, to control for length differences between two populations or categories of treatment (e.g., a habitat treatment of lotic and lentic environments); however, it can be used to assess multiple populations and multiple treatments by simply adding additional dummy variables and the associated interaction terms to equation (10.4). In Box 10.2, we provide an example of ANCOVA based on the two populations of Yellowstone cutthroat trout analyzed in Box 10.1. Both the CI comparisons in Box 10.1 and ANCOVA in Box 10.2 provide results that indicate the slopes of the two lines are not significantly different. However, contrary to interval comparison, the ANCOVA analysis suggests that the intercepts are different. This discrepancy is likely due to two factors. First, the length distributions of the samples are not similar, an important consideration with interval comparison. Second, ANCOVA, which controls for length, and interval comparison ask slightly different questions—the latter asks whether the intercepts of two lines are different when the lines are allowed to float freely or have their own slopes, whereas the ANCOVA test asks whether the intercepts are different when lines are forced to have a common slope.

10.2.3 Weight–Length Regression Line Residual Analyses

Most commonly weight–length regression coefficients are used to describe the relationship between length and weight or to compare differences in body form (condition) at a population level. However, Fechhelm et al. (1995) and Sutton et al. (2000) used magnitude and sign of the individual residuals as an indicator of fish condition (a larger negative residual indicated poorer relative condition) to summarize seasonal and sex-related patterns in condition. This type of analysis is synonymous with condition indices but can overcome some of the limitations associated with testing ratio data.

Box 10.2 Analysis of Covariance (ANCOVA)

The ANCOVA can be used to test for differences between regression parameters (i.e., slopes and intercepts) and is especially appropriate when the length ranges sampled in the populations to be compared are generally unequal. Here, the following SAS program provides results for both the complete (equation [10.5]) and reduced (equation [10.6]) models used to analyze differences in the $\log_{10}$ transformed weight–length regression equations from the low-elevation stream (A) and low-elevation lake (B) Yellowstone cutthroat trout populations presented in Box 10.1.

Program

```
OPTIONS PS=54 LS=75;
DATA TROUT;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
CARDS;
A   129   20
B   254   181
[Input complete data set];
PROC SORT; BY POP;
PROC GLM; CLASS POP; MODEL LOGWT=POP|LOGTL/SS3;
PROC GLM; CLASS POP; MODEL LOGWT=POP LOGTL/SS3 SOLUTION;
RUN;
```

Output

Table The ANCOVA to test for slope differences ($n = 100$). The dependent variable is LOGWT for the two populations (POP).

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	19.83785325	6.61261775	3577.54	<0.0001
Error	96	0.17744331	0.00184837		
Corrected total	99	20.01529656			
R^2	0.991135	Root MSE	0.042993		
CV	1.852261	LOGWT mean	2.321090		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
POP	1	0.00113727	0.00113727	0.62	0.4347
LOGTL	1	5.47809028	5.47809028	2963.74	<0.0001
LOGTL*POP	1	0.00078440	0.00078440	0.42	0.5163

Table The ANCOVA to test for intercept differences. The dependent variable is LOGWT (n = 100).

Class Level Information		
Class	Levels	Values
POP	2	A B

Analysis of Covariance

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	19.83706885	9.91853443	5398.14	<0.0001
Error	97	0.17822771	0.00183740		
Corrected total	99	20.01529656			
R^2	0.991095	Root MSE	0.042865		
CV	1.846757	LOGWT mean	2.321090		

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
POP	1	0.01778497	0.01778497	9.68	0.0024
LOGTL	1	10.20459461	10.20459461	5553.83	<0.0001

Parameter Estimates

Variable	Estimate[a]	SE	*t*-value	*P* > |*t*|
Intercept	–5.209761011 z	0.10539178	–49.43	<0.0001
POP A	0.038319477 z	0.01231671	3.11	0.0024
POP B	0.000000000 z			
LOGTL	3.080368006	0.04133391	74.52	<0.0001

[a] The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter z are not uniquely estimable.

Interpretation

For modeling purposes, the dummy variable (treatment variable POP) value for each fish from the stream and lake samples was 1 and 0, respectively, and the interaction term was calculated as LOGTL times the dummy variable. Thus the interaction term is equal to LOGTL (dummy code * LOGTL) for stream fish and zero for the lake samples. The LOGWT was then regressed against all the independent variables in the complete model (dummy POP, covariate LOGTL, and interaction LOGTL*POP). Of interest is the significance of the interaction term, which indicates whether or not the slopes of the two populations, when controlling for length, are significantly different—in this case they are not (interaction $P = 0.516$). If the slopes had been different, we would have concluded that these two populations had different trends in weight (condition) relative to length (i.e., incremental weight gain for a given increase in length is different) and we would have stopped with our analysis. Further, if we had found a difference between slopes, it would be appropriate to model LOGWT as a function of POP and LOGTL(POP) (length nested in populations).

(Box continues)

Box 10.2 *(continued)*

Because the slopes were not different, we remove the interaction term from the model and regress LOGWT against the remaining independent variables (POP and LOGTL). In this example the adjustment for the dummy or treatment (POP) variable is significantly different from 0 ($P = 0.002$), and we conclude that the intercepts are different. Overall these results suggest that the two populations gain weight incrementally in a similar fashion, but trout in population A are consistently heavier at a given length than are trout in population B.

In the output from the reduced model (i.e., interaction term removed), the coefficient for the dummy variable (POP) is 0.0383. This value represents the magnitude of the difference in intercepts of the linear regressions for the transformed data. Because the lake sample was coded as 0, the population is represented by slope 3.080 for the parameter estimate for LOGTL and intercept –5.210 (see reduced model output), whereas the stream population is represented by slope 3.080 and intercept (–5.2097 + 0.0383) or –5.171.

The residualized weights are the error terms associated with equation (10.2) and can be calculated as the observed transformed weight of an individual fish minus the predicted transformed weight, or

$$e_i = \log_{10}(W_i) - \log_{10}(a) - b \log_{10}(L_i), \tag{10.7}$$

where e_i is the residual value and can be negative; W_i is the weight of fish i; L_i is the length of fish i; and a and b are the regression parameter values for the equation developed from the group of fish of interest (population).

Residual condition uses the weight–length relationship of a discrete, sampled population; thus inferences regarding any individual or group of individuals from that population are relative only to other individuals within the overall sample used to develop the weight–length regression. Larger-scale comparisons of population level residual condition variation would require a single equation developed from all populations under consideration, with the assumption that all individual population weight–length relationships have similar slopes (Jakob et al. 1996; Sutton et al. 2000).

Residual analysis is very similar to the condition indices discussed in the following section (10.3), and the two are often highly correlated, but Fechhelm et al. (1995) suggested that in some cases residuals can be normally distributed in data sets in which condition indices are not, or vice versa. Thus, this technique can provide parametric options that might not otherwise be available with condition indices. Patterson (1992) suggested that in comparing condition to other variables it is inappropriate to use the residuals from a weight–length regression as an index to condition because the residuals are not unbiased estimators of the underlying error of a regression model. Rather, a more complete regression model including all factors that might affect weight should be fitted prior to analyzing residuals (see equation [10.12]).

Raw or standardized residuals can be generated by most statistical software. For example, SAS (SAS Institute 1998) calculates residuals for regression analysis; these values are stored in the variable name RESIDUAL and can be treated like any other SAS variable. Adding the SAS command PLOT RESIDUAL*TL (within the regression procedure command [PROC REG]) to the weight–length regression exercise will produce a plot of residuals as a function of total length (TL); an evident trend in residuals may suggest a lack of fit of the regression model. The residuals can also be used as variables in other common statistical tests, such as mean comparisons, to assess condition level and trends.

10.3 CONDITION INDICES

Condition indices are widely used to assess many facets of fish populations, including the general health of fish stocks, the effects of management actions, community structure, or environmental influences (Bolger and Connolly 1989; Ney 1993; Neumann and Willis 1995; Ward and Zimmerman 1999; Blackwell et al. 2000). Condition indices are intended to estimate physiological condition (e.g., lipid stores) indirectly based on the premise that a fish of a given species and length should weigh as much as a standard or average for its length, and variations from the standard are taken as an indication of the relative wellness of an individual. Measures of fish condition based on a standard weight have been available since the early 1900s and have undergone an evolution in methodology (Murphy et al. 1991) as well as rigorous reviews regarding their correlation with physiological parameters and statistical merit (e.g., Bolger and Connolly 1989; Patterson 1992; Blackwell et al. 2000; Vila-Gispert and Moreno-Amich 2001; Brenden et al. 2003). They remain popular tools because they are simplistic and noninvasive (only weight–length data needed) and are more easily compared than are the regression parameters in weight–length relationships. Murphy et al. (1990) indicated that an ideal condition index should be consistent, that is, maintain similar statistical properties and meaning across length and species; tractable, that is, analyzable by standard statistics; robust, that is, insensitive to data collection and analysis variations; and efficient, that is, provide precision from relatively small sample sizes. Anderson and Neumann (1996) and Blackwell et al. (2000) provided thorough reviews of the history of condition factors, and here we only briefly describe their history and development.

10.3.1 Fulton's Condition Factor

Traditionally, one of the ways to relate fish length to weight was simply to cube the length of the fish (Spencer 1898; Wootton 1990). However, this basic equation is imprecise because it fails to account for allometric growth (i.e., $b \neq 3$; equation [10.1]; Fulton 1904; Martin 1949). Nonetheless, this basic physical principle has been used extensively in fisheries science and is still used today (e.g., Ratz and Lloret 2003; Stone et al. 2003). For example, Fulton's condition factor (Anderson

and Neumann 1996) is calculated as the ratio between observed and expected weight for a fish of given length:

$$K = (W/L^3) \bullet 100{,}000, \tag{10.8}$$

where W is the weight (g), L is length (mm), and 100,000 is a scaling constant. In application, body form changes with length ($b > 3$) and species ($b_1 \neq b_2$), which results in condition factors that are often length and species dependent (Murphy et al. 1991; Jakob et al. 1996; Blackwell et al. 2000). Thus K increases with increasing length, limiting its application to fish of similar length within the same species.

10.3.2 Relative Condition Factor

Le Cren (1951) attempted to solve the deficiencies of K by comparing the actual weight to a standard predicted by the weight–length regression based on the population from which the fish was sampled. Relative condition is calculated as

$$K_n = (W/W') \bullet 100, \tag{10.9}$$

where W is individual fish weight and W' is the predicted length-specific weight based on $\log_{10}$ transformed data. Average fish of all lengths and species have an average K_n value of 100; however, because weight–length relationships can vary among populations and geographic sites, comparisons of K_n must be confined to those populations with homogenous weight–length parameters. Swingle and Shell (1971) indicated that K_n could be useful as an indicator of physiological stress on a population and expanded the concept by establishing species-specific weight–length relationships across a broader geographical range, which allowed comparisons of condition across populations. This broadened application of condition analyses from a population level to regional scale; however, regional differences still existed, making comparison and communication difficult.

10.3.3 Relative Weight

Relative weight (W_r) was proposed by Wege and Anderson (1978) as a condition analysis tool for largemouth bass and represents further evolution of the K_n concept by allowing comparisons of condition across the geographical occurrence of a species, as well as among species. The W_r index is calculated as

$$W_r = (W/W_s) \bullet 100, \tag{10.10}$$

where W is individual fish weight and W_s is a length-specific standard weight predicted from a weight–length regression developed to represent the body form of the species across its geographical range (see Blackwell et al. 2000 for a list of developed standard weight equations). The W_r index uses 100 (or a range, 95–105) as a benchmark for a fish in good condition—a readily identifiable standard

for fisheries scientists. Fish greater than the target are considered in relatively better condition than a standard fish, whereas those less than the target are considered in worse condition with severity depending on the distance from the benchmark. For example, condition values exceeding 105 may indicate abundant prey and favorable environmental conditions (e.g., Marwitz and Hubert 1997; Porath and Peters 1997).

The estimation of a and b in the standard weight equation (note equation [10.2]),

$$\log_{10}(W_s) = a + b(\log_{10}L), \tag{10.11}$$

has undergone several iterations and review of statistical validity (see Anderson and Neumann 1996). The currently accepted technique for development of W_s equations is the 75th regression-line–percentile (RLP) technique proposed by Murphy et al. (1990), which has consistently provided W_s equations with little or no length-related biases, allowing for comparisons within and across species. Gerow et al. (2004), however, suggested this bias has been incorrectly assessed in the past and may be greater than originally reported for most standard W_s equations. Because standard weight equations are developed based on weight–length relationships across the range of the species, comparison and communications of condition analyses are consistent across the species range. Herein lays the value of W_r relative to other condition indices. Whereas a single W_s equation for each species has generally proven adequate, and is preferred for simplicity, differences in body forms between broad habitat types (e.g., lotic versus lentic habitats) has required maintaining multiple standard weight equations or target goals (i.e., something other than 100) for some species (e.g., burbot, Fisher et al. 1996; inland cutthroat trout, Kruse and Hubert 1997).

It is logical that both environmental and genetic factors influence body form and weight, and, by extension, condition as well. Furthermore, it is possible for an individual to increase energetic fitness without a change in body weight (Booth and Keast 1986). Thus, questions remain whether W_r, or any weight to length ratio, is both a valid and interpretable indicator of the physiological condition in fish or a metric sensitive and relevant enough to assess the effects of changed management or environment on fish condition. Numerous studies have investigated the practical limits in the application of W_r. Liao et al. (1995) and Gutreuter and Childress (1990) found W_r a weak indicator of growth, a relationship that seems intuitive based on the assumption that a fish in better condition can devote more energy to growth. Conversely, Brown and Murphy (1991) and Neumann and Murphy (1992) found W_r was correlated with fat composition in the body, an indication that W_r can be a relative measure of individual energy stores. Blackwell et al. (2000) provided excellent discussion regarding the relationships, or in some cases the lack thereof, between W_r and body composition, growth, and reproductive potential, among other things. Brenden et al. (2003) suggested that the lack of clear relationships in some studies attempting to relate W_r to variables that seem intuitively related to individual condition might be the result of an index

that, in most cases, does not satisfy the theoretical assumptions on which the statistical test is founded.

Most analyses of W_r are either mean comparisons among different populations or length categories (e.g., *t*-test, ANOVA, or nonparametric equivalents) or an assessment of the correlation and regression relationship among condition and other independent variables that might influence fitness (e.g., prey density as a good predictor of condition for a given population or size-class of fish). Sections 10.3.4 and 10.6 describe some of the common statistical procedures used to analyze and compare individual and population level condition as measured by an index.

10.3.4 Application and Common Statistical Analysis of Relative Weight

10.3.4.1 Statistical Analysis of Relative Weight Data

The application of W_r has increased over the last decade and is now commonly used as a condition assessment tool in the majority of the USA (Blackwell et al. 2000); thus, we focus our discussion of statistical analyses on W_r. The appropriateness of W_r, which is a ratio, as a variable in statistical testing has been the subject of several reviews. Numerous authors have recommended against the use of ratios to scale biological data because analyses of ratios may point to treatment effects that do not exist or they may fail to detect major differences that do exist (e.g., Tanner 1949; Atchley et al. 1976; Anderson and Lydic 1977; Atchley 1978; Atchley and Anderson 1978; Reist 1985; Packard and Boardman 1988). Bolgor and Connolly (1989) indicated that the potential for greater variability and nonnormal distributions of ratio data such as W_r might make parametric testing of W_r inappropriate. Furthermore, they indicated that ratio data commonly exhibit heteroscedasticity, skewness, and leptokurtosis (a taller distribution with fatter tails as compared with normal), all of which violate the assumptions of common statistical tests (e.g., regression and ANOVA) and weaken the power of these comparisons. Thus, Hyatt and Hubert (2001) concluded that normality for W_r data cannot be assumed and should be assessed before applying parametric tests. Murphy et al. (1990), when evaluating W_r frequencies in walleye populations, suggested that the use of parametric tests to compare differences in W_r data yields conservative results, which Blackwell et al. (2000) interpreted as a greater probability of type II error (failure to reject the null hypothesis when the alternative is true). Contrarily, Bolger and Connolly (1989) stated that while skewness has minimal effect on significance or power, significant leptokurtosis could lead to greater nominal significance values. Sokal and Rohlf (1981) indicated that a nonnormal distribution is only a minor violation of the assumptions for parametric statistics, thus parametric mean-comparison tests are generally robust to departures from normality. If there is concern over violation of assumptions for parametric tests, an alternative is to use a nonparametric test such as Wilcoxon's rank-sum test or Kruskal–Wallis test for comparison.

Patterson (1992) also recognized the problems of skewed distributions of ratios and suggested, as summarized in section 10.2.3, that it is inappropriate to use

weight–length regression residuals because they are biased estimators of regression error. Likewise, Jakob et al. (1996) noted that residuals from the residual index for condition are not comparable across populations. This is germane because individual values of W_r are essentially the de-transformed residuals. As a solution, Patterson (1992) proposed that all variables assumed to affect weight be directly included in the analyses at the same time as length and the coefficient of each parameter used to assess its effect on condition. For example, when testing for mean monthly differences in condition, include month as a variable in the model:

$$\log_{10}(W_i) = \beta(0) + \beta(m) + \beta_1[\log_{10}(L_i)] + e_i, \quad (10.12)$$

where $\beta(0)$ is the overall intercept and $\beta(m)$ are monthly adjustments to the overall intercept. Each parameter coefficient is used to measure the effect on fish condition. This is essentially an extension of the ANCOVA analysis.

More recently, based on a derivation of the statistical properties of the index, Brenden et al. (2003) argued that W_r data are not independent and identically distributed, as required by both parametric and nonparametric tests, because the properties are conditionally dependent on fish length. Conventional tests that assume independence and identical distributions increase the likelihood of a type I error (rejecting the null hypothesis when there is no difference) when applied to W_r data. To alleviate this risk, they proposed an *R*-test as the most appropriate and conservative way to test relative weight data (see Brenden et al. [2003] for a more thorough discussion). Of concern is the relative difficulty of computing the *R*-statistic and its associated significance value, especially when the improvement in testing power is moderate. The application of this recently proposed test is probably greatest for researchers attempting to make definitive conclusions regarding patterns in condition but of less utility for management decisions that might include condition as only one component in a decision-making process.

Given these arguments, it is apparent that care should be taken when statistically analyzing W_r values, and the data should be analyzed to ensure that the assumptions of a chosen statistical test are not violated or that the test is robust enough to handle a violation of the assumptions. Transformations to normalize W_r data and homogenize the variances (e.g., Box–Cox transformation; Box and Cox 1964) have generally proven to be of little value (Murphy et al. 1990; Brenden et al. 2003). Alternatively, nonparametric tests can be used if the data will result in misapplication of parametric tests. However, as mentioned, Brenden et al. (2003) argue that their *R*-test is the most appropriate for testing W_r data. Undoubtedly the statistical merit of W_r comparisons will continue to be debated, leading to a better understanding of the statistical properties of this index, as well as a clearer picture of the potential shortcomings and strengths of using established parametric and nonparametric tests and alternative tests for comparisons. We suggest that mean comparisons (*t*-test, ANOVA, Mann–Whitney, and Kruskal–Wallis) and regression relationships can continue to be adequate methods for testing W_r data, as long as the discussion of comparative results includes reference to the potential

shortcomings of the test in relation to the distribution of the data. Results likely can be clarified and strengthened by comparing the results of multiple tests.

10.3.4.2 Length-Related Patterns in Relative Weight Data

Because environmentally dependent trends in condition across lengths can be averaged out, mean population condition should not be compared unless it can be demonstrated that length-related patterns or differences are absent in the population. Plotting individual or length-group mean W_r values allows a visual assessment of potential or important patterns such as size-related condition trends resulting from, for example, differences in prey availability, gonad maturity, or density. Murphy et al. (1991) suggested that condition data should be summarized by length-group based on Gabelhouse's (1984) five-cell model (stock-, quality-, preferred-, memorable- and trophy-length fish); others have suggested that this model may not be ecologically relevant depending on the relationship being explored and have summarized W_r differently (e.g., 50-cm length-groups; Porath and Peters 1997). Once W_r values have been classified in a fashion relevant to the question of interest (note that this does not preclude the use of individual fish condition as the unit of interest), individuals or groups can be compared with each other to determine whether one is poorer conditioned than another or whether condition as measured by W_r (as the dependent variable) is statistically related to another variable or suite of variables, such as a habitat attribute. Box 10.3 provides examples of tests comparing W_r among multiple populations.

10.3.4.3 Relationship of Relative Weight to Physiological and Environmental Measures

As surrogate indicators of physiological well being, condition index values such as W_r should reflect proximate body composition of individual fish (e.g., lipid content, protein content, or caloric content; Murphy et al. 1991). Strange and Pelton (1987) found a weak relationship between mean condition factor *(K)* and fat percentage in composite samples of prey fishes. However, more recent physiological assessments of W_r have found correlations between W_r and tissue energy content in walleye (Rose 1989), white crappie (Neumann and Murphy 1991), and striped bass and hybrid striped bass (Brown and Murphy 1991). Brown and Murphy (1991) suggested that W_r provided a better estimate of reserve energy than did measures such as the liver–somatic index. Thus, W_r appears to be a reliable index of energy reserves in these species and, as such, might be a good indicator of short-term growth potential or potential for resistance to nutritional stress (Murphy et al. 1991). However, complications such as volume replacement of lipid (fat) reserves by water may confound the relationship between W_r and proximate components (Novinger and Martinez Del Rio 1999).

On the other hand, assessments of relationships between W_r and characteristics that would seem a logical expression of energy use, such as growth, which represents the ultimate expression of individual fitness (Bolger and Connolly 1989), have had mixed results. A common notion is that W_r and other condition indices can be used as indicators of growth: poor condition indicates poor growth and vice versa (e.g., Busacker et al. 1990; Ney 1993). Positive correlations between W_r

Box 10.3 Comparisons of Mean Relative Weight

Murphy et al. (1990) provided a formula for computation of the 95% CI around a mean relative weight ($\overline{W}_r$) value:

$$CI = \overline{W}_r \pm t \cdot (SD/\sqrt{n}), \tag{10.13}$$

where t is the t-value that corresponds to an α-value (usually 0.05) with $n - 1$ df and $\overline{W}_r$ is the mean measure of condition for a specific group (population). The overlap in CIs for mean values from different populations or length-groups can be compared to determine whether they are statistically similar or not. For example, a simple mean calculation for the stream population of Yellowstone cutthroat trout presented in Box 10.1 provides a $\overline{W}_r$ of 94.7 (SD = 9.47), whereas the low-elevation lake population has a $\overline{W}_r$ of 92.9 (SD = 8.45). Thus, the respective CIs would be 94.7 ± 2.69 and 92.9 ± 2.32. Both intervals include the mean value of the other population (see discussion in Box 10.1), indicating that individuals in these two populations are similarly conditioned, but this tells us little about whether there are length-specific differences between populations.

Mean comparison tests such as the two-sample t-test (or the nonparametric equivalent, Mann–Whitney test) or multiple-comparison tests such as ANOVA (or the nonparametric Kruskal–Wallis test) can be used to examine length-related or inter-population trends in W_r. Herein, we discuss how one might test for difference in condition, as indexed by W_r, among length-groups from the same population or among populations. For the Yellowstone cutthroat trout data presented in Box 10.1 the question of interest is whether macro-scale habitat type (stream versus lake and low versus high elevation) has any significant influence on fish condition.

Relative weights were calculated for the three populations described in Box 10.1 based on the lotic ($\log_{10}W_s = -5.189 + 3.099 \cdot \log_{10}[TL]$) and lentic ($\log_{10}W_s = -5.192 + 3.086 \cdot \log_{10}[TL]$) standard-weight equations for cutthroat trout (Kruse and Hubert 1997). An important first step is to assess the distribution of the W_r data to determine whether a parametric or nonparametric test is more appropriate. This can be completed with typical assessments of normality, such as a histogram or box-plot of the data (not shown). In this case, the data appears generally normal, but there is some skewness and outliers for all three populations. It is important to assess whether the outliers (or individuals with extreme values when compared to the mean) are biologically relevant or errors due to measurement or data entry. We retained the outliers in this assessment.

Prior to comparing overall population means, it is prudent to check for length-related patterns in condition within each population (Murphy et al. 1990, 1991; Blackwell et al. 2000). For example, changes in W_r with increasing length for cutthroat trout from the low-elevation stream population can be assessed by grouping cutthroat trout in 50-mm length categories (e.g., group one is 100–149-mm fish and group five is 300–349-mm fish plus the two largest fish). Another way to group the fish is to use the five-cell model (Gabelhouse 1984) for stock- to trophy-length fish (see cutthroat trout length categories in Anderson and Neumann [1996]). The following SAS program calculates W_r values for individual fish and assigns each fish to a length-group for testing differences in $\overline{W}_r$ among length-groups by means of ANOVA, a test that is robust to small departures from normality.

(Box continues)

Box 10.3 *(continued)*

Program

```
OPTIONS PS=54 LS=75;
DATA TROUT;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.189+(3.099*LOGTL));
WR=(WT/WS)*100;
IF TL>99 AND TL<150 THEN GRP=1;
IF TL>149 AND TL<200 THEN GRP=2;
IF TL>199 AND TL<250 THEN GRP=3;
IF TL>249 AND TL<300 THEN GRP=4;
IF TL>299 THEN GRP=5;
CARDS;
A   129   20
A   130   25
[Input complete data set];
PROC ANOVA; CLASS GRP; MODEL WR=GRP;
RUN;
```

Output

Table The ANOVA procedure for comparing differences in $\overline{W}_r$ among length-groups (GRP) in a population of low-elevation stream-dwelling Yellowstone cutthroat trout ($n = 50$).

Class Level Information					
Class	Levels	Values			
GRP	5	1 2 3 4 5			
Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	$P > F$
Model	4	254.915000	63.728750	0.69	0.6005
Error	45	4136.534097	91.922980		
Corrected total	49	4391.449097			
R^2	0.058048	Root MSE	9.587647		
CV	10.12727	$\overline{W}_r$	94.67155		
Source	*df*	SS	Mean square	*F*-value	$P > F$
Group	4	254.9150003	63.7287501	0.69	0.6005

Interpretation

It does appear, if one calculates the means of each 50-mm length-group, that there are some differences in condition. For example, 150–199-mm fish have an average W_r value of 97.5, whereas 200–249-mm fish average only 90.3. This might indicate that the smaller fish have a better prey base than do larger individuals who may be using another food source. However, given the variability in W_r values among individuals within each group, and the differences in sample sizes,

are these mean values significantly different? The test of equality of $\overline{W_r}$ values in each 50-mm length-group (i.e., no differences in means among groups would indicate no length-related patterns) is based on a comparison of two types of variability—within groups (variability of individuals in each category around the mean for that category) and between groups (variability of the mean of each category around the overall mean for the population). The *F*-test value reported in the SAS output is based on the ratio of the variability between groups to the variability within groups. Mean W_r values among length-groups were not different. Thus, there does not appear to be an environmental influence such as prey or habitat selection differentially influencing the condition of Yellowstone cutthroat trout in this population, at least on a length-dependent basis (based on the length-groups we selected). Thus, it may be appropriate to calculate a population-wide $\overline{W_r}$ value for this Yellowstone cutthroat trout population.

For comparison sake, and because the data were somewhat nonnormal, we also provide the SAS program and output for a Kruskal–Wallis test.

Program

```
OPTIONS PS=54 LS=75;
DATA TROUT;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.189+(3.099*LOGTL));
WR=(WT/WS)*100;
IF TL>99 AND TL<150 THEN GRP=1;
IF TL>149 AND TL<200 THEN GRP=2;
IF TL>199 AND TL<250 THEN GRP=3;
IF TL>249 AND TL<300 THEN GRP=4;
IF TL>299 THEN GRP=5;
CARDS;
A   129   20
A   130   25
[Input complete data set];
PROC NPAR1WAY WILCOXON; CLASS GRP; VAR WR;
RUN;
```

Output

Table The NPAR1WAY procedure of SAS for Wilcoxon scores (rank sums) for the variable $\overline{W_r}$ classified by length-group.

Group	*N*	Sum of scores	Expected under H_0	SD under H_0	Mean score
1	11	314.0	280.50	42.699532	28.545455
2	10	307.0	255.00	41.231056	30.700000
3	8	142.0	204.00	37.788887	17.750000
4	14	342.0	357.00	46.281746	24.428571
5	7	170.0	178.50	35.766605	24.285714

(Box continues)

Box 10.3 *(continued)*

Kruskal–Wallis Test

Chi-square	4.1380
df	4
P > chi-square	0.3877

Interpretation

The Kruskal–Wallis test indicates similar results as the ANOVA. The mean ranks suggest that the 150–199-mm fish are slightly better conditioned than are the other length-categories, and the 200–249-mm fish are poorer conditioned; however, the test for differences among length categories is nonsignificant ($P = 0.39$).

Similar to the stream population, there were no length-related patterns in W_r in either of the two lake Yellowstone cutthroat trout populations (results not shown); thus, we can use an ANOVA to determine if there are any differences in fish condition among populations.

Program

```
OPTIONS PS=54 LS=75;
DATA TROUTA;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.189+(3.099*LOGTL));
WR=(WT/WS)*100;
CARDS;
A   129   20
A   130   25
[Input complete data set for population A];

DATA TROUTB;
INPUT POP $ TL WT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.192+(3.086*LOGTL));
WR=(WT/WS)*100;
CARDS;
B   254   181
B   262   186
[Input complete data set for populations B and C];

DATA TROUT; SET TROUTA TROUTB;
PROC SORT; BY POP;
PROC ANOVA; CLASS POP; MODEL WR=POP; MEANS POP;
RUN;
```

Output

Table The ANOVA procedure to compare $\overline{W_r}$ for three populations of Yellowstone cutthroat trout ($n = 150$).

Class Level Information					
Class	Levels	Values			
POP	3	A B C			
Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	276.39380	138.19690	1.75	0.1773
Error	147	11604.08875	78.93938		
Corrected total	149	11880.48255			
R^2	0.023265	Root MSE	8.884784		
CV	9.542586	$\overline{W_r}$	93.10667		
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
POP	2	276.3938029	138.1969014	1.75	0.1773
Population Estimates					
			Relative weight		
Population	*N*		Mean		SD
A	50		94.6715459		9.46685852
B	50		93.2871930		8.45707184
C	50		91.3612615		8.69911855

Interpretation

Mean W_r values for Yellowstone cutthroat trout in the low-elevation stream population (A), the low-elevation lake (B), and the high-elevation lake (C) were 94.7, 93.3, and 91.4, respectively. Even though we might have expected differences either on an elevation gradient or by habitat type, there was no significant difference in $\overline{W_r}$ for these three populations ($P = 0.18$). Thus, on average an individual fish of a given length from these three populations appears to be similarly conditioned. A Kruskal–Wallis test on these data provides similar results ($P = 0.35$). It is important to note that this example included fish from only one of each habitat type; thus, it is inappropriate to conclude that Yellowstone cutthroat trout condition does not vary as a function of elevation or habitat type. To explore the relationship between condition and elevation or habitat type, a different design is needed (i.e., samples of fish are required from multiple low-elevation streams, low-elevation lakes, and high-elevation lakes) because individual water bodies would be the experimental unit of interest, not individual fish within a water body.

and growth have been reported for largemouth bass (Wege and Anderson 1978), northern pike (Willis 1989), yellow perch (Willis et al. 1991), and juvenile striped bass and hybrid striped bass (Brown and Murphy 1991). However, other evidence contradicts the notion that W_r is consistently correlated with growth (Gutreuter and Childress 1990; Gabelhouse 1991). Furthermore, Liao et al. (1995) found no evidence that growth and W_r were correlated for pumpkinseed or golden shiner. Relative weight may reflect growth of some species under certain circumstances, but uncritical use of W_r as a predictor of growth could lead to substantial errors in population assessments.

Another factor commonly linked with W_r is prey availability (Anderson and Gutreuter 1983; Busacker et al. 1990; Flickinger and Bulow 1993; Ney 1993). Poor condition is assumed to reflect prey scarcity, whereas good condition is assumed to reflect an abundance of prey, and both these patterns can be found among size-classes of fish within the same population. Kohler and Kelly (1991) indicated that a quick and cost-effective method for evaluating prey supply was to assess condition of their predators. Porath and Peters (1997) believed that walleye W_r values from standardized fall surveys offer a cost-effective method of detecting prey deficiencies in reservoirs. Small W_r values were reported for lake trout in oligotrophic Wyoming lakes with sparse zooplankton; larger W_r values were found for lake trout in two Wyoming mesotrophic lakes, and the largest W_r values were reported from Flaming Gorge Reservoir, the most productive reservoir in the study (Hubert et al. 1994). Prey availability and W_r values were correlated for pumpkinseed but not for golden shiner; differences in these two species may be related to differences in food habits, with golden shiner having a more flexible and omnivorous diet (Liao et al. 1995). Relative weight may be a good predictor of prey availability especially for species with relatively narrow or specialized diets.

Most of these relationships have been examined through the use of group mean comparisons, bivariate correlations, or linear regression analyses. In Box 10.4 some of these common techniques are applied to the relationship between Yellowstone cutthroat trout W_r and whole-body fat composition.

10.4 PHYSIOLOGICAL MEASURES OF CONDITION

Whereas condition indices attempt to approximate indirectly energetic well-being based on individual whole-body mass, other measures of condition relate directly to the physiological composition of body tissues, thereby providing a more precise measure of actual fitness in terms of stored energy. Physiological measures of condition have used either an index (ratio) of tissue weights or direct measures of tissues such as lipid or protein content. These include the liver–somatic index (hepatosomatic index or ratio of liver weight to body weight minus gonads), body water content, visceral–somatic index, percent composition of body tissues (e.g., percent lipid or fat), and RNA/DNA ratios (Elliott 1976; Heidinger and Crawford 1977; Jensen 1979; Bulow et al. 1981; Adams and McLean 1985; Håkanson 1989; Shackley et al. 1994). These types of measures are typically invasive, lethal, and more costly and time consuming than are indices based on weight–length

information, which has typically been the impetus for developing indexes such as W_r. Statistical procedures and limitations associated with the use of physiological measures of condition are similar to those described above for weight–length relationships and W_r. Physiological information summarized in ratio form (e.g., liver–somatic index) has the same problems of nonnormality, nonindependence, and heteroscedasticity as does W_r. Measures of tissues composition based on percentage body weight vary in synchrony (colinearity or highly related independent variables) by their very nature. For example, if the percentage of fat based on overall body weight increases, than the percentage of another tissue component (e.g., protein or water) must decrease because the total cannot exceed 100%.

Many fisheries scientists simply use these measures in a graphical form to describe the trend in fish condition as measured by tissue weight or composition over time. Others have used simple correlation analysis to relate one measure to another or to some environmental variable. Adams and McLean (1985) used the liver–somatic index as a variable in a regression analysis to predict largemouth bass growth, whereas Delahunty and de Vlaming (1980) determined the organ weight–body weight relationship of goldfish by means of linear regression, tested the seasonal variation of the relationship using ANCOVA, and used ANOVA to determine if lipid values varied by month (see Box 10.4).

10.5 ADDITIONAL MORPHOMETRIC MEASURES OF CONDITION

Morphometric assessments of condition estimate individual fitness based on measurement of body form. Condition indices are a type of morphometric index that measure body form along a single axis, which is used to calculate an average or standard weight for a given length. Instead of the progression of condition indices from K to K_n to W_r, as described in section 10.3, Jones et al. (1999) proposed an alternative condition factor (B) based on two dimensions of fish body form, length and height, in association with weight (building on equation [10.8]) in an attempt to eliminate some of the length and species-related biases associated with Fulton's condition factor (K):

$$B = M/(H \bullet L^2), \tag{10.14}$$

where M is body mass or weight, H represents body height, and L is body length. The premise is that mass is related to body density and form in three dimensions (length, height, and thickness). Jones et al. (1999) suggested that the third dimension, thickness or girth, could be reasonably approximated by length (i.e., thickness is linearly related to length) and reduce regression variability while eliminating substantial handling and measurement time required to assess girth or thickness. Richter et al. (2000) argued that the assumption of a linear relationship between thickness and length was false in most cases and that the effects of allometric growth could be better minimized by the equation

$$B' = M/(H^2 \bullet L). \tag{10.15}$$

Box 10.4 Analysis of Fat Composition Data

In Box 10.3, we tested for differences in W_r within and among populations. Here we examine whether those W_r values are related to whole-body fat content in individual fish and then test whether population mean fat content differs among populations. Fat composition, a direct measure of individual wellness or energy stores, was estimated for the Yellowstone cutthroat trout sampled in stream and lake habitats (see Box 10.1 for data). We compared fat composition to W_r by means of correlation and regression analyses. The question of interest is whether W_r is a good indicator of individual physiological fitness as referenced by tissue fat content. Additionally, we want to know if using fat as the indicator of individual fitness results in a different conclusion regarding the population-level effects that elevation (a surrogate for environmental conditions such as temperature, growing season, and food supply) might have on fish condition. Please note that in this example we did not check for length-related biases (e.g., potential differences among length categories) within each population. The following SAS program regresses wet weight fat percentage against individual W_r (all populations combined into one data set) and compares mean percent fat composition among the three Yellowstone cutthroat trout populations by means of ANOVA.

Program

```
OPTIONS PS=54 LS=75;
DATA TROUTA;
INPUT POP $ TL WT FAT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.189+(3.099*LOGTL));
WR=(WT/WS)*100;
CARDS;
A   129   20   5.91
A   130   25   12.88
[Input complete data set for population A];
DATA TROUTB;
INPUT POP $ TL WT FAT;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
WS=10**(-5.192+(3.086*LOGTL));
WR=(WT/WS)*100;
CARDS;
B   254   181   10.59
B   262   186   9.08
[Input complete data set for populations B and C];
DATA TROUT; SET TROUTA TROUTB;
PROC SORT; BY POP;
PROC REG; MODEL FAT=WR;
PROC ANOVA; CLASS POP; MODEL FAT=POP; MEANS POP/TUKEY;
RUN;
```

Output

Table Linear regression of wet weight fat percentage against individual W_r.

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	857.46498	857.46498	158.72	<0.0001
Error	148	799.53931	5.40229		
Corrected total	149	1657.00428			
r^2	0.5175	Root MSE	2.32428		
Adjusted r^2	0.5142	Dependent mean	6.57033		
CV	35.37542				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	–18.44302	1.99447	–9.25	<0.0001
Relative weight	1	0.26865	0.02132	12.60	<0.0001

Interpretation

Fat composition and W_r are significantly correlated with each other ($r = 0.719$, $P < 0.001$) and the regression F-test ($P < 0.0001$) indicates that the slope (0.27) of the relationship between these variables is significantly greater than zero. Thus, it is apparent that W_r does reflect whole-body fat content (as a percentage of whole body weight) in individual fish to some degree. However, the coefficient of determination (r^2), or the proportion of the variability in percent fat explained by the linear relationship with W_r, is 0.52, suggesting only moderate explanatory power and providing evidence that other factors are influencing the weight and ultimately W_r of individual Yellowstone cutthroat trout.

Additional Output

Table Comparison of percent fat among populations ($n = 150$). The Tukey's studentized range (HSD) test controls the type I experimentwise error rate, but it generally has a higher type II error rate than the Ryan–Einot–Gabriel–Welsch multiple-range test.

Class Level Information		
Class	Levels	Values
POP	3	A B C

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	414.859585	207.429793	24.55	<0.0001
Error	147	1242.144698	8.449964		
Corrected total	149	1657.004283			
R^2	0.250367	Root MSE	2.906882		
CV	44.24254	Fat mean	6.570333		

(Box continues)

Box 10.4 *(continued)*

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
POP	2	414.8595853	207.4297927	24.55	<0.0001

Tukey's Studentized Range (HSD) Test for Fat

Alpha	0.05
Error *df*	147
Error mean square	8.449964
Critical value of studentized range	3.34848
Minimum significant difference	1.3765

Tukey grouping[a]	Mean fat	*N*	Population
A	8.8376	50	A
B	5.9782	50	B
B	4.8952	50	C

[a] Means with the same letter are not significantly different.

Interpretation

When comparing percent fat scores among populations, the *F*-test for the ANOVA was significant ($P < 0.0001$), indicating that at least one of the populations had significantly different overall mean percent fat. However, the ANOVA does not provide information regarding which or how many populations are significantly different; thus, a post hoc multiple-comparisons test is needed. There are several post hoc multiple comparisons that can be used to determine which group mean(s) are statistically different, such as Tukey's studentized range test, Duncan's multiple-range test, least significant differences, and Scheffé's statistic. Carmer and Swanson (1973) provide a good decision tree regarding which multiple-comparison test is most appropriate.

In this case, we used Tukey's test to determine which populations were different. The Tukey's grouping shows that Yellowstone cutthroat trout in both lake populations (B and C) had significantly lower percent body fat than did Yellowstone cutthroat trout in the low-elevation stream population (A). Further, percent body fat for Yellowstone cutthroat trout in the two lakes are not significantly different from each other (either there are no real differences or there was enough uncertainty or variance in the percent fat values that the multiple-comparison test could not differentiate B from C). These results seem different than those of the similar ANOVA we ran in Box 10.3, which indicated that mean population condition as measured by W_r was not statistically different among populations. However, we must remember that W_r estimates were calculated using different W_s equations, which were designed to account for general body-form differences between cutthroat trout in lotic and lentic systems. Even so, there is some question, as illustrated by both the linear regression relationship between W_r and fat in this example and the ANOVA in the previous example (Box 10.3), as to whether W_r provides a true reflection of fish condition in these Yellowstone cutthroat trout populations, at least as measured by fat reserves in the body.

Both B and B' provide improvement over K when comparing the regression relationship between actual body mass and the body mass back-calculated from the condition factors (Jones et al. 1999; Richter et al. 2000). These are an appropriate modification to the condition factor concept, allowing broader condition comparisons across size ranges and populations, especially for those species for which W_s equations have not been developed. Statistical tests similar to those discussed for condition indices in section 10.3 can be applied to B or B' to provide rigorous comparisons.

Similarly, measurements of body form dimensions other than length, such as distances between anatomical landmarks, can be used in lieu of weight to assess condition. This approach may be especially useful when individual measurements of weight are highly imprecise, such as with small fish. Box 10.5 describes an example of condition assessment in juvenile largemouth bass based on body depth (height) and length in an ANCOVA. The use of one or two anatomical distances to assess condition is a simplistic form of truss analysis.

Truss analysis has been in use for several decades (Humphries et al. 1981; Strauss and Bookstein 1982) but primarily for morphometric comparisons of differences in body form among different types or stocks of fish. This type of analysis involves systematic measurement between multiple pairs of landmarks across the body in order to differentiate body shapes computationally. These measurements, often based on discrete juxtapositions such as fin insertion points (Fitzgerald et al. 2002), form polygons across the body, which give rise to the term truss analysis. Fitzgerald et al. (2002) applied truss analysis to quantify changes in fish condition by using a 10-point truss system to assess the effect of differing feed rations. Eigenvector coefficients from principle component analysis (PCA), a multivariate data reduction technique, were successful in demonstrating that key truss measurements change as condition changes and can be used to describe differences in body form between groups of better- and lesser-conditioned fish. The PCA approach is a common analytical tool for truss comparisons (Toline and Baker 1993; Moore and Bronte 2001).

Truss analysis can be used to compare the condition of fish among groups (populations), habitats, or sampling time. Fitzgerald et al. (2002) argued that although truss measurements may currently be more time and cost consuming than traditional condition indices, they provide a much clearer picture of the effect condition, or lack thereof, has on body form and allow for precise comparisons over time. Truss comparisons may prove to be more ecologically, morphologically, genetically, and physiologically revealing than are more popular and traditional numerical constructs of fish condition. As digital imaging techniques and computer analytical software continue to evolve and advance, truss analysis likely will become a common technique for analyzing fish condition.

10.6 FACTORS AFFECTING CONDITION DATA

Seasonal changes occur in fish condition due to changes in fish behavior and physiology that are influenced by many factors (e.g., changes in temperature,

Box 10.5 Morphological Assessment of Juvenile Condition

The following data are used to assess effects of starvation on body condition of largemouth bass juveniles. For most fishes, standard condition indices (e.g., W_r) are applicable to only adults and large juveniles because weight measurements are imprecise for small fish. A controlled experiment was conducted to determine if simple morphological measurements could be used to determine condition of juvenile largemouth bass (partial data set from Smith et al. [2005]). Hatchery-reared largemouth bass were raised until completion of fin development and then divided into two experimental groups of fed and unfed fish. Differences in body morphology existed after only 3 d of food deprivation, and a simple bivariate ratio of body depth at the anus to standard length was almost as efficient and robust at classifying fed and unfed largemouth bass as a multivariate index based on 23 morphometric characters. Here we provide an assessment of differences in the body depth after 6 d of food deprivation.

Table Standard length (SL; mm) and body depth (BD; mm) of juvenile largemouth bass. Fed largemouth bass were provided brine shrimp; unfed largemouth bass were deprived food for 6 d.

Fed		Unfed	
SL	BD	SL	BD
9.237	1.706	11.934	2.427
9.267	1.730	10.482	2.164
9.500	1.895	10.605	1.907
9.291	1.811	10.604	1.903
9.291	1.814	13.024	2.811
12.296	2.680	12.215	2.390
12.575	2.585	12.660	2.324
12.296	2.388	12.984	2.419
12.707	2.495	11.047	1.875
11.329	2.328	11.853	2.259
12.659	2.489	11.531	2.296
9.842	2.148	12.136	2.390
10.237	1.981	11.651	2.196
8.818	1.791	11.167	2.032
8.500	1.707	12.216	2.358
10.105	2.129	12.054	2.229
11.344	2.530	12.821	2.290
8.053	1.454	12.581	2.194
8.474	1.621	11.653	1.969
9.503	2.105	12.342	2.196
10.422	2.127	11.540	2.103
9.212	1.961	13.638	2.583
10.848	2.425	11.168	1.872
8.369	1.537	11.490	2.003
10.925	2.316	11.651	1.907
12.448	2.674	11.697	2.097

Program

The following SAS program tests for differences among the body depth of fed and unfed large-mouth bass by means of ANCOVA to remove the confounding effect of fish size.

```
OPTIONS PS=54 LS=75;
DATA LMB;INPUT FOOD $ SL BD @@;CARDS;
F   9.237   1.706      U11.9342.427
F   9.267   1.730      U10.4822.164
[Input complete data set];
PROC SORT; BY FOOD;
PROC GLM; CLASS FOOD; MODEL BD=FOOD|SL/SS3;PROC GLM; CLASS FOOD; MODEL
BD=FOOD SL/SS3 SOLUTION;RUN;
```

Output

Table An ANCOVA to test for slope differences in body depth (BD) of fed and unfed fish ($n = 52$).

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	4.06177300	1.35392433	73.52	<0.0001
Error	48	0.88395183	0.01841566		
Corrected total	51	4.94572483			
R^2	0.821270	Root MSE	0.135704		
CV	6.321951	BD mean	2.146558		
Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
FOOD	1	0.00498644	0.00498644	0.27	0.6052
SL	1	2.64331777	2.64331777	143.54	<0.0001
SL*FOOD	1	0.00007687	0.00007687	0.00	0.9488

Table An ANCOVA to test for intercept differences in BD of fed and unfed fish ($n = 52$).

Class Level Information		
Class	Levels	Values
Food	2	F U

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	2	4.06169612	2.03084806	112.57	<0.0001
Error	49	0.88402870	0.01804140		
Corrected total	51	4.94572483			
R^2	0.821254	Root MSE	0.134318		
CV	6.257381	BD mean	2.146558		

(Box continues)

Box 10.5 *(continued)*

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
FOOD	1	0.61792239	0.61792239	34.25	<0.0001
SL	1	3.91445980	3.91445980	216.97	<0.0001

Parameter Estimates

Variable	Estimate[a]	SE	*t*-value	*P* > |*t*|
Intercept	–0.5671680765 z	0.18968254	–2.99	0.0044
FOOD-FED	0.2627719497 z	0.04490009	5.85	<0.0001
FOOD-UNFED	0.0000000000 z			
SL	0.2330097751	0.01581879	14.73	<0.0001

[a] The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter z are not uniquely estimable.

Interpretation

Differences existed in the body depth between fed and unfed largemouth bass (slopes were not different [$P = 0.95$]; intercepts were different [$P < 0.0001$]; see discussion in Box 10.2), with greater body depth for fed fish (see figure below). Thus, body morphology is related to nutritional status in juvenile largemouth bass. Therefore, distances between anatomical landmarks or trusses (see Strauss and Bookstein [1982] for a discussion of trusses) may be used in some instances to quantify fish condition. This approach may be especially useful for assessing condition of larval and juvenile fishes; however, careful consideration must be given to ontogenetic stage, size, and species (Suthers 1992; Ferron and Leggett 1994). In addition, changes in fish condition in response to changes in food availability is likely greatest at intermediate abundances of prey (Ferron and Leggett 1994). That is, no change in condition will occur with an increase in prey abundance if a larval fish is already consuming the biological maximum amount of food (i.e., food intake is limited by handling and digestion). Likewise, little change in condition is expected for a starved larval fish that is provided a small amount of food, especially if the fish is near the threshold for irreversible starvation (also called the point-of-no-return). Thus, our statistical ability to detect differences in fish condition will vary as a function of food abundance and period of assessment.

turbidity, food supplies, and photoperiod; Pope and Willis 1996). Condition is a short-term indicator of fish health status and is primarily influenced by resource availability and gonadal growth. Typically with spring spawners, fish condition is greatest in the spring just before spawning, declines immediately after spawning, and then increases through the summer and into the fall. Obviously, the seasonal trend in condition for fish species that spawn in the summer (e.g., bluegill) or fall (e.g., brook and brown trout) should be different than spring spawners. Furthermore, differences in gonadal development between males and females may show gender differences in seasonal condition trends. Finally, fish size may also affect the seasonal trend in fish condition (see Pope and Willis 1996 for detailed examples of related studies). Le Cren (1951) noted that seasonal changes in condition of mature fish are often due to changes in gonad weight. However,

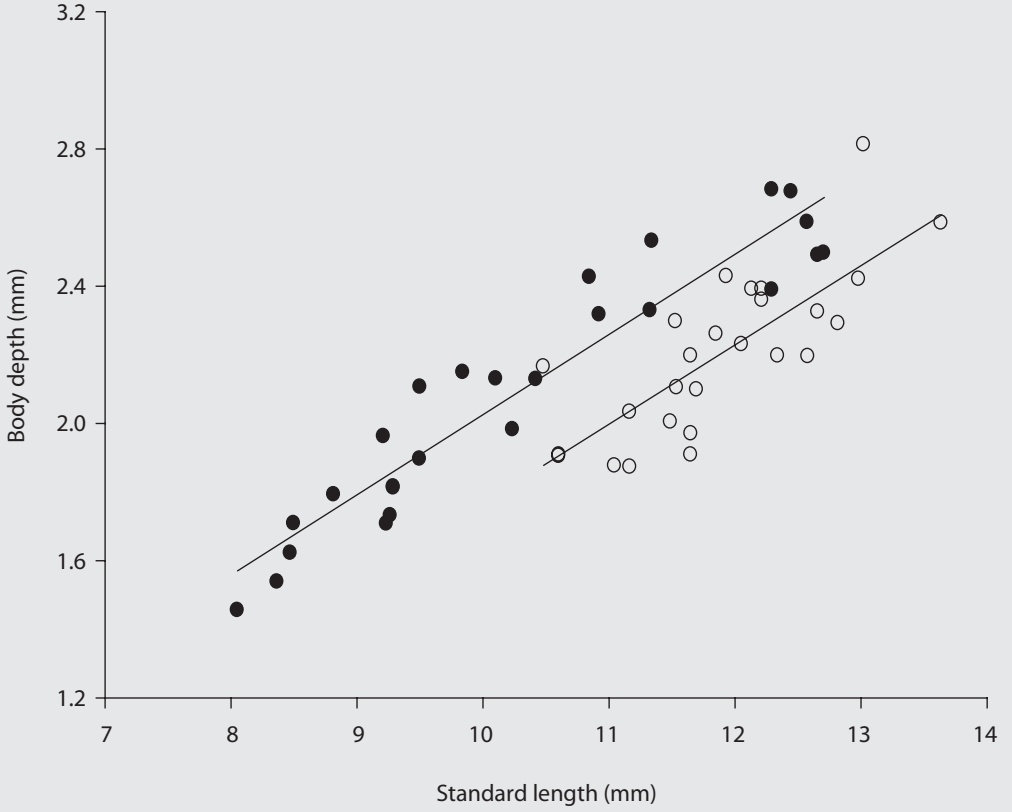


Figure Body depth (mm) as a function of size (standard length; mm) for fed (solid circles) and unfed (open circles) juvenile largemouth bass.

seasonal changes in the condition of immature fish may be attributable to feeding conditions throughout the winter and spring. For example, Brown (1993) reported that smaller (125–300-mm TL) largemouth bass in Aquilla Lake, Texas, came out of the winter with a low W_r (i.e., 85), and condition remained low until late spring–early summer, when W_r increased (i.e., 105). Gabelhouse (1991) found that small white crappies (130–199-mm TL) in Melvern Reservoir, Kansas, exhibited the greatest W_r in July and that W_r continued to decline throughout the fall and winter. He speculated that the summer peak condition of small white crappies reflected the feeding conditions associated with peak spawning of gizzard shad in mid to late May. Thus, it is inappropriate to combine condition data across seasons. Furthermore, condition data should be reported separately for mature and immature fish and may need to be separated by gender for mature individuals. Generally, W_s

equations are reported for combined sexes; however, Neumann and Willis (1994) provided separate W_s equations for male and female muskellunge (slopes of these two equations were different).

Although general seasonal trends in condition of fish are observed, more specific spatial and temporal patterns of variation in W_r also exist. For example, Liao et al. (1995) observed spatial and temporal differences among lakes for pumpkinseed and golden shiner. Temporal variations in condition have been reported for black crappie (Gabelhouse 1991; Guy and Willis 1991), burbot (Pulliainen and Korhonen 1990), northern pike (Guy and Willis 1991), walleye (Guy and Willis 1991), and yellow perch (Le Cren 1951; Guy and Willis 1991). Many of these studies have resulted in the common practice of sampling during "standard" periods for assessing condition of fishes. However, the temporal asynchrony of pumpkinseed and golden shiner W_r suggests that standard sampling periods might not be as comparable among lakes or among years as previously believed (Liao et al. 1995). This temporal asynchrony illustrates some of the biotic and abiotic variability that fisheries scientists must deal with when assessing fisheries.

Fisheries scientists primarily use condition assessments as a measure of the quality of fish populations, ideally with respect to local environmental and climatic conditions and species potential, and as a means of measuring changes in population quality resulting from management practices (Childress 1991). Thus, comparisons of condition are made on many different scales. Comparisons can be made within populations to assess differences across length-groups or to conduct spatiotemporal comparisons. Theoretically, data on the condition of various sizes of fish within a population can be accumulated over many years to establish a norm for a specific water body. Any deviation from the norm would indicate some fluctuation within the population or some physical or chemical condition interacting with a segment of the population (Swingle and Shell 1971). Comparisons can also be made among populations to evaluate temporal and spatial differences or to evaluate influences of factors (such as parasites) that affect portions of populations (in effect, creating two populations: a population of affected individuals and a population of unaffected individuals; Box 10.6). Prentice (1987) used ANCOVA to test differences in species-specific weight–length relations among river systems and ecological regions within the state of Texas. He found differences among river systems and ecological regions for all species assessed. He also found differences between genders for many of the species he assessed. If a common currency is used to assess condition (such as W_r), comparisons can also be made among species. Condition indices can also indicate changes in environment and ecological processes (e.g., Gabelhouse 1991; Hubert et al. 1994; Liao et al. 1995). Finally, condition assessments are often important in manipulative studies to determine if treatments affect condition.

10.7 CONCLUSION

Condition data have been and will continue to be an important component of ecological assessment in aquatic systems. When combined with other information

Box 10.6 Use of Fulton's Condition to Assess the Effects of Parasites

Parasites may negatively affect the condition of fish. Here we determine if condition of Arkansas River shiners (29–60 mm TL) is reduced when fish are parasitized by anchor worm, a cosmopolitan cyclopoid copepod. Arkansas River shiners were captured with a seine (see Hayes et al. [1996] for a discussion of this gear), measured (TL; mm), weighed (0.1 g), and inspected to determine the presence of the parasite (partial data set from Durham et al. 2002). Differences in condition among fish with and without the parasite were assessed using ANOVA to test differences in Fulton's condition (*K*), an appropriate assessment metric as the fish are from a single population over identical size ranges. Individual fish from this experiment were treated as the experimental unit because our research question asked if differences in condition existed between two populations of Arkansas River shiners (population PRESENT contained parasites and population ABSENT contained no parasites).

Table Total length (TL; mm) and weight (WT; g) of Arkansas River shiners with and without anchor worm.

With parasite				Without parasite			
TL	WT	TL	WT	TL	WT	TL	WT
29	0.210	46	0.707	29	0.175	45	0.654
29	0.187	46	0.656	31	0.254	45	0.710
34	0.286	47	0.810	31	0.228	46	0.757
35	0.356	48	0.813	31	0.201	46	0.828
38	0.420	48	0.697	31	0.219	47	0.788
38	0.460	48	0.624	32	0.269	47	0.833
39	0.448	49	0.962	33	0.278	48	0.940
39	0.252	49	0.778	35	0.356	49	0.986
40	0.514	51	1.136	36	0.356	51	1.097
42	0.555	52	1.216	39	0.478	51	1.105
43	0.412	53	0.903	39	0.505	51	1.063
44	0.589	53	1.388	40	0.505	52	1.158
44	0.664	55	0.996	40	0.604	52	1.273
45	0.739	56	1.065	41	0.535	57	1.573
45	0.646	60	1.081	43	0.610	60	1.686

Program

The following SAS program provides output to compute length and weight summary statistics. Differences in condition were tested using ANOVA to test differences in Fulton's condition (*K*).

```
OPTIONS PS=54 LS=75;
PROC FORMAT;
VALUE PARACODE 0='ABSENT' 1='PRESENT';
DATA PARASITE;
INPUT TL WT PARASITE @@;
LOGTL=LOG10(TL);
LOGWT=LOG10(WT);
K=(WT/(TL*TL*TL))*100000;
FORMAT PARASITE PARACODE.;
CARDS;
```

(Box continues)

Box 10.6 *(continued)*

```
29   0.21    1   290.175  0
29   0.187   1   310.254  0
[Input complete data set];
PROC SORT; BY PARASITE;
PROC MEANS MEAN STDERR; BY PARASITE; VAR TL WT K;
PROC REG; BY PARASITE; MODEL LOGWT=LOGTL;
PROC ANOVA; CLASS PARASITE; MODEL TL=PARASITE;
PROC ANOVA; CLASS PARASITE; MODEL WT=PARASITE;
PROC ANOVA; CLASS PARASITE; MODEL K=PARASITE;
RUN;
```

Output

Table Descriptive statistics.

	Parasite absent		Parasite present	
Variable	Mean	SE	Mean	SE
TL	42.6000000	1.5585139	44.8333333	1.3838837
WT	0.7008000	0.0747452	0.6857133	0.0566121
K	0.8020460	0.0102421	0.7195553	0.0220817

Table Regression analysis of $\log_{10}$ transformed weight (LOGWT) on $\log_{10}$ transformed length (LOGTL) in the absence of parasite.

Analysis of Variance

Source	*df*	SS	Mean square	*F*-value	$P > F$
Model	1	2.27177	2.27177	2782.11	<0.0001
Error	28	0.02286	0.00081656		
Corrected total	29	2.29463			

r^2	0.9900	Root MSE	0.02858
Adjusted r^2	0.9897	Dependent mean	–0.23476
CV	–12.17232		

Parameter Estimates

Variable	*df*	Parameter estimate	SE	*t*-value	$P > \|t\|$
Intercept	1	–5.30889	0.09634	–55.11	<0.0001
LOGTL	1	3.13084	0.05936	52.75	<0.0001

Table Regression analysis of LOGWT on LOGTL in the presence of parasite.

Analysis of Variance					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	1.29852	1.29852	215.29	<0.0001
Error	28	0.16888	0.00603		
Corrected total	29	1.46740			
r^2	0.8849	Root MSE	0.07766		
Adjusted r^2	0.8808	Dependent mean	–0.21406		
CV	–36.28050				

Parameter Estimates					
Variable	*df*	Parameter estimate	SE	*t*-value	*P* > |*t*|
Intercept	1	–4.69465	0.30570	–15.36	<0.0001
LOGTL	1	2.72350	0.18562	14.67	<0.0001

Table An ANOVA to test for differences in TL, WT, and *K* in the presence versus absence of the parasite ($n = 60$).

Class Level Information		
Class	Levels	Values
PARASITE	2	ABSENT PRESENT

ANOVA for Total Length					
Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	74.816667	74.816667	1.15	0.2884
Error	58	3779.366667	65.161494		
Corrected total	59	3854.183333			
R^2	0.019412	Root MSE	8.072267		
CV	18.46496	TL mean	43.71667		

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Parasite	1	74.81666667	74.81666667	1.15	0.2884

(Box continues)

Box 10.6 *(continued)*

ANOVA for Weight

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	0.00341411	0.00341411	0.03	0.8727
Error	58	7.64884655	0.13187666		
Corrected total	59	7.65226067			
R^2	0.000446	Root MSE	0.363148		
CV	52.38295	WT mean	0.693257		

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Parasite	1	0.00341411	0.00341411	0.03	0.8727

ANOVA for *K*

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	1	0.10207080	0.10207080	11.48	0.0013
Error	58	0.51547644	0.00888752		
Corrected total	59	0.61754724			
R^2	0.165284	Root MSE	0.094274		
CV	12.39138	*K* mean	0.760801		

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Parasite	1	0.10207080	0.10207080	11.48	0.0013

Interpretation

Mean ± SE TL, WT, and *K* values for Arkansas River shiners (20–60 mm TL) not parasitized were 42.6 ± 1.6, 0.70 ± 0.07, and 0.80 ± 0.01, respectively. Mean ± SE total TL, WT, and *K* values for Arkansas River shiners (20–60 mm TL) parasitized by anchor worm were 44.8 ± 1.4, 0.69 ± 0.06, and 0.72 ± 0.02, respectively. When analyzed separately (ANOVA), no differences were found in length ($P = 0.29$) or weight ($P = 0.87$) of Arkansas River shiners with and without anchor worm ($P > 0.28$). However, differences ($P = 0.001$) were noted when *K* was assessed. Thus, it appears that parasitism by anchor worm causes condition to decrease in Arkansas River shiners. Note that visual examination of data (see figure below) suggests that about one-third of Arkansas River shiners parasitized by anchor worm have suppressed condition values, suggesting to us that about one-third of the Arkansas River shiners collected with anchor worm had been parasitized for a relatively long period (long enough to decrease condition), whereas the other two-thirds had been recently parasitized. This interpretation is not possible from the statistical assessment and illustrates the need to examine data visually.

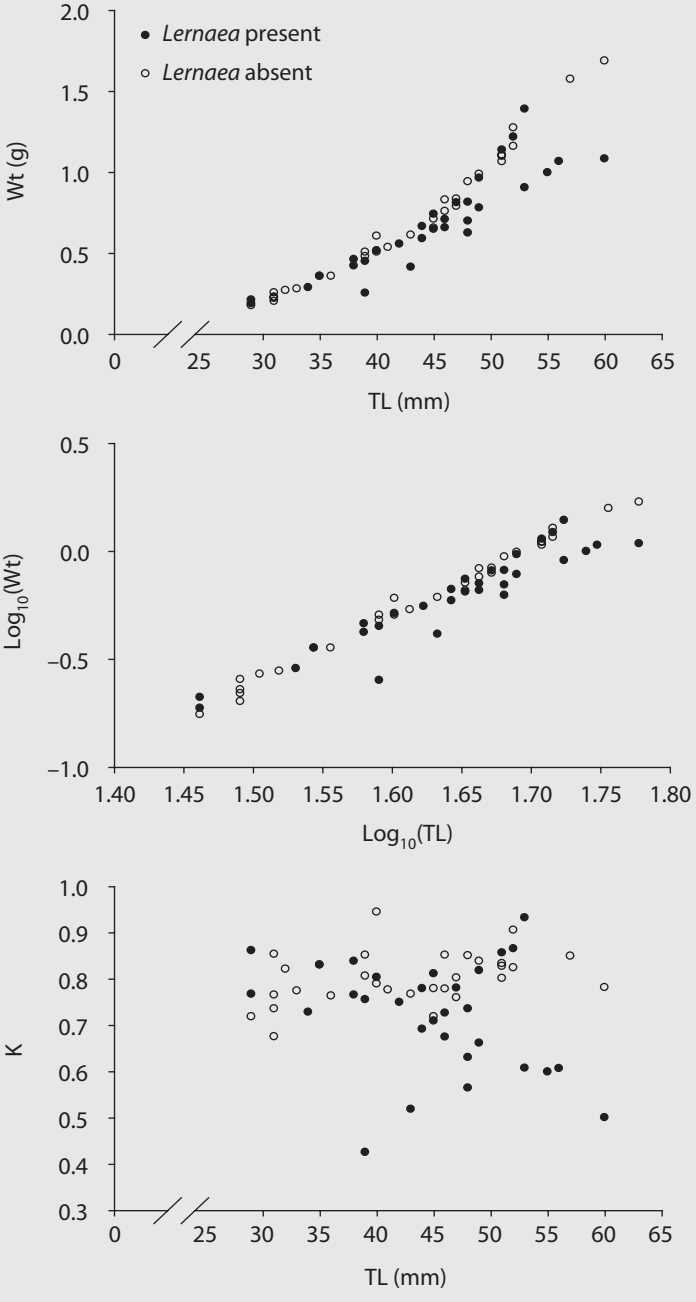


Figure Graphical depiction of relationship between weight and length (top panel), $\log_{10}$ transformed weight and $\log_{10}$ transformed length (middle panel), and Fulton's condition (*K*) and length (bottom panel) for Arkansas River shiners with and without anchor worm *Lernaea cyprinacea*.

(e.g., density, prey availability, size structure, community composition, and exploitation), condition data provide fisheries scientists a more complete understanding of population dynamics (recruitment, growth, and mortality) and environmental influences. Several techniques have been used to assess fish condition, and it is clear that there is much debate regarding the most appropriate way to analyze and present condition data, mostly centered on statistical shortcomings of analysis techniques. Appropriately, analytical techniques continue to evolve, as demonstrated by the most recent critique of W_r provided by Brenden et al. (2003).

Because of the relative ease of computation and use, the popularity of condition indices will continue to increase. Condition indices offer fisheries scientists a tool to evaluate effects of various management strategies and, indirectly, ecological interactions in fish populations and communities (Murphy and Willis 1991). More research is necessary to determine both the statistical appropriateness and relativity (to proximate factors and other expressions of fitness) of the condition measure. However, it is apparent that condition indices are useful for assessing fish condition (Blackwell et al. 2000).

Given the limitations discussed herein, controversy about assessment of condition will likely continue as fisheries scientists attempt to separate effects of fish condition from effects of fish size. Detailed assessments of various measures of fish condition that are tested with multiple statistical analyses will provide a clearer picture of relationships among measures of condition and help clarify the usefulness and shortcomings of various techniques. In the meantime, morphometric assessments of condition can be assessed appropriately using graphical display of data in a bivariate plot and ANCOVA with length as a covariate. Further, ratios can be used for descriptive purposes.

10.8 REFERENCES

Adams, S. M., and R. B. McLean. 1985. Estimation of largemouth bass, *Micropterus salmoides* Lacepede, growth using the liver somatic index and physiological variables. Journal of Fish Biology 26:111–126.

Agresti, A., and B. Finlay. 1986. Statistical methods for the social sciences, 2nd edition. Dellen Publishing, San Francisco.

Anderson, R. O., and S. J. Gutreuter. 1983. Length, weight, and associated structural indices. Pages 283–300 *in* L. A. Nielsen and D. L. Johnson, editors. Fisheries techniques. American Fisheries Society, Bethesda, Maryland.

Anderson, D. E., and R. Lydic. 1977. On the effect of using ratios in the analysis of variance. Biobehavioral Reviews 1:225–229.

Anderson, R. O., and R. M. Neumann. 1996. Length, weight, and associated structural indices. Pages 447–482 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Atchley, W. R. 1978. Ratios, regression intercepts, and the scaling of data. Systematic Zoology 27:78–83.

Atchley, W. R., and D. Anderson. 1978. Ratios and the statistical analysis of biological data. Systematic Zoology 27:71–78.

Atchley, W. R., C. T. Gaskins, and D. Anderson. 1976. Statistical properties of ratios. I. Empirical results. Systematic Zoology 25:137–148.

Blackwell, B. G., M. L. Brown, and D. W. Willis. 2000. Relative weight *(W_r)* status and current use in fisheries assessment and management. Reviews in Fisheries Science 8:1–44.

Bolger, T., and P. L. Connolly. 1989. The selection of suitable indices for the measurement and analysis of fish condition. Journal of Fish Biology 34:171–182.

Booth, D. J., and J. A. Keast. 1986. Growth energy partitioning by juvenile bluegill sunfish, *Lepomis macrochirus* Rafinesque. Journal of Fish Biology 38:37–45.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B (Methodological) 26:211–243.

Brenden, T. O., B. R. Murphy, and J. B. Birch. 2003. Statistical properties of the relative weight *(W_r)* index and an alternative procedure for testing W_r differences between groups. North American Journal of Fisheries Management 23:1136–1151.

Brown, M. L. 1993. Temporal genetic structure and energy dynamics of an intergrade largemouth bass population. Doctoral dissertation. Texas A&M University, College Station.

Brown, M. L., and D. J. Austin. 1996. Data management and statistical techniques. Pages 17–61 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Brown, M. L., and B. R. Murphy. 1991. Relationship of relative weight *(W_r)* to proximate composition of juvenile striped bass and hybrid striped bass. Transactions of the American Fisheries Society 120:509–518.

Bulow, F. J., M. E. Zeman, J. R. Winningham, and W. F. Hudson. 1981. Seasonal variations in RNA–DNA ratios and in indicators of feeding, reproduction, energy storage, and condition in population of bluegill. Journal of Fish Biology 18:237–244.

Busacker, G. P., I. A. Adelman, and E. M. Goolish. 1990. Growth. Pages 363–377 *in* C. B. Shreck and P. B. Moyle, editors. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Carlander, K. D. 1969. Handbook of freshwater fisheries biology, volume 1. Iowa State University Press, Ames.

Carmer, S. G., and M. R. Swanson. 1973. An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. Journal of the American Statistical Association 68:66–74.

Childress, W. M. 1991. Standard-weight curves derived from Texas length and weight data. North American Journal of Fisheries Management 11:480–482.

Cone, R. S. 1989. The need to reconsider the use of condition indices in fishery science. Transactions of the American Fisheries Society 118:510–514.

Craig, J. F. 1977. The body composition of adult perch, *Perca fluviatilis* in Windermere, with reference to seasonal changes and reproduction Journal of Animal Ecology 46: 617–632.

Delahunty, G., and V. L. de Vlaming. 1980. Seasonal relationships of ovary weight, liver weight and fat stores with body weight in the goldfish, *Carassius auratus* (L.). Journal of Fish Biology 16:5–13.

Durham, B. W., T. H. Bonner, and G. R. Wilde. 2002. Occurrence of *Lernaea cyprinacea* on Arkansas River shiners and peppered chubs in the Canadian River, New Mexico and Texas. Southwestern Naturalist 47:95–98.

Elliott, J. M. 1976. Body condition of brown trout (*Salmo trutta* L.) in relation to temperature and ration size. Journal of Animal Ecology 45:273–289.

Fechhelm, R. G., W. B. Griffiths, W. J. Wilson, B. J. Gallaway, and J. D. Bryan. 1995. Intra- and interseasonal changes in the relative condition and proximate body composition of broad whitefish from the Prudhoe Bay region of Alaska. Transactions of the American Fisheries Society 124:508–519.

Ferron, A., and W. C. Leggett. 1994. An appraisal of condition measures for marine fish larvae. Advances in Marine Biology 30:217–303.

Fisher, S. J., D. W. Willis, and K. L. Pope. 1996. An assessment of burbot (*Lota lota*) weight–length data from North American popuplations. Canadian Journal of Zoology 74: 570–575.

Fitzgerald, D. G., J. W. Nanson, T. N. Todd, and B. M. Davis. 2002. Application of truss analysis for the quantification of changes in fish condition. Journal of Aquatic Ecosystem Stress and Recovery 9:115–125.

Flickinger, S. A., and F. J. Bulow. 1993. Small impoundments. Pages 469–492 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America. American Fisheries Society, Bethesda, Maryland.

Fulton, T. W. 1904. The rate of growth of fishes. Twenty-Second Annual Report to the Fishery Research Board for Scotland 3:141–241.

Gabelhouse, D. W., Jr. 1984. A length-categorization system to assess fish stocks. North American Journal of Fisheries Management 4:273–285.

Gabelhouse, D. W., Jr. 1991. Seasonal changes in body condition of white crappies and relations to length and growth in Melvern Reservoir, Kansas. North American Journal of Fisheries Management 11:50–56.

Gerow, K. G., W. A. Hubert, and R. C. Anderson-Sprecher. 2004. An alternative approach to detection of length-related biases in standard weight equations. North American Journal of Fisheries Management 24:903–910.

Gutreuter, S., and W. M. Childress. 1990. Evaluation of condition indices for estimation of growth of largemouth bass and white crappies. North American Journal of Fisheries Management 10:434–441.

Guy, C. S., and D. W. Willis. 1991. Seasonal variation in catch rate and body condition for four fish species in a South Dakota natural lake. Journal of Freshwater Ecology 6: 281–292.

Håkanson, J. L. 1989. Analysis of lipid components for determining the condition of anchovy larvae *(Engraulis mordax)*. Marine Biology 102:143–151.

Hayes, D. B., C. P. Ferreri, and W. W. Taylor. 1996. Active fish capture methods. Pages 193—220 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Heidinger, R. C., and S. D. Crawford. 1977. Effect of temperature and feeding rate on the liver–somatic index of the largemouth bass, *Micropterus salmoides*. Journal of the Fisheries Research Board of Canada 34:633–638.

Hubert, W. A., R. D. Gipson, and R. A. Whaley. 1994. Interpreting relative weights of lake trout stocks. North American Journal of Fisheries Management 14:212–215.

Humphries, J. M., F. L. Bookstein, B. Chernoff, G. R. Smith, R. L. Elder, and S. G. Poss. 1981. Multivariate discrimination by shape in relation to size. Systemic Zoology 30:291–308.

Hyatt, M. W., and W. A. Hubert. 2001. Statistical properties of relative weight distributions of four salmonid species and their sampling implications. North American Journal of Fisheries Management 21:666–670.

Jakob, E. M., S. D. Marshal, and G. W. Uetz. 1996. Estimating fitness: a comparison of body condition indices. Oikos 77:61–67.

Jensen, A. J. 1979. Energy content analysis from weight and liver index measurements of immature pollock *(Pollachius virens).* Journal of the Fisheries Research Board of Canada 36:1207–1213.

Jones, R. E., R. J. Petrell, and D. Pauly. 1999. Using modified length–weight relationships to assess the condition of fish. Aquaculture Engineering 20:261–276.

Keys, A. B. 1928. The weight–length relation in fishes. Proceedings of the National Academy of Sciences 14:922–925.

Kohler, C. C., and A. M. Kelly. 1991. Assessing predator–prey balance in impoundments. Pages 257–260 *in* J. L. Cooper and R. H. Hamre, editors. Warmwater fisheries symposium I. U.S. Forest Service General Technical Report RM-207, Fort Collins, Colorado.

Kruse, C. G., and W. A. Hubert. 1997. Proposed standard weight equations for interior cutthroat trout. North American Journal of Fisheries Management 17:784–790.

Le Cren, E. D. 1951. The length–weight relationship and seasonal cycle in gonad weight and condition in the perch *(Perca fluviatilis).* Journal of Animal Ecology 20:201–219.

Liao, H., C. L. Pierce, D. H. Wahl, J. B. Rasmussen, and W. C. Leggett. 1995. Relative weight *(W_r)* as a field assessment tool: relationships with growth, prey biomass, and environmental conditions. Transactions of the American Fisheries Society 124:387–400.

Martin, W. R. 1949. The mechanics of environmental control on body form in fishes. Publications of the Ontario Fisheries Research Laboratory 70:1–91.

Marwitz, T. D., and W. A. Hubert. 1997. Trends in relative weight of walleye stocks in Wyoming reservoirs. North American Journal of Fisheries Management 17:44–53.

Moore, S. A., and C. R. Bronte. 2001. Delineation of sympatric morphotypes of lake trout in Lake Superior. Transactions of the American Fisheries Society 130:1233–1240.

Murphy, B. R., M. L. Brown, and T. A. Springer. 1990. Evaluation of the relative weight index with new applications to walleye. North American Journal of Fisheries Management 10:85–97.

Murphy, B. R., and D. W. Willis. 1991. Application of relative weight *(W_r)* to western warmwater fisheries. Pages 243–248 *in* J. L. Cooper and R. H. Hamre, editors. Warmwater fisheries symposium I. U. S. Forest Service General Technical Report RM-207, Fort Collins, Colorado.

Murphy, B. R., D. W. Willis, and T. A. Springer. 1991. The relative weight index in fisheries management: status and needs. Fisheries (Bethesda) 16(2):30–38.

Neter, J., W. Wasserman, and M. H. Hunter. 1989. Applied linear regression models. Irwin Publishing, Boston.

Neumann, R. M., and B. R. Murphy. 1991. Evaluation of the relative weight index for assessment of white crappie and black crappie populations. North American Journal of Fisheries Management 11:543–555.

Neumann, R. M., and B. R. Murphy. 1992. Seasonal relationships of relative weight to body composition in white crappie, *Pomoxis annularis* Rafinesque. Aquaculture and Fisheries Management 23:243–251.

Neumann, R. M., and D. W. Willis. 1994. Relative weight as a condition index for muskellunge. Journal of Freshwater Ecology 9:13–18.

Neumann, R. M., and D. W. Willis. 1995. Seasonal variation in gill net sample indexes for northern pike collected from a glacial prairie lake. North American Journal of Fisheries Management 15:838–844.

Ney, J. J. 1993. Practical use of biological statistics. Pages 137–158 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management of North America. American Fisheries Society, Bethesda, Maryland.

Novinger, D. C., and C. Martinez Del Rio. 1999. Failure of total body electrical conductivity to predict lipid content of brook trout. North American Journal of Fisheries Management 19:942–947.

Packard, G. C., and T. J. Boardman. 1988. The misuse of ratios, indices, and percentages in ecophysiological research. Physiological Zoology 61:1–9.

Patterson, K. R. 1992. An improved method for studying the condition of fish, with an example using Pacific sardine, *Sardinops sagax* (Jenyns). Journal of Fish Biology 40:821–831.

Pope, K. L., and D. W. Willis. 1996. Seasonal influences on freshwater fisheries sampling data. Reviews in Fisheries Science 4:57–73.

Porath, M. T., and E. J. Peters. 1997. Use of relative weight *(W_r)* to assess prey availability. North American Journal of Fisheries Management 17:628–637.

Prentice, J. A. 1987. Length–weight relationships and average growth rates of fishes in Texas. Texas Parks and Wildlife Department, Inland Fisheries Data Series 6, Austin.

Pulliainen, E., and K. Korhonen. 1990. Seasonal changes in condition indices in adult mature and non-maturing burbot, *Lota lota* (L.), in the northeastern Bothnian Bay, northern Finland. Journal of Fish Biology 36:251–259.

Ratz, H. J., and J. Lloret. 2003. Variation in fish condition between Atlantic cod *(Gadus morhua)* stocks, the effect on their productivity and management implications. Fisheries Research 60:369–380.

Reist, J. D. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. Canadian Journal of Zoology 63:1429–1439.

Richter, H., C. Luckstadt, U. Focken, and K. Becker. 2000. An improved procedure to assess fish condition on the basis of length–weight relationships. Archive of Fishery Marine Research 48:255–264.

Rose, C. J. 1989. Relationship between relative weight *(W_r)* and body composition in immature walleye. Master's thesis. Texas A&M University, College Station.

SAS Institute. 1998. SAS/Stat user's guide, version 6: volumes 1 and 2. SAS Institute, Cary, North Carolina.

Shackley, P. E., C. Talbot, A. Cowan, and A. Watt. 1994. The use of body water, sodium, potassium and calcium content to investigate the nutritional status of first year Atlantic salmon parr in two Scottish Highlands streams. Journal of Fish Biology 44:693–706.

Smith, C. D., C. L. Higgins, G. R. Wilde, and R. E. Strauss. 2005. Development of a morphological index to the nutritional status of juvenile largemouth bass. Transactions of the American Fisheries Society 134:120–125.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. Freeman, New York.

Spencer, H. 1898. The principles of biology, volume 2. Williams and Norgate, London.

Steel, R. G. D., and J. H. Torrie. 1980. Principles and procedures of statistics a biometrical approach, 2nd edition. McGraw-Hill Publishing, New York.

Strange, R. J., and J. C. Pelton. 1987. Nutrient content of clupeid forage fishes. Transactions of the American Fisheries Society 116:60–66.

Strauss, R. E., and F. L. Bookstein. 1982. The truss: body form reconstruction in morphometrics. Systematic Zoology 31:113–135.

Stone, N., E. McNulty, and E. Park. 2003. The effect of stocking and feeding rates on growth and production of feeder goldfish in pools. North American Journal of Aquaculture 65:82–90.

Suthers, I. M. 1992. The use of condition indices in larval fish. Pages 49–58 *in* D. A. Hancock, editor. Larval biology. Australian Society for Fish Biology Workshop, Australian Government Publishing Service, Bureau of Rural Resources Proceedings 15, Canberra.

Sutton, S. G., T. P. Bult, and R. L. Haedrich. 2000. Relationships among fat weight, body weight, water weight, and condition factors in wild Atlantic salmon parr. Transactions of the American Fisheries Society 129:527–538.

Swingle, W. E., and E. W. Shell. 1971. Tables for computing relative conditions of some common freshwater fishes. Auburn University Agricultural Experiment Station, Circular 183, Auburn, Alabama.

Tanner, J. M. 1949. Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. Journal of Applied Physiology 2:1–15.

Toline, C. A., and A. J. Baker. 1993. Foraging tactic as a potential selection pressure influencing geographic differences in body shape among populations of dace *(Phoxinus eos)*. Canadian Journal of Zoology 71:2178–2184.

Vila-Gispert, A., and R. Moreno-Amich. 2001. Fish condition analysis by a weight least squares procedure: testing geographical differences of an endangered Iberian cyprinodontid. Journal of Fish Biology 58:1658–1666.

Ward, D. L., and M. P. Zimmerman. 1999. Response of smallmouth bass to sustained removals of northern pikeminnow in the lower Columbia and Snake rivers. Transactions of the American Fisheries Society 128:1020–1035.

Wege, G. J., and R. O. Anderson. 1978. Relative weight (W_r): a new index of condition for largemouth bass. Pages 79–91 *in* G. Novinger and J. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, North Central Division, Special Publication 5, Bethesda, Maryland.

Willis, D. W. 1989. Proposed standard length–weight equation for northern pike. North American Journal of Fisheries Management 9:203–208.

Willis, D. W., C. S. Guy, and B. R. Murphy. 1991. Development and evaluation of a standard weight (W_s) equation for yellow perch. North American Journal of Fisheries Management 11:374–380.

Wootton, R. J. 1990. Ecology of teleost fish. Chapman and Hall, London.

Zar, J. H. 1984. Biostatistical analysis, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

11 Assessment of Diets and Feeding Patterns

Steven R. Chipps and James E. Garvey

11.1 INTRODUCTION

Quantitative assessment of food habits is an important aspect of fisheries management. Successful management of sport fishes often hinges on our ability to manage prey resources (Noble 1981; DeVries and Stein 1990). As a result, knowledge of prey resources can help guide management efforts aimed at increasing fish production. Accurate description of fish diets and feeding habits also provides the basis for understanding trophic interactions in aquatic food webs (Garvey et al. 1998a; Vander Zanden et al. 2000). Diet composition analysis or other techniques, such as stable isotope analysis, can be used to evaluate effects of ontogeny, habitat, or the establishment of exotic species.

Diets of fishes represent an integration of many ecological components that include behavior, condition, habitat use, energy intake and inter- and intraspecific interactions. As a result, food habit studies can be incorporated in a variety of different research objectives. In the simplest case, a food habits study might be conducted to determine the most frequently consumed prey or determine whether a particular food category is present in the stomach of fishes. In other instances, we may be interested in more complex questions, such as (1) determining the relative importance of different food types to fish nutrition, (2) quantifying the consumption rate of individual prey types (Chapter 12), or (3) understanding foraging trade-offs associated with predator avoidance (Chapter 16). Each of these questions requires information on fish diets but necessitates different approaches in how we collect and analyze data. In this chapter, we outline quantitative techniques used to describe food habits and feeding patterns of fishes.

11.2 QUANTIFYING DIET COMPOSITION

11.2.1 Sampling and Identifying Stomach Contents

Most studies of fish diets rely on examination of stomach contents to quantify prey abundance. This information characterizes foraging choices made over a relatively short time scale (e.g., usually <24 h). Hence, time of day, sampling location, prey

availability, and even the type of collecting gear used need to be considered before initiating a diet study or analyzing diet data. Investigators using historical diet samples or processed data must be aware of the sampling protocols, laboratory procedures, and preservation techniques used. Failure to understand how diet data were collected may preclude accurate interpretation of foraging patterns.

Stomach contents can be collected from live fish by means of the lavage technique (Seaburg 1957), whereby food items are flushed from the stomachs by use of pressurized water. Similarly, emetics can be used to induce regurgitation in live fish (Jernejcic 1969; see Bowen 1996 for review). Regardless of the method, investigators should ensure that the removal technique effectively samples all items in the gut. Otherwise, data will be skewed toward items that are more easily displaced from the stomach. Alternatively, fish can be sacrificed and stomach contents removed for analysis. If fish are to be sacrificed, they should be preserved immediately either by freezing or by fixing in formalin (Bowen 1996). Stomach contents will continue to digest, rendering rapid preservation of the fish or removed contents necessary to prevent loss of resolution. Various taxa digest at different rates (Sutela and Huusko 2000; Kim and DeVries 2001). As such, recently consumed taxa may be present in the foregut, but only resistant items remain in the hindgut. Investigators must consider the relative digestibility of prey when deciding on the section of the alimentary tract to sample. To avoid bias when both easily digested prey and resistant prey are present, only the immediate foregut (i.e., stomach) should be sampled (Sutela and Huusko 2000).

Prey items in fish stomachs are often not intact. Otoliths or other relatively indigestible hard parts, such as scales, pharyngeal teeth, cleithra, or backbones, have diagnostic, species-specific characteristics useful for identifying prey (Garman 1982; Holland-Bartels et al. 1990). Alternatively, partially digested prey may be identified using biochemical signatures, such as allozyme electrophoresis (Hartman and Garton 1992), immunoassays (Feller 1992; Schultz and Clarke 1995), or promising new techniques based on fatty acid analysis (Raclot et al. 1998).

Hard structures are often used to determine lengths or weights of prey items by regressing the dimension of an indigestible hard part (e.g., head capsule of an insect) against whole-body length or mass (least-squares regression models; Trippel and Beamish 1987; Scharf et al. 1997). Combining back-calculated estimates in this fashion may compound error in estimates of total prey weight (or volume). Thus, it is imperative that biometric relationships and measurements of hard parts used to reconstruct diet items are precise and not biased.

The proper taxonomic resolution for identifying stomach contents largely depends on the research question. Coarse taxonomic resolution is appropriate when quantifying ontogenetic changes in diet composition. Presence of fish in the diet may prove adequate for determining the size or time at which fish switch to piscivory. In other instances, finer taxonomic resolution may be needed, such as determining seasonal or spatial differences in diet composition or comparing percent composition of native versus exotic species.

Often, it is pragmatic to reduce the number of variables involved in the analysis by pooling diet items into categories based on taxonomy or habitat. Three types

of data pooling can be considered for prey items in fish stomachs: (1) necessary, (2) intuitive, and (3) statistical (Crow 1979). Necessary pooling occurs when unidentified prey are present in stomachs. If three categories of fish prey and one category of unidentified fish prey arise, then we should consider either pooling fish prey or dropping the unidentified category. An analysis with both identified and unidentified fish may be misleading because we do not know what proportion of unidentified fish were components of items we could successfully identify (Crow 1979). Intuitive pooling is based on taxonomic or ecological similarities among prey. Three species of calanoid copepods might be pooled into a single category (e.g., copepods) given similar morphological and behavioral characteristics. Similarly, we could pool species by habitat so that categories represent benthic, pelagic, or littoral prey. Finally, statistical pooling uses quantitative statistical procedures as a basis for pooling prey categories. Here, the investigator hypothesizes that two or more prey categories act as a single resource (Crow 1979). This hypothesis is tested using a 2×2 contingency table to identify whether prey are either positively or negatively associated (Box 11.1). Positive association implies that prey are acting as a single resource and may be pooled (Crow 1979).

11.2.2 Designing Appropriate Sampling Designs

11.2.2.1 Conducting Field Studies

Feeding patterns of fishes may be quantified in the field or with carefully designed experiments. In either case, the sampling design should be well considered before data are collected. As with other field studies, appropriate sampling designs for diet analysis include (1) simple random sampling, (2) stratified random sampling, (3) systematic sampling, and (4) multistage sampling (see Chapter 3). The choice of a particular sampling design depends on a variety of factors that include the research question, logistics, accessibility, and costs.

Prior to collecting diet data, attention should be given to factors that influence the quantity and quality of stomach contents. One important consideration in diet studies is that foraging behavior of fishes often varies with time of day (Shepard and Mills 1996). Hence, sampling plans should incorporate a diel component to determine how stomach contents change through time. Failure to standardize measurement times among sites or lakes may lead to erroneous conclusions about foraging patterns (see section 11.3.4).

Moreover, sampling approaches, such as electrofishing or gillnetting, may cause loss of stomach contents through regurgitation (Bowen 1996). Similarly, high-speed tow nets can eviscerate larval fish resulting in a loss of information (K. Arend, Ohio State University, personal communication). The use of active or passive gear types can also affect inferences about stomach fullness. Fish collected with passive gears can have more food in their stomachs than do fishes collected with active gears because passive gears often collect actively feeding fish (Hayward et al. 1989). Careful consideration should be given to sampling time and gear type to help reduce variability among samples.

Box 11.1 Pooling Prey Items as a Single Resource

Prey items in fish stomachs can sometimes be pooled prior to analysis. To determine whether two (or more) prey items act as a single resource, we can use chi-square contingency table analysis. In the example below, we are interested in whether prey *i* and prey *j* can be pooled prior to analysis.

We construct a 2 × 2 contingency table by totaling the number of fish that contain both prey types, either prey *i* or prey *j*, or neither prey type in the diet. In this example, the diets of 70 fish have been examined for prey *i* and prey *j*.

	Prey *j* present	Prey *j* absent	Total
Prey *i* present	18	9	27
Prey *i* absent	18	25	43
Total	36	34	70

Resulting output was obtained using the PROC FREQ procedure in SAS (SAS Institute1999). Here, the likelihood ratio chi-square value (*G*-statistic), 4.145, is larger than the critical value of a chi-square distribution (i.e., 1 df, $P = 0.041$), implying that the prey are either positively or negatively associated. Prey can be pooled only if they are positively associated. To determine association (*A*), we calculate the cross-product ratio of the contingency table as

$$A = (\text{cell } 11)(\text{cell } 22)/(\text{cell } 12)(\text{cell } 21)$$
$$= (18)(25)/(18)(9) = 2.8.$$

If *A* is greater than 1, then prey types are positively associated, implying that they are acting as a single resource and can be pooled (Crow 1979). If *A* were less than 1, then prey types would be negatively associated and should not be pooled.

Before initiating a field study, it is desirable to know how many samples are needed to describe the diet. Cumulative prey curves are useful for determining when a sufficient number of stomachs have been sampled. In this approach, the cumulative number of prey types is plotted against the cumulative number of pooled stomachs (Cortés 1997). The point at which the curve becomes asymptotic provides a minimum number of stomachs needed to characterize prey composition.

It is not uncommon to find empty stomachs. However, investigators must be cautious about how increasing sampling effort to find fish containing food affects their estimates. To our knowledge, the impact of this practice remains unexplored. Presumably, greater sample sizes arising when empty guts are frequent would affect variance estimates relative to other samples. Often, investigators restrict their analyses to the subset of individuals containing diet items (i.e., dropping individuals with empty guts) to explore diet preference (see section 11.3.7). This practice also must be approached cautiously. Diet characteristics of fish populations for which empty stomachs were frequent may be quite different than those for which empty stomachs were rare.

When sampling diets in the field, large numbers of fish are typically encountered, requiring subsampling across sizes. Stratifying samples as a function of body size is important because size often affects both the quantity and composition of items within diets (Schael et al. 1991; Bremigan and Stein 1994). The number of subsamples taken can be stratified by the relative proportion of individuals within each size-class or as some set number of individuals per size-class. Subsamples taken randomly in proportion to the actual number within each size-class reasonably reflect size-based patterns within the whole sample (Kimura 1977). However, this sampling design may poorly represent the diets of the largest individuals within the population, which are infrequently encountered. To remedy this, many sampling designs incorporate the set number per size-class approach. In a similar example with length-at-age data, Bettoli and Miranda (2001) demonstrated how simply pooling data from such a stratified sampling distribution poorly reflects the distribution within the overall sample. As such, extrapolating diet patterns within each size-class to that of the whole sample requires weighting the stratified diet data by the relative proportion of individuals within each size-class.

11.2.2.2 Conducting Experimental Studies

Food can be a limiting resource to fish populations. As a result, we are often interested in how competition for prey affects foraging success. Field data on stomach contents are inadequate to address competition questions. Rather, competition studies are generally performed under controlled experimental settings. Several approaches can be used to determine whether one species affects the foraging behavior of another. In all cases, it is important that densities of species be manipulated within the range of those in the environment to determine how variation in abundance affects competition. Three approaches are generally recognized in competition experiments: (1) substitutive, (2) density-gradient, and (3) response surface experiments (Goldberg and Scheiner 2001). The substitutive experimental approach involves varying treatment levels by substituting individuals of one species with an equal number of the other (Figure 11.1A). Total density is kept constant across all treatments. This approach tests for only the relative strengths of intra- and interspecific competition. The absolute magnitude of interspecific competitive effects is not isolated. A density-gradient approach involves holding the density of one focal species constant while varying that of another (Figure 11.1B). A problem with this approach is that foraging responses of the focal species are potentially confounded by an increase in frequency of the competitor and an overall increase in density (see Welker et al. 1994). A response surface experiment, which includes all density combinations of both competitors, avoids potential confounding effects but requires a large number of treatment combinations (Figure 11.1C). Clearly, designing an experiment to determine how competition affects foraging requires foresight about potential responses. The design of any experiment in which diet is a response variable requires careful consideration of the hypotheses being tested.

Field-derived patterns of foraging preference are by nature correlative. Only experiments definitively show how changes in food quality or quantity affect dietary

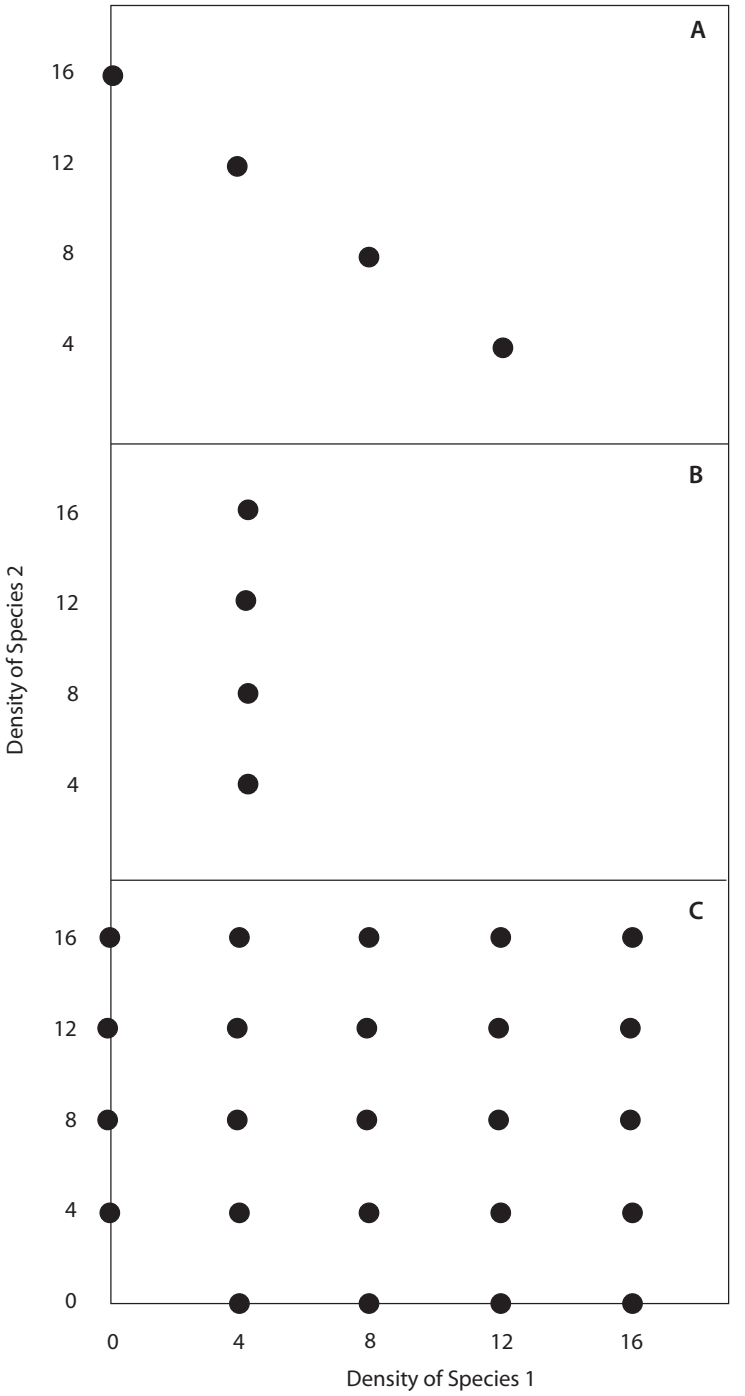


Figure 11.1 Three potential designs for competition experiments in which the impact of species 1 on the diet composition of species 2 is assessed. These experiments are **(A)** substitutive, **(B)** density gradient, and **(C)** response surface. Each point represents a single experimental treatment (adapted from Goldberg and Scheiner 2001).

choice. Size-dependent selection by larval fish for zooplankton prey in reservoirs (Bremigan and Stein 1997), selection of spot by piscivorous southern flounder in marine estuaries (Wright et al. 1993), and foraging preference for snails by pumpkinseed sunfish in natural lakes (Mittelbach et al. 1999) are examples of carefully designed experiments that provided insight into field dietary patterns. The outcome of foraging experiments such as these can be affected by many conditions. Using the same prey or predators across experimental trials may influence learning, which may cause foraging patterns to change through time (Reiriz et al. 1998). Using naïve consumers and prey in each replicate will remedy this confounding problem. Interactions among prey items within experimental units may cause different patterns of vulnerability relative to prey being exposed to consumers independently (Huang and Sih 1991). Similarly, changes in prey frequency as individuals are removed may influence their relative conspicuousness, thereby affecting selection patterns through time (Werner and Hall 1974). Hence, investigators may replenish prey throughout an experiment to keep densities as constant as feasible. Hunger levels of the consumer and prey and the size and realism of the experimental arena also may alter responses. In summary, foraging experiments can be insightful but also greatly misleading. Their design and interpretation must be carefully considered.

11.3 ANALYZING FOOD HABITS DATA

11.3.1 Selecting a Diet Index

A variety of measures have been used to quantify diet composition of fishes (Bowen 1996). Selecting an appropriate diet measure is strongly dependent on the research question; no single index is likely to provide a useful measure of prey importance under all conditions (Bowen 1996). For questions regarding the seasonal use of a prey resource, simple indices, such as frequency of occurrence, are usually adequate. Alternatively, we may want to quantify the energetic contribution of different prey types—a process that requires data on the abundance, weight, and caloric content of prey.

Traditional indices used for stomach content analysis include percent composition by number (N_i), percent composition by weight (W_i), and frequency of prey occurrence (O_i) (Bowen 1996; Table 11.1). It is important to recognize that each index emphasizes different information about the diet of fishes (Hyslop 1980; Cortés 1997). When evaluating percent composition by number, small prey can represent a dominant component of the diet. In contrast, percent composition by weight tends to emphasize the relative contribution of larger prey. Frequency of occurrence can provide information on how often a particular prey item was eaten but provides no indication of the relative importance of prey to the overall diet.

When calculated from the entire sample, N_i and W_i represent single measures with no corresponding variance estimate. If interest lies in evaluating the potential impact of predators on prey populations, then calculating N_i and W_i for the entire sample is appropriate. However, if diet data are to be used for statistical

Table 11.1 Equations for calculating diet indices (adapted from Pope et al. 2001). Symbols in equations are food or prey item (subscript *i*); fish (subscript *j*); number of fish (*J*); number of fish with food in their stomachs (*P*); number of fish containing prey *i* (J_i); number in food category *i* (N_i); number of food types (*Q*); weight of prey type *i* (W_i); weight of fish *j* (F_j); volume (mL) of food category *i* (V_i);. caloric density ($J \cdot g^{-1}$ wet weight) of food type *i* (X_i); and stomach capacity (mL) of fish *j* (C_j).

Diet index	Index symbol	Computational equation
Frequency of occurrence	O_i	$= \frac{J_i}{P}$
Proportion by number	N_i	$= \frac{N_i}{\sum_{i=1}^{Q} N_i}$
Proportion by weight	W_i	$= \frac{W_i}{\sum_{i=1}^{Q} W_i}$
Mean proportion by number	MN_i	$= \frac{1}{P}\sum_{j=1}^{P}\left(\frac{N_{ij}}{\sum_{i=1}^{Q} N_{ij}}\right)$
Mean proportion by weight	MW_i	$= \frac{1}{P}\sum_{j=1}^{P}\left(\frac{W_{ij}}{\sum_{i=1}^{Q} W_{ij}}\right)$
Mean proportion body weight	MBW_i	$= \frac{1}{P}\sum_{j=1}^{P}\left(\frac{W_{ij}}{F_j}\right)$
Mean stomach fullness	MSF_i	$= \frac{1}{P}\sum_{j=1}^{P}\left(\frac{V_{ij}}{C_j}\right)$
Prey importance index	PII_i	$= \frac{1}{P}\sum_{j=1}^{P}\left(\frac{W_{ij}X_i}{\sum_{i=1}^{Q} W_{ij}X_i}\right)$
Index of relative importance	IRI_i	$(\%N_i + \%W_i)(\%O_i)$
Relative importance index	RI_i	$\frac{100\mathrm{AI}_i}{\sum \mathrm{AI}_i}$, where $\mathrm{AI}_i = O_i + N_i + W_i$

comparisons then N_i and W_i should be calculated for individual fish and then averaged for each prey type (see MN_i and MW_i in Table 11.1). In this way, we treat individual fish as the sampling unit and assume that they represent a random sample (Table 11.2). Diet items in the stomachs of individual fish are not independent and generally should be measured to provide proportional data for individual fish (Hurlbert 1984; Krebs 1989).

When one is evaluating diet composition, prey weights are often more useful than are prey counts because weights are measured in comparable units. Consider

Table 11.2 Summary of prey weights for 10 bluegills. All weights are given in grams, and values in parentheses represent prey proportions for each fish. Mean proportion by weight (MW_i) and frequency of occurrence (O_i) are given in the last two rows. Note that dipteran larvae had the highest frequency of occurrence but contributed the least to the overall diet by weight—illustrating some of the problems associated with interpreting different diet measures.

		Prey weight and proportion				Total
Bluegill	Fish weight	Amphipods	Larval fish	Dipteran larvae	Mayfly nymphs	prey weight
A	150	0.3 (0.52)	0.24 (0.42)	0.02 (0.03)	0.016 (0.03)	0.576
B	91	0.09 (0.78)	0 (0)	0.018 (0.16)	0.008 (0.07)	0.116
C	99	0.11 (0.66)	0 (0)	0.024 (0.14)	0.032 (0.19)	0.166
D	123	0.03 (0.25)	0 (0)	0.052 (0.43)	0.04 (0.33)	0.122
E	210	0 (0)	0.12 (0.83)	0.001 (0.01)	0.024 (0.17)	0.145
F	102	0.22 (0.92)	0 (0)	0.003 (0.01)	0.016 (0.07)	0.239
G	124	0 (0)	0 (0)	0.006 (0.10)	0.056 (0.90)	0.062
H	199	0.015 (0.09)	0.12 (0.71)	0.003 (0.02)	0.032 (0.19)	0.170
I	101	0.45 (0.91)	0 (0)	0.015 (0.03)	0.032 (0.06)	0.497
J	111	0.26 (0.39)	0.36 (0.53)	0.054 (0.08)	0 (0)	0.674
MW_i		45%	25%	10%	20%	
O_i		80%	40%	100%	90%	

the difficulty in determining the relative importance of 1,500 zooplankton versus 1 fish. When measured as dry weight, we can directly compare 0.06 g of zooplankton to 0.2 g of fish in the diet. For this reason, prey weights are more appropriate when interest lies in comparing the energetic importance of different prey types (Bowen 1996). To correct for effects of fish size, it is often useful to express prey weight as a percentage of predator body mass.

Other indices used for diet analysis include mean stomach fullness and the prey importance index (Table 11.1). Early methods for measuring stomach fullness in fishes included subjective techniques such as the points method whereby food items were awarded points proportional to their estimated contribution to stomach volume (Swynnerton and Worthington 1940; Hynes 1950). Although easy to apply, these techniques have been criticized for their subjectivity (Windell and Bowen 1978). A more objective approach is to calculate the ratio of observed prey volume to estimated stomach capacity (Kimball and Helm 1971; Knight and Margraf 1982). Here, total volume of prey in each stomach is estimated either directly by water displacement or indirectly by means of geometric measurements. Maximum total prey volume is then regressed against fish size to estimate maximum stomach volume as

$$V = aL^b, \tag{11.1}$$

where V = maximum stomach capacity, a = regression coefficient, L = total length, and b = instantaneous rate of change (Knight and Margraf 1982). The ratio of observed prey volume (v) to maximum stomach volume (V) provides an index of

stomach fullness that accounts for fish length. The mean stomach fullness index (MSF_i) has several desirable advantages including it (1) eliminates subjectivity associated with the points method, (2) is relatively quick and easy to apply, (3) can be obtained from preserved or live fish, and (4) can be analyzed by a variety of statistical procedures (Knight and Margraf 1982). Furthermore, the MSF_i correlates well with prey caloric contribution, providing a robust index for evaluating the energetic contribution of different prey types (Pope et al. 2001).

The prey importance index (PII_i) combines information on the abundance, weight, and caloric content of prey (Table 11.1). Given sufficient information on prey assimilation efficiencies, the caloric densities of prey can be adjusted to account for energy actually metabolized by fishes (Probst et al. 1984). In most cases, this type of information is not readily available, so that total energy of prey is used. The usefulness of a caloric-based index such as the PII_i is that it provides a quantitative measure of the nutritional benefit of individual prey rather than relative importance based on numbers, weight, or occurrence in the diet (i.e., N_i, W_i, or O_i).

Diet measures each provide unique information about relative prey importance. In an attempt to consolidate the desirable properties of individual diet measures (e.g., N_i, W_i, and O_i), compound indices were developed that combine two or more diet measures into a single index (Table 11.1). The belief is that compound indices capture more information than do single, component measures. Several authors, however, argue that compound indices, such as the index of relative importance (IRI_i) and the relative importance index (RI_i), provide little or no additional information than that provided by single indices (MacDonald and Green 1983; Hansson 1998). Proponents of compound indices, on the other hand, have argued that (1) compound indices provide a more balanced view of fish diets because they capture all of the unique properties affecting individual measures (e.g., N_i, W_i, or O_i), and (2) there is a need for a standardized method for reporting relative prey importance (Cortés 1997). This argument has been criticized on the basis that the addition or multiplication of percentages has no biological meaning because both quantities are dimensionless ratios (Bowen 1996).

The usefulness of compound indices is constrained by several limitations. Comparisons with single measures indicate that compound indices can be a redundant source of information (MacDonald and Green 1983). A more significant problem, however, is that compound indices can be affected by the taxonomic resolution of prey items (Hansson 1998). At different taxonomic resolutions, the importance of different prey types can change, rendering the IRI_i a poor choice for a standard index in diet analyses (Hansson 1998).

The search for an index that best describes relative prey importance has led to much controversy over which diet index is best (Hyslop 1980; Cortés 1997; Hansson 1998). No doubt much of this confusion stems from the fact that relative prey importance is context specific and can be defined in a variety of ways. If we intend to evaluate energy flow, prey composition by weight (or volume) would be a better choice than composition by number. On the other hand, prey numbers could be used to assess prey preference if corresponding information on in situ prey

abundance was available (Hansson 1998). A general framework for selecting diet measures is given in Figure 11.2.

11.3.2 Presenting Data with Graphical Techniques

Diet measures such as N_i, W_i, and O_i are usually presented in tabulated format, making it difficult to interpret two or more indices simultaneously. Graphical techniques attempt to overcome this problem by combining two or more diet measures in two-dimensional space (i.e., bivariate plots). By examining relationships between different diet measures, graphical techniques can be used to interpret (1) predator feeding strategies, (2) relative prey importance, and (3) diet variability.

A graphical technique that relates prey abundance (N_i or W_i) to frequency of occurrence (O_i) was developed by Costello (1990) and later modified by Amundsen et al. (1996). In the Amundsen method, prey-specific abundance is plotted against frequency of occurrence, where prey-specific abundance is defined as the proportion a prey item constitutes of all prey items in only predators that contain prey i (Amundsen et al. 1996). The equation used to calculate prey-specific abundance (P_i) is

$$P_i = (\Sigma S_i / \Sigma S_{ti})100, \tag{11.2}$$

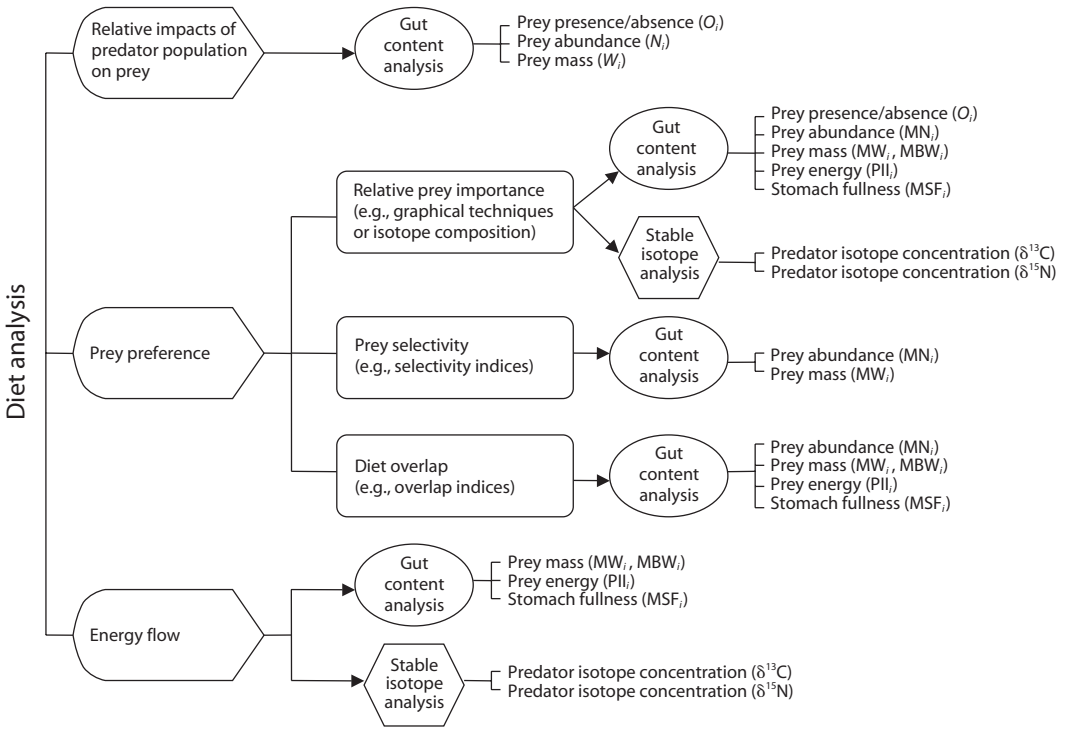


Figure 11.2 Diet measures commonly used to address questions about predator impacts, prey preference, or energy flow. See Table 11.1 for calculation of different diet measures.

where P_i equals prey-specific abundance (numbers, mass, or volume) of prey i, S_i equals the abundance of prey i in stomachs, and S_{ti} equals the total abundance of prey in predators that contain prey i. As an example, consider three fish, respectively with 1, 2, and 3 g of prey i and 8, 7, and 5 g of total stomach contents. We calculate P_i as follows:

$$P_i = \frac{1+2+3}{8+7+5}(100) = \frac{6}{20}(100) = 0.3(100) = 30\%.$$

When plotted against frequency of occurrence, prey-specific abundance can be used to evaluate three important aspects of the fish diet: (1) feeding strategy (specialized versus general), (2) prey importance (dominant versus rare), and (3) niche width (Figure 11.3). In practice, four interpretations can be made by relating prey-specific abundance to frequency of occurrence that could otherwise not be determined from single diet measures (Box 11.2).

Because prey-specific abundance and frequency of occurrence are calculated for the entire sample of fish, graphical techniques that use these indices represent analysis at the population level. To assess feeding patterns at the individual level, graphical methods have been developed that incorporate the use of prey diversity and number of prey in individual stomachs (Bridcut and Giller 1995). In this approach, individual prey diversity for each fish is calculated using a diversity index and then plotted against the total number of prey in the stomach. A generalist feeding strategy is characterized by high prey diversity and low abundance of each prey type, whereas a specialist strategy is represented by low prey diversity and high utilization of a few prey types. Methods for defining high prey diversity, however, have not been developed. As a result, this technique involves subjective interpretation but can be useful for examining patterns of diet specialization across time or space (Bridcut and Giller 1995).

11.3.3 Exploring Variation in Prey Size

Often investigators are interested in the relationship between prey size and predator size, particularly as it relates to gape limitation in fishes. Hence, the maximum linear dimensions of each diet item are plotted against predator length (Juanes 1994). Resulting distributions are often wedge-shaped because small fish are generally limited to small prey, whereas large fish can incorporate a variety of prey sizes in their diet (see example in Box 11.3). In many cases, identifying maximum and minimum prey sizes, rather than the average size, is desirable. Although the maximum and minimum edges of these bivariate scatter plots can be described using least squares regression (LSR), the choice of what edge data to include in the analysis is arbitrary. In addition, LSR is sensitive to the effects of outliers in the chosen edge distribution. A promising method involves the use of a quantile regression technique called least absolute values regression (Scharf et al. 1998; Cade and Noon 2003), in which the sum of the absolute values of the residuals are minimized (rather than the sum of squares of residuals as in conventional LSR).

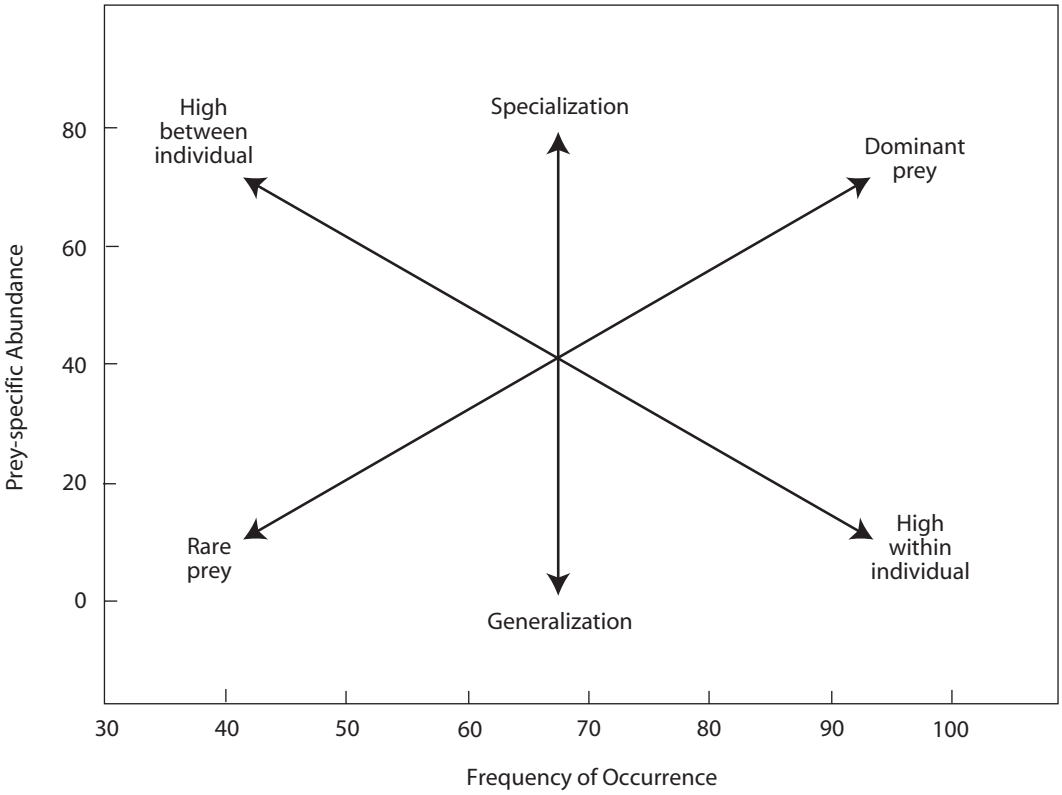


Figure 11.3 Graphical model that depicts feeding strategy (specialized or generalized), relative prey importance (dominant or rare), and niche variation (individual versus population patterns) based on the distribution of individual prey types. Prey-specific abundance is calculated from only those predators that contain prey *i* and is plotted against frequency of occurrence for each prey (O_i). Prey points located in the upper left of the plot indicate prey that are consumed by few individuals displaying specialization; points located in the lower right reveal prey items that have been eaten occasionally by most individuals (Amundsen et al. 1996). Figure adapted from Amundsen et al. (1996) as first described by Costello (1990).

Estimates are obtained through minimization of

$$\sum_i |y_i - \beta_0 - \beta_1 X_i| h_i, \tag{11.3}$$

where h_i is a multiplier equal to a chosen quantile value (e.g., 0.5 for the median) if the residual within the absolute value symbols is positive or one minus the quantile value if the residual is negative (Scharf et al. 1998, Cade and Noon 2003). This technique was quite robust in identifying upper and lower bound slopes in scatter diagrams (Scharf et al. 1998; Cade and Noon 2003) and is very useful for characterizing prey size–predator length relationships (Box 11.3).

Box 11.2 Presenting Diet Measures Graphically

By combining different diet measures in two-dimensional space, graphical techniques can relay important information about feeding behavior of fishes. Using Figure 11.3, we can interpret feeding strategies of each predator population in the graphs presented below.

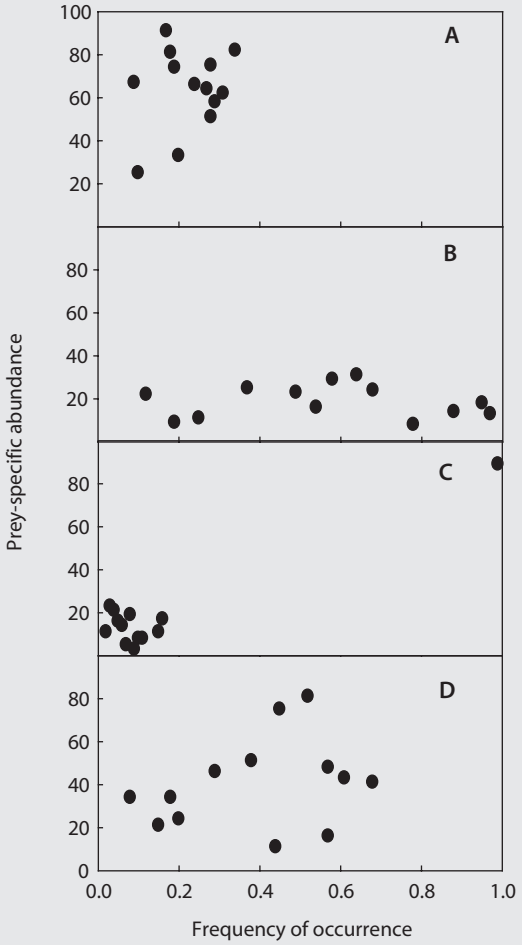


Figure Graphs showing abundance of 13 prey types for four predator populations—A, B, C, and D. Each point represents a different prey type and is expressed as prey-specific abundance plotted against frequency of occurrence (adapted from Admundsen et al. 1996).

We see that fish from population A specialize on individual prey types. As a result, these fish show a high degree of between-individual variation in diet breadth. In population B, predators have a more generalized diet and higher within-individual variation in diet breadth. In population C, the predator population is specializing on a single prey type while occasionally consuming other prey. Finally, population D represents a mixed feeding strategy in which some individuals have a specialized diet and other fish have a more generalized feeding strategy. Graphical techniques, such as the one illustrated here, provide insight about fish feeding patterns that might not be inferred from single diet indices.

Box 11.3 Determining the Minimum and Maximum Sizes of Prey

The maximum size of prey in fish diets often increases with body size. However, the minimum size of prey may change relatively little. In addition to determining the mean or median size of prey consumed by use of bivariate plots, investigators may want to characterize the maximum and minimum sizes consumed (i.e, the edges of the scattergrams). Least absolute values regression (LAV), also called least absolute deviations regression, can be used to evaluate these types of diet data (Scharf et al. 1998). An extension of LAV, quantile regression, fits any specified quantile as a linear regression model. The LAV is the 50th percentile (median) in quantile regression. Such an analysis is available in the Blossom Statistical Software Package (Cade and Richards 2000; available at http://www.mesc.usgs.gov). This program generates test statistics by permutations of the original data through re-randomization.

For example, we want to characterize the upper and lower bounds of lengths of fish prey in age-0 largemouth bass diets from Tappan Lake, Ohio, during 1994 through 1996 (data from Garvey et al. 1998b).

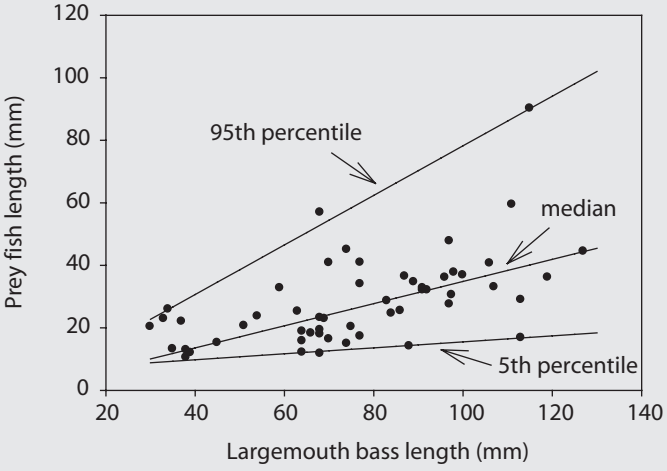


Figure Prey length versus age-0 largemouth bass length (from Garvey et al. 1998b).

Using quantile regression, we first determine the median regression model that minimizes the least absolute differences between the observed values and the residuals. We then determine the quantile regression models that fit the 5th and 95th percentiles of the data. The form of each linear regression model is $y = \beta_0 + \beta_1 x$. The test statistic generated for the LAV regression (i.e., quantile = 0.5) is equivalent to that of a typical least-squares regression comparing the proportional reduction in deviations when passing from a reduced to a full model. Because quantile regression involves weighted absolute deviations (see equation [11.3]), we cannot assume identical error distributions across the independent variables. As such, Cade and Richards (2000) recommend using a rank-sums test for quantile regression (i.e., quantile $\neq$ 0.5), in which the statistic is based on the sign of the residual from the reduced parameter null model.

(Box continues)

Box 11.3 *(continued)*

Table Model values for the full LAV (median), 95th percentile, and 5th percentile regressions. The absolute values of the residuals for the 95th and 5th quantile regression are weighted, and the *P*-value is based on a rank-sums test.

Quantile	β_0	β_1	Sum of absolute values of residuals	*P*
0.5	–0.5748	0.3541	433.07	0.0002
0.95	–1.0507	0.7930	112.92	0.077
0.05	5.9543	0.0957	41.73	0.055

The full LAV regression model (quantile = 0.50) was significantly different than the reduced model. The quantile regression models describing the 95th and 5th percentiles had slopes greater than 0, and only 7.7% and 5.5% of the corresponding test statistics generated by the permutation procedure had more significant values. We conclude that this technique effectively characterizes the median as well as the upper and lower bounds of prey sizes consumed by age-0 largemouth bass in Tappan Lake.

11.3.4 Evaluating Nonindependence of Diet Data

The compositional nature of diet measures (i.e., proportions) has important implications for data analysis. Interpretations about the relatedness of prey items or sites can be very different when using compositions relative to unstandardized (e.g., raw) data (Jackson 1997). In practice, arcsine transformations are often applied to compositional data prior to analysis in attempts to normalize the data. Such transformations should not be applied arbitrarily; rather, data should be examined (for normality) to verify that transformations are needed. Traditional statistical techniques (e.g., *t*-test and analysis of variance [ANOVA]) can be applied when assumptions of normality are met or large sample sizes are obtained. In cases in which compositional data are not normally distributed, nonparametric rank procedures can be useful for detecting differences in individual prey proportions.

Fisheries scientists must be aware of the nonindependence trait of diet data. A well-considered experimental design will avoid the pitfall of pseudoreplication (Hurlbert 1984). Diets from individual fish often contain multiple items that cannot be treated independently. In addition, fish diets are usually sampled either repeatedly through time or at the same location. A variety of statistical techniques can be used that account for autocorrelation within diet data. When designing a study, fisheries scientists should determine if the assumptions of these tests are met.

Temporal and spatial variation in diet data may be analyzed using conventional parametric statistical techniques such as ANOVA if stomachs of individual fish are collected from independent experimental units. For example, temporal dietary patterns in an experiment may be analyzed using ANOVA if they derive from

independent mesocosms or aquaria sampled only once during an experiment. In this case, each stomach would represent a one date–replicate combination. Obviously, meeting this assumption may require a large number of replicates in an experiment because variation among individuals within treatments will likely be high.

In field studies, stomach contents are often collected from groups of fish at the same location during multiple sampling trips. Diet data also may be collected from the same live fish multiple times during an experiment. In these cases, diet data are not independent. For fish captured at the same location, time, or both, stomach samples will likely be more similar than those collected at other times and locations. These potentially confounding problems of spatial or temporal autocorrelation may be addressed statistically using several techniques including repeated-measures ANOVA.

Repeated-measures designs use the same subject (e.g., site or fish) for each of the treatments in a study (Neter et al. 1990). The subject is considered a block, and the treatment(s) are applied to each subject in random order. In a randomized-complete-block repeated-measures ANOVA, each subject receives all of the treatment combinations. If the subject is being followed through time, then time is the repeated measure within each subject (i.e., the within-subject effect; Box 11.4). It is assumed that the variance within each subject (i.e., individual) will be less than that among subjects (Neter et al. 1990). The randomized-complete-block approach is often difficult to employ because it may be impossible to apply all treatments to all subjects. A split-plot ANOVA is a special case of the repeated-measures design that allows subjects to be included in only some of the treatments (see Maceina et al. 1994). To illustrate, let us explore temporal variation in fish diets both within days and among weeks. If we consider each fish to be a fixed subject, then biomass consumed by each may be quantified during morning for half of the fish, while the remaining half is sampled during afternoon. All diets are quantified on a weekly basis. In this case, variation in biomass consumed must be partitioned due to (1) individual fish, (2) time of day, (3) week, and (4) interactions among fish, time of day, and week. Fish with similar characteristics are blocked in pairs, and each is randomly assigned a morning or afternoon sampling time (Table 11.3). Samples are then taken for several weeks. We perform an ANOVA exploring the effects of block, time of day, and their interaction, called the main-plot effects. We also determine the subplot effects of week and the week × time of day interaction. The allure of this approach is that individual fish can be followed through time, and not all fish need to be handled twice each sampling date.

Randomized-complete-block repeated-measures ANOVAs involve strict assumptions about the sphericity of the variance–covariance matrix of the within-subject factor (e.g., time). For the matrix to be spherical, the variance of the difference between any two levels of the within-subject factors must be constant. This property is tested by determining the sphericity of the variance–covariance matrix, such as with a Mauchly's test of sphericity (SAS Institute 1999). If assumptions of sphericity are not met, then the likelihood of rejecting the null hypothesis of no within-subject effect (e.g., time) is inflated, and an adjusted test must be used (Box 11.4).

Box 11.4 Analyzing Diet Data with Repeated-Measures Analysis of Variance (ANOVA)

Diet data often arise from multiple samples within the same system or from multiple observations of the same individual. A repeated-measures ANOVA approach is useful for teasing apart variation as a function of independent effects (e.g., between subject) and nonindependent effects (e.g., within subject).

In this hypothetical example, we conducted a 5-week experiment to determine the effect of gizzard shad on the mean percent by weight (MW_i) of zooplankton in diets of bluegill. Four bluegills were sampled from each replicate once weekly. The resulting data were MW_i for each replicate and date. Data were arcsine(x) transformed prior to analysis.

Table Zooplankton (MW_i) in diets of bluegill ($n = 4$ per replicate per sampling period) in treatments with and without gizzard shad. Data in the table are untransformed.

Treatment (gizzard shad) and replicate	Week				
	1	2	3	4	5
Absent					
1	0.6	0.57	0.59	0.68	0.67
2	0.36	0.37	0.4	0.41	0.49
3	0.43	0.39	0.48	0.49	0.48
4	0.55	0.54	0.6	0.58	0.52
Present					
1	0.72	0.45	0.4	0.32	0.29
2	0.65	0.4	0.38	0.27	0.1
3	0.53	0.46	0.38	0.29	0.23
4	0.45	0.4	0.29	0.23	0.25

Program

```
data one;
input treat $ rep week1-week5;
cards;
[input data];
proc glm;
class treat;
model week1-week5=treat;
repeated time/ printe;
run;
```

Other techniques are available for analyzing autocorrelation in temporal or spatial diet data. Long-term observations of diets may be analyzed using time-series techniques, which are particularly useful in unreplicated systems such as lakes or reservoirs. Autoregressive integrated moving average (ARIMA) models and related techniques may be used to identify nonrandom patterns through time (Rasmussen et al. 1993), assuming that observations are available in discrete, evenly spaced intervals. These models can be extended to compare the treatment response

Interpretation
The test for the gizzard shad effect (between subjects) was significant at $P = 0.057$.

Table The effect of gizzard shad on the MW_i of zooplankton in diets of bluegill.

Effect	*df*	Mean square	*F*-value	*P*
Treatment	1	0.222	5.51	0.0572
Error	6	0.040		

When testing for a time effect (within subjects), a Mauchly's test for sphericity was rejected ($P = 0.0021$), indicating that the variance–covariance matrix was not circular. This is typical for data that are sampled repeatedly through time. A test such as the Greenhouse–Geisser epsilon (G–G) must be used to adjust the error rate. These tests are automatically computed by the SAS procedure.

Table Test for time effect and time*treatment effect of gizzard shad on zooplankton in bluegill diet.

Effect	*df*	Mean square	*F*-value	G–G adjusted *P*
Time	4	0.0350	10.06	0.0040
Time*treatment	4	0.0746	21.42	0.0002
Error	24	0.0034		

We conclude from this analysis that MW_i changed in both treatments through time. The time*treatment effect indicates that the treatments changed in different ways through time, probably because zooplankton increased in diets in the absence of gizzard shad but declined in treatments with gizzard shad.

of a single, unreplicated experimental system to that of a reference system. A limitation of these moving average techniques is that they usually require large sample sizes (>50 dates; Rasmussen et al. 1993).

Autocorrelated spatial patterns in diet data can be analyzed in a variety of ways. The Mantel test is a randomization test that determines whether differences between two $n \times n$ distance matrices are random (Fortin and Gurevitch 1993). Spatial variation among individuals (distance matrix 1) can be compared to the

Table 11.3 Split-plot repeated-measures ANOVA design for biomass consumed by individual fish ($n = 6$) blocked into pairs and then randomly selected to be sampled in the morning or evening (effect A). Consumption of each fish was quantified once a week (effect B).

Block and fish	Time of day (effect A)	Week (effect B)		
		1	2	3
Block 1				
1	AM	$A_{AM}B_1$	$A_{AM}B_2$	$A_{AM}B_3$
2	PM	$A_{PM}B_1$	$A_{PM}B_2$	$A_{PM}B_3$
Block 2				
3	PM	$A_{PM}B_1$	$A_{PM}B_2$	$A_{PM}B_3$
4	AM	$A_{AM}B_1$	$A_{AM}B_2$	$A_{AM}B_3$
Block 3				
5	AM	$A_{AM}B_1$	$A_{AM}B_2$	$A_{AM}B_3$
6	PM	$A_{PM}B_1$	$A_{PM}B_2$	$A_{PM}B_3$

relative proportion of a specified diet item in the stomachs (distance matrix 2). The Mantel test will determine if nonrandom spatial patterns in diet composition exist. Alternatively, two or more bivariate plots of spatial distributions of prey occurrence in diets of individual fish may be compared using a multiway, two-dimensional Kolmogorov–Smirnov test (2DKS; Garvey et al. 1998c). The 2DKS test can also be used to determine if spatial distributions within single plots differ significantly from randomly generated ones. An example of using 2DKS to assess spatial variation in predator diets is given in Box 11.5.

Diel changes in diet have important implications for choosing sampling times or understanding gastric evacuation patterns for fishes (see section 11.2.2.1). Using analysis of covariance (ANCOVA), the content of fish stomachs (e.g., weight) can be regressed against fish weight (the covariate) using conventional least-squares regression during each sampling time. If these relationships can be transformed such that they are linear, the slopes of each line can be compared (Box 11.6). If slopes are the same (i.e., parallel), then the intercepts among the regression lines can be compared. Significant among-intercept differences indicate a diel pattern in the abundance of prey in stomach contents.

11.3.5 Comparing Diet Composition

11.3.5.1 Employing Multivariate Analysis of Variance with Randomization

The multivariate nature of fish diets often requires approaches other than univariate statistics when interest lies in the simultaneous evaluation of all prey categories. By example, consider the decisions we make when buying fishing gear. Before making a purchase, we (e.g., anglers) often consider at least four factors: (1) price, (2) quality, (3) brand name, and (4) style. One angler may rate their decision on (1) style, (2) quality, (3) brand name, and (4) price, whereas another angler may consider the purchase based on (1) price, (2) quality, (3) style, and (4) brand name. Here, we are interested in asking whether anglers use the same

Box 11.5 Assessing Spatial Patterns in Diet with the Two-Dimensional Kolmogorov–Smirnov Test

Several statistical methods are available to relate diet patterns to the distribution of habitat in aquatic systems. Mantel and partial-Mantel tests are powerful techniques that test whether spatial patterns are random or due to some treatment (or time). These tests are not specifically discussed here. More information can be obtained in Fortin and Gurevitch (1993) and Chapter 18. If spatial data can be arranged into bivariate spatial coordinates, a two-dimensional Kolmogorov–Smirnov (2DKS) test can be used to (1) identify whether a single distribution has arisen by random effects or (2) compare two bivariate distributions (see Garvey et al. 1998c for a review). This nonparametric test finds the maximum difference, D_{bks} (where bks represents bivariate Kolmogorov–Smirnov), in integrated probabilities for four quadrants around each point in a plane. If the maximum D_{bks} between two distributions exceeds that expected randomly, we conclude that they differ. The significance of the test statistic D_{bks} is determined by rerandomizing the original data 5,000 times and then comparing this randomly generated distribution to the observed value.

In the following hypothetical example, we want to know how vegetation in a large lake affects piscivory in age-0 smallmouth bass. We partition the bottom of a shallow lake into 80 habitat quadrants (20 × 4) and determine whether each contains vegetation. Within each quadrant, we sample smallmouth bass diets by means of gastric lavage and note whether piscivory is present or absent.

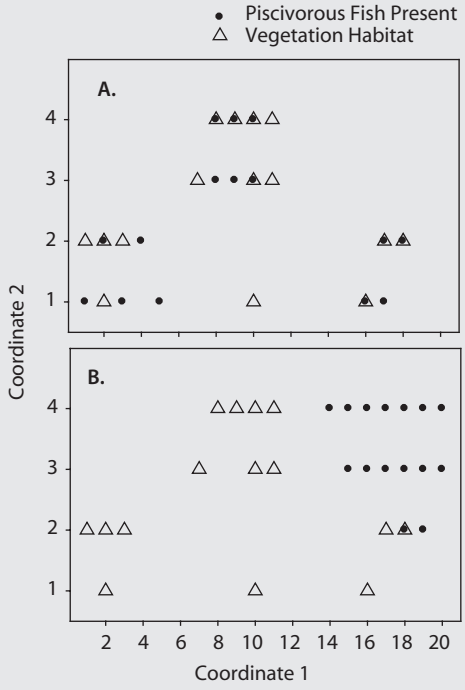


Figure Two hypothetical scenarios of smallmouth bass in a shallow lake. The bottom of the lake is partitioned into 80 habitat quadrants (20 × 4), and it is determined whether each contains vegetation. Within each quadrant, smallmouth bass diets are sampled to determine whether piscivory is present or absent.

(Box continues)

Box 11.5 *(continued)*

In the scenario depicted in the upper panel **(A),** piscivorous smallmouth bass (closed circles) appeared to be closely associated with quadrants containing vegetation (open triangle). In confirmation, a 2DKS test comparing the bivariate distributions of vegetation and piscivorous smallmouth bass revealed no difference ($D_{bks} = 0.143$, P = 0.999). In the scenario depicted in the lower panel **(B),** the distribution of vegetation is identical to that in A. However, the spatial distribution of piscivorous smallmouth bass appears to be associated with some other factor. The 2DKS test detected a difference between the spatial distributions of vegetation and piscivorous fish ($D_{bks} = 0.714$, P = 0.002).

Of course, the 2DKS test is useful for determining only presence and absence in this example. We also must assume strong site-fidelity of fish within habitats and that fish are homogeneously distributed among vegetated and nonvegetated sites. Cells with missing data are acceptable. It is important to note that Mantel tests incorporate quantities within each cell, allowing us to compare other responses such as the frequency of occurrence of fish in diets.

decision factors before purchasing fishing gear. To address this issue, we treat the purchase as a multivariate response by evaluating these decisions simultaneously. In the same way, we can treat the diet of fish as a multivariate response defined by the abundance of different prey items in the stomach.

There are a variety of approaches for analyzing multivariate diet data. For the convenience of modeling and performing statistical tests, techniques such as multivariate analysis of variance (MANOVA) require the assumption of multivariate normality (Khattree and Naik 1999). Other multivariate methods, such as cluster analysis and ordination techniques, are largely distribution free in nature and are useful for generating biologically meaningful patterns from multivariate data.

When diets are expressed as prey weight or volume, MANOVA can be useful for testing differences in diet composition. The assumptions of MANOVA require that prey proportions have a multivariate normal distribution and a similar variance–covariance structure among samples. Prior to performing MANOVA, tests for multivariate normality should be applied to data to evaluate this assumption (Khattree and Naik 1999). When diet composition data do not meet the assumption of multivariate normality (as is often the case), a nonparametric-based randomization procedure can be applied to test for differences in diet composition between samples (Crow 1979; Somerton 1991). In this approach, MANOVA is combined with a randomization procedure. Randomization procedures are not new to ecological analysis but have received little attention in the analysis of fish diets (but see Somerton 1991). Randomization procedures are relatively straightforward and proceed as follows.

1. Combine diet proportion data from time or area samples.
2. Randomly sort data into n new samples equal in size to the original data.
3. Calculate a test statistic based on the new samples.
4. Repeat steps 2 and 3 a large number of times (e.g., 5,000).

Box 11.6 Determining Diel Patterns in Diet Data with Analyis of Covariance (ANCOVA)

We often want to determine if diel patterns in diet data occur. This has important implications for designing sampling protocols and interpreting diet data. One way to determine whether diel variation in feeding occurs is by sampling fish during different times of the day.

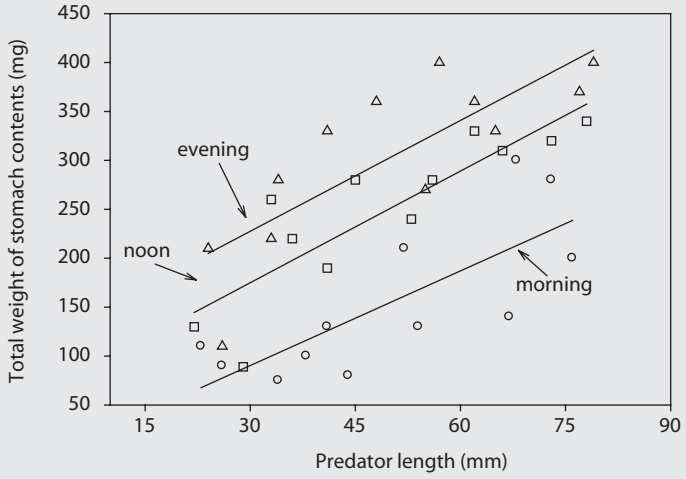


Figure The total weight (mg) of food found in the diets of different size fish are given for three time periods—morning, noon, and evening.

Program

An ANCOVA was used to test the null hypothesis that the three regression lines are equal. For an ANCOVA to be valid, the slopes of the regression lines must be parallel ($slope_{dawn} = 0.0322$; $slope_{noon} = 0.0380$; and $slope_{dusk} = 0.0377$). If this assumption holds, then we can use the general linear model (GLM) procedure in SAS in which length is the continuous covariate and time is the categorical variable. The assumption of parallel slopes is rejected if a length*time interaction is detected.

```
data one;
input time $ length diet;
cards;
[input data];
proc glm;
class time;
model diet = time length;
run;
```

Interpretation

Table Effect of time of day and length of fish on weight of items in fish stomachs.

Effect	*df*	Mean square	*F*-value	*P*
Time	2	6.770	28.65	0.0001
Length	1	14.431	61.06	0.0001
Error	32	0.236		

No length × time interaction occurred in the initial model. We then assumed that slopes were parallel and dropped the interaction from the model. The ANCOVA revealed that total weight of items increased in guts with increasing body size. In addition, the amount of food varied with time of day, suggesting that sampling time be carefully considered when developing a protocol.

From these data, a probability distribution of the randomized test statistic is generated. If the observed test statistic is within the upper (or lower) 5% tail of the randomized distribution, then the result is significant at the 5% level. Similarly, if our observed value falls within the 1% tail, then we can conclude that the difference is significant at the 1% level and so on. The choice of a test statistic depends on the research question being addressed and the characteristics of the test statistic. In a two-sample case, we could perform a randomization procedure on the Hotelling's T^2-statistic and test for a difference between sample means. Similarly, the *F*-statistic could be used to test for a treatment effect among three or more factors (e.g., lakes, seasons, and sites). An example of a nonparametric MANOVA that tests for diet differences is given in Box 11.7.

11.3.5.2 Examining Prey Numbers with Log-Linear Contingency Tables

When diet data are expressed as prey numbers, a multiway contingency table analysis can be used to assess diet variation (Cortés 1997). In this approach, data are arranged in an $R \times C$ contingency table, where R is the number of prey categories and C is the number of predator categories. Each cell in the table contains the total number of the ith prey category found in the stomachs of the jth predator category. One limitation of contingency table analysis is that large samples sizes are needed so that less than 20% of the cells have an expected frequency less than five. One way to remedy this situation is to pool prey species, so that we reduce the total number of categories and increase the sample size for the remaining categories (see section 11.2.1; Crow 1979).

Contingency table analysis begins by testing for significant interactions. In this way, we are testing a hierarchy of models starting with the most complex. In a three-way contingency table, we would start by examining the three-way interaction. If this term were not significant, we would delete it from the model and then proceed to test all the two-way interactions. The advantage of this approach is that by proceeding with posthoc tests, we can readily identify the rows (prey types) and columns (predators) that contribute the most to diet variation (Cortés 1997). An example of a three-way contingency table is given in Box 11.8.

11.3.5.3 Applying Ordination Techniques

Ordination techniques, such as principal components analysis (PCA), are widely used in ecological data analysis. Because diet data are often measured as proportions, analytical techniques are affected by the constant-sum constraint (i.e., as the abundance of one taxa increases, one or more taxa must decrease; Jackson 1997). To deal with compositional data, two alternative ordination methods have been proposed. The first approach is a log-ratio analysis performed on the logarithms of the percentages; this approach is most appropriate when compositional data do not contain zeros (Aitchison 1983). Although not new to the ecological literature, log-linear PCA techniques have only recently been applied to fish diet data (De Crespin De Billy et al. 2000). Termed %PCA, this technique is based on a PCA performed on a proportion table in which each column is defined by a prey

Box 11.7 Comparing Diet Data from Different Locations or Times with Multivariate Analysis of Variance (MANOVA)

Because of the multivariate nature of diet data, we are often interested in determining whether diet composition differs among fishes sampled from different locations or at different times. When diet data are measured as prey mass (or volume), MANOVA can be useful for testing an overall location (or time) effect.

Table Hypothetical diet data for three bluegill populations. Data are presented as mean percent composition by weight (MW_i) for four different prey items.

	Prey type			
Bluegill	Chironomids	Amphipods	Odonates	Copepods
		Lake A		
1	0.12	0.35	0.44	0.09
2	0.09	0.22	0.63	0.06
3	0.12	0.35	0.5	0.03
4	0.26	0.38	0.22	0.14
5	0.27	0.29	0.27	0.17
		Lake B		
6	0.49	0.01	0.38	0.12
7	0.36	0.04	0.59	0.01
8	0.34	0.05	0.57	0.04
9	0.42	0.03	0.24	0.31
10	0.57	0.11	0.21	0.11
		Lake C		
11	0.08	0.34	0.49	0.09
12	0.06	0.27	0.59	0.08
13	0.02	0.33	0.58	0.07
14	0.11	0.57	0.28	0.04
15	0.01	0.57	0.31	0.11

Program

Here, we are interested in testing for an overall lake effect in diet composition and perform a MANOVA analysis. The MANOVA procedure was used in the following SAS program to generate output.

```
data test;
input lake $ fish chiro amph odon zoo;
cards;
[input data];
proc glm;
class lake;
model chiro amph odon zoo=lake;
manova h=lake / printe printh;
title 'Manova of diet data';
run;
```

(Box continues)

Box 11.7 *(continued)*

Interpretation

Results from the MANOVA test for the hypothesis of no overall lake effect are presented below. Several statistics are produced from MANOVA analysis and provide similar results. For randomization procedures, we will consider the Wilk's lambda test statistic. From the output shown below, we would conclude that there is an overall lake effect on bluegill diets.

Table Test statistics comparing MANOVA results of diet composition for bluegills among three lakes.

Statistic	Value	*F*-value	*P*
Wilk's lambda	0.079	8.52	0.0001
Pillais trace	0.92	3.16	0.021
Hotelling–Lawley trace	11.58	17.38	0.0001
Roy's greatest root	11.58	42.46	0.0001

To determine which prey types vary among bluegill populations, individual ANOVAs are performed on each prey type. Because these tests are a posteri, an appropriate alpha level can be obtained using the Bonferroni inequality by which adjusted alpha levels are equal to the overall alpha divided by *n*. In this case, we would consider individual ANOVAs to be significant at 0.0125. In the table below, we see that chironomids and amphipods vary significantly among lakes.

Table Analysis (ANOVA) of which prey types vary among the three bluegill populations.

Prey item	Source of variation	*df*	Sum of squares	*F*-value	*P*
Chironomid	Lake	2	0.379	31.38	0.0001
Amphipod	Lake	2	0.363	20.98	0.0001
Odonate	Lake	2	0.007	0.13	0.87
Copepod	Lake	2	0.004	0.34	0.71

type and each row represents an individual fish. An advantage of this technique is that individual fish and their prey are analyzed simultaneously and can be displayed on the same graph (De Crespin De Billy et al. 2000). A second approach offered as a solution to the problem of compositional data is correspondence analysis (CA; Jackson 1997). This approach is particularly well suited to handle compositional data and provides advantages over other methods (e.g., PCA; Digby and Kempton 1987; Jackson 1997). Furthermore, unlike log-ratio analysis, CA is not constrained by the presence of zeros in the data, providing a robust approach for analyzing compositional data.

Although the mathematical derivation of these techniques is beyond the scope of this chapter, a user-friendly program for running these analyses is CANOCO

Once we obtain our observed test statistic (e.g., Wilk's lambda = 0.079), we then perform a randomization test (Crystal Ball 7.0, Decisioneering, Inc., Denver, Colorado; http://crystalball.com/crystal_ball/) to evaluate the significance of our statistic. After performing 4,999 simulations, we obtain the following probability distribution of the test statistic based on our data.

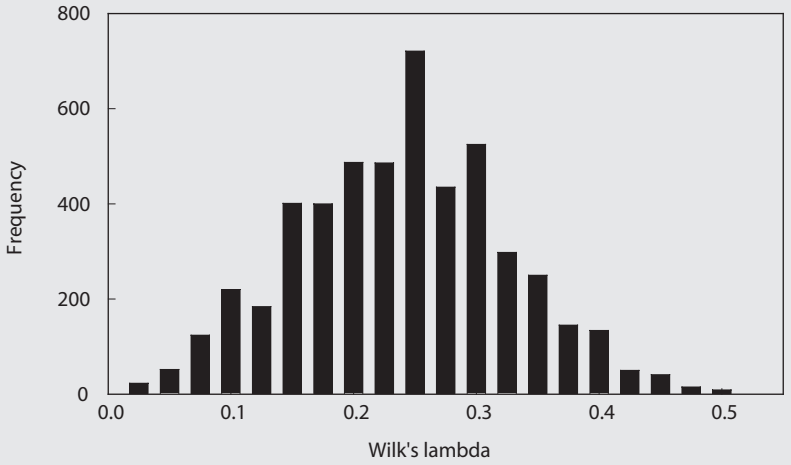


Figure Randomized frequency distribution for Wilk's lambda based on 4,999 simulations.

From the randomized frequency distribution, we see that our observed value (0.079) easily falls in the lower 5% of the observations. In fact, only 75 observations were less than our observed Wilk's value (0.079). We can estimate a *P*-value as 75/5,000 or 0.015. Hence, it is unlikely that we would obtain a value of 0.079 if the null model is true, and we would conclude that diets are significantly different among lakes. Similarly, randomization procedures could be performed on the individual ANOVAs (e.g., *F*-value) to confirm that chironomids and amphipods account for these differences.

for Windows (ter Braak and Smilauer 1998). To demonstrate the usefulness of log-ratio analysis (%PCA), we used diet data provided in Table 11.2 to assess individual variation in bluegill food habits (Box 11.9).

11.3.6 Estimating Diet Overlap

Niche overlap indices are often used to measure the magnitude of resource overlap among different species. Although these indices are sometimes used to infer competition, we should recognize that high resource overlap between two species may not indicate competitive bottlenecks. Rather, it may be indicative of high resource abundance, such as seasonal peaks in prey availability.

Box 11.8 Testing Prey Counts with Multiway Contingency Table Analysis

When diet data are measured as prey counts, multiway contingency table analysis can be used to test for treatment effects.

Table The following data represent numbers of prey for two different life stages of fish collected from two different environments. A three-way contingency table is used to test for differences among the three different levels: (1) life stage, (2) habitat, and (3) prey type.

	Prey type			
Habitat and life stage	Amphipods	Chironomids	Mayflies	Ostracods
Littoral				
Adult	29	69	9	10
Juvenile	19	43	4	6
Pelagic				
Adult	6	21	6	4
Juvenile	1	8	5	4

Program
The following SAS program was used to generate output.

```
data test;
input prey $ stage $ habitat $ number;
cards;
[input data];
proc catmod;
weight number;
model prey*stage*habitat=_response_ / pred=freq;
loglin prey|stage|habitat;
run;
```

The summary statistics below show that the three-way interaction, prey*stage*habitat, is not significant. If this term were significant, there would be no reason to examine two-way interactions or main effects.

Table Summary statistics for multiway contingency table analysis.

Source	*df*	χ^2	*P*
Prey	3	71.27	0.0001
Stage	1	8.99	0.002
Habitat	1	26.75	0.0001
Prey*stage	3	1.70	0.637
Prey*habitat	3	12.88	0.0049
Stage*habitat	1	0.18	0.674
Prey*stage*habitat	3	3.09	0.377

Deleting the prey*stage*habitat term from the model, we obtain a significant interaction for prey*habitat ($P = 0.0095$). To determine which prey items are responsible for the significant prey*habitat interaction, we can delete individual prey categories and revaluate the interaction term. Below, we see that by deleting individual prey types, we are unable to obtain a nonsignificant interaction term for amphipods, chironomids, or ostracods. However, when we delete two groups of prey from the analysis we find that amphipods and chironomids are responsible for the significant interaction observed in the prey*habitat term.

Table Analysis to determine which prey items are responsible for the significant prey*habitat interaction.

Prey type deleted	*P*-value for interaction term
Single prey	
Amphipods	0.024
Chironomids	0.007
Mayflies	0.115
Ostracods	0.006
Combined prey	
Amphipods & chironomids	0.377
Amphipods & ostracods	0.010
Chironomids & ostracods	0.002

By tabulating the observed and expected frequencies (in parentheses) for amphipods and chironomids, we can make inferences about how these prey types differ across fish life stages and habitats. Here we see higher than expected numbers of chironomids in adult diets from both habitats. For both prey types, adult fish also showed higher than expected values compared to juvenile fish.

Table Comparison of prey type across life stage and habitats. Given are observed and expected (in parentheses) frequencies of prey type in diet; note that observed and expected frequencies are not equal because other prey types are not shown.

Prey type and life stage	Littoral habitat	Pelagic habitat
Amphipod		
Adult	29 (22)	1 (5)
Juvenile	19 (22)	6 (5)
Chironomid		
Adult	69 (57)	8 (12)
Juvenile	43 (57)	21 (12)

Box 11.9 Exploring Diet Data with Principal Component Analysis (PCA)

Traditional multivariate techniques, such as PCA, can be constrained by the compositional nature of diet data in so much as the row sums must equal one. Log-ratio analysis, such as %PCA (see text), is performed on the logarithm of proportions and can be useful for exploring individual variation in diet data. For values equal to zero, very small numbers (e.g., 0.00001) are entered prior to analysis as recommended by Aitchison (1983). A %PCA analysis was performed on the diet composition data given in Table 11.2. The first two components accounted for 94% (%PC1 = 60%; %PC2 = 34%) of the total variation in diet data.

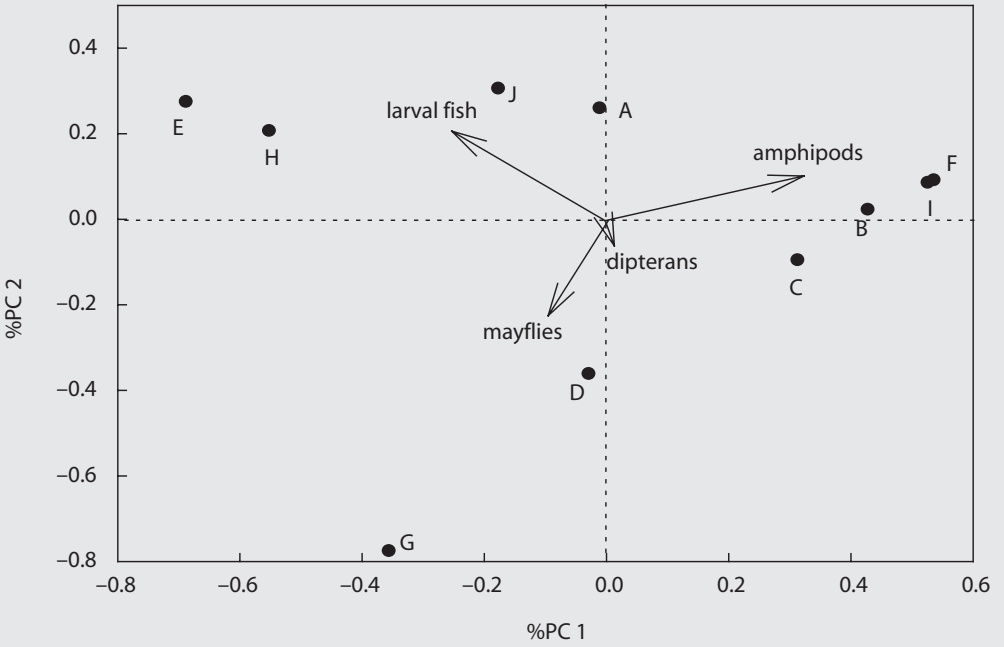


Figure The graph shows each prey type linked to an arrow for which the length of the arrow is proportional to the relative abundance of the prey. The %PCA results for individual bluegills (A–J) are then superimposed on the prey distribution to show individual variation in diet composition. Amphipods, larval fish, and mayflies accounted for much of the variation in individual diets, whereas dipterans accounted for little variation and were distributed near the population centroid (origin).

The patterns represented in the figure above can be compared to physical or biological characteristics to help identify factors affecting diet variability. For example, we might be interested in whether fish size accounts for variation in stomach contents. Correlation analysis reveals a significant relationship between the first axis scores (%PC1) and fish size ($r = -0.84$; $P = 0.002$). Smaller fish (i.e., B, C, I, and F) consumed more amphipods, whereas larger fish (i.e., H and E) were more likely to contain larval fish in their diet. Similarly, this approach could be used to assess a variety of factors such as habitat characteristics, limnological parameters, or fish community attributes.

A variety of indices have been proposed to quantify diet overlap, and there is controversy as to which index is best (Krebs 1989). In cases where prey numbers are available, Morista's index has been recommended as the most robust index (Smith and Zaret 1982). Morista's index is calculated using the equation

$$M = \frac{2\Sigma_i^n p_{ij} p_{ik}}{\Sigma_i^n p_{ij}[(n_{ij}-1)/(N_j-1)] + \Sigma_i^n p_{ik}[(n_{ik}-1)/(N_k-1)]}; \quad (11.4)$$

M = Morista's index of niche overlap between species j and k;
p_{ij} = proportion resource i is of the total resources used by species j;
p_{ik} = proportion resource i is of the total resources used by species k;
n_{ij} = number of individuals of species j that use resource category i;
n_{ik} = number of individuals of species k that use resource category i; and
N_j, N_k = total number of individuals of each species in sample.

If data are not expressed as prey numbers (e.g., biomass or volume), then Horn's index is recommended (Krebs 1989) and is calculated as

$$H = \frac{\Sigma(p_{ij}+p_{ik})\log(p_{ij}+p_{ik}) - \Sigma p_{ij}\log p_{ij} - \Sigma p_{ik}\log p_{ik}}{2\log 2}, \quad (11.5)$$

where H = Horn's index of overlap between species j and k. In equation (11.5), any base of logarithms may be used.

Confidence limits or tests of significance can be calculated for diet overlap values. One way to estimate confidence limits on diet overlap values is to use bootstrapping techniques. Bootstrap techniques are relatively simple and proceed as follows.

1. Using the original data with n observations, randomly select n diet overlap values with replacement. Because we are sampling with replacement, some values may be selected one or more times or not at all. Repeat this step at least 100 times (preferably 1,000).
2. Calculate a mean diet overlap value for each bootstrap sample.
3. Estimate the mean and standard error from the replicate bootstrap values.

Because bootstrap procedures estimate the sample mean, rather than the population mean, they contain a bias that can be corrected using the equation,

$$\text{Bootstrap mean}_{\text{adj}} = 2\bar{x}_s - \bar{x}_B, \quad (11.6)$$

where $\bar{x}_s$ = observed mean of original sample and $\bar{x}_B$ = bootstrap estimate of mean (Krebs 1989).

11.3.7 Estimating Prey Preference

When given a variety of prey types, most fishes select some food categories over others. To measure this selectivity, a variety of indices have been developed that incorporate measures of prey use and prey availability (see review in Bowen 1996). While prey use can be easily determined from gut content analysis, accurate description of prey availability can be problematic. What we quantify as prey availability may be quite different than what fish perceive under natural conditions. Furthermore, because different prey can occupy different habitats, a single sampling technique may not adequately quantify the relative abundance of different prey items in the environment. This is important because we cannot use volumetric estimates of zooplankton abundance (e.g., number/L) and real densities of benthic invertebrates (e.g., number/m^2) as simultaneous measures of prey availability. Only in cases where prey are collected with the same gear type, such as open-water zooplankton, can we begin to compare use versus availability.

Like diet and overlap indices, there is much controversy over which preference index is best (Wallace and Ramsey 1983). Comparisons of different indices have revealed that the Manly–Chesson (Chesson 1983) and the linear (Strauss 1979) indices are good choices for quantifying prey preference (Smith and Zaret 1982; Wallace 1981; Krebs 1989). The Manly–Chesson index is frequently used to quantify prey preference and can be calculated for two scenarios (Krebs 1989).

Constant prey abundance. This form of the Manly–Chesson index is used when the number of prey eaten is very small relative to that prey item's total population or when prey are replaced, as in laboratory studies. The equation for the Manly–Chesson index under constant prey abundance is

$$\alpha_i = \frac{r_i}{n_i} \frac{1}{\Sigma (r_j / n_j)}; \tag{11.7}$$

α_i = Manly's alpha for prey type i;
r_i, r_j = proportion of prey type i or j in the diet;
n_i, n_j = proportion of prey type i or j in the environment; and
m = total number of prey types.

Values of α_i are normalized so that $\sum_{i=1}^{m} \alpha_i = 1.0$.

Prey preference is indicated when α_i values are greater than $1/m$. Conversely, α_i values less than $1/m$ imply that prey species i is avoided in the diet because it is used in lower proportion than its availability in the environment.

Variable prey abundance. This form of the Manly–Chesson index is used when the number of prey eaten is large relative to that prey item's total population in the environment or when, in experimental studies, prey are not replaced after being eaten. The Manly–Chesson index for variable prey populations is calculated using the equation

$$\alpha_i = \frac{\log P_i}{\sum_{j=1}^{m} P_j}; \tag{11.8}$$

α_i = Manly's alpha for variable prey populations;
P_i, P_j = proportion of prey i or j remaining at the end of the experiment (e_i/n_i);
e_i = number of prey type i remaining at the end of experiment;
n_i = number of prey type i at the beginning of the experiment; and
m = total number of prey types.

In equation (11.8), any base of logarithms can be used.

It is recommended when using the Manly–Chesson index for variable prey populations that the number of prey eaten and the number of prey remaining are greater than 10 (Manly 1974; Chesson 1983; Krebs 1989). In practice, indices such as the Manly–Chesson can be used to test for differences in prey selectivity providing important information about preferred (or vulnerable) prey types (Box 11.10).

Box 11.10 Assessing Prey Preference

Differences in prey selectivity provide important insight about foraging patterns of fishes. In many cases, these type of data are collected under controlled, experimental settings in which changes in the absolute abundance of prey can be accurately determined.

Catalano et al. (2001) examined the effects of tag color on vulnerability to predation. Age-0 bluegills were marked with either brightly colored fluorescent tags or cryptic tags and then exposed to largemouth bass predators in a series of tank experiments. Manly's alpha was calculated using the equation for variable prey populations (equation [11.8]).

Table Vulnerability of age-0 bluegills with differently colored tags to predation by largemouth bass (data from Catalano et al. 2001).

Trial and tag color	Initial number of prey	Final number of prey	Proportion remaining	Manly's alpha
1				
Bright	120	22	0.183	0.618
Cryptic	60	21	0.350	0.382
2				
Bright	96	26	0.271	0.529
Cryptic	48	15	0.313	0.471
3				
Bright	150	64	0.427	0.528
Cryptic	75	35	0.467	0.472

Here, we are interested in whether marking pattern (bright versus cryptic tags) influences prey selectivity. A Student's *t*-test reveals a significant difference in selectivity between brightly colored and cryptic tags ($t = 2.76, P = 0.05$); fish marked with brightly colored tags had a higher selectivity ($\bar{x} = 0.56$) than did fish marked with cryptic tags ($\bar{x} = 0.44$).

11.4 TRACKING ENERGY FLOW WITH STABLE ISOTOPE ANALYSIS

Stomach content analysis provides a high degree of taxonomic precision but is limited in many ways because it provides only a snapshot in time of consumer diets. Stable isotope analysis is an alternative approach to traditional food habit studies that provides time-integrated information useful for tracking energy flow in aquatic food webs (Fry and Sherr 1984; Peterson and Fry 1987). Stable isotope analysis has several advantages as a method for quantifying feeding patterns: (1) it reflects materials actually assimilated by fish, enhancing our ability to detect subtle but important feeding interactions that might go undetected by traditional gut content analysis; (2) it allows more efficient use of sampled fish because there is no loss of information when stomachs are empty; and 3) it can be used to evaluate within-population variation in fish feeding habits.

Carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) are the most commonly used isotopes in aquatic food web studies. In general, $\delta^{13}C$ signatures of consumers are similar to those of their prey and can be used to identify carbon sources at the base of the food chain. Conversely, $\delta^{15}N$ signatures exhibit a step-wise increase from prey to predator. A 3–4‰ enrichment of the heavy nitrogen isotope represents a typical trophic level increment (e.g., zooplankton to fish). Hence, $\delta^{15}N$ signatures can be used to identify important feeding relationships and energy pathways.

Stable isotope ratios are expressed in delta (δ) notation, defined as the parts per thousand (ppt; ‰) deviation from a standard material (Peterson and Fry 1987). The formula for calculating $\delta^{13}C$ or $\delta^{15}N$ is

$$\delta^{13}C \text{ (or } \delta^{15}N) = ([R_{sample}/R_{standard}] - 1) \times 1{,}000, \tag{11.9}$$

where R equals the ratio of $^{13}C/^{12}C$ (or $^{15}N/^{14}N$). Standard materials are represented by Pee Dee belemnite limestone for $\delta^{13}C$ or atmospheric nitrogen for $\delta^{15}N$, where both standards have a ppt value set to 0. A positive (or less negative for carbon) isotopic value indicates the sample is "isotopically" enriched and contains more of the heavy stable isotope (^{13}C or ^{15}N; Vander Zanden et al. 2000).

Samples for stable isotope analysis are usually collected from white dorsal muscle tissue (1–2 g wet weight) of individual fish and frozen until analysis. Although samples are usually collected from sacrificed fish, biopsy punches (6–8 mm) are useful for obtaining nonlethal samples in the field where fish can be quickly treated with an antibiotic ointment and released. For invertebrates or larval fish, whole samples are obtained in the field and then frozen. Prior to freezing invertebrates and larval fish, it is recommended that they be placed in filtered water for up to 12 h to allow gut evacuation. Prey items in the guts of small invertebrates and larval fish can affect $\delta^{15}N$ signatures (Yoshioka et al. 1994). Because 0.1 g dry weight is usually required to analyze stable isotopes, samples should consist of about 1–2 g wet weight. Samples are then dried at 70°C to a constant weight, ground into a fine powder, and packed into 4 × 6-mm tin capsules for isotopic analyses. Isotope analysis of $\delta^{13}C$ and $\delta^{15}N$ is performed using a mass spectrometer.

11.4.1 Applying Stable Isotope Data

A promising new technique for assessing energy flow in aquatic ecosystems involves the calculation of fish trophic position, a continuous variable that quantifies the average energy pathway to a consumer (Vander Zanden and Rasmussen 1999). Trophic position is useful for assessing feeding patterns because it incorporates the relative contribution of different trophic levels to fish diets. Both dietary data and stable isotope ratios can be used to calculate trophic position of fish. Because trophic position incorporates omnivorous feeding behavior, it provides an advantage over food chain studies that fail to consider omnivorous trophic interactions and food web studies that fail to weight food links according to their energetic importance (Polis 1991; Gaedke et al. 1996; Vander Zanden and Rasmussen 1999).

To demonstrate how trophic position can be estimated from dietary data, consider a lake trout population that has a diet consisting of 20% herbivorous zooplankton (trophic level = 2) and 80% planktivorous fish (trophic level = 3). These data, usually obtained from numerous fish within a size-class, can be used to calculate trophic position (TP_{diet}) as

$$TP_{diet} = \Sigma(V_i T_i) + 1, \tag{11.10}$$

where V_i is the percent volumetric contribution of the ith prey item (e.g., 0.2 or 0.8) and T_i is the trophic position of the ith prey item (e.g., 2 or 3; Vander Zanden et al. 1997). Hence, the size-specific trophic position for this lake trout population is 3.8.

Alternatively, trophic position can be estimated for individual fish from stable isotope ratios as

$$TP_{fish} = [(\delta^{15}N_{fish} - \delta^{15}N_{baseline})/3.4] + 2, \tag{11.11}$$

where $\delta^{15}N_{fish}$ is the isotope signature of the fish, $\delta^{15}N_{baseline}$ is the "corrected" isotope signature of the fish, 3.4 is the assumed per mil increase in $\delta^{15}N$ per trophic level, and 2 represents the number of trophic levels involved (Vander Zanden and Rasmussen 1999).

Because $\delta^{15}N$ values vary greatly among organisms at the base of the food chain, the $\delta^{15}N$ value of a consumer cannot be regarded as an absolute measure of trophic position. Hence, it is necessary to correct the $\delta^{15}N$ signatures of fish to account for $\delta^{15}N$ variation among primary consumers (e.g., zooplankton, chironomids, and amphipods; Angradi 1994; Vander Zanden and Rasmussen 1999). To accomplish this, bivariate plots of $\delta^{15}N - \delta^{13}C$ are used to describe the relationship between nitrogen and carbon signatures for primary consumers. This relationship can then be used to calculate baseline conditions ($\delta^{15}N_{baseline}$) that are used to correct $\delta^{15}N$ values of secondary consumers (see Vander Zanden and Rasmussen 1999).

Isotopically derived measures of fish trophic position can be used to assess diet variability within a population. Bivariate plots that depict trophic position–body

size relationships are first constructed to assess variation in energy flow among different-sized fish. Because trophic position normally increases with fish size, variance estimates may be higher for populations with steep trophic position–body size slopes. To remedy this, variance estimates can be estimated as the mean absolute residual value from trophic position–body size relationships (Box 11.11). This variation is independent of body size and can be used to assess factors affecting fish trophic position (Vander Zanden et al. 2000).

Variation in trophic position reflects the magnitude of two diet components: (1) diet breadth–the overall range of prey consumed, and 2) diet consistency–the

Box 11.11 Determining Trophic Position of Fishes with Stable Isotope Analysis

Stable isotope data is often used to estimate the trophic position of fishes (Vander Zanden et al. 2000). Variation in trophic position can then be used to evaluate factors affecting fish foraging patterns across space or time.

In this example, isotope data were used to calculate the following trophic position estimates (TP) for different size walleyes. The relationship between TP and walleye size was then used to develop the equation

$$TP_{predicted} = 2.797 + 0.001445(L),$$

where predicted trophic position ($TP_{predicted}$) is estimated as a function of walleye length (L = total length in mm).

Table Trophic position (TP) versus walleye size. From this relationship, predicted TP and residuals are calculated.

Walleye size (mm)	TP	$TP_{predicted}$	|Residual|
152	3.5	3.0166	0.4834
254	3.6	3.1640	0.4360
305	3.4	3.2377	0.1623
355	3.5	3.3100	0.1900
381	3.61	3.3475	0.2625
457	3.42	3.4574	0.0374
508	3.45	3.5311	0.0811
584	3.61	3.6409	0.0309
609	3.59	3.6770	0.0870
660	3.7	3.7507	0.0507
Mean residual			0.1821

Residual values were calculated from the trophic position–body size relationship as the difference between TP – $TP_{predicted}$. Absolute residual values are then averaged to obtain a measure of trophic position variation that is independent of fish body size (i.e., 0.1821). Variation in trophic position can then be compared across time or space or correlated with biotic or abiotic variables to assess factors affecting diet variation of walleyes.

degree to which an individual fish repeatedly consumes the same prey type. High levels of variation indicate high diet breadth and high diet consistency, whereas low variation can represent either (a) high diet breadth and low consistency (all individuals consume similar proportions of a wide range of prey), or (b) low diet breadth (e.g., all individuals specialize on a few prey types; Vander Zanden et al. 2000). Variables, such as lake area, prey diversity, number of competitor species, food chain length, and lake productivity, are just a few parameters that can be compared with trophic position variation in an attempt to understand factors affecting feeding patterns. Similarly, variation in trophic position can be compared across seasons as a method for evaluating temporal changes within a population.

11.5 CONCLUSIONS

Food habit assessments are an integral part of many research and management plans. While specific goals of food habit studies vary, the usefulness of diet data relies on the accurate quantification of diet composition. Factors such as time of year and sample location can profoundly affect prey availability and diet composition of fishes. At smaller scales, time of day, habitat characteristics, and collecting gear can influence diet composition. The degree to which these factors affect interpretation of fish diets largely depends on the research question. As a result, it is important in diet studies to have well-defined research objectives that account for factors affecting diet composition.

Because diets can be quantified in many different ways, it is unlikely that a single diet index will be useful in all circumstances. Rather, we should rely on the unique properties of individual measures to select a method that is most appropriate for our study. As previously discussed, single indices based on prey number, weight (or volume), or occurrence each have their place in studies of diet composition. Similarly, graphical techniques that incorporate two or more single indices can provide important insight into feeding strategies, niche breadth, and relative prey importance. For questions concerning energy flow in aquatic ecosystems, techniques such as stable isotope analysis can provide powerful tools for quantifying important energy pathways to fishes.

Diet data have several important characteristics that affect analysis and interpretation: data are usually (1) multivariate in nature, (2) proportional, (3) variable at the individual level, and (4) autocorrelated across space and time. By appreciating and understanding these characteristics, we can design appropriate studies and select sound analytical techniques for assessing food habits and feeding patterns of fishes.

11.6 REFERENCES

Aitchison, J. 1983. Principal component analysis of compositional data. Biometrika 70:57–65.

Amundsen, P. A., H. M. Gabler, and F. J. Staldvik. 1996. A new approach to graphical analysis of feeding strategy from stomach contents data—modification of the Costello method. Journal of Fish Biology 48:607–614.

Angradi, T. R. 1994. Trophic linkages in the lower Colorado River: multiple stable isotope evidence. Journal of the North American Benthological Society 13:479–495.

Bettoli, P. W., and L. E. Miranda. 2001. Cautionary note about estimating mean length at age with subsampled data. North American Journal of Fisheries Management 21: 425–428.

Bowen, S. H. 1996. Quantitative description of the diet. Pages 513–522 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Bremigan, M. T., and R. A. Stein. 1994. Gape-dependent larval foraging and zooplankton size: implications for fish recruitment across systems. Canadian Journal of Fisheries and Aquatic Sciences 51:913–922.

Bremigan, M. T., and R. A. Stein. 1997. Experimental assessment of the influence of zooplankton size and density on gizzard shad recruitment. Transactions of the American Fisheries Society 126:622–637.

Bridcut, E. E., and P. S. Giller. 1995. Diet variability and foraging strategies in brown trout *(Salmo trutta)*: an analysis from subpopulations to individuals. Canadian Journal of Fisheries and Aquatic Sciences 52:2543–2552.

Cade, B. S., and B. R. Noon. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1:412–420.

Cade, B. S., and J. D. Richards. 2000. User manual for Blossom Statistical Software. Midcontinent Ecological Science Center, U.S. Geological Survey, Fort Collins, Colorado.

Catalano, M. J., S. R. Chipps, M. A. Bouchard, and D. H. Wahl. 2001. Evaluation of injectable fluorescent tags for centrarchid fishes: mark retention rate and effects on vulnerability to predation. North American Journal of Fisheries Management 21:911–917.

Chesson, J. 1983. The estimation and analysis of preference and its relationship to foraging models. Ecology 64:1297–1304.

Cortés, E. 1997. A critical review of methods of studying fish feeding based on analysis of stomach contents: application to elasmobranch fishes. Canadian Journal of Fisheries and Aquatic Sciences 54:726–738.

Costello, M. J. 1990. Predator feeding strategy and prey importance: a new graphical analysis. Journal of Fish Biology 36:261–263.

Crow, M. E. 1979. Multivariate statistical analysis of stomach contents. Pages 87–96 *in* S. J. Lipovsky and C. A. Simenstad, editors. Fish food habits studies: proceedings of the second Pacific Northwest technical workshop. Washington Sea Grant Program, University of Washington, Seattle.

De Crespin De Billy, V., S. Doledec, and D. Chessel. 2000. Biplot presentation of diet composition data: an alternative for fish stomach contents analysis. Journal of Fish Biology 56:961–973.

DeVries, D. R., and R. A. Stein. 1990. Manipulating shad to enhance sport fisheries in North America: an assessment. North American Journal of Fisheries Management 10:209–223.

Digby, P. G. N., and R. A. Kempton. 1987. Multivariate analysis of ecological communities. Chapman and Hall, New York.

Feller, R. J. 1992. Potential applications of immunoassays in studies of flatfish recruitment. Proceedings of the First International Symposium on Flatfish Ecology 29:1–3.

Fortin, M., and J. Gurevitch. 1993. Mantel tests: spatial structure in field experiments. Pages 342–359 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Fry, B., and E. B. Sherr. 1984. $\delta^{13}C$ measurements as indicators of carbon flow in marine and freshwater ecosystems. Contributions in Marine Science 27:13–47.

Gaedke, U., D. Straile, and C. Pahl-Wostl. 1996. Trophic structure and carbon flow dynamics in the pelagic community of a large lake. Pages 60–71 *in* G. A. Polis and K. O. Winemiller, editors. Food webs: integration of patterns and dynamics. Chapman and Hall, New York.

Garman, G. C. 1982. Identification of ingested prey fish based on scale characteristics. North American Journal of Fisheries Management 2:201–203.

Garvey, J. E., N. A. Dingledine, N. S. Donovan, and R. A. Stein. 1998a. Exploring spatial and temporal variation within reservoir food webs: predictions for fish assemblages. Ecological Applications 8:104–120.

Garvey, J. E., E. A. Marschall, and R. A. Wright. 1998c. From star charts to stoneflies: detecting relationships in continuous bivariate data. Ecology 79:442–447.

Garvey, J. E., R. A. Wright, R. A. Stein, K. H. Ferry, and S. M. Micucci. 1998b. Assessing the influence of size on overwinter survival of largemouth bass in Ohio on-stream impoundments. Ohio Division of Wildlife, Final Report, F-69-P, Columbus, Ohio.

Goldberg, D. E., and S. M. Scheiner. 2001. ANOVA and ANCOVA: field competition experiments. Pages 77-98 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Oxford University Press, New York.

Hansson, S. 1998. Methods of studying fish feeding: a comment. Canadian Journal of Fisheries and Aquatic Sciences 55:2706–2707.

Hartman, K. J., and D. W. Garton. 1992. Electrophoretic identification of partially digested prey of piscivorous fish. North American Journal of Fisheries Management 12:260–263.

Hayward, R. S., F. J. Margraf, C. T. Knight, and D. J. Glomski. 1989. Gear bias in field estimation of the amount of food consumed by fish. Canadian Journal of Fisheries and Aquatic Sciences 46:874–876.

Holland-Bartels, L. E., S. K. Littlejohn, and M. L. Huston. 1990. A guide to the larval fishes of the upper Mississippi River. U.S. Fish and Wildlife Service Publication, LaCrosse, Wisconsin.

Huang, C., and A. Sih. 1991. Experimental studies on direct and indirect interactions in a three-trophic level stream system. Oecologia 85:530–536.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Hynes, H. B. N. 1950. The food of freshwater sticklebacks (*Gasterosteus aculeatus* and *Pygosteus pungitius*) with a review of methods used in studies of the food of fishes. Journal of Animal Ecology 19:36–58.

Hyslop, E. J. 1980. Stomach contents analysis: a review of methods and their application. Journal of Fish Biology 17:411–429.

Jackson, D. A. 1997. Compositional data in community ecology: the paradigm or peril of proportions? Ecology 78:929–940.

Jernejcic, F. 1969. Use of emetics to collect stomach contents of walleye and largemouth bass. Transactions of the American Fisheries Society 98:698–702.

Juanes, F. 1994. What determines prey size selectivity in piscivorous fishes? Pages 79–100 *in* D. J. Stouder, K. L. Fresh, and R. J. Feller, editors. Theory and application in fish feeding ecology. University of South Carolina Press, Columbia, South Carolina.

Khattree, R., and D. N. Naik. 1999. Applied multivariate statistics with SAS software, 2nd edition. SAS Institute, Cary, North Carolina.

Kim, G. W., and D. R. DeVries. 2001. Adult fish predation on freshwater limnetic fish larvae: a mesocosm experiment. Transactions of the American Fisheries Society 130: 189–203.

Kimball, D. C., and W. T. Helm. 1971. A method of estimating fish stomach capacity. Transactions of the American Fisheries Society 100:572–575.

Kimura, D. K. 1977. Statistical assessment of the age–length key. Journal of the Fisheries Research Board of Canada 34:317–324.

Knight, R. L., and J. F. Margraf. 1982. Estimating stomach fullness in fishes. North American Journal of Fisheries Management 2:413–414.

Krebs, C. J. 1989. Ecological methodology. Harper Collins, New York.

MacDonald, J. S., and R. H. Green. 1983. Redundancy of variables used to describe importance of prey species in fish diets. Canadian Journal of Fisheries and Aquatic Sciences 40:635–637.

Maceina, M. J., P. W. Bettoli, and D. R. DeVries. 1994. Use of a split-plot analysis of variance design for repeated-measures fisheries data. Fisheries (Bethesda) 19(3):14–20.

Manly, B. F. J. 1974. Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

Mittelbach, G. C., C. W. Osenberg, and P. C. Wainwright. 1999. Variation in feeding morphology between pumpkinseed populations: phenotypic plasticity or evolution? Evolutionary Ecology Research 1:111—128.

Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied linear statistical models: regression, analysis of variance, and experimental designs. Irwin Publishing, Boston.

Noble, R. L. 1981. Management of forage fishes in impoundments of the southern United States. Transactions of the American Fisheries Society 110:738–750.

Peterson, B. J., and B. Fry. 1987. Stable isotopes in ecosystem studies. Annual Review of Ecology and Systematics 18:293–320.

Polis, G. A. 1991. Complex trophic interactions in deserts: an empirical critique of food web theory. American Naturalist 138:123–155.

Pope, K. L., M. L. Brown, W. G. Duffy, and P. H. Michaletz. 2001. A caloric-based evaluation of diet indices for largemouth bass. Environmental Biology of Fishes 61:329–339.

Probst, W. E., C. F. Rabeni, W. G. Covington, and R. E. Marteney. 1984. Resource use by stream-dwelling rock bass and smallmouth bass. Transactions of the American Fisheries Society 113:283–294.

Raclot, T., R. Groscolas, and Y. Cherel. 1998. Fatty acid evidence for the importance of myctophid fishes in the diet of king penguins, *Aptenodytes patagonicus*. Marine Biology 132:523–533.

Rasmussen, P. W., D. M. Heisey, E. V. Nordheim, and T. M. Frost. 1993. Time-series intervention analysis: unreplicated large-scale experiments. Pages 138–158 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Reiriz, L., A. G. Nicieza, and F. Brana. 1998. Prey selection by experienced and naive juvenile Atlantic salmon. Journal of Fish Biology 53:100–114.

SAS Institute. 1999. SAS procedures guide, version 8. SAS Institute, Cary, North Carolina.

Schael, D. M., L. G. Rudstam, and J. R. Post. 1991. Gape limitation and prey selection in larval yellow perch *(Perca flavescens),* freshwater drum *(Aplodinotus grunniens),* and black crappie *(Pomoxis nigromaculatis).* Canadian Journal of Fisheries and Aquatic Sciences 38:1919–1925.

Scharf, F. S., J. A. Buckel, F. Juanes, and D. O. Conover. 1997. Estimating piscine prey size from partial remains: testing for shifts in foraging mode by juvenile bluefish. Environmental Biology of Fishes 49:377–388.

Scharf, F. S., F. Juanes, and M. Sutherland. 1998. Inferring ecological relationships from the edges of scatter diagrams: comparison of regression techniques. Ecology 79:448–460.

Schultz, D. R., and M. E. Clarke. 1995. An immunological study of predation on hatchery-reared, juvenile red drum *(Sciaenops ocellatus)*: preparation and assays of a "red drum-specific" protein for predator–prey experiments. Journal of Experimental Marine Biology and Ecology 189:1–2.

Seaburg, K. G. 1957. A stomach sampler for live fish. Progressive Fish-Culturist 19:137–139.

Shepard, W. C., and E. L. Mills. 1996. Diel feeding, daily food intake, and *Daphnia* consumption by age-0 gizzard shad in Oneida Lake, New York. Transactions of the American Fisheries Society 125:411–421.

Smith, E. P., and T. M. Zaret. 1982. Bias in estimating niche overlap. Ecology 63:1248–1253.

Somerton, D. A. 1991. Detecting differences in fish diets. U.S. National Marine Fisheries Service Fishery Bulletin 89:167–169.

Strauss, R. E. 1979. Reliability estimates for Ivlev's electivity index, the forage ratio, and a proposed linear index of food selection. Transactions of the American Fisheries Society 108:344–352.

Sutela, T., and A. Huusko. 2000. Varying resistance of zooplankton prey to digestion: implications for quantifying larval fish diets. Transactions of the American Fisheries Society 129:545–551.

Swynnerton, G. H., and E. B. Worthington. 1940. Notes on the food of fish in Haweswater. Journal of Animal Ecology 9:183–187.

ter Braak, C. J. F., and P. Smilauer. 1998. CANOCO for Windows, version 4.0. Center for Biometry, Wageningen, The Netherlands. (Distributed by Microcomputer Power, 113 Clover Lane, Department C8, Ithaca, New York, 14850-4930, USA. Available at http://www.microcomputerpower.com/.)

Trippel, E. A., and F. W. H. Beamish. 1987. Characterizing piscivory from ingested remains. Transactions of the American Fisheries Society 116:773–776.

Vander Zanden, M. J., G. Cabana, and J. B. Rassmussen. 1997. Comparing the trophic position of littoral fish estimated using stable nitrogen isotopes ($\delta^{15}N$) and dietary data. Canadian Journal of Fisheries and Aquatic Sciences 54:1142–1158.

Vander Zanden, M. J., and J. B. Rasmussen. 1999. Primary consumer $\delta^{15}N$ and $\delta^{13}C$ and the trophic position of aquatic consumers. Ecology 80:1395–1404.

Vander Zanden, M. J., B. J. Shuter, N. P. Lester, and J. B. Rasmussen. 2000. Within- and among-population variation in the trophic position of a pelagic predator, lake trout *(Salvelinus namaycush).* Canadian Journal of Fisheries and Aquatic Sciences 57:725–731.

Wallace, R. K. 1981. An assessment of diet-overlap indexes. Transactions of the American Fisheries Society 110:72–76.

Wallace, R. K., and J. S. Ramsey. 1983. Reliability in measuring diet overlap. Canadian Journal of Fisheries and Aquatic Sciences 40:347–351.

Welker, M. T., C. L. Pierce, and D. H. Wahl. 1994. Growth and survival of larval fishes: roles of competition and zooplankton abundance. Transactions of the American Fisheries Society 123:703–717.

Werner, E. E., and D. J. Hall. 1974. Optimal foraging and the size selection of prey by the bluegill sunfish *(Lepomis macrochirus)*. Ecology 55:1042–1052.

Wright, R. A., L. B. Crowder, and T. H. Martin. 1993. The effects of predation on the survival and size-distribution of estuarine fishes—an experimental approach. Environmental Biology of Fishes 36:291–300.

Windell, J. T., and S. H. Bowen. 1978. Estimating food consumption rates of fish populations. Pages 227–254 *in* T. Bagnel, editor. Methods for assessment of fish production in fresh waters, 3rd edition. Blackwell Scientific Publications, Oxford, UK.

Yoshioka, T., E. Wada, and H. Hayashi. 1994. A stable isotope study on seasonal food web dynamics in a eutrophic lake. Ecology 75:835–846.

12 Bioenergetics

Kyle J. Hartman and Robert S. Hayward

12.1 INTRODUCTION

Bioenergetics is the study of pathways and mechanisms through which energy enters and is then stored, used, and ultimately lost from living organisms to allow for their maintenance, growth, and reproduction. Bioenergetics investigations have ranged from those focused at the molecular and cellular levels to those focused at population, community, and ecosystem levels. Within fisheries and aquatic ecology, bioenergetics investigations have focused primarily at the organismal level but have often been extrapolated to populations and communities. However, as our demand increases for more accurate and detailed estimates and predictions of fish energetics responses, bioenergetics studies at finer organizational levels will likely become necessary.

Studies of bioenergetics in fisheries have sought mainly to develop estimators of energy consumption and growth for the "average" fish. These estimators have as their underpinning the following balanced energy equation:

$$C = G + (M + \text{SDA}) + F + U. \tag{12.1}$$

Consumption (C) represents the amount of energy ingested by a fish over some common unit of time. Growth (G) is the increase, or decrease, in energy content of the fish that occurs after various energy costs and losses have been subtracted from C. Metabolism (M) comprises both a basal and active component, and specific dynamic action (SDA) is the energy lost as heat during the chemical transformation of food into utilizable energy. Egestion (F) and excretion (U) are waste products of ingestion. Bioenergetics models are based on the balanced energy equation for which knowledge of the values for any five variables permits solving for the sixth. Values of the energy equation variables (usually C or G) are often estimated from the field whereas values of M, SDA, F, and U are typically determined from laboratory studies.

Although fisheries bioenergetics studies have arisen from seemingly limited objectives, a broad and ever increasing array of questions have been addressed. At the heart of these objectives has remained an interest in estimating the growth or

consumption of fishes. Traditional methods include in situ studies in which consumption is estimated by sampling fish over diel periods and applying gastric evacuation information in a variety of model forms (Jobling 1986). During the 1970s and 1980s, development and availability of microcomputers permitted the ready use of bioenergetics models that estimated consumption or growth from mathematical models of fish energetics parameters and commonly collected field data (Kitchell et al. 1974, 1977; Hewett and Johnson 1987). We begin this chapter by discussing the individual components of fish energy budgets and the experimental designs and data analyses commonly used to develop estimators of these components. Approaches for combining the component predictors into whole bioenergetics models and applications of bioenergetics models, as well as approaches for evaluating their predictive accuracies, are considered. We also provide suggestions for improving and standardizing procedures in fish energetic studies and highlight areas where additional research is most warranted.

12.2 FISH ENERGY BUDGETS

Fish energy budgets, such as the one shown in Figure 12.1, often provide the basis for indirect estimation of fish consumption and growth rates in field settings by means of bioenergetics models. For example, estimation of consumption rate (C) by fish over a given period would require equation inputs of fish growth over that period (Chapter 5). In addition, other values would be required to enable the estimation of the energy cost and loss terms M, SDA, F, and U.

Most bioenergetics terms have been derived from measures on captive fishes; however, recent technological advances have permitted limited estimates on fishes in natural settings. Metabolic costs and SDA are typically measured as depletion of oxygen concentrations in closed respirometry chambers (Brett 1970, 1976; Beamish 1974, 1990; Brett and Groves 1979; Cech 1990), with the resulting oxygen consumption converted into units of energy by means of oxycalorific coefficients (e.g., 13,730 J/g, Elliott and Davidson 1975). Telemetry and videography have been applied to fishes in freshwater systems to estimate active metabolism (Lucas et al. 1991; Krohn and Boisclair 1994). Active metabolism has also been estimated using an algebraic solution of the balanced energy equation where consumption and growth are known and laboratory or field estimates exist for the other parameters (F, U, and SDA); the model is solved for the activity level (ACT) required to balance the equation (Boisclair and Leggett 1989; Hartman and Brandt 1995a). Egestion and excretion are most often estimated on captive fish for which collection of feces and monitoring of nitrogenous wastes can be conducted. Growth in the bioenergetics models is often constrained by the maximum consumption function. This function dictates the maximum rate of consumption based upon fish size and temperatures experienced by the fish and are typically developed from ad libitum feeding experiments.

A common extension of energy budgets is to calculate a scope for growth for a given species and size of fish (Warren and Davis 1967; Brett 1970; Kitchell et al. 1977). Scope for growth is the potential growth rate of a fish under the range of

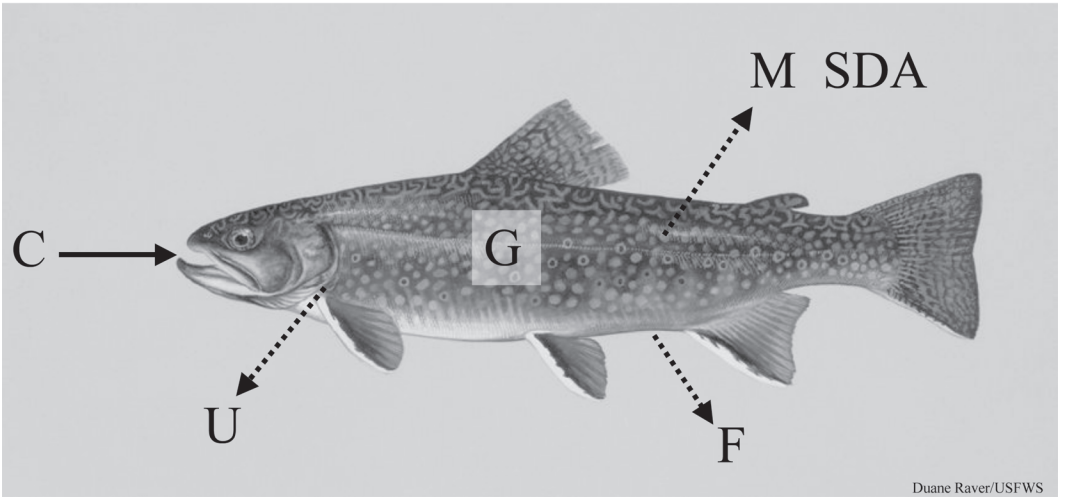


Figure 12.1 A conceptual visualization of the bioenergetics components in the balanced energy equation for fishes, where *C* is consumption, *G* is growth, *M* is metabolism, SDA is specific dynamic action, *F* is egestion, and *U* is excretion.

feeding levels and environmental temperatures it may encounter (Brett 1976). The growth scope allows a fisheries scientist to assess rapidly how well a particular species may do if stocked into a novel environment, or it may be useful as a diagnostic tool in fisheries that appear to be growth limited. Providing the fisheries scientist has information on the water temperatures available to fish in a body of water, and given assumptions about food availability, scope for growth can be used to evaluate a water body for introductions of species (see Box 12.1).

12.3 BIOENERGETIC LABORATORY EXPERIMENTS

Laboratory experiments have been used to measure feeding levels and energetics components of fishes under a variety of controlled conditions. This section begins by describing key energetics components and the factors that affect them. Because similarity exists among experiments used to measure metabolic costs and losses in egestion and excretion, many of the aspects of design, analysis, and descriptive model development for laboratory energetics studies have been handled together. Therefore, this section is further organized into experimental analysis and design sections with descriptions of the various energetics components and factors that affect them.

Discussion of laboratory consumption experiments is complicated by the fact that they are often run for different objectives that require distinct designs and analyses. Often, experiments are aimed at defining the maximum consumption (C_{max}) level for a given-size fish under different thermal conditions. Models or functions that describe relationships among C_{max}, fish size, and water temperature have been widely used in bioenergetics model development (Hanson et al. 1997) as well as in applications, where they are often used to estimate the predatory

Box 12.1 Scope for Growth

Scope for growth is a mathematical expression of the possible growth of a fish given the energetic constraints at different temperatures and fish sizes. It is most often presented graphically, showing how the relationships change with temperature for a given size fish. Below is an example of scope for growth for a 10-g fish. Growth is bounded by C_{max} (maximum consumption) and the sum of all the energetic costs and losses (M + SDA + F + U; see Figure 12.1 for abbreviations). Here, greatest growth is possible from 13 to 18°C. Studies of wild fish populations find that most wild fish feed at 40–60% of C_{max}, or at a P-value (proportion of C_{max}) of 0.4 to 0.6. This level of feeding can be assumed by fisheries scientists in assessing how well a species may do, or may be doing, in a system based on the scope of growth model. Thus, a similar graph could be constructed to evaluate potential suitability of a site for stocking a novel species by replacing C_{max} below with a more realistic expectation of wild fish feeding, such as 0.6 (C_{max}). If the resulting scope for growth was positive across most of the temperatures available to the fish in this site, then the site would likely have suitable thermal conditions for that species.

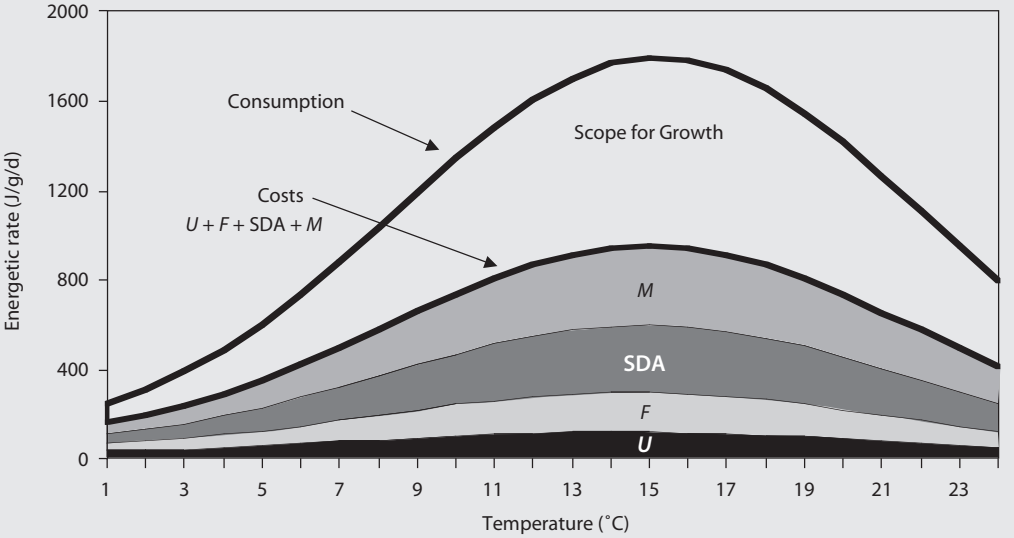


Figure Scope for growth for a 10-g fish. Shaded areas under each curve represent specific rates at each temperature for the cost terms *U, F,* SDA, and *M* (see Figure 12.1 for abbreviations).

demand of wild fish populations on their prey (Stewart et al. 1981; Brandt and Hartman 1993; LaBar 1993). Other consumption experiments have aimed to evaluate optimum temperatures for fishes in cultured and wild settings or to estimate maintenance ration requirements by fishes under various conditions.

12.3.1 Sources of Variation in Bioenergetics Experiments

Because fish are poikilotherms, the prevailing water temperatures dictate physiological rates. In addition to temperature, fish size or mass also influences metabolism and consumption parameters. Any factors that increase stress or alter

activity may also affect bioenergetic rates. Stressors such as decreased oxygen concentrations may reduce feeding levels and resultant growth (Chiba 1988; Crocker and Cech 1997), and fish may alter activity levels as a means of escaping or waiting out the perturbation. Prey types may also affect maximum consumption rates (Grove et al. 1978; Hartman and Brandt 1993), and social interaction among conspecific fish can profoundly influence consumption rate and cost terms through antagonistic interactions or stress, even at low densities (Hartman and Brandt 1993). Well-fed fish or those in high condition may have less motivation to feed during experiments than those in poor condition, whereas fish in poor health may reduce or cease feeding and activity. Hyperphagic responses have also been noted in a growing list of species, and it is likely this phenomena is more the rule than the exception in fishes (Dobson and Holmes 1984; Quinton and Blake 1990; Russell and Wootton 1992; Jobling et al. 1993, 1994; Paul et al. 1995; Bull and Metcalfe 1997; Hayward et al. 1997; Whitledge et al. 1998; Gaylord and Gatlin 2000; Hayward and Wang 2001). Evidence also suggests that seasonal cues may influence metabolism and activity in some fishes (Olla and Studholme 1971; Chipps et al. 2000). This great range of factors that influence bioenergetic parameters makes controlling them an obstacle in model development. The best advice to give persons conducting bioenergetic experiments is that while measuring bioenergetic parameters every attempt should be made to match the conditions under which the models will be applied.

12.3.2 Design of Fish Energetics Experiments

12.3.2.1 Experiment Duration

Metabolism experiments.—Although typically of short duration, metabolism experiments include an acclimation period to allow the influence of prior feeding to fade, to permit acclimation to a new exposure temperature, and to allow time for the fish to become acclimated to the experimental chamber. Measurements of oxygen uptake rates (g O_2/g fish/h) should be taken 3–5 times daily over a 3–7-d period. These initial data indicate fish acclimation to the test tanks via inspection of the pattern of oxygen uptake over time for evidence of stabilization. A standard curve, such as a negative power function, can be fit to oxygen consumption data over time, and a standard point at which the curve asymptotes can be used to define fish acclimation. After observations on several fish at each temperature, the acclimation protocol is established and followed on subsequent measurements. Metabolism measurements begin following the acclimation period, with several measures being made for each fish. The average of these measures is used to represent that fish at the specific combination of fish size and water temperature. The duration of a metabolism experiment will vary depending upon the gastric evacuation rate (required to clear the gut from previous meals) and the time required for acclimation. Typically this is less than 4 d per experiment; longer times may induce metabolic compensation due to lack of feeding and give erroneous results.

C_{max} experiments.—Maximum consumption experiments involve estimating the average daily ration (g/g fish/d) of captive fish under ad libitum feeding. The duration of such experiments has varied greatly from just a few days to hundreds of days. The choice of duration is important in the analysis and interpretation of the data in that fish offered unlimited food tend to eat less over time as their size and energy reserves increase; thus, the average maximum consumption rate will necessarily decline in longer-duration experiments.

The length of C_{max} experiments should be matched to the intended use of the data and resulting bioenergetics model. In aquaculture applications, where fish are often fed at, or near, ad libitum ration, experiments should be of durations that simulate periods under which production or production models are to be used. For C_{max} experiments relating to the development of bioenergetics models for free-living fish (particularly for estimation of consumption from growth data), durations can be shorter (e.g., 4–14 d). Longer periods of unlimited food supply are not typical in nature; moreover, using shorter experiment periods permits additional replication, provides time savings, and allows better experimental control of size- and stress-related effects on ration levels. Due to the effect of feeding history on C_{max}, it is recommended that feeding regime be standardized as part of acclimation for consumption experiments. For example, an average feeding level of 50% ration during acclimation can be achieved by feeding fish ad libitum every other day, thereby establishing a uniform feeding pattern and hunger level prior to experimentation.

12.3.2.2 Treatment Levels and Replication

Many studies have shown the importance of water temperature and body size upon consumption and metabolism of fishes (Winberg 1956; Brett and Groves 1979; Penczak 1990). Therefore, at a minimum, these two variables are usually considered in the design of energetics experiments. In most studies to date, more attention has been paid to the temperature effect than to the size effect upon C_{max} or metabolism. Studies commonly employ 4–8, and as many as 15, temperature treatments, whereas body size treatments are typically 1–5 in number (Table 12.1). The reason for the disparity in treatment levels among temperature and mass is not that size is unimportant. Rather, size is more difficult to control than is temperature, and patterns of change with size are more consistent than with temperature: slopes can both increase and decrease along the temperature continuum. Nearly all energetics studies that have looked at size as a treatment have found the relationship between specific consumption and size to be a negative power function (see Figure 12.2, upper panel). Based upon this apparent relationship it would be best to have a minimum of four size treatments over a wide range of sizes to define the relationship. At a minimum, attention should be paid to ensure that both small and large fish are included in the experiments, as prior study has shown size dependence to be a simple power function relationship (linearized with a $\log_{10}$ transformation).

The number of treatment levels for water temperature should depend upon the range of environmental temperatures occupied by the fish. Most researchers

Table 12.1 Review of some previously published bioenergetics laboratory studies that have applied treatment levels and replication. Size range is the range in weights of fish studied, temperature range is the range in temperature treatments, temperature treatments is the number of temperature treatments, temperature replicates is the number of replicates for each temperature treatment (or number of observations [obs.]), duration is the length of the experiment in days, and model type is the form of equation to describe the relationship, where 1 represents an exponential temperature function and power function of size and 3 represents the Thornton and Lessem (1978) algorithm relating temperature effects and a power function of size. These model type numbers correspond with those depicted in Figure 12.2, lower panel. Sources of each study are as follows: (1) Hartman and Brandt 1995a; (2) Stewart et al. 1983; (3) Wahl and Stein 1991; (4) Vigg and Burley 1991; (5) Elliot 1976; (6) Beamish 1990; (7) Paul et al. 1990; (8) Smith et al. 1988; (9) Moser and Hettler 1989; (10) Paul et al. 1988; and (11) Lezama and Guenther 1992.

Species	Size range (g)	Temperature range (°C)	Temperature treatments	Temperature replicates	Duration (d)	Model type	Source
			Maximum consumption				
Striped bass	5–2,700	2–30	7–8	4–24	3–7	3	1
Bluefish	26–1,013	10–29	5	5–11	3–7	3	1
Weakfish	6–1,251	15–28	4	5–11	3–7	3	1
Lake trout	40–1,200	3.5–20.0	1–8		84	1	2
Escoids	7–44	5–30	7	9–15	14	1[a]	3
Northern pikeminnow	501–2,000	8–21.5	4		30	1, 3	4
Brown trout	10–306	3.8–21.6	15	11–20		1	5
			Metabolism				
Striped bass	7–2,886	6–30	6–8	6–24	2	1	1
Bluefish	19–1,013	15–29	4	8–16	2	1	1
Weakfish	5–1,318	12–29	4	8–16	2	1	1
Lake trout	20–1,800	3.5–15.0	5	578 obs.	1	1	2
Walleye	5–7	5–23.5	5	12			6
Escoids	9.5–53.2	5–30	5	13		1	3
Yellowfin sole	3–603	1.5–10.0				Linear	7
Walleye pollock	345–750	1.6–7.1	4				8
Spot	3–24	21–32	2	≥4		[b]	9
Pacific cod	600–1,300	2–15		122 obs.		1	10
Jaguar guapote	6–800	22–32	3	26–37	<1		11

[a] Model with a third-order polynomial.

[b] Multiple regression with temperature, salinity, and wet weight fish.

have selected an upper and lower treatment temperature and then evenly distributed the treatments within that range. The reason for this is the response of C_{max} or metabolism to temperature varies across species and may also vary with ontogeny (Hartman and Brandt 1995a). In many species, C_{max} and metabolism increase to a point, levels off, and then declines as temperatures become stressful (see Figure 12.2, lower panel). Defining the optimum temperature for consumption and the upper temperature ranges at which consumption often declines has required researchers to use a larger number of treatments for temperature than for size.

The number of replicates (individual fish or fish groups) for each treatment level or combination of treatments has also varied greatly in the literature (Table 12.1). The number of replicates used appears more often determined from availability of laboratory space or fish than from any statistical rationale. In designing C_{max} or metabolism experiments, researchers should consider statistical power as well as logistical constraints in the determination of replication levels. Metabolism and C_{max} experiments aimed at determining size and temperature effects are well suited to analysis of variance (ANOVA) designs, as long as size and temperature have no interactive effect (see section 12.3.3).

Within size treatments, the number of replicates needed to describe the effects of temperature upon C_{max} or metabolism will vary with species and other factors. It will also vary depending upon the level of accuracy desired in estimating the dependent variable. Among-individual variability in energetics responses tends to be highest where rates are highest. For example, small fish tend to exhibit higher specific rates and more variation than do large fish; and with the exception of low temperatures where some individuals may not feed, within a size-class variability is often highest near the optimum temperature for consumption (see Figure 12.3).

For example, based on the variability in C_{max} and metabolism data for three species of fish (striped bass, weakfish, and brook trout), we can determine a range of replicates (number of individual fish) needed to achieve a given level of accuracy (Table 12.2). Sample size can be determined by the equation,

$$n = (1.96^2)(\sigma^2)/A^2, \tag{12.2}$$

where n is the required sample size to be within A units of the population mean (determined by the level of accuracy) with 95% confidence, and σ^2 is the variance in the estimates (Devore and Peck 1986). For example, if the mean was 0.100 g/g/d and we wanted to be within 10% of the true value with 95% confidence, A would be 10% of the mean, in this case, 0.010 g/g/d. This value of A would be used in calculating the required replication level.

In the data sets for striped bass, weakfish, and brook trout, the mean C_{max} levels at optimum temperatures were 0.112, 0.227, and 0.120 g/g/d, respectively. Variances were generally low, but the level of replication needed for a given level of accuracy in estimating C_{max} varies substantially among these data sets. To be 95% confident that our estimate of C_{max} will fall within 25% of the true population mean would require from 2 to 6 fish (Table 12.2). The same confidence to be within 10% of the population C_{max} would require 9–33 fish and to be within 5% of

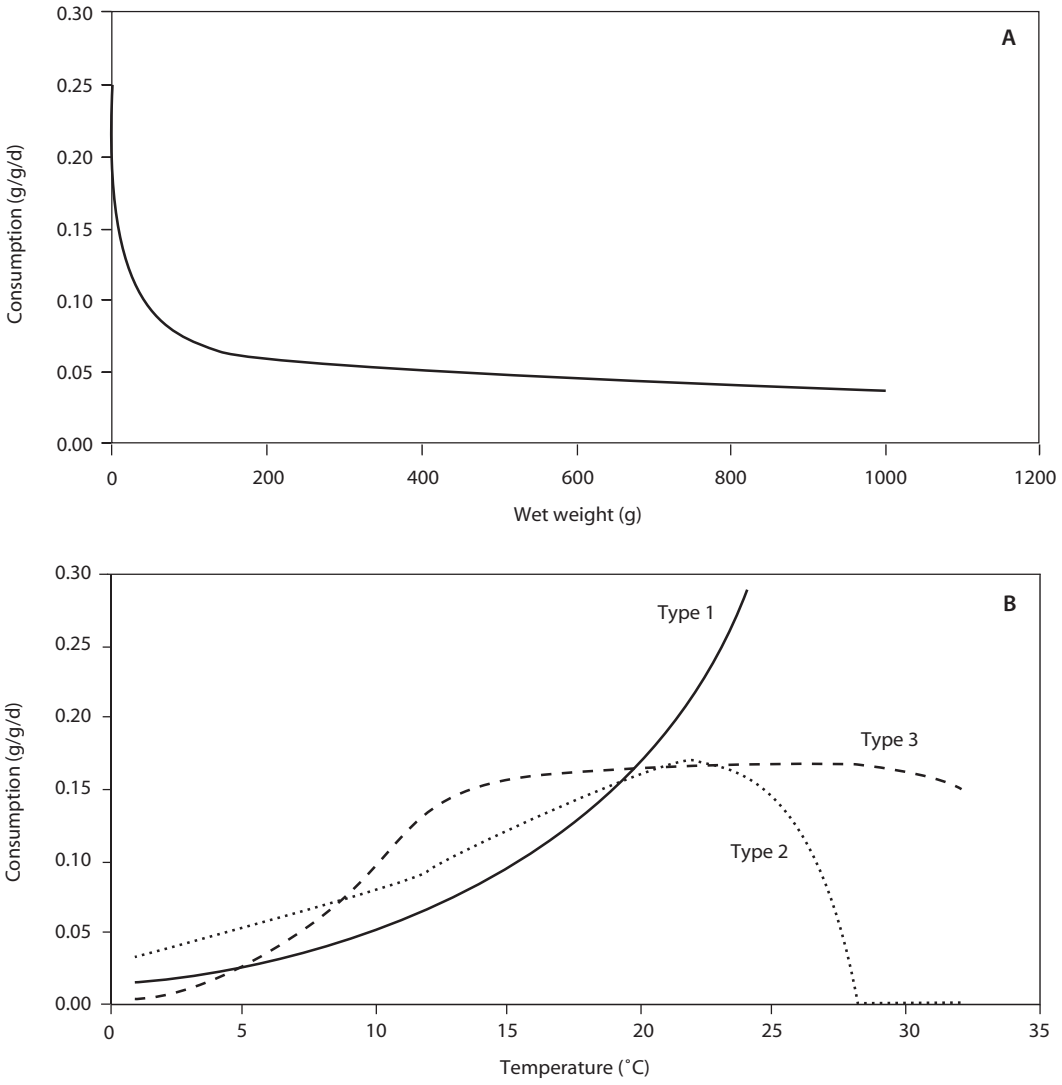


Figure 12.2 Within a species, the two most important influences on C_{max} are fish mass (top panel) and temperature (bottom panel). Specific consumption is a negative power function of weight. Temperature has been modeled as one of three types as defined in Hanson et al. (1997). Type 1 is an exponential function, whereas type 2 (Kitchell et al. 1977) and type 3 (Thornton and Lessem 1978) are different models that fit one form to the increasing limb and another to the decreasing limb of the data. Types 2 and 3 are most often used where fish may experience temperatures above their thermal optimum.

the population mean with 95% confidence would require 33–131 replicates. To be within 1% of the true population parameter for C_{max} would require over 800 replicates. The number of replicates needed for a given level of accuracy in metabolism experiments yield estimates of 1–5 fish for 25% and 6–28 fish for 10% accuracy with 95% confidence (Table 12.2).

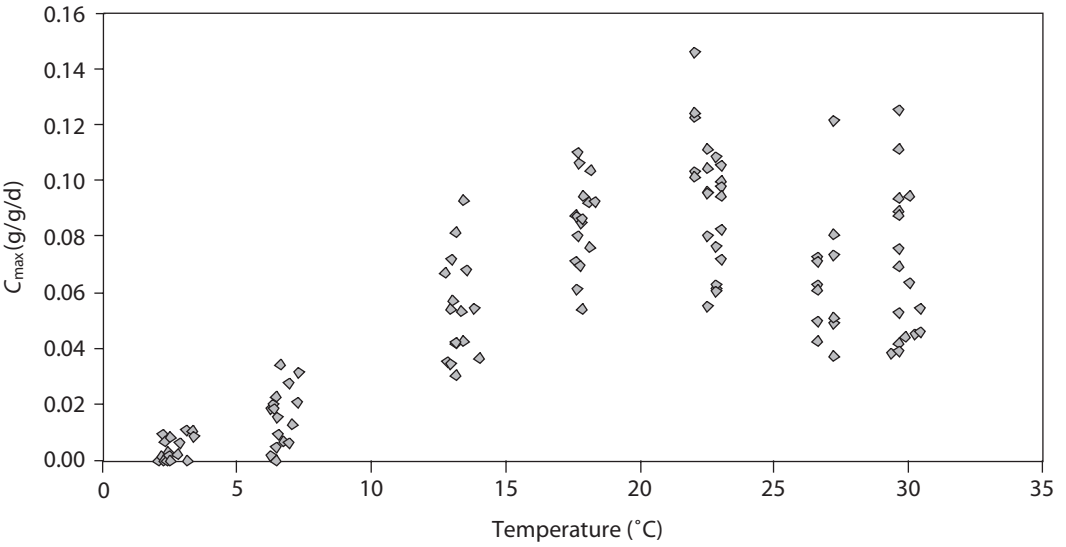


Figure 12.3 Variability in C_{max} estimates among individual striped bass is greatest near the optimum temperature for consumption—23°C. In assessing sample size requirements for energetics studies, pilot studies near the expected optimum temperature for consumption (or metabolism) can be run to provide a measure of variability for calculation of sample size requirements (figure from Hartman and Brandt 1995a).

Given the rapid increase in the number of replicates required for increasingly modest improvements in estimate accuracy, it is wise to choose an acceptable level of accuracy for estimates a priori and then proceed with experiments near the optimum temperature for the species to gain a measure of variance. With this measure of variance and the mean C_{max} or metabolism value, the researcher can then use equation (12.2) to evaluate the number of replicates needed for the desired level of accuracy. We recommend the use of a priori data to establish replication levels in experiments, but in the absence of such data we recommend 10–15 replicates as this generally will place the energetic estimate within about 10% of the true mean value with 95% confidence based upon data in Table 12.2.

12.3.2.3 Repeated Measures Versus Random Factorial Design

Energetics experiments, particularly consumption and metabolism experiments, have been conducted both as a repeated-measures design and a completely random design with factorial treatments. There are pros and cons associated with either approach. The completely random with factorial treatment (CRFT) design is a purer statistical design as each fish contributes only a single observation. The disadvantage of the factorial approach is that it requires the evaluation of many individual fish and, consequently, substantial laboratory space and equipment. For the example outlined above (Table 12.2) a C_{max} experiment of 4 size-classes of fish to be evaluated at 5–6 temperatures with 10–20% accuracy would require 12

Table 12.2 Data on the mean rates and variability in estimates of C_{max} and metabolism for three species of fish. The sample size (number of replicates) needed to achieve a given level of accuracy (25, 10, 5, or 1%) with 95% confidence is also provided. These values were obtained using equation (12.2). The mean given is either the mean C_{max} (g/g/d) or the mean metabolism (g O_2/g fish/h), respectively.

Species	Optimum temperature (°C)	Weight range (g)	Mean	Variance	Replicates needed for accuracy			
					25%	10%	5%	1%
				C_{max}				
Striped bass	22–23	10.4–32.9	0.1173	0.00117	5	33	131	3,267
Weakfish	24–25	14.9–20.4	0.2272	0.00111	1	8	33	826
Brook trout	16–21	6.1–22.7	0.1202	0.00031	1	8	33	824
				Metabolism				
Striped bass	27–30	15.2–53.4	0.00632	$2.9222 \cdot 10^{-6}$	5	28	112	2,811
Weakfish	24–25	5.3–25.1	0.00626	$5.4270 \cdot 10^{-7}$	1	5	21	532
Bluefish	26–30	248–1,013	0.00780	$2.0926 \cdot 10^{-6}$	2	13	53	1,321

replicates; in total, 240–288 fish would be required for a CRFT design. Often laboratories are not equipped to hold this number of fish simultaneously. Researchers have often attempted to sidestep the space limitation issue by using a repeated-measures approach by which each individual receives multiple measurements (e.g., temperature or ration), or by conducting studies on small fish and applying the results to larger fish, often with assumed reductions in the thermal optima range. One disadvantage of the repeated-measures approach is that individual fish's responses carry substantial weight such that an atypical fish could greatly influence overall results. Also, true repeated-measures designs require that each fish be exposed to the various treatment levels in a randomly selected order that is often not possible due to space constraints. Repeated-measures analysis is often performed without meeting this condition. Nonetheless, the repeated-measures approach is usually preferable to eliminating treatment levels (such as fish size) from analyses because of the impractical number of observations required for CRFT designs.

12.3.2.4 *Experimental Treatments*

Most energetics experiments have similar experimental treatment designs. In metabolism and C_{max} experiments, size and water temperature are the typical treatments. For experiments aimed at defining egestion, excretion, and specific dynamic action, temperature and ration level effects are often evaluated. With two variables (weight and temperature or temperature and ration), the experimental design will vary slightly depending upon the intended objectives of the study. For example, if the experiments are truly aimed at identifying the influence of temperature and fish size upon metabolism or consumption, then the preferred design should be an $n \times m$ completely randomized factorial design with n size treatments and m temperature treatments. Ideally, in such a design, the treatments and replicates are all run simultaneously. Again, in most laboratory facilities space and temperature control limitations conspire to make this impossible. A statistician might argue that if all experiments are not run simultaneously then they must be randomly selected for experimental temperature treatment sequence to ensure that time (or treatment level order) is not influencing the results. Thus, in this experimental application, all size-classes should be run simultaneously with the temperature treatments randomly selected for each series of experiments until the block is filled. In this case, analysis would consist of ANOVA with treatment sequence as the blocked variable and temperature and weight as the categorical variables.

Often, energetics experiments are conducted with the intent of defining parameters for bioenergetics models. Thus, the objective in these types of experiments is to develop predictive models that will define values of key energetics parameters for fish under various environmental conditions. Regression analysis is commonly used to define relationships between a given energetic parameter and size and temperature. Analysis for these experiments can vary from the ANOVA detailed above to a mixed regression model in which temperature is a categorical explanatory variable and weight is a continuous explanatory variable. Regardless

of the design, the aim is to construct an equation that describes the influences of temperature and weight upon specific rates of consumption or metabolism—typically using regression.

12.3.3 Analysis of Fish Energetics Data

Here, we consider data from C_{max} experiments as an example, but similar procedures can be used for analyzing metabolic rate or other energetics component data. A reasonable first step in analyzing energetics data is to construct plots to determine whether relationships appear linear or curvilinear. If the plot of the energetic variable (e.g., C_{max} or metabolism) versus water temperature, as well as the plot of the energetic variable versus fish weight is linear (or can be linearized with transformation), then a simple linear regression model can be employed for analysis. If the relationship between the dependent variable and weight or temperature is curvilinear, then a quadratic equation should be fitted to the data. Regression will still be used to analyze the data, but the interaction between variables becomes difficult to assess. Alternatively, a better fit to the data may be obtained by using the Thornton and Lessem (1978) algorithm, which is used more commonly to fit parameters for bioenergetics models. It should be noted that although standard types of curves have been established for common relationships such as the allometric (power function) relationship for weight-dependent responses, these standard curves may not always fit new data for previously unstudied species, and alternative polynomial models should be explored to determine best fit.

Perhaps the best way to describe the analysis of energetics data is to present an example based on a data set of C_{max} for striped bass (from Hartman 1993; see Box 12.2). Although these are C_{max} data, the procedures for analyzing metabolism or other components of the energy budget would proceed in the same manner. The data are arranged in the columns weight, temperature, and C_{max}. Before analyzing the data we should verify the relationship between C_{max} and weight and determine whether the C_{max} versus temperature relationship is linear (or can be linearized with transformation) or curvilinear. The data are plotted on each of two graphs: each has C_{max} as the dependent (y-axis) variable with either temperature or weight as the independent (x-axis) variable. On the weight-dependence graph (step 1; Box 12.2, Figure A), different lines are used for each temperature treatment. On the temperature-dependence graph, all the data are plotted, but different symbols are used for each size range of fish (Box 12.2, Figure B). From these plots, it is apparent that the size dependence of consumption varies with temperature (Box 12.2, Figure A). The slope of the relationship is steepest at the highest temperatures and declines to virtually no slope at the lower temperatures. This suggests a possible interaction between temperature and fish size, meaning that a single parameter for size dependence of C_{max} at all temperatures may not be appropriate (step 2). In this case, different size-dependent model parameters may need to be determined for different temperature ranges. (This presently cannot be done with the Fish Bioenergetics Model 3.0 software [Hanson et al. 1997].).

Box 12.2 Sample Analysis of Striped Bass Data

Presented here is a step-by-step account of an approach to analyzing common energetics data such as C_{max} or metabolism. In the example provided here, data are from maximum consumption experiments for striped bass conducted by Hartman (1993). The data set (available in the Chapter 12 compact disk [CD] folder) is arranged by the variables wet weight, temperature, and ration.

Step 1 Graph data.
Data are graphed to determine the shape of any relationships between C_{max} and temperature or weight.

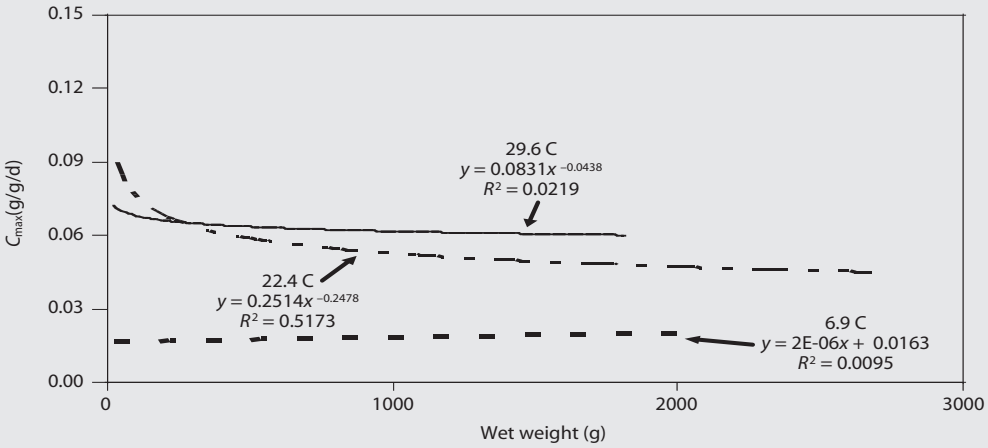


Figure A Plot of C_{max} versus wet weight for different temperature groups from the striped bass C_{max} data set. For simplicity, only data from four temperature ranges are shown. Curves appear similar for weight effects on C_{max}, but the relationship becomes insignificant at temperature extremes, especially at 6.9°C. This suggests that at thermal extremes the size dependence is not significant and may lead to a significant interaction term in subsequent analyses of variance (ANOVAs).

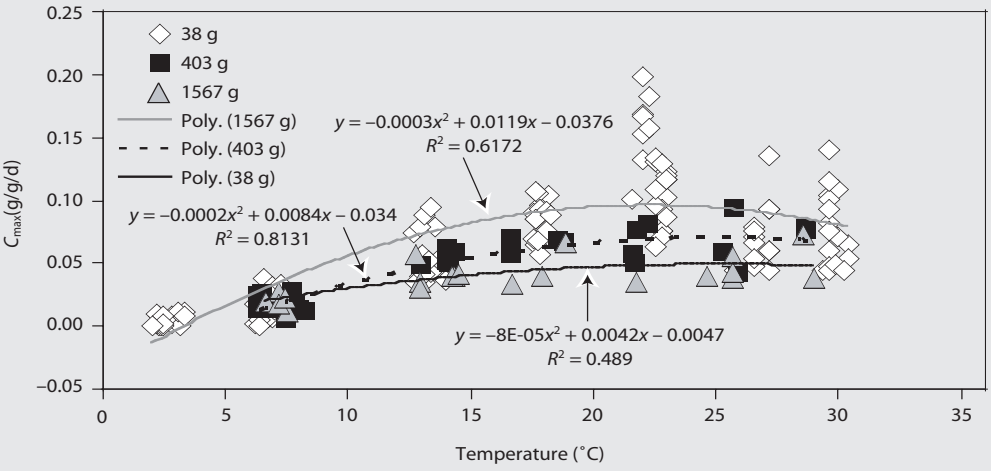


Figure B Plot of C_{max} versus temperature for small (<77g, mean 38 g), medium (131–725 g, mean 403 g), and large (>809 g, mean 1,567g) striped bass. The regression results shown on the graph represent polynomial regressions (poly.) performed in Excel describing C_{max} for each size-group as a function of temperature and temperature squared. Notice that even given the wide range of sizes used in the groups, these equations describe 49–81% of the variability in the data.

It is apparent from the graphs above that there is a significant effect of temperature and weight upon C_{max}. However, the effect of temperature is not linear and will require a polynomial regression to fit (Figure B). Weight effects on C_{max} can be linearized with $\log_{10}$ transformation, but the significance of weight effects appears to decline as temperature declines (Figure A).

Step 2 Test for significant interaction.
Because there appears to be no weight effect on C_{max} at the coldest temperature, we need to verify that interaction between the two independent variables, weight and temperature, does not exist. To do this we run an ANOVA in SAS (SAS Institute 2004) on consumption (cons) data with weight (WW), temperature (tem), and the interaction term as model variables. The SAS program to do this is given below. The data set and SAS code are contained in the Chapter 12 CD folder.

```
Data One;
input ww tem cons;

data two;
set one;
if cons=0 then delete;
Lww=LOG10(ww);
Lcons=LOG10(cons);

proc glm;
model Lcons = Lww tem lww*tem;
run;
```

This SAS program $\log_{10}$ transforms the data after removing consumption values of 0, which cannot be $\log_{10}$ transformed. The program produces the following output.

Table Results of general linear model (GLM) procedure of SAS with $\log_{10}$ transformed values of consumption (Lcons) as dependent variable. Abbreviations are sum of squares (SS), coefficient of variation (CV), mean square error (MSE), $\log_{10}$ wet weight (Lww), and temperature (tem).

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	16.21097039	5.40365680	69.43	<0.0001
Error	160	12.45300437	0.07783128		
Corrected total	163	28.66397476			
R^2	0.565552	Root MSE	0.278983		
CV	−20.59431	Lcons Mean	−1.354659		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Lww	1	1.46877402	1.46877402	18.87	<0.0001
tem	1	14.51104974	14.51104974	186.44	<0.0001
Lww*tem	1	0.23114662	0.23114662	2.97	0.0868

(Box continues)

Box 12.2 *(continued)*

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Lww	1	0.06050657	0.06050657	0.78	0.3793
tem	1	2.41835685	2.41835685	31.07	<0.0001
Lww*tem	1	0.23114662	0.23114662	2.97	0.0868

Examination of the type III SS indicates that the interaction term is not significant ($P = 0.09$). Had this interaction term been significant we would be forced to break the data into smaller subsets to describe the temperature and size effects such that their interaction would be insignificant. For example, we would develop statistical models describing temperature effect for different size ranges of fish or would use different size-dependent model values for different ranges of temperature (e.g., <6.9 and >6.9 in Figure A). These methods are important in attempting to model C_{max} accurately for use in bioenergetics models.

Here the interaction term was not significant so we may continue with the process of developing a single statistical model to describe the relationship between weight, temperature, and C_{max}.

Step 3 Develop a statistical model.
In our example the C_{max} of striped bass appears to increase with temperature to a point, and then it declines slightly with increasing temperatures. A good start for such curvilinear data is to use a quadratic model that includes wet weight (WW), temperature (T), and a quadratic form such as temperature × temperature (T^2), for example:

$$\log_{10}[C_{max}\,(\text{g/g/d})] = a + B_1 \cdot \log_{10}\text{WW} + B_2T + B_3T^2 + e\,,$$

where a is the intercept, e is the error term, and B_1, B_2, and B_3 are the parameter estimates for WW, T, and T^2, respectively. This model is easily run in SAS using the following code.

```
data one;
input ww tem cons;

data two;
set one;
if cons=0 then delete;
Lww=LOG10(ww);
Lcons=LOG10(cons);

proc glm;
model lcons = lww tem tem*tem;
run;
```

This program returns output on a polynomial equation describing C_{max} as a function of WW, T, and T^2. Line 5 removes lines of data where $C_{max} = 0$ so that $\log_{10}$ transformation of the weight and C_{max} data can occur (lines 6 and 7). Running the SAS program yields the following output.

Table Results of GLM procedure based on polynomial equation describing C_{max} as a function of WW, T, and T^2.

GLM Procedure

Source	*df*	SS	Mean square	*F*-value	*P* > *F*
Model	3	23.01195018	7.67065006	217.14	<0.0001
Error	160	5.65202458	0.03532515		
Corrected total	163	28.66397476			
R^2	0.802818	Root MSE	0.187950		
CV	–13.87433	Lcons Mean	–1.354659		

Source	*df*	Type I SS	Mean square	*F*-value	*P* > *F*
Lww	1	1.46877402	1.46877402	41.58	<0.0001
tem	1	14.51104974	14.51104974	410.79	<0.0001
tem*tem	1	7.03212641	7.03212641	199.07	<0.0001

Source	*df*	Type III SS	Mean square	*F*-value	*P* > *F*
Lww	1	0.87579819	0.87579819	24.79	<0.0001
tem	1	11.76947748	11.76947748	333.18	<0.0001
tem*tem	1	7.03212641	7.03212641	199.07	<0.0001

Parameter Estimates

Parameter	Estimate	SE	*t*-value	*P* > |*t*|
Intercept	–2.412543486	0.07306064	–33.02	<0.0001
Lww	–0.119482562	0.02399631	–4.98	<0.0001
tem	0.145070569	0.00794772	18.25	<0.0001
tem*tem	–0.003296181	0.00023362	–14.11	<0.0001

All three parameters in the model yield significant results. The derived equation is

$$\text{Lcons} = -2.412 - 0.1195\text{Lww} + 0.1451T - 0.0033T^2,$$

which transforms to

$$C_{max}\ (\text{g/g/d}) = 10^{(-2.412 + 0.1455T - 0.0033T^2)} \cdot \text{WW}^{-0.1195}.$$

The model explains 80% of the variability in the data. A final step is to evaluate how well this statistical model fits the data upon which it was derived.

Step 4 Evaluate model fit.
Although the model explains a significant portion of the variability in the data we still must evaluate how well it fits all the data before deciding to accept this model. It is still possible that the

(Box continues)

Box 12.2 *(continued)*

model fits poorly in places. To evaluate this model we plot the residuals of the model versus temperature using different symbols for fish within each of the three size-groups described in Figure B of this box: for small (<77g, mean 38 g), medium (131–725 g, mean 403 g), and large (>809 g, mean 1,567g) striped bass. This plot will identify if the model fits poorly at different temperatures, sizes, or combinations thereof. In Figure C there does appear to be some "funneling" of residuals, particularly for the small fish as temperature increases. However, examination of the CVs of residuals for the same data does not show this pattern, suggesting a good model fit (Figure D). However, even when the CV varies with the independent variable, resulting submodels may still be used in bioenergetics models provided the resulting model is calibrated and tested.

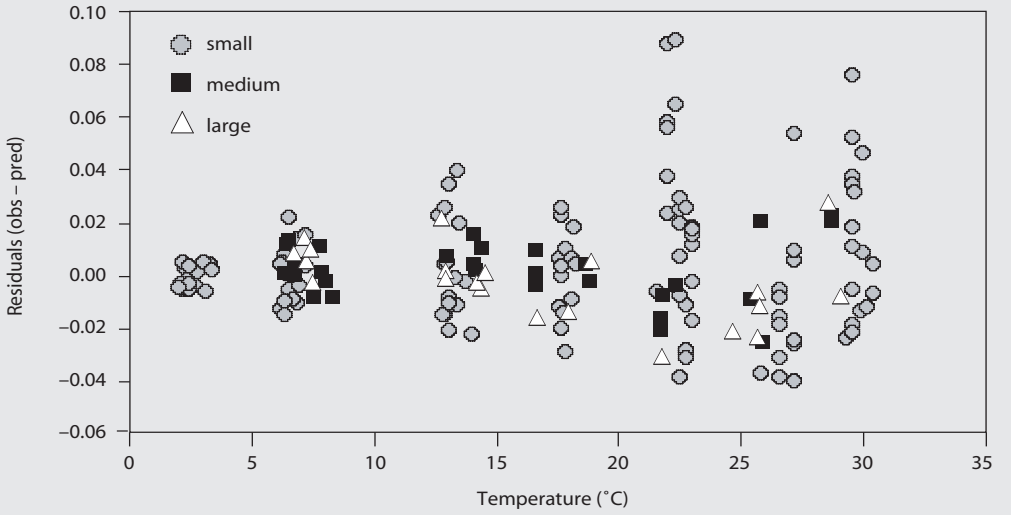


Figure C Plot of the residuals (observed – predicted) of the polynomial model describing C_{max} as functions of WW, T, and T^2 shows the model seems to do a good job of fitting the observed data. The model fits less well for medium and large fish at 22–31°C, but the residuals appear relatively normally distributed, suggesting a fairly good model fit.

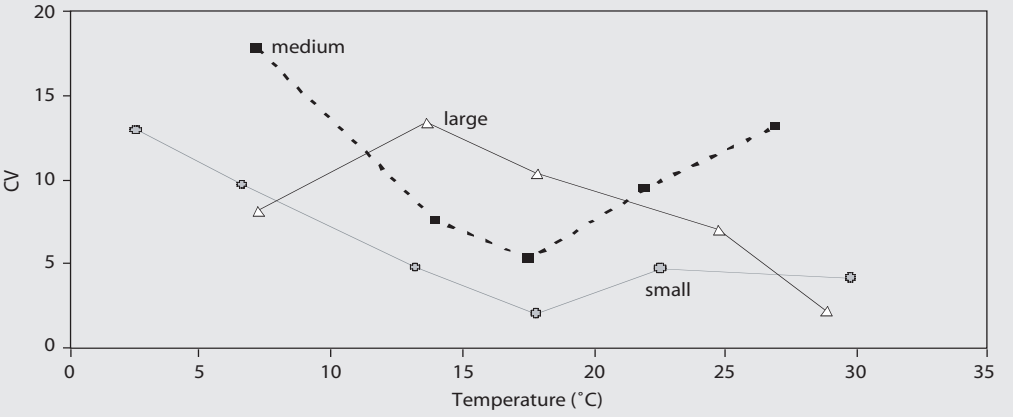


Figure D Plots of CV versus temperature for small, medium, and large striped bass show no clear pattern in variation relative to the mean, suggesting the funneling of residuals in Figure C is due to the variation of the mean with temperature.

Some success has been had fitting metabolic responses over different temperatures and fish sizes by means of multiple regression analysis with a model form of

$$\log_{10} M = a + bT + c(\log_{10} \mathrm{W\,W}), \tag{12.3}$$

where $\log_e M$ is the natural log of metabolic rate, a, b, and c are empirically derived constants, T is water temperature, and $\log_{12}$WW is the natural log of fish wet weight. This model accommodates a power function relationship with body weight and an exponential relationship with temperature. Regardless of the model form, plots of the temperature versus energetics data will reveal the nature of temperature effects.

In our example, the temperature–C_{max} relationship is curvilinear (Box 12.2, Figure B). In past bioenergetics model fitting, two different nonstatistical methods were used to fit a curve to the data subjectively (Kitchell et al. 1977; Thornton and Lessem 1978; Hanson et al. 1997). However, if a statistical model is desired then a quadratic equation in a multiple regression model is required (step 3). The rule here is to keep the model as simple as possible. A reasonable starting point for curvilinear data is a model that includes temperature in the linear (T) and quadratic forms (T^2):

$$\log_{10} C_{max}\ (\mathrm{g/g/d}) = a + b(\log_{10}\mathrm{WW}) + cT + dT^2, \tag{12.4}$$

where a, b, c, and d are empirically derived constants and WW is wet weight. We may consider adding interaction terms between weight and temperature (e.g., INT1 = [$\log_{10}$WW] · T and INT2 = [$\log_{10}$WW] · T^2) and using stepwise regression to evaluate whether there is a significant interaction that should be accounted for in the resulting model.

Regardless of whether interaction is included, to see if the quadratic form captures the essence of the temperature relationship with C_{max}, we regress the resulting model for a single discrete size range of fish (without $\log_{10}$WW) and inspect plotted residuals (step 4, Box 12.2, Figure C). Plotting the residuals for each size- or age-class of fish will suggest whether different models and parameters are required for each size- or age-class. Regardless of the relationship or model form decided upon, all models should be evaluated by plotting the residuals versus the independent variables to verify goodness of fit. Figure C in Box 12.2 shows the plot of residuals for the model in equation (12.4), describes potential shortcomings, and discusses funneling of residuals (heteroscedasticity), which is common in energetics data. Figure D in Box 12.2 shows plots of the coefficient of variation of residuals with temperature for the three size-classes of striped bass; these curves suggest heteroscedasticity does not prevent us from a good model fit. The complete process of analysis of the striped bass C_{max} data is presented in Box 12.2.

In the event that the relationship between the dependent variable and temperature is not curvilinear, analysis of covariance (ANCOVA) can be used to evaluate the interaction between the fish size and temperature variables. Multiple regression is conducted to examine whether the interaction term ([$\log_{10}$WW] · T) is significant. If the type-III sum of squares is significant ($P < 0.05$), the interaction

is significant and should be evaluated in the model. Evaluation can consist of the plotting methods described above or regression analysis of weight effects on the dependent variable at each temperature level (see Box 12.2 for SAS code, output, and analysis). Again, if the interaction term contributes significantly to the model, one may need to develop model parameters for different size-groups with parameters changing as fish grow into new size ranges.

12.4 FISH BIOENERGETICS MODELS

12.4.1 Model Platforms and Model Execution

A number of different platforms have been developed for running fish bioenergetics models. The most common and popular is the Fish Bioenergetics 3.0 (Hanson et al. 1997; available from University of Wisconsin Sea Grant Institute, Communications Office, 1800 University Avenue, Madison, WI, 53705–4094, USA, for a nominal fee of US$75 in 2005), which was preceded by versions 1.0 and 2.0 (Hewett and Johnson 1987, 1992). This model is commonly referred to as the Wisconsin model due to its development, production, and support at University of Wisconsin and Wisconsin Sea Grant Institute. The platform runs on all Windows-based operating systems from Windows 3.1 through Windows XP. It has the advantage of a menu-driven operation, and parameterized models for 40 species and life stages are currently included in the model.

Although the Fish Bioenergetics 3.0 model is widely popular, it is not without limitations. Many experiments associated with fish bioenergetics have sought to derive parameters for fish bioenergetics model software as these models are "data hungry," with typical models involving 15–30 parameters (Ney 1990, 1993). One limitation of Fish Bioenergetics 3.0 software (Hanson et al. 1997) is that it is limited to the model functions and styles that were in vogue at the time of publication. There are presently no options in the Fish Bioenergetics 3.0 software for constructing models of a different form than that included in the model settings. The ease of applying the Fish Bioenergetics 3.0 software has led some users to compromise the use of alternate models that may better fit experimental energetics data in favor of model forms included with Fish Bioenergetics 3.0 software. An example of this problem would be data that have a significant interaction term between size and temperature effects which is best described by a polynomial equation. Although we can accurately model this relationship, it cannot be incorporated into the Fish Bioenergetics 3.0 software, which is constrained by a model structure that does not include any interaction terms or other polynomial forms. In such situations, the modeler can use alternate model platforms.

In fitting models to data for construction of bioenergetics models, trade-offs emerge among the perceived superiority of using statistical models such as regression, the need to enhance model performance by fitting functions over a range of variables for which no interaction of terms occurs, and the limitations of the popular software Fish Bioenergetics 3.0 in which most bioenergetics applications are compiled. In cases in which energetics data can be statistically fit through regression,

this is the preferred method of developing a model. However, despite high correlation, models may not fit particularly well at all temperatures (e.g., failure of size dependence of C_{max} at lower temperatures). In such cases, separate sets of model parameters may need to be established for particular size- or age-classes of fish or temperature ranges.

If the data conform to the models provided in the Fish Bioenergetics 3.0 model, then it may be adequate to use that application for modeling. The Fish Bioenergetics 3.0 model software includes three basic functions for C_{max} and metabolism and two functions for egestion and excretion (Hanson et al. 1997). Metabolism and C_{max} functions include an exponential model (type 1) and two different forms of a convex model (types 2 and 3) that fit one curve to the increasing part of the relationship and another to the declining limb of the relationship between C_{max} or metabolism and temperature (see Figure 12.2, lower panel).

As described in section 12.3.3, after a component submodel is constructed, it should be evaluated to see how well it fits the original data before incorporating it in a bioenergetics model. Evaluation of model fit should be done with individual submodel components such as C_{max} and metabolism. These submodels can be evaluated by calculating the predicted value from the submodel for each data point in the original dataset, similar to the example in Box 12.2, Figure C (step 4).

Another potential issue with the Wisconsin model is the way in which it constrains consumption, a constraint that can result in unrealistic patterns of growth in the fish. The Wisconsin model estimates consumption as a proportion *(P)* of maximum ration (C_{max}) varying between 0 and 1. Consumption over a time interval is estimated from growth curves generated from observed initial and final weights by solving for P as the independent variable. Different values of P are run iteratively in the model with a daily time step until the resultant growth trajectory matches the final weight. The resulting value of P is inserted into the model to estimate consumption over the interval. The estimated consumption is the sum of all daily rations as estimated by the best fit of P. The model's use of P to estimate consumption is appealing in that it forces the model to agree with the initial and final weights of the fish. However, depending upon the length of the growth interval, the actual growth trajectory between initial and final weights can be quite different than observed or expected. The result is over- or underestimation of final weights within the interval and then converging with the final weight by the end of the period. These patterns may occur due to the length of the time interval, even when actual fish growth is nearly linear. In these circumstances, the growth period should likely be broken into shorter periods for better model results.

Other model platforms for fish bioenergetics models include SAS (SAS Institute 2004), or a simulation modeling package such as STELLA (available at http://www/iseesystems.com/softwares/Education/StellaSoftware.aspx), STELLA, MATLAB (available from MathWorks, Inc., Natick, Massachusetts), and spreadsheets such as Microsoft Excel. The disadvantage of using these other platforms is that the user must construct the models and write the code. However, this does allow greater flexibility in being able to model interactions between temperature and size or to code other model functions, such as quadratics, that are not currently

possible in the Wisconsin model platform. The model can be easily coded into spreadsheets. An example in Excel is provided in the Chapter 12 compact disk [CD] folder.

Alternative approaches do exist for the indirect estimation of fish growth and consumption and, to some extent, prediction of fish responses to changing energetic conditions. In general, these alternative models require fewer input parameters than do bioenergetics models. While this may be beneficial in terms of number of potential error sources, there is also elevated risk in that high sensitivity is attached to the few input variables (Ney 1993). Ney (1993) described a simplified, bioenergetics model-like approach and provided assessments of (1) a simplified predictor based on the equation of Winberg (1956) as well as (2) models that predict whole population or cohort consumption from estimates of production combined either with conversion efficiency (Eck and Wells 1983) or biomass (Ney 1990). A distinct approach, the nitrogen balance method, estimates food consumption of fish by determining their nitrogen loss over a period of postcapture confinement (Davis and Warren 1971). Nitrogen losses are added to nitrogen gains through growth to estimate food mass consumed. In a number of cases, laboratory-derived relationships between food consumption and growth rates have been used to estimate consumption in the field from growth estimates (Davis and Warren 1971; Allen and Wootton 1982; Brafield 1985). However, it has been observed from laboratory data that such relationships may perform substantially better when estimating fish growth from observed consumption than when estimating consumption from observed growth (R. Hayward, unpublished data). Estimates of food consumption have also been made by measuring changes in fish's body concentrations of radioisotopes over time periods (Davis and Foster 1958; Gingras and Boisclair 2000); increases in body concentrations of radioisotopes over relatively short periods (e.g., a month) are related to prey consumption. Artificial neural networks (ANNs), a relatively new and rapidly proliferating modeling approach, have shown potential to predict fish daily consumption patterns in aquaculture settings (Ruohonen 1999). The ANN approach may provide better predictions of day-to-day consumption than is currently possible with bioenergetics models and other fish consumption predictors. These alternative approaches for estimating consumption and growth have, in general, received less attention than the currently popular bioenergetics models, with rigorous evaluations being limited or nonexistent in some cases. Further consideration of these less frequently used approaches would be of value because it is likely that some will hold advantages over currently popular bioenergetics models in certain situations.

12.4.1.1 Calibration of Bioenergetics Models

The balanced nature of the energy equation permits users to calibrate bioenergetics models to fit observed data better. This feature can be used to force models to provide the correct values of C or G. Calibration experiments can be conducted in which C and G, temperatures experienced, and energy content of fish are estimated for the duration of the experiment. When these data are used in

the bioenergetics model, discrepancies between observed and predicted *G* or *C* can be corrected by altering parameters such as ACT, *F, U,* or SDA. This results in calibrated models. However, such forcing may result in errors in the apportionment of energy between cost and loss terms.

One way in which models have been fit to arrive at correct values of *C* and *G* has been with respect to estimating the activity (ACT) multiplier needed to balance the model when both *C* and *G* have been measured on captive fish. In the spreadsheet example included in the Chapter 12 CD folder; an ACT estimator allows users to estimate the ACT multiplier needed to balance growth and consumption with known values. A similar approach can be used to fit other parameters when the remaining components of the energy budget have been measured.

Attempts to calibrate bioenergetics models from field-derived data are not recommended because of the potential for substantial error in the estimation of key input and output parameter values (e.g., *C, G,* temperature, and caloric densities), which would lead to inappropriate calibration. A new form of model correction determines and corrects for systematic error in bioenergetics models based on intensive laboratory evaluations (see Bajer et al. 2003).

12.4.1.2 Site- and Species-Specific Input Data

At a minimum, bioenergetics models require two basic types of information: species-specific energetics information from which to model energy dynamics, and site-specific information. Energetics information has already been gathered and developed into bioenergetics models for many popular species (Hanson et al. 1997). For species lacking a bioenergetics model, information can often be gathered from the literature or obtained in experiments if needed. Site-specific information includes data that can typically be gathered as part of a fisheries investigation, such as age and growth, diet composition, energy content of predators and prey, and thermal history of the fish to be modeled. Growth inputs to the model are generally conducted by age-class such that we are essentially modeling the growth and consumption dynamics of the average individual of that age-class. In cases in which multiple cohorts within an age-class exist (Hartman and Brandt 1995b) or in which growth differs by sex, these differences can be accounted for by running distinct groups as individual cohorts.

Recommendations on energy content measures and thermal history measures are limited. Despite the fact that these parameters may have large effects on bioenergetics output, most studies assume constant values for the energy content of predators and prey (Ney 1993), which are known to vary seasonally as well as with size and sex (Adams et al. 1982). For bioenergetics studies that require precision, energy density of predators and prey should be estimated directly. This can be done rather easily by drying samples of fish and using the strong generalized relationships between the dry weight percentage of wet weight (DW%) and energy content (J/g wet weight) reported by Hartman and Brandt (1995c) and others (see Table 12.3). Alternatively, one can apply bioelectrical impedance analysis models that permit nonlethal estimation of body composition and hence energy content in fishes (Cox and Hartman 2005).

Table 12.3 Equations describing energy content (J/g wet weight) as a function of the dry weight (DW) percentage of wet weight for many common taxonomic groups used in bioenergetics modeling. All equations are highly significant ($P < 0.001$) and have r^2 values greater than 0.80. Models are of the form J/g wet weight = $a + b \cdot$ DW. Equations are reproduced from Hartman and Brandt 1995c, Table 1.

	Model coefficient	
Model taxon	*a*	*b*
Clupeiformes	–2,532	328.6
Cypriniformes	–1,265	262.2
Perciformes	–1,875	309.5
Pleuronectiformes	–1,832	286.1
Salmoniformes	–3,386	379.0
Clupeidae	–2,532	328.6
Cottidae	–1,498	306.0
Cyprinidae	–981	251.1
Percichthyidae	–2,533	349.1
Salmonidae	–3,632	386.7
Sciaenidae	–1,936	309.9
Alewife	–2,086	323.7
Atlantic menhaden	–2,695	309.0
Bloater	–2,424	336.2
Bluefish	–3,792	372.4
Coho salmon	–3,207	367.8
Lake trout	–3,809	397.9
Muskellunge	–1,939	294.5
Rainbow smelt	–1,094	303.2
Rainbow trout	–2,735	357.5
Striped bass	–1,460	313.9
Weakfish	–1,997	319.4
Yellow perch	–2,873	313.1
Zander	–2,011	309.4
Combined (all species)	–3,419	375.0

Accurate knowledge of fish's thermal history is important to bioenergetics model estimates because temperature effects on energetics parameters are typically stronger than are other influences such as body size. In stratified or thermally heterogeneous environments, fish are usually assumed to reside in the most optimal temperatures available. However, a recent study with white crappie has demonstrated that this assumption can lead to erroneous determinations of fish's thermal histories. (P. Bajer, University of Minnesota, unpublished data). In heterogeneous environments, sound approximations of thermal history may be obtained through technological advances such as radio tags with thermal sensors, which provide descriptions of the temperatures actually occupied by fishes in the wild.

12.4.1.3 Estimation of Consumption and Growth

Bioenergetics models can be used to estimate any unknown parameter in the balanced energy equation provided we have estimates or measures of the other

components. It is this principle that allows us to estimate consumption or growth with a bioenergetics model. If we have estimated growth, then we can input that into the model and estimate how much the average fish had to eat in order to grow in the observed manner given thermal, size, energy content, and other constraints. Similarly, we can estimate growth with an estimate of consumption, although this is less often done in bioenergetics models due to the difficult nature of estimating ration in the field. It was, in part, the difficulties and expense of estimating ration in the field that led to the development of the bioenergetics models (Kitchell et al. 1974) and other low-effort consumption estimation methods (Hayward and Hiebert 1993).

12.4.2 Model Evaluations

Evaluations of bioenergetics models' predictive accuracies during the past 25 years have been far less common than their applications, which have increased geometrically with time. Early evaluations, most being field based, considered only small portions of the full array of conditions under which these models were applied. More recently, a substantial number of laboratory-based evaluations have been completed (largemouth bass, Whitledge and Hayward 1997; hybrid sunfish [green sunfish ♀ × bluegill ♂], Whitledge et al. 1998; lake trout, Madenjian and O'Connor 1999; yellow perch, Bajer et al. 2003; subadult and adult smallmouth bass, Whitledge et al. 2003), some covering broad ranges of ration level, fish size, and temperature. Results of laboratory evaluations have been mixed, showing good model performance under some sets of conditions but much poorer performance under others.

Efforts to improve predictive accuracy of bioenergetics models have been particularly rare because evaluations have tended not to identify sources of model error. One of the few examples of an effort to improve a bioenergetics model is that of Karas and Thoresson (1992), who incorporated effects of body weight on optimum and maximum temperatures for consumption and respiration of Eurasian perch into an existing model for yellow perch (Kitchell et al. 1977). Recent laboratory evaluations have identified an important form of systematic error that appears common to many bioenergetics models. This error is strongly linked to consumption level, becoming greater at extreme (high or low) levels of consumption (Madenjian and O'Connor 1999; Bajer et al. 2003; Bajer et al. 2004a). The connection of this widespread bioenergetics model error to a particular model variable (consumption level) suggests that energy cost parameters within bioenergetics models that are consumption level dependent are potential sources of the model error. Approaches to remove or diminish this error include the use of error-correction equations (Bajer et al. 2004b), and the redevelopment of equations that determine the magnitude of energy cost parameters within bioenergetics models, particularly those parameters that are consumption level dependent.

Strong efforts to improve predictive accuracy of bioenergetics models are warranted. Although these models may have been originally intended to serve only as inductive machines for elucidating causes of observed trophic dynamics, in fact

they now are being increasingly applied in settings where inaccurate prediction may have serious consequences for aquatic populations and communities and humans. Moreover, models that are substantially inaccurate may be ineffective as inductive machines! Recent findings (Bajer et al. 2004a, 2004b) suggest a way to improve bioenergetics models' predictive accuracies. Equally critical to improving these models may be a philosophical adjustment whereby model users, as much as model builders, bear responsibility for model performance when they apply them to research and management endeavors. If a bioenergetics model has not been well evaluated under the conditions to which it will be applied, users should consider conducting an evaluation and possibly improving the model if warranted. If users were more often required to substantiate the accuracy of a bioenergetics model that they have used, not only would these models be used more judiciously, but their broad evaluation and improvement would likely progress more expeditiously than has previously occurred.

12.4.2.1 Two Types of Model Evaluation

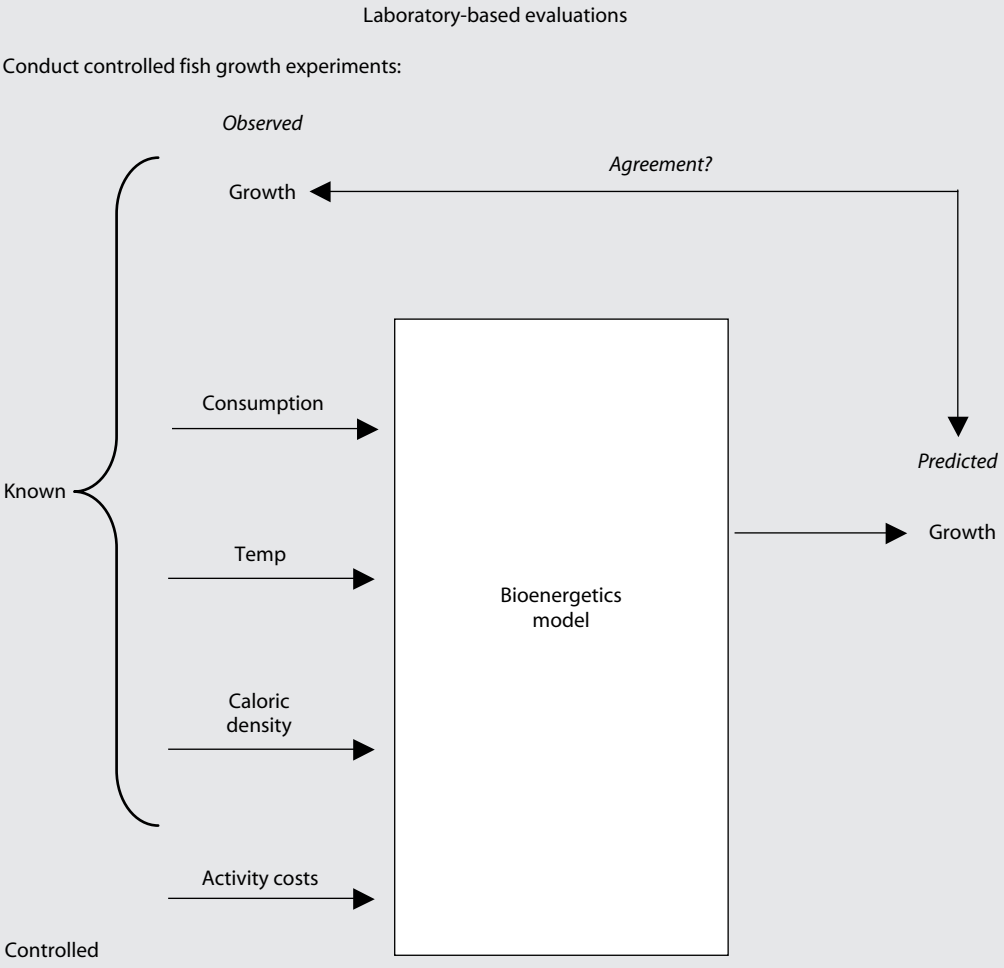
Laboratory- and field-based approaches have been used to evaluate bioenergetics models' abilities to predict consumption and growth rates of fishes accurately. In laboratory evaluations, simultaneous values of fish growth rate, food consumption, thermal experience, and caloric density of both the consumer and its food can be directly determined and essentially known over a time period. Activity cost, a notoriously hard-to-estimate model input variable in field settings, can usually be controlled or closely estimated in the laboratory. Consequently, laboratory evaluations can provide highly accurate test data sets against which bioenergetics model predictions can be rigorously tested (Box 12.3)

When laboratory evaluations are conducted over ranges of the variables that influence bioenergetic responses (including ration level, temperature, and fish weight), inaccuracies in these models can be identified. For example, if a controlled laboratory evaluation shows that an otherwise well-performing model predicts growth and consumption poorly for larger fish, a problem with the allometric (body mass-dependent) portion of the standard metabolism equation would be suggested. This error source is suggested because standard metabolism is often the only energy cost–loss parameter within bioenergetics models that has body-weight dependence. Examination of the original data used to formulate the body mass influence on standard metabolism might reveal that large fish were not well considered, suggesting that the poor growth prediction may arise from faulty extrapolation of metabolic rates for larger fish. Improvement of this relationship by conducting additional standard metabolism experiments that include larger fish would be warranted.

Sources of model error have been investigated on the basis of sensitivity analyses that indicate which model parameters and input variables most influence model predictions (e.g., Kitchell et al. 1977; Bartell et al. 1986; Adams and Breck 1990). Low-sensitivity parameters and variables receive little attention. However, it should be considered that high-sensitivity model parameters are not the only potential

Box 12.3 Laboratory- versus Field-Based Evaluations of Bioenergetics Models

For laboratory evaluations of bioenergetics models (BEMs), corresponding values of fish consumption, temperature experience (Temp), growth, and caloric density (for predator and prey) are accurately determined in laboratory experiments. Activity cost can be controlled to be negligible. With such data, the BEM can be run with known, corresponding input and output variable values. In this way, a BEM's predictive performance can be accurately determined by comparing model predictions of growth or consumption with known values. In contrast, field evaluations of BEMs rely on field-predicted input and output variables, any of which may be inaccurately estimated. Consequently, if model predictions of consumption or growth do not agree with field estimates, it is not known whether the lack of agreement is due to BEM error or inaccurate field estimation of input or output variables.



sources of model error; moderate- and even low-sensitivity parameters can cause substantial model error if their values are sufficiently inaccurate. Another approach for identifying sources of bioenergetics model error examines whether model predictive error is associated with key input–output variables including consumption level, fish body weight, and temperature. Each of these variables has known linkages to internal model parameters. Bajer et al. (2003) found that prediction error in two bioenergetics models was strongly correlated with consumption level. Because consumption level is involved in the computation of only three internal parameters in these models (egestion, excretion, and SDA), it was suggested that the observed systematic model error likely owed to inaccurate internal representation of at least one of these parameters. These findings suggest that the existing equations for egestion, excretion, and SDA should be reevaluated through laboratory studies. This finding, in turn, is interesting because these three internal model parameters have been considered to be of low sensitivity, which has promoted "borrowing" (sensu Ney 1993) of nearly identical equation forms for calculating egestion, excretion, and SDA values across bioenergetics models for many fish species.

A new approach for reducing systematic error in bioenergetics models may not require data collection beyond that used initially to evaluate a model. Bajer et al. (2004b) evaluated a white crappie bioenergetics model (Zweifel 2000). Laboratory data sets were developed for which concurrent values of white crappie daily consumption, thermal experience, growth, and other model input variables were essentially known for individual fish. Collectively, substantial ranges of fish weight, temperature, and consumption level were covered by the laboratory data sets for model testing. Plots of model prediction error versus fish weight, consumption level, and temperature indicated relationships between magnitude of model error and the levels of certain input variables. These relationships were confirmed by simple regression analyses. Multiple regression equations that predicted bioenergetics model error according to levels of fish weight, consumption, and temperature were then incorporated into the white crappie model to correct model predictions on a daily time-step basis. Cross validation was used to fairly evaluate improvement in the bioenergetics model's predictive accuracy from incorporating the correcting equations for both growth and consumption predictions. Through this approach, predictive accuracy of the original white crappie model was both evaluated and improved from the same data set. Because systematic error appears to be common in bioenergetics models (Hartman and Brandt 1993; Bajer et al. 2003; Bajer et al. 2004a) it is reasonable to explore associations that both confirm systematic error and identify its likely sources when bioenergetics models are evaluated.

In field evaluations of bioenergetics models, corresponding values of the previously described model inputs and outputs are estimated over time periods to provide a test data set. Accuracy of field estimates are far less certain than are values determined in the laboratory, even when good estimate precision is indicated (Ver Hoef and Cressie 1993). Consequently, when field evaluations are conducted without the benefit of prior, well-controlled laboratory evaluations, indications

concerning model predictive accuracy are truly inconclusive. Poor agreement between model predictions and field estimates of consumption and growth raises doubt as to whether the bioenergetics model or the field estimates are most in error. Good field-to-model corroboration, although taken to indicate model soundness can, in fact, occur when both model predictions and field estimates of key input and output variables are erroneous. For this reason, controlled laboratory evaluations should first be carried out to verify that a model is sound as a predictor of consumption and growth over a range of conditions. With this established, field evaluations can then provide valuable indication of how well bioenergetics model input and output variables have collectively been estimated in a field setting.

12.4.2.2 Design of Laboratory Evaluations

The strength of laboratory evaluations of bioenergetics models lies in their capacity to determine accurately concurrent values of consumption, growth, thermal experience, caloric density of the consumer and its food, and activity costs of fish over a period of time and to provide accurate data sets for evaluating bioenergetics models. Accordingly, effort should be put forth to ensure accurate determinations of these variables.

In general, water temperature should be tightly regulated and monitored daily. Daily consumption should be directly measured as the difference between food amounts provided to test animals and that which remains unconsumed after 24 h. Fish body weight should be determined frequently enough during an experiment to portray growth trajectories adequately but not so frequently as to disrupt fish behavior or to cause prolonged stress. Longer intervals between body weight measurements can be used for laboratory experiments of longer durations. For example, Whitledge et al. (2003) made weekly determinations of fish body mass when developing a 63-d laboratory data set to evaluate a new subadult–adult smallmouth bass bioenergetics model. On the other hand, Bajer et al. (2003) determined fish weights only every 2 weeks when developing 120-d laboratory data sets to evaluate two yellow perch bioenergetics models.

Caloric densities of the fish and their food will preferably be determined directly by bomb calorimetry and not by relying on published values from other studies as these can vary widely. When possible, caloric density estimates should be made throughout laboratory experiments, with consideration given to when values are likely to change (as when new batches of live food are introduced or when fish's ration levels are increased). Alternatively, the use of established relationships that predict caloric densities of the test fish (or their food) from readily determined variables such as fish condition (Neumann and Murphy 1991; Bajer et al. 2003) or water content (Diana and Salz 1990; Hartman and Brandt 1995c) can produce reasonably accurate measures. However, it is recommended that published relationships be tested for accuracy in the particular study setting.

The range of conditions (e.g., temperature, fish size, and ration level) under which a bioenergetics model is evaluated in a laboratory study is typically based on the array of conditions under which an investigator plans to apply the model

(Hartman and Brandt 1993; Madenjian and O'Connor 1999). However, an ultimate goal should be to evaluate bioenergetics models under the full array of conditions under which they may be applied. The possibility of seasonal shifts in bioenergetic responses of some fish species (Chipps et al. 2000) warrants further elucidation as this may hold important implications for accurate modeling.

Laboratory experiments intended to provide data sets for evaluating bioenergetics models must be run long enough to allow observed growth and consumption trajectories to become established and, for some experimental designs to allow adequate statistical power to be achieved. However, laboratory experiments for model evaluation are labor intensive and should not be run longer than is necessary. If short, efficient laboratory experiments can be run for each set of growth conditions, multiple sets of conditions can often be evaluated (e.g., Whitledge et al. 2003). Longer experiment durations are required when a single group of fish is used without replication to produce a data set for model testing; such designs are common in field-based studies (e.g., Rice and Cochran 1984) but are sometimes seen in laboratory assessments as well. Longer experimental durations are needed in unreplicated settings because analysis of model prediction accuracy is based on the single observed and model-predicted trajectories, and having sufficient observations along each trajectory is needed for statistical power. Appropriate experiment duration for designs involving unreplicated trajectories is difficult to address because a variety of statistical and analytical procedures are used, and we are aware of no sample size or power analyses that have been conducted for these analyses. In contrast, shorter experiment durations for each set of growth conditions (levels of consumption, fish body weight, and temperature) are possible when replication is incorporated in laboratory-based experimental designs, either by using individually held fish or multiple fish groups. Here, variation is measured among the multiple observed and predicted trajectories rather than along one observed and one predicted trajectory. In addition, duration of laboratory experiments used to evaluate bioenergetics models can influence the magnitude of determined model error rates. This is particularly true if systematic model error is present because model error often increases with trajectory length. Consequently, it is reasonable from the perspective of comparing performance among models or across conditions, that laboratory experiment durations be standardized. We suggest that experimental designs for developing laboratory test data sets involve replication and that experiment durations of 21 d be used for each distinct set of growth conditions applied. This duration would allow four weekly measurements of body weight and 21 daily measurements of consumption per sample unit.

12.4.2.3 Design of Field Evaluations

As in the laboratory, field evaluations of bioenergetics models warrant strong efforts to determine concurrent values of model input and output variables accurately over a time period. In field settings, the set of growth-influencing conditions that fish experience is less well known than those in the laboratory. Whereas most of the input and output variables for bioenergetics models can be directly

measured throughout a laboratory experiment, in field settings most must be estimated (Box 12.3). Growth conditions in field settings cannot be controlled as in the laboratory, and substantial temporal variation in these conditions can occur between sampling dates. Temporal variation in growth conditions in field settings is problematic because sampling date intervals are typically substantial (weeks to months) due to high sampling costs, and it must be assumed that values of certain variables (e.g., consumption rate, predator and prey caloric density, and temperature experience) follows a linear trajectory between sampling dates. Consequences of linear interpolation between field sampling dates to estimate trajectories of variables whose values can change markedly from day-to-day were illustrated by Whitledge and Hayward (2000). By applying Monte Carlo simulations to in situ estimates of daily consumption made over 30 successive days for two fish populations (in a stream and small impoundment), they found that short sampling date intervals not exceeding 5 d were needed to estimate consumption trajectories adequately through linear interpolation between sampling dates. Similarly, a sample date interval of less than 7 d was found necessary to estimate fish's consumption levels accurately in an estuarine environment (Hartman 2000). Sampling date intervals of less than 7 d are uncommon for in situ studies of fish daily ration. However, it is noted that a similar study (Trudel and Boisclair 1993) found less day-to-day variation in fish's daily consumption in a natural lake, indicating that broader sampling date intervals may be possible at some times in some settings.

Fish sampled on distinct dates at a given field location also may not be continually exposed to conditions at that sampling site throughout a sampling date interval. This can occur from fish's substantial movements in both the vertical and horizontal dimensions and contribute to inaccurate field evaluations of bioenergetics models. Of course, in laboratory settings fish have little choice but to show high fidelity to the "sampling site." This problem can apply to field estimates of fish's thermal experience, consumption, growth, and caloric densities of both predator and prey. In general, field estimation errors due to temporal variation in conditions as well as from fish changing locations should diminish as sampling date intervals are reduced. If fidelity to a sampling site cannot be reasonably assumed for a population and exposure to disparate environmental conditions is possible, expanded spatial coverage of the field evaluation should be considered. Because in situ daily consumption estimation is particularly labor intensive, the use of low-cost, in situ consumption estimators may be appropriate to allow increased spatial coverage (Boisclair and Leggett 1988; Hayward and Hiebert 1993; Madon 1998).

12.4.2.4 *Analysis of Model Evaluation Data*

General considerations.—Evaluations of bioenergetics models typically involve assessments of how accurately they predict fish growth or estimate consumption over a time period, using test data sets developed either in the laboratory or field. Occasionally, bioenergetics models have been evaluated for their ability to predict either consumption or growth alone, under the expectation that the other

variable will be predicted with similar accuracy in the opposite direction. For example, a model observed to overestimate growth by 15% may be assumed to underestimate consumption by about 15% under the same conditions. However, recent evaluations of bioenergetics models under broad ranges of growth conditions have shown that predictive accuracy for consumption and growth can be quite disparate (Bajer et al. 2004b). Consequently, it is recommended that evaluations of bioenergetics models consider predictive accuracy for both consumption and growth.

Data used to evaluate bioenergetics models typically include concurrently derived observed trajectories of fish body weight and daily or cumulative consumption over a period of time (Figure 12.4). These observed data have either been directly determined from laboratory experiments or estimated in a field setting. Paralleling observed data are corresponding trajectories of predicted consumption and growth from a bioenergetics model. Analysis of a bioenergetics model's predictive accuracy typically involves (1) visually portraying the likeness of corresponding predicted and observed trajectories (Figure 12.4), (2) representing the lack of agreement between predicted and observed trajectories as some form of percent error, and (3) applying one or more statistical tests or indices to facilitate decisions of whether or not apparent lack of agreement between predicted and observed trajectories is statistically different from nil or otherwise acceptable.

Predictive accuracy for fish growth.—Accuracy of a model's predictions of fish growth is usually evaluated by entering into a model a series of daily values of observed consumption and temperature experience, based on daily observations in the laboratory or linear interpolation between observed values in the field. Caloric densities of the consumer and its food, as well as an activity cost value, are also entered. Values of fish body weight on each day of a model run are required by internal equations and are provided by model estimates of daily growth, which are sequentially added to a set starting weight. In this way, a model-predicted growth trajectory is produced to be compared to an observed growth trajectory (Figure 12.4A). Differences between these two trajectories provide the basis for visualizing, quantifying, and statistically testing the model's ability to predict growth under a given set of conditions. In cases in which replicate observed and predicted growth trajectories are produced, as when individual fish serve as the sample unit, growth trajectories can be represented as either absolute (g/d) or relative (g/g/d) growth rates upon which statistical analyses are conducted (e.g., Whitledge et al. 2003). Representing predicted and observed growth trajectories in terms of relative versus absolute growth diminishes the problem of variance inflation among individual fish weights with time (that is, growth depensation), which tends to reduce statistical power.

Predictive accuracy for consumption.—Evaluating a bioenergetics model's ability to estimate consumption over a period typically involves entering fixed daily consumption levels into the model in iterative fashion until finding the fixed level that produces a growth trajectory that matches the initial and final observed fish body weights in an observed growth trajectory. For the Fish Bioenergetics Model 3.0 software described by Hanson et al. (1997), this iterative process involves inputs of a constant percentage of maximum daily consumption and is termed a "*p*-fit" (see

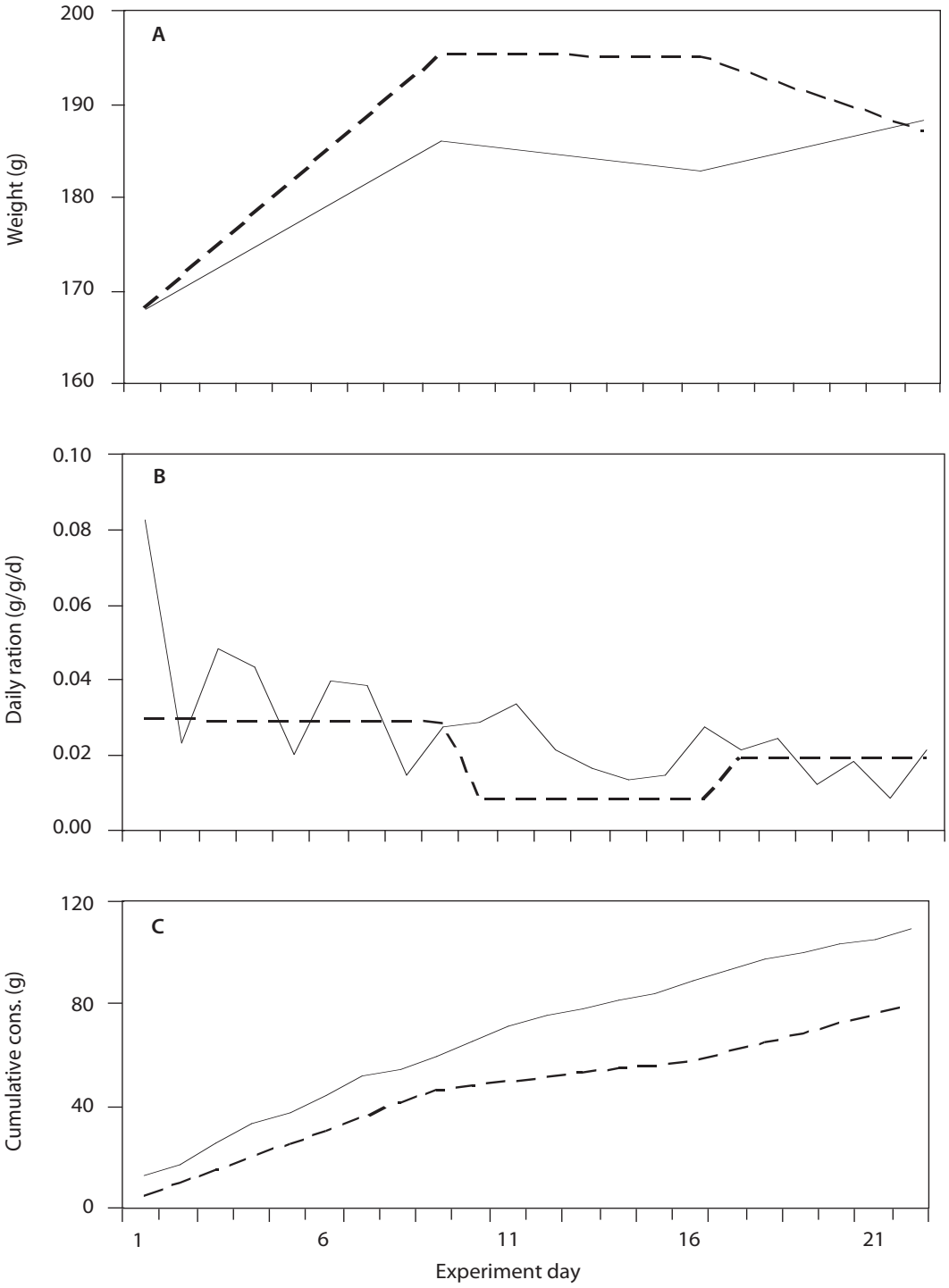


Figure 12.4 Corresponding trajectories of (A) growth, (B) daily ration, and (C) cumulative consumption (cons.) observed from a laboratory experiment (solid lines) and predicted from a bioenergetics model (broken lines) for largemouth bass. Data are from Whitledge and Hayward (1997).

section 12.4.1). Predicted and observed trajectories of daily consumption (Figure 12.4B) or cumulative consumption over the period (Figure 12.4C) are considered in evaluations of model predictive accuracy. When using the *p*-fit approach, by which fish consumption rates are estimated over multi-day periods in a single step, it is unlikely that the predicted trajectory of daily consumption will fully match short-term fluctuations in the observed daily consumption trajectory, which are more pronounced at higher consumption levels (Figure 12.4B). Because this approach promotes error in model predictions of daily consumption, the use of cumulative consumption trajectories instead (Figure 12.4C), which dampen these fluctuations, is preferred.

Ambiguity in representing model error.—Percent error has been used commonly to represent predictive accuracy of bioenergetics models (Table 12.4). This measure has been used to compare a single model's performance across different sets of conditions and to compare performance among models. To an extent, percent error has also served as a barometer of acceptable model performance (see performance comparison of six bioenergetics models by Ney 1993). Consequently, it seems important that percent error for model predictions would be determined in a manner that is both accurate and consistent across model evaluations.

However, it is often unclear in studies what form of predicted and observed values were used in evaluating models. Often the predicted and observed trajectories of weight, cumulative consumption, or daily consumption have been final-day values. Although the use of final-day values may be appealing due to simplicity, this expression of percent error is often not representative of model performance over the complete trajectories. For example, consider the predicted and observed growth trajectories in Figure 12.4A, which are substantially separated for the most part but converge near the final day. Determination of percent error for model predictions of growth based on final-day values would clearly be misleading, as a near-zero value would not reflect overall predictive error.

Recommended portrayal of model error.—A more appropriate representation of model error when predicting growth trajectories is mean absolute daily growth error or, more succinctly, the mean growth error (MGE) between model-predicted and observed trajectories (Table 12.4). This measure of error is representative of the full predicted and observed growth trajectories and approximates the area between the two trajectories. By standardizing for the number of days in a model-evaluation trajectory, it is less influenced by the length of model evaluation trajectories, unlike calculations of percent error based solely on final-day error. Absolute values of daily differences between predicted and observed values are used because predicted trajectories can cross over observed trajectories leading to positive and negative errors (e.g., Bajer et al. 2003). Summing of positive and negative daily errors would underrepresent total error, whereas use of absolute values avoids this. One drawback associated with the use of MGE is that direction of the model error (positive versus negative) is not indicated. For this reason it is recommended that the predominant direction of the error be stated or that graphic portrayals of predicted versus observed trajectories be shown. In addition to MGE, it is recommended that the maximum absolute daily growth error (MaxGE) also be reported

Table 12.4 There are several commonly used metrics for evaluation of bioenergetics models with test (growth and consumption trajectory) data sets. Below we list these metrics and provide instances in which they may be properly used to report model assessment. In reporting new model evaluations we recommend use of multiple metrics to identify where models may be weakest and their volatility. Abbreviations are as follows: predicted value (P), observed value (O), growth (G), and consumption (C).

Metric	Equation	Use
Percent error (PE)	$PE = (P - O)/O \times 100$	Generic approach for quantifying error in predicted trajectories of G or C. What is represented by P and O has frequently been unclear, but these likely have most often been final-day values in trajectories.
Mean growth error (MGE) or mean consumption error (MCE)[a]	$MGE \text{ or } MCE = (\sum_{i=1}^{n} \lvert (P_i - O_i)/O_i \rvert)/n \times 100$	Recommended for representing model prediction error for G or daily C. Accommodates positive and negative errors and permits comparisons of error across modeled trajectories of different durations.
Maximum absolute daily growth (MaxGE) or consumption error (MaxCE)[a]	$MaxGE \text{ or } MaxCE = \max \lvert (P_i - O_i)/O_i \rvert$	Represents model volatility.
Mean cumulative consumption error (MCCE)[b]	$MCCE = (P_f - O_f)/O_f/n \times 100$	Consumption prediction error based on cumulative consumption. Permits comparisons of error across modeled trajectories of different durations.

[a] Here P_i and O_i are the predicted and observed values for G or C on day i (units of g/g/d), and n is the number of days in the trajectory.
[b] Here P_f and O_f are the predicted and observed final-day values of cumulative consumption (g).

when models are evaluated under particular sets of growth conditions. The MaxGE provides indication of model volatility that may not be evident from MGE and would be particularly valuable in a situation in which a bioenergetics model gives close predictions of observed growth over much of a trajectory, but predictions then depart substantially over the final few days.

While use of the MGE and MaxGE approach would also be logical for representing model prediction error associated with daily consumption trajectories (Figure 12.4B), recall that it is more appropriate that model prediction error for consumption be based on cumulative consumption (Figure 12.4C). An appropriate representation of error for model predictions of cumulative consumption is based on final-day cumulative consumption values, but this should also be standardized for the number of days in a modeled trajectory. This measure of error is

termed mean cumulative consumption error (MCCE). Standardizing MCCE for the number of days in a modeled trajectory reduces the effect of trajectory length on the representation of model prediction error (Table 12.4). Crossing over of predicted and observed trajectories of cumulative consumption can occur; however, in such an event MCCE would still be appropriately calculated as in Table 12.4.

Analyses of modeling error.—Several statistical analyses are commonly applied to test for differences between bioenergetics-model-predicted and observed values of growth and consumption in model evaluation studies. The choice of statistical procedure is determined by whether the growth or consumption trajectories are replicated. Historically, some of the commonly applied statistical procedures and model performance indicators (Leggett and Williams 1981; Mayer and Butler 1993) were used because they accommodated model evaluations without replicated trajectories. Unreplicated trajectories of consumption and growth based on mean responses are most common in field evaluations of bioenergetics models whereas replicated trajectories are more common in laboratory evaluations.

Multivariate profile analysis has often been used to test whether a predicted and observed trajectory of consumption and growth are parallel with differences of zero along their full course (Rice and Cochran 1984). It has also been used in a model evaluation with replicated trajectories (Whitledge and Hayward 1997). The analysis is based on Hotellings's T^2-statistic:

$$T^2 = n \cdot Y' \cdot S^{-1} \cdot Y, \tag{12.5}$$

where n is the number of replicates; Y is a column vector of average model deviations between successive observation times; Y' is the transpose of Y; and S^{-1} is the inverse of the variance–covariance matrix (Timm 1975). In another commonly used statistical approach, degree and sources of prediction error are evaluated by decomposing the mean square error (MSE) associated with a linear least-squares regression of predicted on observed values (Rice and Cochran 1984; Wahl and Stein 1991; Whitledge and Hayward 1997; Zweifel 2000). The MSE (variance around the predicted–observed line) is decomposed into a mean component (m, error due to differences in the means of predicted and observed values), a slope component (s, error due to the slope differing from unity), and a residual component (r, the portion of MSE due to random error). Values of $m = 0$, $s = 0$, and $r = 1$ are most favorable because they indicate that predictive error is totally random and therefore asystematic, meaning that the model does not contain significant bias. Bonferroni joint confidence intervals are then applied to determine whether the slope and intercept of a regression of predicted on observed values are different from 1 and 0, respectively (Neter et al. 1990).

In contrast to field evaluations, replicated trajectories of fish consumption and growth are more readily produced in laboratory evaluations of bioenergetics models by measuring responses of multiple, individually held fish or fish groups under imposed combinations of growth-influencing conditions. Model evaluations using experimental designs that produce replicated trajectories of

growth and consumption permit application of more well-known statistical tests such as ANOVAs and *t*-tests or their nonparametric analogs.

Whitledge et al. (1998) evaluated the capacity of a bluegill bioenergetics model (Hanson et al. 1997) to predict cumulative consumption and growth of hybrid sunfish (green sunfish ♀ × bluegill ♂) fed ad libitum (control group) or according to three distinct schedules (treatment groups) that each elicited compensatory growth (Hayward et al. 1997). Each of the four groups comprised seven individually held hybrid sunfish; each fish's daily consumption and weekly growth was determined for 105 d at constant temperature. The bioenergetics model was applied to predict cumulative consumption (CC) and absolute growth rate (AGR) for each fish ($N = 28$) across groups. Differences between model-predicted and observed responses of AGR and CC for individual fish were tested within control and treatment groups using a paired *t*-test to determine if mean differences were nonzero. No difference between mean model-predicted and observed values of AGR and CC were found for the control group but differences were found for some of the treatment groups. Results indicated that the bioenergetics model predicted hybrid sunfish AGRs and CCs better under ad libitum feeding conditions than under conditions where compensatory growth occurred. In retrospect, it would have been preferable to have applied one-way ANOVA to test simultaneously whether mean differences between model-predicted and observed values differed among the four groups. Determinations of whether within-group means differed from zero could then have followed using, for example, the least-squares means (LSMEANS) procedure in SAS (SAS Institute 2004).

Besides evaluating performance of a single bioenergetics model, there can be interest in determining whether one of two or more models performs best. These situations can arise, for example, when (1) multiple versions of a newly constructed model are being evaluated, (2) adjustments have been made to an existing model and there is interest in whether model performance has been improved; or (3) performances of two distinct models applicable to the same species are compared. Bajer et al. (2003) used replicated growth and cumulative consumption trajectories from laboratory experiments involving a number of individually held yellow perch to compare predictive accuracy of two bioenergetics models, one for yellow perch and another for Eurasian perch. Observed laboratory growth and consumption for each fish was predicted by each of the two models. Values of MGE and MaxGE (portraying growth prediction error) as well as MCCE (portraying cumulative consumption prediction error) were computed for each model's prediction of growth or consumption for each fish (see Table 12.4 for equations). Differences between MGE values for each model's prediction of each fish's growth were tested for a nonzero mean difference through paired *t*-tests. A significant outcome indicated that predictions by the two bioenergetics models differed on average, in which case the model with the lower mean MGE value was considered to perform better. This same analytical process was repeated for MaxGE and MCCE.

In a more complex laboratory-based evaluation, Whitledge et al. (2003) simultaneously tested the relative performances of three smallmouth bass bioenergetics models, each applied to fish growth under three sets of conditions.

Daily consumption and growth rates of seven individually held smallmouth bass (100–270 g) were directly determined while they were simultaneously subjected to three successive 3-week periods involving different combinations of temperature and applied daily ration level. Each fish's growth rate (expressed as relative growth rate, RGR) and CC were predicted by each of the three bioenergetics models over each 3-week period. Absolute values of predicted minus observed RGR and CC values for each of the seven fish, as modeled by the three bioenergetics models over each of the three periods, provided the test data set. Absolute values were used because predicted minus observed RGR and CC values were both positive and negative and combining these values would have resulted in the inappropriate cancellation of error. Differences in model performance were evaluated by applying a completely randomized 3 (model) × 3 (growth condition) factorial ANOVA design with blocking done on the individual fish and comparing the absolute predicted minus observed RGR values and likewise for the CC values. Following indication from the ANOVA that significant differences existed, intermodel performance differences within each of the three sets of growth conditions were clarified using the LSMEANS procedure in SAS. The LSMEANS procedure was also used to determine whether mean predicted minus observed values of RGR and CC for each model–condition combination differed from zero.

12.4.3 Alternative Models

Currently popular bioenergetics models, including Wisconsin-type models (Kitchell et al. 1977; Hewett and Johnson 1987, 1992; Hanson et al. 1997) are fundamentally only predictors of fish growth and estimators of consumption. However, they are appealing and being applied with rapidly increasing frequency in part because their flexible design permits desktop insights into complex fish bioenergetic responses to changes in growth–influencing factors. Despite their appeal, it is probably also fair to say that their current popularity also owes, in part, to their high availability in user-friendly software (Hanson et al. 1997).

Direct approaches for in situ estimation of food consumption of fish populations are well known (e.g., Popova and Sytina 1977; Elliott and Persson 1978; Eggers 1979) and typically involve substantial field efforts where fish are collected at 3–4-h intervals over 24-h sampling periods (e.g., Hayward and Margraf 1987; Pedersen 2000). High effort requirements associated with in situ estimation procedures often cause deviations from ideal experimental designs such that sampling procedures are not well standardized. Despite the drawbacks, in situ estimation of consumption remains common, likely because of unavailability of indirect consumption estimation models for many fish species and life stages and the ability to gain direct insight into feeding dynamics of a population with these methods. Three categories of in situ consumption models and associated procedures are described by Adams and Breck (1990): chronology-of-feeding methods, carnivore feeding models, and continuous feeding models. Many in situ consumption models require the estimation of fish's gastric evacuation rates. Recent work indicates that gastric evacuation rates may often be underestimated leading to substantial

underestimation of in situ consumption (Richter et al. 2002). Also, Bochdansky and Deibel (2001) concluded that gastric evacuation patterns for many fishes are actually linear over time although appearing to be curvilinear when plotted against postfeeding time. They demonstrate that food consumption estimates may be biased twofold when a standard exponential gastric evacuation model is applied instead of a linear evacuation model. Approaches to reduce amounts of field effort required to make in situ estimates of daily food consumption of fishes have been developed by Boisclair and Leggett (1988) and Hayward and Hiebert (1993) (see also Madon 1998). A few studies indicate that it may be possible to gain accurate and precise in situ estimates of fish population daily consumption based on fewer within-day samples and fewer fish per sample by measuring food amounts in fish's whole guts versus only stomachs (Boisclair and Leggett 1988; Heroux and Magnan 1996).

12.5 INFORMATION NEEDS FOR BIOENERGETICS STUDIES

Because in situ estimation of fish consumption rates remains common, further evaluation and improvement of commonly used approaches is warranted. Laboratory-based evaluations of widely used in situ consumption models are needed for fishes of different feeding modes and life stages. Laboratory data sets that can include known 24-h consumption levels as controls would permit rigorous evaluations of consumption models, associated estimates of gastric evacuation, and effects of using whole-gut versus exclusively stomach-content weight as the basis for consumption estimation. Further investigation of predictive accuracy of reduced-effort procedures for in situ consumption estimation is also needed (e.g., Boisclair and Leggett 1988; Hayward and Hiebert 1993). Low-effort approaches may be used exclusively for certain applications (e.g., Whitledge and Hayward 2000), but their greatest value may be to expand spatial coverage of studied population's consumption rates or to increase estimate frequency over whole study periods. Further evaluation and development of radioisotopic approaches (Gingras and Boisclair 2000) is also needed.

Applications of bioenergetics models based on energy balance equations (Winberg 1956; Kitchell et al. 1977) are increasing geometrically and will likely continue to do so. Laboratory-based evaluations of soundness of existing versions of this class of model are much needed over broad ranges of conditions of fish size, temperature, and ration level. Follow-up efforts to correct detected areas of model weakness are critical, once model weaknesses have been identified. Recent findings (Bajer et al. 2003; Bajer et al. 2004a) show that substantial, consumption-level-dependent systematic error is likely widespread in bioenergetics models. These results also point toward consumption-dependent parameters in bioenergetcis models, including egestion, excretion, and SDA as likely sources of this error. Efforts to evaluate rigorously and improve current submodels that calculate values of energy costs–loss parameters within bioenergetics models' components are encouraged, even for some components for which sensitivity has been considered low (Bajer et al. 2003). Data and equation borrowing from other species to construct

new bioenergetics models should be more critically assessed (Ney 1993). Efforts should also continue to quantify fish activity costs from all sources in natural environments accurately and to evaluate the true practical importance of this parameter in bioenergetic model applications more thoroughly. The potential to estimate fish activity costs from readily measured correlates or indicators should continue to be explored, as should efforts to determine the accuracy of measurements of fish activity in the field obtained through physiological telemetry that monitors cardiac and muscle contraction rate. Fuller understanding of seasonal effects and also effects of exposure to periods of low food supply on fish bioenergetic rates including consumption and metabolism may be important toward improving predictive accuracy of bioenergetics models. Also, while bioenergetics models can often provide reasonable estimates of a fish's average daily consumption rates over weeks to months, they tend not to predict the substantial day-to-day fluctuations in consumption, which may be particularly pronounced when food availability is high (Whitledge and Hayward 1997). Development of capacity to predict daily fluctuations in food consumption would be valuable both in ecological applications (e.g., to predict short-term shifts in predatory impacts) as well as in aquaculture (e.g., to avoid over- and under-feeding of cultured fish). If relationships between fish growth in weight and corresponding skeletal growth could be defined, bioenergetics models could be used to predict fish growth in length as well as change in condition under various scenarios. Such capacity was recently developed (Bajer and Hayward 2006) and should be highly useful in fisheries management applications that tend to be more length than weight based. An ability to model fish condition under varying growth conditions would likewise be very useful.

Finally, evaluations of less commonly used in situ consumption estimation approaches described in section 12.4.3 would be valuable toward broadening awareness of the feasibility and reliability of different methods. In many cases low use rates of some approaches may not be due to their ineffectiveness but rather because other avenues have been more emphasized.

12.6 REFERENCES

Adams, M. S., and J. E. Breck. 1990. Bioenergetics. Pages 389–415 *in* C. B. Schreck and P. B. Moyle, editors. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Adams, M. S., R. B. McLean, and M. M Huffman. 1982. Structuring of a predator population through temperature-mediated effects on prey availability. Canadian Journal of Fisheries and Aquatic Sciences 39:1175–1184.

Allen, J. R. M., and R. J. Wootton. 1982. Age growth and rate of food consumption in an upland population of the three-spined stickleback *Gasterosteus aculeatus*. Journal of Fish Biology 21:95–106.

Bajer, P. B., and R. S. Hayward. 2006. A combined multiple-regression and bioenergetics model for simulating fish growth in length and condition. Transactions of the American Fisheries Society 135:695–710.

Bajer, P. B., R. S. Hayward, G. W. Whitledge, and R. D. Zweifel. 2004b. Simultaneous identification and correction of systematic error in bioenergetics models: demonstration with a white crappie *Pomoxis annularis* model. Canadian Journal of Fisheries and Aquatic Sciences 61:2168–2182.

Bajer, P. B., G. W. Whitledge, and R. S. Hayward. 2004a. Widespread consumption-dependent systematic error in fish bioenergetics models and its implications. Canadian Journal of Fisheries and Aquatic Sciences 61:2158–2167.

Bajer, P. B., G. W. Whitledge, R. S. Hayward, and R. D. Zweifel. 2003. Laboratory evaluation of two bioenergetics models applied to yellow perch *(Perca flavescens)*. Journal of Fish Biology 62:436–454.

Bartell, S. M., J. E. Breck, R. H. Gardner, and A. L. Brenkert. 1986. Individual parameters perturbation and error analysis of bioenergetics models of fish growth. Canadian Journal of Fisheries and Aquatic Sciences 43:16–168.

Beamish, F. W. H. 1974. Apparent specific dynamic action of largemouth bass, *Micropterus salmoides*. Journal of the Fisheries Research Board of Canada 31:1763–1769.

Beamish, F. W. H. 1990. Swimming metabolism and temperature in juvenile walleye, *Stizostedion vitreum vitreum*. Environmental Biology of Fishes 27:309–314.

Bochdansky, A. B., and D. Deibel. 2001. Consequences of model specification for the determination of gut evacuation rates: a case for simplicity. Canadian Journal of Fisheries and Aquatic Sciences 58:1032–1042.

Boisclair, D., and W. C. Leggett. 1988. An *in situ* experimental evaluation of the Elliott Persson and Eggers models for estimating fish daily ration. Canadian Journal of Fisheries and Aquatic Sciences 45:138–145.

Boisclair, D., and W. C. Leggett. 1989. The importance of activity in bioenergetics models applied to actively foraging fishes. Canadian Journal of Fisheries and Aquatic Sciences 46:1859–1867.

Brafield, A. E. 1985. Laboratory studies of energy budgets. Pages 257–282 *in* P. Tytler and P. Calow, editors. Fish energetics: new perspectives. Johns Hopkins University Press, Baltimore, Maryland.

Brandt, S. B., and K. J. Hartman. 1993. Innovative approaches with bioenergetics models—future applications to fish ecology and management. Transactions of the American Fisheries Society 122:731–735.

Brett, J. R. 1970. Temperature. Pages 515–560 *in* O. Kline, editor. A comprehensive integrated treatise of life in oceans and coastal waters. Wiley-Interscience, New York.

Brett, J. R. 1976. Scope for metabolism and growth of sockeye salmon, *Oncorhynchus nerka*, and some related energetics. Journal of the Fisheries Research Board of Canada 33: 307–313.

Brett, J. R., and T. D. D. Groves. 1979. Physiological energetics. Pages 279–352 *in* W. S. Hoar, D. J. Randall, and J. R. Brett, editors. Fish physiology, volume VIII. Academic Press, New York.

Bull, C. D., and N. B. Metcalfe. 1997. Regulation of hyperphagia in response to varying energy deficits in overwintering juvenile Atlantic salmon. Journal of Fish Biology 50: 498–510.

Cech, J. J., Jr. 1990. Respirometry. Pages 335–362 *in* C. B. Schreck and P. B. Moyle, editors. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Chiba, K. 1988. The effect of dissolved oxygen on the growth of young striped bass. Bulletin of the Japanese Society of Scientific Fisheries 54(4):599–606.

Chipps, S. R., D. F. Clapp, and D. H. Wahl. 2000. Variation in routine metabolism of juvenile muskellunge: evidence for seasonal metabolic compensation in fishes. Journal of Fish Biology 56:311–318.

Cox, M. K., and K. J. Hartman. 2005. Nonlethal estimation of proximate composition in fish. Canadian Journal of Fishes and Aquatic Sciences 62:269–275.

Crocker, C. E., and J. J. Cech, Jr. 1997. Effects of environmental hypoxia on oxygen consumption rate and swimming activity in juvenile white sturgeon, *Acipenser transmontanus,* in relation to temperature and life intervals. Environmental Biology of Fishes 50: 383–389.

Davis, G. E., and C. E. Warren. 1971. Estimation of food consumption rates. Pages 227–248 *in* W. E. Ricker, editor. Methods for assessment of fish production in fresh waters, 2nd edition. Blackwell Scientific Publications, Oxford, UK.

Davis, J. J., and R. F. Foster. 1958. Bioaccumulation of radioisotopes through aquatic food chains. Ecology 39:530–535.

Devore, J., and R. Peck. 1986. Statistics: the exploration and analysis of data. West Publishing, St. Paul, Minnesota.

Diana, J. S., and R. Salz. 1990. Energy storage, growth, and maturation of yellow perch from different locations in Saginaw Bay, Michigan. Transactions of the American Fisheries Society 119:976–984.

Dobson, S. H., and R. M. Holmes. 1984. Compensatory growth in rainbow trout, *Salmo gairdneri* Richardson. Journal of Fish Biology 25:649–656.

Eck, G. W., and L. Wells. 1983. Biology, population structure, and estimated forage requirements of lake trout in Lake Michigan. U.S. Fish and Wildlife Service Technical Papers 111, Ann Arbor, Michigan.

Eggers, D. M. 1979. Comments on some recent methods for estimating food consumption by fish. Journal of the Fisheries Research Board of Canada 36:1018–1019.

Elliot, J. M. 1976. The energetics of feeding, metabolism and growth of brown trout (*Salmo trutta* L.) in relation to body weight, water temperature and ration size. Journal of Animal Ecology 45:923–948.

Elliott, J. M., and W. Davidson. 1975. Energy equivalents of oxygen consumption in animal energetics. Oecologia 19:195–201.

Elliott, J. M., and L. Persson. 1978. The estimation of daily rates of food consumption for fish. Journal of Animal Ecology 47:977–991.

Gaylord, T. G., and D. M. Gatlin, III. 2000. Assessment of compensatory growth in channel catfish *Ictalurus punctatus* R. and associated changes in body condition indices. Journal of the World Aquaculture Society 31:326–336.

Gingras, J., and D. Boisclair. 2000. Comparison between consumption rates of yellow perch *(Perca flavescens)* estimated with a digestive tract model and with a radioisotope approach. Canadian Journal of Fisheries and Aquatic Sciences 57:2547–2557.

Grove, D. J., L. G. Loizides, and J. Nott. 1978. Satiation amount, frequency of feeding and gastric emptying rate in *Salmo gairdneri*. Journal of Fish Biology 12:507–516.

Hanson, P. C., T. B. Johnson, D. E. Schindler, and J. F. Kitchell. 1997. Fish bioenergetics 3.0. University of Wisconsin System, Sea Grant Institute, WISCU-T-97-001, Madison.

Hartman, K. J. 1993. Striped bass, bluefish, and weakfish in the Chesapeake Bay: energetics, trophic linkages, and bioenergetics model applications. Doctoral dissertation. University of Maryland, College Park.

Hartman, K. J. 2000. Variability in daily ration estimates of age-0 striped bass in the Chesapeake Bay. Transactions of the American Fisheries Society 129:1181–1186.

Hartman, K. J., and S. B. Brandt. 1993. Systematic sources of bias in a bioenergetics model—examples for age-0 striped bass. Transactions of the American Fisheries Society 122: 912–926.

Hartman, K. J., and S. B. Brandt. 1995a. Comparative energetics and the development of bioenergetics models for sympatric estuarine piscivores. Canadian Journal of Fisheries and Aquatic Sciences 52:1647–1666.

Hartman, K. J., and S. B. Brandt. 1995b. Predatory demand and impact of striped bass, bluefish, and weakfish in the Chesapeake Bay—applications of bioenergetics models. Canadian Journal of Fisheries and Aquatic Sciences 52:1667–1687.

Hartman, K. J., and S. B. Brandt. 1995c. Estimating energy density of fish. Transactions of the American Fisheries Society 124:347–355.

Hayward, R. S., and J. F. Hiebert. 1993. Low-effort regression estimation of average daily food weight in fishes: application to larval gizzard shad. Transactions of the American Fisheries Society 122:834–844.

Hayward, R. S., and F. J. Margraf. 1987. Eutrophication effects on prey size and food available to yellow perch in Lake Erie. Transactions of the American Fisheries Society 116:210–233.

Hayward, R. S., D. B. Noltie, and N. Wang. 1997. Use of compensatory growth to double hybrid sunfish growth rates. Transactions of the American Fisheries Society 126: 316–322.

Hayward, R. S., and N. Wang. 2001. Failure to induce over-compensation of growth in maturing yellow perch. Journal of Fish Biology 59:126–140.

Heroux, D., and P. Magnan. 1996. In situ determination of fish daily ration: review and field evaluation. Environmental Biology of Fishes 46:61–74.

Hewett, S. W., and B. L. Johnson. 1987. A generalized bioenergetics model of fish growth for microcomputers, version 1.0. University of Wisconsin Sea Grant Institute, Madison.

Hewett, S. W., and B. L. Johnson. 1992. A generalized bioenergetics model of fish growth for microcomputers, version 2.0. University of Wisconsin Sea Grant Institute, Madison.

Jobling, M. 1986. Mythical models of gastric emptying and implications for food consumption studies. Environmental Biology of Fishes 16:35–50.

Jobling, M., E. H. Jorgensen, and S. I. Siikavuopio. 1993. The influence of previous feeding regime on the compensatory growth response of maturing and immature Arctic charr, *Salvelinus alpinus*. Journal of Fish Biology 43:409–419.

Jobling, M., O. H. Meloy, J. dos Santos, and B. Christiansen. 1994. The compensatory growth response of Atlantic cod: effects of nutritional history. Aquaculture International 2:75–90.

Karas, P., and G. Thoresson. 1992. An application of a bioenergetics model to Eurasian perch *(Perca fluviatilis)*. Journal of Fish Biology 41:217–230.

Kitchell, J. F., J. F. Koonce, R. V. O'Neill, H. H. Shugart, Jr., J. J. Magnuson, and R. S. Booth. 1974. Model of fish biomass dynamics. Transactions of the American Fisheries Society 103:786–798.

Kitchell, J. F., D. J. Stewart, and D. Weininger. 1977. Application of a bioenergetics model to yellow perch *(Perca flavescens)* and walleye *(Stizostedion vitreum vitrum)*. Journal of the Fisheries Research Board of Canada 34:1922–1935.

Krohn, M. M., and D. Boisclair. 1994. Use of a stereo-video system to estimate the energy expenditure of swimming fish. Canadian Journal of Fisheries and Aquatic Sciences 51:1119–1127.

LaBar, G. W. 1993. Use of bioenergetics models to predict the effect of increases lake trout predation on rainbow smelt following sea lamprey control. Transactions of the American Fisheries Society 122:942–950.

Leggett, R. W., and L. R. Williams. 1981. A reliability index for models. Ecological Modelling 13:303–312.

Lezama, E. and J. Guenther. 1992. The routine metabolism of the guapote, *Cichlasoma managuense* (Guenther 1869), related to body weight and temperature. Journal of Fish Biology 41:373–380.

Lucas, M. C., I. G. Preide, J. D. Armstrong, A. N. Z. Gindy, and L. De Vera. 1991. Direct measurements of metabolism, activity and feeding behaviour of pike, *Esox lucius* L., in the wild by the use of heart rate telemetry. Journal of Fish Biology 39:325–345.

Madenjian, C. P., and D. V. O'Connor. 1999. Laboratory evaluation of a lake trout bioenergetics model. Transactions of the American Fisheries Society 128:802–814.

Madon, S. P. 1998. Low-effort regression estimation of daily ration in young walleye. Canadian Journal of Fisheries and Aquatic Sciences 55:2058–2066.

Mayer, D. G., and D. G. Butler. 1993. Statistical validation. Ecological Modelling 68:21–23.

Moser, M. L. and W. F. Hettler. 1989. Routine metabolism of juvenile spot, *Leiostomus xanthurus* (Lacepede), as a function of temperature, salinity and weight. Journal of Fish Biology 35:703–707.

Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied linear statistical models, 3rd edition. Irwin, Homewood, Illinois.

Neumann, R. M., and B. R. Murphy. 1991. Evaluation of the relative weight (W_r) index for assessment of white crappie and black crappie populations. North American Journal of Fisheries Management 11:543–555.

Ney, J. J. 1990. Trophic economics in fisheries: assessment of demand–supply relationships between predators and prey. Aquatic Sciences 2:55–81.

Ney, J. J. 1993. Bioenergetics today: growing pains on the cutting edge. Transactions of the American Fisheries Society 122:736–748.

Olla, B. L., and A. L. Studholme. 1971. The effects of temperature on the activity in the bluefish, *Pomatomus saltatrix*. Biological Bulletin (Woods Hole) 141:337–349.

Paul, A. J., J. M. Paul, and R. L. Smith. 1988. Respiratory energy requirements of the cod *Gadus macrocephalus* relative to body size, food intake, and temperature. Journal of Experimental Marine Biology and Ecology 122:83–89.

Paul, A. J., J. M. Paul, and R. L. Smith. 1990. Rates of oxygen consumption of yellowfin sole (*Limanda aspera* [Pallas]) relative to body size, food intake, and temperature. ICES (International Council for the Exploration of the Sea) Journal of Marine Science 47: 205–207.

Paul, A. J., J. M. Paul, and R. L. Smith. 1995. Compensatory growth in Alaska yellowfin sole, *Pleuronectes asper*, following food deprivation. Journal of Fish Biology 46:442–448.

Pedersen, J. 2000. Food consumption and daily feeding periodicity: comparison between pelagic and demersal whiting in the North Sea. Journal of Fish Biology 57:402–416.

Penczak, T. 1990. Bioenergetic trends in ecological studies on fish in Poland. Polish Ecological Studies 16:157–168.

Popova, O. A., and L. A. Sytina. 1977. Food and feeding relations of Eurasian perch *(Perca fluviatilis)* and pikeperch *(Stizostedion lucioperca)* in various waters of the USSR. Journal of the Fisheries Research Board of Canada 34:1559–1570.

Quinton, J. C., and R. W. Blake. 1990. The effect of feed cycling and ration level on the compensatory growth response in rainbow trout, *Oncorhynchus mykiss*. Journal of Fish Biology 37:33–41.

Rice, J. A., and P. A. Cochran. 1984. Independent evaluation of a bioenergetics model for largemouth bass. Ecology 65:732–739.

Richter, H., C. Leuckstaedt, C. Schlechtriem, F. Ulfert, and K. Becker. 2002. Evidence of violation of a basic assumption in models currently used for food consumption estimation in fish. Archive of Fishery and Marine Research 49:171–188.

Ruohonen, K. 1999. Modeling fish food consumption with artificial neural networks. Aquaculture Research 30:545–548.

Russell, N. R., and R. J. Wootton. 1992. Appetite and growth compensation in the European minnow, *Phoxinus phoxinus* (Cyprinidae), following short periods of food restriction. Environmental Biology of Fishes 34:277–285.

SAS. 2004. SAS, version 9.1. SAS Institute, Cary, North Carolina.

Smith, R. L., A. J. Paul, and J. M. Paul. 1988. Aspects of energetics of adult walleye pollock, *Theragra chalcogramma* (Pallas), from Alaska. Journal of Fish Biology 33:445–454.

Stewart, D. J., J. F. Kitchell, and L. B. Crowder. 1981. Forage fish and their salmonid predators in Lake Michigan. Transactions of the American Fisheries Society 110: 751–763.

Stewart, D. J., D. Weininger, D. V. Rottiers, and T. A. Edsall. 1983. An energetics model for Lake Trout, *Salvelinus namaycush*: application to the Lake Michigan population. Canadian Journal of Fisheries and Aquatic Sciences 40:681–698.

Thornton, K. W., and A. S. Lessem. 1978. A temperature algorithm for modifying biological rates. Transactions of the American Fisheries Society 107:284–287.

Timm, N. H. 1975. Multivariate analysis with applications in education and psychology. Brooks/Cole Publishing, Monterey, California.

Trudel, M., and D. Boisclair. 1993. An in situ evaluation of the day-to-day variation in the quantity of food consumed by fish. Canadian Journal of Fisheries and Aquatic Sciences 50:2157–2165.

Ver Hoef, J. M., and N. Cressie. 1993. Spatial statistics: analysis of field experiments. Pages 319–341 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Vigg, S. and C. C. Burley. 1991. Temperature-dependent maximum daily consumption of juvenile salmonids by northern squawfish (*Ptychocheilus ovegonesis*) from the Columbia River. Canadian Journal of Fisheries and Aquatic Sciences 48:2491–2498.

Wahl, D. H., and R. A. Stein. 1991. Food consumption and growth of three esocids: field tests of a bioenergetic model. Transactions of the American Fisheries Society 120: 230–246.

Warren, C. E., and G. E. Davis. 1967. Laboratory studies on the feeding, bioenergetics, and growth of fish. Pages 175–214 *in* S. D. Gerking, editor. The biological basis of freshwater fish production. Wiley, New York.

Whitledge, G. W., and R. S. Hayward. 1997. Laboratory evaluation of a bioenergetics model for largemouth bass at two temperatures and feeding levels. Transactions of the American Fisheries Society 126:1030–1035.

Whitledge, G. W., and R. S. Hayward. 2000. Determining sampling date interval for precise in situ estimates of cumulative fish consumption by fishes. Canadian Journal of Fisheries and Aquatic Sciences 57:1131–1138.

Whitledge, G. W., R. S. Hayward, D. B. Noltie, and N. Wang. 1998. Testing bioenergetics models under feeding regimes that elicit compensatory growth. Transactions of the American Fisheries Society 127:740–746.

Whitledge, G. W., R. S. Hayward, C. F. Rabeni, and R. D. Zweifel. 2003. Development and laboratory evaluation of a bioenergetics model for subadult and adult smallmouth bass. Transactions of the American Fisheries Society 132:316–325.

Winberg, G. G. 1956. Rate of metabolism and food requirements of fishes. Belorussian University, Minsk. Translated from Russian: Fisheries Research Board of Canada Translation Series 164, 1960, Ottawa.

Zweifel, R. D. 2000. Development and evaluation of a bioenergetics model for white crappie. Master's thesis. University of Missouri, Columbia.

13 Fish Population Bioassessment

Michael Power

13.1 INTRODUCTION

Approaches to detecting, interpreting, and reporting the significance of stressors on fish populations vary. The tools described in Chapters 4 through 10 may all be used alone, or in combination, to yield insights into how fluctuations in environmental variables or anthropogenic stressors such as exploitation, habitat degradation, and pollution—or some combination of both—may affect abundance, reproductive potential, or growth. Despite general agreement in the fisheries literature that an understanding of the population level effects of stressors is of paramount importance, relatively few studies have attempted to estimate stressor effects systematically (Barnthouse 1993) or to link changes explicitly at lower levels of biological organization to changes at the population level (Shuter and Regier 1989). Most investigations have followed the pioneering lead of Selye (1976), who defined stress as the measurable biological response of an individual to an external stimulus (stressor), and are based on laboratory experimentation aimed at determining individual responses to acute or chronic stressors (Adams 2002). Although valuable for achieving the specific objectives of determining whether a measurable response exists, studies focused at the individual level lack ecological realism and do not necessarily predict or facilitate understanding the consequent effects of stress at the population level (Power and McCarty 1997).

The problem of quantifying the effect of stressor perturbations on populations is not new and has formed the core of wildlife and fisheries management research for many years (Barnthouse 1993). As with most population-based investigations, the objective of stress-related population research is to infer from individual sample data the characteristics of a well-defined grouping of like organisms (the population) and the probable dynamic responses of the group to stress. In practice, fisheries scientists have tended to equate stress with anthropogenic disturbances that induce responses outside the normal range of variation for the selected measurement endpoint (Evans et al. 1990). Although adoption of the normative range concept (Odum et al. 1979) dealt adequately with the duality of possible beneficial and harmful stressor effects at a theoretical level, inherent variability caused

by the action of environmental variation and the operation of natural compensation mechanisms has confounded the simple interpretation of stress–response relationships (Evans et al. 1990).

Understanding the responses of populations to stressors (population bioassessment), therefore, requires knowledge of both the ways in which individuals respond to stress and the effect of collective individual responses on the processes that govern population dynamics. Individual responses to stress may be observed directly from experimental or field studies. A detailed compendium of technical and theoretical considerations, much of it directly related to fishes, is given in Adams (1990, 2002). Collective individual responses will sum to determine population level impacts, and these can be inferred only indirectly using individual response data to estimate sample means and variances. In practice, sample statistics have been found to have low predictive power because of the effect of population regulating feedback mechanisms operating through compensatory adjustments to birth and death rates (e.g., density-dependent recruitment success, compensatory growth, and fecundity changes [McFadden 1977]). The likelihood that stressors will interact spatially and temporally with population compensation mechanisms has further complicated attempts to evaluate stressor effects and has necessitated understanding the ways in which interactions between population regulating mechanisms and stressors might influence observed population level traits (Minns 1992).

Ecological theory has tended to view each population as unique and so tightly integrated into its own particular ecosystem that the data collection programs necessary to produce credible assessments of stressor and stressor interaction effects would themselves significantly alter the populations under study (Rigler 1982). Furthermore, the sensitivity of manipulative experiments at the population level can be severely compromised because the number of replicate populations available for use in experimentation is limited and local uniqueness among replicates is so high that variance estimates remain large (Walters et al. 1989). Accordingly, even with the many documented cases of population collapses in fish stocks (e.g., Great Lakes), it has typically been impossible to ascertain unambiguously the specific causes of decline (Barnthouse et al. 1990). Nevertheless, attention will continue to focus on stressor effects at the population level because of public concern for the fate of highly valued populations and legislative mandates requiring attempts to ensure the continued viability of threatened or endangered species and no net loss of habitat.

Against this background of ecological complexity, this chapter will examine population bioassessment methods for determining the possible consequences of stressor action on fish populations. The emphasis is on identifiable quantitative methods whose objective is the determination of a population level measure of success (e.g., abundance) or process (e.g., operation of density dependence). In that regard, field-based assessment techniques that combine consideration of sampling protocols with routine statistical analysis of obtained data are not considered, except where such approaches have an identifiable analytical framework. Thus generic approaches, such as before–after comparison of impact (BACI)

methods, are not considered here, but graded-exposure–response techniques (e.g., Adams et al. 1993) are considered. Also not considered here are the numerous ecosystem process or energy flow modeling frameworks that have been developed in recent years. These include bioenergetic-based techniques, habitat supply models, individual-based models, and systemwide simulation frameworks (e.g., Ecopath and Ecosim). Many of the considered techniques are not specific to fisheries, having been developed by entomologists, terrestrial ecologists, and others. All approaches herein attempt to incorporate consideration of population regulating mechanisms that confound the direct influences of the biotic and abiotic environmental factors critical to determining population level characteristics (e.g., intrinsic rate of population increase or mean fecundity).

For purposes of discussion, a population is defined as a group of individuals of the same species that occupy a definable geographic range and are reproductively self-sustaining. A population need not be isolated from the effects of immigration or emigration, but for simplicity, such isolation is often assumed. Populations are governed by the fundamental processes of reproduction, mortality, and the somatic increases that render individuals capable of reproduction (Shuter 1990). These critical processes are moderated by density-independent and density-dependent adjustments that compensate for abnormal levels of numerical abundance. Ecologists have recognized for some time that populations can persist only if some form of compensatory response exists (e.g., Nicholson 1933). The compensatory processes that allow populations to persist, however, are also partially capable of counteracting the adverse effects of stress at the population level (Nicholson 1954), thereby complicating attempts at population level bioassessment. The exact nature of density-dependent factors in the control of populations has been debated for many years (see reviews by Clark et al. 1967; Begon et al. 1990), but there is now widespread agreement that fluctuations in the abundance of persisting populations are the result of both density-dependent and density-independent processes (Hassell 1986; Elliott 1994). Discussions of density-dependent processes, their possible effects on fish populations, and the complications density-dependent feedbacks pose for the measurement and interpretation of fisheries data are given in Goodyear (1980), Hassell (1986), Evans et al. (1990), Meyers (2002), and Power (2002).

One of the biggest problems in population bioassessment has been differentiating the effects of natural variability from the operation of density-dependent population regulation. Natural variability may arise for reasons of demographic or environmental stochasticity. Sources of variability in demographic parameters can be broadly classified into those that depend on trophic interactions and those that depend on physical processes (Shuter 1990). These diverse considerations include biotic factors affecting variability in the energy available for growth and reproduction, the abundance and availability of prey items, the predator field that is encountered, and other factors such as disease and parasitism (Fogarty et al. 1991). Natural variability may also be mistaken for density-dependent processes resulting from within-generation heterogeneity arising from mechanisms that render some individuals more susceptible to mortality than others (Hassell 1986).

Such mechanisms include nonrandom mortality factors, the existence of refugia protecting individuals from a mortality-causing agent, and processes that give rise to temporal asynchrony (e.g., differential spawning or hatch dates).

Although much of the mainstream fisheries literature has concerned itself with addressing the implications of density-dependent factors for determining population success, traditional contaminant-based studies of possible population level stressor effects have tended to ignore the issue altogether (Power and McCarty 1997). For example, density-dependent effects are not recognized among the traditional list of significant biotic modifying factors (e.g., individual size or nutritional status), knowledge of which is necessary to interpret toxicity test results appropriately. The omission assumes density-related factors can be treated as experimental constants and denies the importance of population regulating mechanisms as determinants of the status of individuals in the population.

Realization of the importance of population processes and life history strategies for conclusions about fish population status has driven much of the development of traditional fisheries assessment and management practices. A listing of common measures used for assessing fish population status is given in Table 13.1. Although developed largely to assess exploitation pressures, the techniques are useful for assessing the effects of other anthropogenic and natural stressors on population status (e.g., Healey 1978; Mills 1985; Mohr et al. 1990; Berlinsky et al. 1995). Reviews of the basic techniques required to estimate the measures listed in Table 13.1 are given in Chapters 4 through 10 and will not be discussed in detail here.

In its broadest form, population level bioassessment can be taken to include any measurement-based observational or experimental study that is directed toward the estimation of a summary population characteristic (e.g., intrinsic rate of growth, abundance, or mean fecundity) from sample data. Thus, many of the methods developed specifically for fisheries ecology have been included in comprehensive monitoring programs attempting to combine information on stressors and their measurable biological effects for making judgments about risk (e.g., USEPA 1992). Furthermore, attempts have been made to integrate disparate field-based measurements into frameworks that integrate stressor exposure and effects data by placing specific emphasis on the causal mechanism and quantification of the cause–effect relationship. As an alternative, sample data are used to estimate population vital statistics from which life history accounting (e.g., life tables) or projection models (e.g., Leslie matrix or stock–recruitment) may be constructed. Accordingly, there is a dichotomy between the largely field-based (section 13.2) and modeling-based (section 13.3) approaches to population level bioassessment.

13.2 FIELD-BASED APPROACHES TO POPULATION LEVEL BIOASSESSMENT

The field-based approaches may be grouped into three main categories: compare and contrast, graded-exposure–response, and sequential sampling methods. A full discussion of these techniques, and their short-comings, is given in Power (2002) and will not be repeated in detail here. Briefly, compare and contrast approaches to assessing population level impacts have a long history (Cairns et al.

Table 13.1 Common methods or indicators used for assessing fish population status as described in Chapters 4 through 10. Each is based on or derived from field measurements and describes a single aspect of the growth, survival or reproductive processes that regulate population abundance or biomass. To infer possible stressor effects on fish populations appropriately, at least one metric from each category listed below must be included in any bioassessment exercise. Metrics within each category are not necessarily independent of others within the same or other categories. For example, changes in specific growth rates hold direct implications for fecundity and may affect age-specific survival rates in subsequent generations.

Growth-Related Metrics	
Mean weight- or length-at-age	Specific growth rates
Allometric relationships	Proximate body condition
Population size structure	Production, production:biomass ratios
Condition factor	Biomass indices
Survival-Related Metrics	
Age-specific mortality rates	Density or abundance
Year-class strength	Mean or maximum age
Population age structure	Recruitment indices
Catch per unit effort	Mean life expectancy
Intrinsic rate of population increase	Generational cycle length
Reproduction-Related Metrics	
Age-at-maturity	Egg size
Reproductive life span	Spawning frequency
Gonad somatic index	Net reproductive rate
Age-specific fecundity rates	Intrinsic rate of population increase

1984; Ryder and Edwards 1985; Munkittrick and Dixon 1989; Shuter 1990) and stemmed from the need to develop reliable methods capable of detecting the adverse effects of a wide variety of environmental stressors. A basic assumption of the compare and contrast approach is that the population is the best indicator of its own status and that consequent changes in population sample characteristics (e.g., age, fecundity, and condition factor) will be functionally related to the presence or absence of a putative stressor. Accordingly, compare and contrast methods are typically executed as before and after and reference site comparisons. Before and after comparisons monitor critical population characteristics before and after the application of a possible stressor to determine if changes have occurred (e.g., Munkittrick and Dixon 1989). The approach assumes that the stressor is the cause of any measured change in monitored characteristics and that characteristics would not have varied in the absence of the stressor. Reference site comparisons use data on population characteristics from affected and unaffected areas that are ecologically similar (e.g., Swanson et al. 1994) to make inferences about cause and effect by assuming area differences are attributable to the differential presence of the stressor rather than to other possible systematic differences.

Compare and contrast frameworks have several limitations (Munkittrick and Dixon 1989), the most serious of which are their static nature (Power 2002).

Attempts to validate compare and contrast frameworks by means of modeling (Jaworska et al. 1997) and multiple reference site (van den Heuvel et al. 1999) approaches have met with limited success. As van den Heuvel et al. (1999) have concluded, the physiological response indices used by static frameworks may not be the most sensitive indicators of nonlethal stressor exposure because impacts can be masked by population level compensatory responses.

Graded-exposure–response approaches compare groups of individuals (possibly populations) in a series of graded stressor exposures and seek to establish a correlative link between the measured exposure (e.g., concentrations of a contaminant or degree of exploitation) and biological characteristics of each group (e.g., Adams et al. 1994). The approach assumes that any gradient observed in sample characteristics is attributable to known differences in stressor intensity alone. That is, the approach makes an "all other things equal" assumption about the possible differences in the systematic operation and effect of other population level regulating factors or stresses. The approach is well described in a series of papers by Adams (Adams et al. 1993; Adams and Ryon 1994; Adams et al. 1994) and others (Karas et al. 1991; Sandstrom 1994) and is noted for its use of sample information on multiple response characteristics. Use of information on multiple characteristics assumes environmental complexity is such that it is unlikely that single characteristic measures will accurately reflect responses to stress (Adams and Ryon 1994). Although the procedure is useful in helping to identify factors that impair fish populations and the degree of difference among these factors, it is a weak diagnostic tool in the sense that it cannot detect the cause of the problem (Adams et al. 1993). However, further detailed validation studies are being conducted to determine on a case-specific basis the levels of evidence required for inferring causal linkages.

Sequential sampling methods use traditional field sampling approaches to study the long-term effects of exposure to a single stressor incident (e.g., Mills and Chalanchuk 1987; Mills et al. 2000). The assumptions here are that critical response variables (e.g., age at maturity or fecundity) relevant to determining stressor effects may be identified a priori and that sufficiently long temporal data series will prove adequate for the quantification of observed associations between population level responses and the stressor. Few examples of sequential sampling studies on the effects of stress at the population level exist. Among the more innovative sets of sequentially sampled data are those assembled from the Experimental Lakes Area of northwestern Ontario to study the effects of controlled lake acidification and recovery (Mills and Chalanchuk 1987; Mills et al. 2000) and whole lake fertilization (Mills 1985; Mills and Chalanchuk 1987; Mills et al. 1998). Results from these experiments point to the importance of multi-trophic sampling and increased sampling frequency in the detection of critical population level responses to stress and the validation of response predictions. Although important for increasing understanding of ecosystem structure and function as it pertains to fish populations of interest, lessons from such experiments are difficult to generalize, and it is probably unreasonable to expect adoption of the approach for anything but specialized research needs.

As a group, field-based attempts to assess population level responses to stress and their possible ecological significance follow from the observation that fishes in their natural environment are typically subjected to a number of stressors that alone, or in combination, are capable of triggering measurable physiological responses having population level implications (Adams et al. 1993). Field-based measures are viewed as integrative and as one means of directly capturing the consequences of the complex interaction of environmental factors for studied populations, while at the same time avoiding the difficulties associated with the extrapolation of single stressor–response laboratory test data (Munkittrick and Dixon 1989). As a result, field-based studies have become a popular means of assessing stressor impacts when questions about the larger scale or longer-term implications of stressors arise. A more detailed summary of the main field-based approaches is given in Power (2002).

13.3 MODELING-BASED APPROACHES TO POPULATION LEVEL BIOASSESSMENT

The many field-based attempts to quantify population level stress responses have taught us much about the difficulties associated with measuring the magnitude of stressor effects and understanding the ways in which measurable population characteristics may be functionally related to known stressor intensity. The number of biotic and abiotic factors and mechanisms capable of modifying the effect of a single stressor on measured population level responses suggests that the interpretation, or prediction, of population level responses based on models alone will never be an easy task and should not be separated from the basic ecological work necessary to describe affected populations in their natural environment.

Numerous analytical modeling frameworks suitable for determining possible population level effects of a wide variety of stressors have been developed. Although modeling studies of populations can provide descriptions of the effects of stress, they rely on making assumptions about causal mechanisms and cannot establish the existence of actual cause–effect linkages (Maltby 1999). Nevertheless, Minns (1992) argues that the complexity of ecosystems is such that modeling provides one of the few systematic means within which the dynamics of population level responses to a suite of interacting anthropogenic and natural stimuli can be appropriately analyzed. Modeling is also less expensive and permits investigation of potential population fluctuations over much longer time scales (Landahl et al. 1997). Furthermore, modeling has progressed to the point at which identifying vulnerable life stages, ranking sources of stress on an effect basis, and comparing alternative mitigating strategies can now be easily accomplished (Vaughan et al. 1984; DeAngelis et al. 1990; Evans et al. 1990).

The suitability of a model type for predicting possible population level effects will depend on the complexity of the question being asked, the suitability and availability of data for model development, the availability of stressor–effect data, the skill of the modeler in conceptualizing and representing complex processes within the constraints of simplifying model assumptions, and the perceived usefulness of

model output to decision makers (Vaughan et al. 1984; Chambers 1993; Power and McKinley 1997). To that end, care should be taken when selecting between available model types. Emlen (1989) has also argued that to address the effects of stressors at the population level appropriately, models must produce outputs describing population endpoints of regulatory relevance (e.g., intrinsic rate of population increase) and must incorporate available scientific information on stressor effects. With these criteria in mind, three categories of model types have been selected for discussion on the basis of their historical utility for estimating fish population level responses to stress: life tables, matrix models, and stock assessment frameworks. Two further categories are discussed because of their possible future utility for summarizing temporal data or determining the life history phases most critical for the regulation of population abundance: variance and key-factor analysis.

13.3.1 Life Table Analysis

Life tables have a long history of use in the insurance industry because life table models provide useful summaries of changes in population characteristics for a given set of conditions summarized in population-specific mortality and natality data. The life table technique can be readily adapted to fish populations to determine the net effect of environmentally induced changes in mortality and natality rates on summary parameters (e.g., intrinsic rate of population increase) defining the ability of populations to sustain themselves. As a consequence, life tables have been widely applied as a means of summarizing population level effects of stress (Walthall and Stark 1997).

Two strong assumptions dominating life table analysis are environmental stability and the lack of a limit placed on population growth. Violation of either or both assumptions can lead to large predictive errors when attempting to determine potential stressor effects on population abundance. Any component of the environment that affects age-specific fecundity or mortality will also affect the life table summary parameters. Accordingly, key life table parameters can be computed for only a given set of environmental conditions. Under natural conditions, however, environmental constancy is rarely the case. Populations vary in response to environmental influences acting on natality and mortality and to the feedback effects environmental variation has on the natality- and mortality-derived demographic schedules from which summary life table parameters are computed (Birch 1953). Nevertheless, life table data and derivative age distributions can be used to judge the status of a population. Care should be taken to ensure that the summary parameters are calculated for a series of population densities and environmental scenarios.

13.3.1.1 Life Table Types and Construction

Life tables proceed by summarizing population mortality and natality schedules over an arbitrary time interval. The chosen interval should be related to a relevant biological time scale (e.g., annual reproductive cycle) and for most fishes

will equal a year. The chosen interval may be shorter, but this comes at the cost of increasing the data detail needed to construct tables. Furthermore, it is conventional to include only females in the construction of life tables. Accordingly, population data derived from field observation or sampling must be adjusted by sex ratio information to obtain sex-specific numbers. Finally, census data are preferred for life table construction.

In the case of field-based studies it is often not possible to study the entire population or to age all individuals accurately. As a result, the development and application of life tables in fisheries studies does not rely on the use of census data. Instead, static or cohort-specific life tables are constructed using survey data. A static life table is constructed from cross-sectional sampling of a population at a specific point in time and presents derived population parameters as averages under the assumption that age-specific fecundity and mortality rates have remained constant from year to year (i.e., the population is stationary). A cohort-specific life table is constructed from temporal information obtained from following a single age-class from birth to death. The cohort-specific life table is considered to be the more reliable of the life table types for use in ecological study (Krebs 1999). If the environment is static, the two approaches will be equivalent, and the population under study will be in equilibrium. If the environment is variable, the static table will combine age-specific variations in vital rates with inter-annual variations in vital rates, confounding accurate estimation of summary table parameters (e.g., net reproductive rate, R_0, and the intrinsic rate of population increase, r).

Life table descriptions come in many forms (e.g., Wootton 1990; Krebs 1999), and there is some variation in the use of symbols. Generally, the following notation is used to describe the age-specific data computed for use in life-table analysis:

x = age or time interval used in table computations;
n_x = number of individuals in a cohort or age-class alive at the start of age interval x;
l_x = proportion of individuals in a cohort or age-class surviving from age 0 to age x (note that conversion to proportions allows among-population comparisons of survival rates);
d_x = number of individuals in a cohort or age-class dying in the interval x to $x + 1$ (values may be summed to define total mortality over any defined period of time);
q_x = finite per capita mortality rate in the interval x to $x + 1$;
p_x = finite per capita survival rate in the interval x to $x + 1$; and
e_x = mean expectation of life for individuals alive at the start of age interval x.

Once a single column of information for a life-table is known (typically n_x or d_x), relationships among the information columns allow the remaining columns to be computed using the formulae given in Table 13.2 (Caughley 1977) and below for e_x (Krebs 1999).

$$e_x = T_x / n_x, \text{ where } T_x = \sum_{i=x}^{m} L_i \text{ and } L_x = \frac{n_x + n_{x+1}}{2}. \quad (13.1)$$

Table 13.2 Formulae for converting between elements of the life table. The left-hand column defines the data type requiring conversions. Formulae to the right define the conversion relationship for the type of elemental data required. For example, n_x is converted to d_x using the relationship $(n_{x+1} - n_x)$ (source Caughley 1977). Symbols are as follows: x = age or time interval used in table computations; n_x = number of individuals in a cohort or age-class alive at the start of age interval x; l_x = proportion of individuals in a cohort or age-class surviving from age 0 to age x; d_x = number of individuals in a cohort or age-class dying in the interval x to $x + 1$ (values may be summed to define total mortality over any defined period of time); q_x = finite per capita mortality rate in the interval x to $x + 1$; and p_x = finite per capita survival rate in the interval x to $x + 1$.

Element to be converted	Conversion relationship				
	n_x	l_x	d_x	q_x	p_x
n_x		$\frac{n_x}{n_0}$	$n_{x+1} - n_x$	$\frac{n_{x+1}}{n_0} - 1$	$\frac{n_{x+1}}{n_0}$
l_x	$l_x \cdot n_0$		$(l_{x+1} - l_x)n_0$	$1 - \frac{l_{x+1}}{l_x}$	$\frac{l_{x+1}}{l_x}$
d_x	$\sum_{y=x}^{\infty} d_y$	$\frac{\sum_{y=x}^{\infty} d_y}{n_0}$		$\left(\frac{d_x}{\sum_{y=x}^{\infty} d_y}\right)$	$1 - \left(\frac{d_x}{\sum_{y=x}^{\infty} d_y}\right)$
q_x	$n_0 \prod_{y=0}^{x-1} (1 - q_y)$	$\frac{\prod_{y=0}^{x-1} (1 - q_y)}{n_0}$	$q_x \prod_{y=0}^{x-1} (1 - q_y)$		$1 - q_x$
p_x	$n_0 \prod_{y=0}^{x-1} p_y$	$\frac{\prod_{y=0}^{x-1} p_y}{n_0}$	$1 - q_x \prod_{y=0}^{x-1} (1 - q_y)$	$1 - p_x$	

The average number of individuals alive in the interval x to $x + 1$ is defined by L_x, and T_x is the sum of all L_x values over the interval 0 to m (maximum observed age).

For both static and cohort-specific life tables, the number of individuals with a common birth period dying (d_x) or surviving (n_x) in successive intervals of time must be determined. For fisheries work such data may be difficult to collect except when long-term studies are undertaken, such as those described by McFadden et al. (1967) for brook trout populations in Hunt Creek, Michigan. Even when long-term data are available, life table construction may require application of indirect estimation procedures including the construction of age–length keys and the completion of length-specific population estimates (McFadden et al. 1967). Nevertheless, there are numerous possible data sources for use in fish population life table construction, including those listed below.

1. Direct observation of the population, usually under controlled experimental conditions if data are being collected from longitudinal studies, where the number of individuals alive, or dead, in successive intervals is recorded for a cohort. The observed data are the n_x or d_x columns of the life table.
2. Observed age at death, usually from creel surveys or other cross-sectional survey methods. Data allow the estimation of mortality through the use of catch curves that may be used to determine the l_x column of the life table.
3. Direct observation of population age structure at a known point in time obtained from cross-sectional sampling methods that must be controlled for known sampling biases (e.g., net selectivity). Numbers must be counted directly rather than be based on estimates of the proportion of individuals in each age-class. Data allow the number of individuals alive at age x to be compared with those that die before reaching $x+1$ to derive the number of deaths in a given age interval, d_x, and a direct estimate of q_x.

There is little guidance in the literature about the sample sizes needed to construct accurate life tables (Krebs 1999). On the basis of experience with terrestrial animal ecology, Caughley (1977) recommended a minimum of 150 individuals when age distributional data are used in life table construction. For fish species with highly variable age-at-first maturity or fecundity, a sample size of 150 individuals is probably too few for the accurate estimation of life table parameters. An average of 30 or more individuals in each age-class ought to be included for reasons of ensuring the statistical adequacy of age-specific parameters used to derive the key n_x, d_x, l_x, or q_x schedules in a life table.

13.3.1.2 *Summary Parameters Derived from Life Tables*

Critical to life table construction are the summary computations for parameters used to characterize population dynamics. Computed parameters include both individual and population values. Gross reproductive rate (GRR), generational cycle length (G), and mean life expectancy (e_0) are among the useful summary parameters, with GRR being the most useful for reproductive assessments and G and e_0 being most useful for describing population cycle lengths and expectations of individual longevity. Net reproductive rate (R_0) is used as an indicator of population status. The intrinsic and finite rates of population increase, *(r)* and (λ) respectively, are similarly useful population indicators that define the potential rates of change in population abundance per unit of time. Formulae detailing the computation of these parameters are given in Leslie et al. (1955), Pianka and Parker (1975), Caughley (1977), Wootton (1990) Krebs (1999) and in the example discussed in Box 13.1.

The intrinsic and finite rates of population increase are commonly used in the literature to compare populations and draw inferences about relative success. The intrinsic rate of population increase *(r)* in a population with fixed l_x and b_x schedules (Box 13.1) may be computed when the population has achieved a stable age distribution via iteration based on the Euler–Lotka equation:

$$l = \sum_{x=0}^{\infty} e^{(-rx)} l_x b_x. \tag{13.2}$$

Box 13.1 Life Table Analysis

Consider the following data obtained for the 1954 cohort of brook trout from Hunt Creek, Michigan (McFadden et al. 1967). Details of sampling methodology and estimation of the life table column data are given in McFadden et al. (1967).

Table Data for the 1954 cohort of brook trout from Hunt Creek, Michigan (McFadden et al. 1967). Symbols are defined in the text that follows. Computations in this box are carried to five decimal places.

Age	n_x	l_x	d_x	q_x
0	52,000	1.00000	49,882	0.95927
1	2,118	0.04073	1,339	0.63220
2	779	0.01498	620	0.79589
3	159	0.00306	146	0.91824
4	13	0.00025	13	1.00000
5	0	0.00000		

The column n_x defines numbers of individuals alive at each age, with the proportions of the population surviving from age 0 to age x (l_x) schedule being computed directly from the n_x schedule by means of the relationships given in Table 13.2. The d_x column defines the number of individuals dying in the interval x to $x + 1$ and is combined with n_x to compute the finite per capita mortality rate (q_x) in the interval x to $x + 1$ (Table 13.2). Because population growth will depend on both natality and mortality rates, the fecundity schedule (b_x) must be determined before life table summary parameters describing population dynamics can be estimated. Modification of the life table to include fecundity data given in McFadden et al. (1967) yield the table below from which summary measures for the gross reproductive rate (GRR) and net reproductive rate (R_0) may be computed as shown (Krebs 2002).

Table Summary measures for gross (GRR) and net (R_0) reproductive rates based on fecundity schedule (b_x) for 1954 cohort of brook trout.

Age and summary measure	n_x	l_x	b_x	$l_x b_x$
0	52,000	1.00000	0.0	0.0
1	2,118	0.04073	0.0	0.0
2	779	0.01498	43.0	0.64414
3	159	0.00306	122.6	0.37516
4	13	0.00025	346.2	0.08655
5	0	0.00000		
Summary equation	$\text{GRR} = \sum_{x=0}^{5} b_x$	$R_0 = \sum_{x=0}^{5} l_x b_x$		
and value	GRR = 511.8	$R_0 = 1.10585$		

The average number of daughters produced by a female living to maximum age is defined by GRR, whereas R_0 defines the average number of age-0 female offspring produced by the average newborn female during its life. The important difference between GRR and R_0 is that GRR is an individual measure of reproductive potential, whereas R_0 defines the multiplication rate of the population per generational cycle. Thus, if $R_0 > 1$ the population is increasing; if $R_0 = 1$ the population is stable; and if $R_0 < 1$ the population is decreasing.

Once R_0 is known, mean generational cycle length (G) may be estimated as follows:

$$G = \frac{\sum_{x=0}^{X} l_x b_x x}{R_0},$$

where G is best interpreted as the "average" time between the birth of parents and progeny. In an iteroparous population, G will only be approximate, whereas in a semelparous population with a fixed life cycle, G will be exact (Krebs 2001). Using the data above, G may be calculated as

$$G = \frac{\sum_{x=0}^{X} l_x b_x x}{R_0} = \frac{2.75996}{1.10585} = 2.49578.$$

Another parameter closely related to G is the mean life expectancy at a birth (e_0), as defined by the equation (13.1). Leslie et al. (1955) derived a variance for e_0 computed as

$$\text{var}(e_0) = \sum_{x=0}^{m-1} W_x = \sum_{x=0}^{m-1} \left[\frac{S_{x+1}^2 q_x}{p_x (n_x - 0.5 a_x)} \right],$$

where l_x, n_x, p_x, q_x, and x are as defined in section 13.3.1; a_x is the number of accidental deaths or removals resulting from experimental or observational handling during the interval x to $x + 1$; $S_x = l_x + l_{x+1} + \ldots + l_{m-1} + 0.5 l_m$; and m is the number of age-groups in the data. Note that in many cases a_x values will equal 0. Once the variance has been computed, confidence limits on e_0 can be obtained as

$$e_0 \pm t_\alpha \sqrt{\text{var}(e_0)},$$

where t_α is the tabular value for Student's t-value with $n_0 - 1$ df. More generally, for any e_x the degrees of freedom are $(n_x - 1)$, where n_x is the number of individuals alive at the start of interval x. Continuing with the sample data given above for the 1954 cohort of brook trout in Hunt Creek, age-specific estimates of the finite rates of mortality (q_x) and survival (p_x) may be computed.

Table Age-specific estimates of the finite rates of mortality (q_x) and survival (p_x) for the 1954 cohort of brook trout.

Age	n_x	l_x	q_x	p_x
0	52,000	1.00000	0.95927	0.04073
1	2,118	0.04073	0.63220	0.36780
2	779	0.01498	0.79589	0.20411
3	159	0.00306	0.91824	0.08176
4	13	0.00025	1.00000	0.00000
5	0	0.00000	0.00000	0.00000

Using the above estimates, the interim values required for computing e_0 and the variance of e_0 may be estimated.

(Box continues)

Box 13.1 *(continued)*

Table Interim values required for computing e_0 and the variance of e_0 for the 1954 cohort of brook trout. The average number of individuals alive in the interval x to $x + 1$ is given by L_x.

Age	L_x	T_x	S_x	W_x
0	27,059.0	29,069.0	1.05902	1.578×10^{-6}
1	1,448.5	2,010.0	0.05902	2.715×10^{-7}
2	469.0	561.5	0.01829	5.484×10^{-8}
3	86.0	92.5	0.00331	4.415×10^{-9}
4	6.5	6.5	0.00025	

Mean life expectancy and associated variance may then be computed as

$$e_0 = \frac{T_0}{n_0} = \frac{29{,}069}{52{,}000} = 0.559,$$

and

$$\begin{aligned}\text{var}(e_0) &= \frac{0.05902^2 \times 0.95927}{0.04073 \times 52{,}000} + \frac{0.01829^2 \times 0.63220}{0.36780 \times 2{,}118} + \ldots \\ &= 1.908 \times 10^{-6}.\end{aligned}$$

The confidence limits on mean life expectancy with $n_x = 52{,}000$ and $t = 1.96$ are

$$\begin{aligned}e_0 \pm t_\alpha \sqrt{\text{var}(e_0)} &= 0.559 \pm 1.96 \sqrt{1.908 \times 10^{-6}} \\ &= 0.556, 0.562\end{aligned}$$

To compare population rates of change per unit of time, the intrinsic rate of population increase (r) is computed using equation (13.2) and an approximate starting value for the iteration that is defined by the parameter estimates for R_0 and G obtained above. Accordingly, the starting value for the iterative solution of the Euler–Lotka equation is

$$r = \frac{\log_e R_0}{G} = \frac{\log_e 1.10585}{2.49577} = 0.04031,$$

which in turn yields the values in the table below.

Table Euler-Lotka values for the 1954 cohort of brook trout based on a starting value of $r = 0.04031$.

Age and summary measure	$l_x b_x$	Euler–Lotka values
0	0.0	0.0
1	0.0	0.0
2	0.64414	0.59425
3	0.37516	0.33243
4	0.08655	0.07366
5	0.0	0.0
Summary measure		$\sum_{x=0}^{\infty} e^{(-rx)} l_x b_x = 1.0034$

Although the *r*-value yields a sum close to 1, the sum may be improved by incrementing *r* to 0.04044.

Table Adjusted Euler-Lotka values for the 1954 cohort of brook trout based on $r = 0.04044$.

Age and summary measure	$l_x b_x$	Euler–Lotka values
0	0.0	0.0
1	0.0	0.0
2	0.64414	0.59409
3	0.37516	0.33230
4	0.08655	0.07362
5	0.0	0.0
Summary measure	$\sum_{x=0}^{\infty} e^{(-rx)} l_x b_x = 1.00001$	

Accordingly, the value $r = 0.04044$ satisfies the Euler–Lotka equation, and the finite rate of population increase (λ) may then be determined as

$$\lambda = e^{0.04044} = 1.04127,$$

where *r* is interpreted as the production of 0.04044 females per female per year, and λ is 4.127 %. These values will be useful for predicting period to period changes in abundance only if the age-specific fecundity and mortality schedules defined for the 1954 cohort remain fixed. Population doubling time (DT) is sometimes computed from *r* as

$$\mathrm{DT} = \frac{\log_e(2)}{r} = \frac{0.6931}{0.04044} = 17.1401.$$

Finally, for any pair of unchanging l_x and b_x schedules there will be a stable age distribution for the population (Lotka, 1922) in which the population percentage in each age-class remains constant whatever the total population abundance. The proportion of the population in any age-class (C_x) can be computed to obtain a stable age distribution for any set of unchanging l_x and b_x schedules as follows once λ is known.

$$C_x = \frac{\lambda^{-x} l_x}{\sum_{x=0}^{\infty} \lambda^{-i} l_i}$$

For example, using the 1954 cohort data, the proportion of age-2 individuals may be computed as

$$C_2 = \frac{\lambda^{-2} l_2}{\sum_{i=0}^{5} \lambda^{-i} l_i} = \frac{1.04127^{-2} \times 0.01498}{1.05585} = 0.01309.$$

(Box continues)

Box 13.1 *(continued)*

Repeating the computations for all possible age-classes yields the following table.

Table Computations for determination of proportion of the population in any age-class (C_x) to obtain a stable age distribution for the 1954 cohort of brook trout.[1]

Age and sum	l_x	$\lambda^{-x}(l_x)$	C_x
0	1.00000	1.00000	0.94710
1	0.04073	0.03912	0.03705
2	0.01498	0.01382	0.01309
3	0.00306	0.00271	0.00257
4	0.00025	0.00021	0.00020
5	0.00000	0.00000	0.00000
Sum		1.05585	1.00000

1. Please check your λ^{-x} values and hence $\lambda^{-x}(l_x)$ values. The values are correct. The problem is the 0.04044 value. This should be quoted to six decimal places to get the results in the table. the simplest thing is to change the value in lines 22 and 25 to 0.040444. then using the new value, you get the values in the table rounded to five decimal places.

Further details of computations for reproductive values at age x and the decomposition of reproductive potential into current and future progeny may be found in Pianka and Parker (1975).

When solving via iteration, values of r are successively varied in the Euler–Lotka equation until the sum on the right-hand side of the equation approaches unity to within some prespecified level of precision. An approximate starting value for the iteration can be obtained using R_0 and G as follows:

$$r = \log_e(R_0) / G. \tag{13.3}$$

If the resulting value for r is greater than 1, population abundance is increasing per unit of time; if r equals 1, population abundance is stable; and if r is less than 1, population abundance is decreasing per unit of time. The intrinsic rate of population increase may be related to the finite rate of population increase ($\lambda = N_{t+1}/N_t$) as follows:

$$\lambda = e^r. \tag{13.4}$$

The difference between the two measures is that r is a rate computed over an immeasurably short (infinitesimal) period of time, whereas the λ is computed over a measurable time period (e.g., days, weeks, or months) more useful for summarizing periodic responses in fish populations. Methods for approximating the variance of λ, which facilitates statistical testing, are given by Lenski and Service (1982) and Rago and Dorazio (1984). The former method is a conservative approximation for which only large differences in λ are likely to prove significant. The penalty for

conservatism is that type II statistical errors become large, thus the more precise methods discussed in Rago and Dorazio (1984) are recommended.

The intrinsic rate of population increase is argued to integrate the age-at-first reproduction, age-specific fecundity and survivorship, brood frequency, and longevity effects of all environmental factors acting on a population, including stress-related factors (Walthall and Stark 1997), and to account for the apparently contradictory effects elicited by compensatory effects (Daniels and Allan 1981). For population level bioassessments, Forbes and Calow (1999) concluded that r was a better summary measure of responses to stressors than the weakly predictive measures of individual level effect endpoints typically obtained from laboratory experimentation (e.g., acute survival and chronic survival). This is because r integrates potentially complex interactions among life history variables and provides a more ecologically relevant measure of population level impacts. As a result, life table analysis can provide important insights into the population level consequences of stressors and can be used to generate testable hypotheses to explain why certain species are dominant in stressed habitats while others disappear (Forbes and Calow 1999).

13.3.1.3 *Examples of Life Table Studies*

McFadden et al. (1967) provide an early example of the use of life tables in fish population analysis. Using 14 years of data tracking numerical changes in brook trout populations in Hunt Creek, Michigan, McFadden and coworkers were able to construct life tables and age-specific fecundity schedules for 11 successive cohorts of brook trout. The study demonstrated the relative importance of variability in mortality over natality for the determination of population abundance, indicating that changes in mortality rates rather than fecundity was the proximate cause of observed fluctuations in population abundance.

Jensen (1971a) subsequently used data from the literature to construct life tables for seven trout populations as a means of studying mechanisms responsible for the re-establishment of the birth–death equilibrium upset by fishing. Although fishing was found to alter population age structure by increasing juvenile abundance, increased fishing did not directly affect age-0 mortality. The analysis suggested fishing mortality could be compensated for by higher age-specific fecundity, with reproductive adjustments being the apparent compensatory mechanism determining the dynamics of brook trout population responses to fishing.

Detailed studies of Eurasian perch in Lake Windermere from 1955 to 1972 were summarized by Craig (1980), who used life table analysis to estimate a time series of R_0 for populations in the north and south basins of the lake. Net reproductive rates were then related to environmental factors in the form of an exponential model that could be estimated by linear regression:

$$\log_e(R_0) = b_0 + b_1 x_1 + b_2 x_2, \tag{13.5}$$

where b_0, b_1, and b_2 were estimated model coefficients, and x_1 and x_2, respectively, were temperature in degree-days above 14°C and biomass of Eurasian perch already

present measured in energy equivalence terms (kJ/m^2), both for the year of hatch. A good fit to available data ($r > 0.95$) was achieved in both basins, suggesting other environmental factors, including anthropogenic stressors, might be used to explain variability in net reproductive rates given sufficient data.

13.3.2 Matrix Models

Where populations have been affected by disturbances such as overfishing or chemical exposure, age-structured models may be a more appropriate means of predicting the consequences of the stressor for future population dynamics. Although age-specific rates of fecundity and mortality are used in the construction of life tables, the rates are summed in such a way that derived population parameters implicitly assume every individual within the population is identical (the average). Thus, individual differences are not allowed to influence population dynamics. For population bioassessment the suppression of age-specific effects on population dynamics is a potential problem. Chronic and acute stressors are known to be size (age) dependent (Adams 2002), and lack of knowledge of age-specific responses can result in misplaced management efforts even when population protection is a priority (e.g., Crouse et al. 1987). One means of overcoming the problem of assuming a population is composed of replicate individuals is to use age-structured models that divide the population into age-classes. Individuals within an age-class are still treated as identical, but age-class averages are allowed to differ.

13.3.2.1 Age-Based Matrix Models

A popular analytical tool that overcomes the problems associated with assuming every individual in the population is identical is the Leslie matrix approach (Leslie 1945). Leslie noted that if a population was closed (no immigration or emigration), the number of individuals alive at time period $t + 1$ would depend directly on the age-specific fecundity, aging, and mortality rates at time t. Formulating the basic information on age-specific vital rates in the form of a matrix equation allowed Leslie to predict future population abundances as a function of what was known about existing vital rates. Assume an initial age distribution for a population is given by the vector $\mathbf{N}_t$ divided into age-classes $n_k(t)$ representing the number of individuals of age k at time t as follows:

$$\mathbf{N}_t = [n_0(t),\ n_1(t),\ \ldots,\ n_k(t)]. \tag{13.6}$$

The initial age distribution can be multiplied by a matrix containing age-specific survival probabilities and fecundities to determine the age distribution of the survivors and descendants in the next time period $(t + 1)$ as follows:

$$\mathbf{N}_{t+1} = \mathbf{M}\mathbf{N}_t. \tag{13.7}$$

The elements of the Leslie matrix, **M,** are the age-specific survival probabilities (P_i) and fecundity rates (b_i) arranged as follows:

$$\mathbf{M} = \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{k-1} & b_k \\ P_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & P_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & P_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & P_{k-1} & 0 \end{bmatrix}. \tag{13.8}$$

It follows that the number of individuals alive in different age-groups at an arbitrary time *(t)* depends on the number initially alive, $\mathbf{N}_0$ as follows (Leslie 1945; Caswell 1989):

$$\mathbf{N}_t = \mathbf{M}^t \mathbf{N}_0. \tag{13.9}$$

The changing size and age distribution of the population is thus calculated by a process of successively multiplying the age distribution by the Leslie matrix, on the assumption that the age-specific survival and fecundity rates remain constant from period to period. As the population size increases, the relative number of individuals in each age-class will vary until a distribution (i.e., stable age distribution) is reached (Manly 1990). Although the number of time periods required to reach a stable age distribution will depend on the initial age distribution, a stable age distribution will be reached regardless of the initial number of individuals in each age-class.

Once the stable age distribution is reached, the proportion of individuals in each age-class will remain constant, and the ratio of population sizes at successive intervals, $(\mathbf{N}_{t+1}/\mathbf{N}_t)$, becomes constant and is known as the finite rate of population increase (λ). The ratio is used to define the proportion by which the population will change in each successive time period and may be related to the intrinsic rate of population increase (r) as $\lambda = e^r$. Values of λ greater than 1 are indicative of an exponentially expanding population; values of λ equal to 1 define an equilibrium population; and values of λ less than 1 are characteristic of populations decreasing to extinction. The finite rate of population increase, therefore, determines the long-term behavior of the population and may be equated to a measure of fitness for the population (Roff 1992). The value is derived mathematically as the dominant eigenvalue of the Leslie matrix (Caswell 1989). Further details on the underpinning mathematical theory, construction, and estimation of matrix models can be found in Caswell (1989) or Manly (1990). A simple working example is presented in Box 13.2.

When building a Leslie matrix model it is important to select a time step (e.g., days, weeks, or years) for which adequate biological data are available. Any time step may be used, but the selected time step must be kept identical for all age-classes and should correspond to biologically meaningful time increments. Where incomplete biological data are available, the closely related stage-based modeling frameworks developed by Lefkovitch (1965) and Usher(1966, 1969) may be used.

One advantage of building any model to assess population level effects is that the model may be used as a surrogate experimental framework. Thus, with a Leslie matrix model it is possible to test the sensitivity of the population growth rate to

Box 13.2 Leslie Matrix Model Analysis

Consider a simple two-age-class population (age-0 and age-1) with age-class abundances as follows:

$N_0 = 10$, and
$N_1 = 5$.

Age-specific fecundities are

$b_0 = 10$, and
$b_1 = 25$,

and age-specific mortality rates expressed as survival probabilities as follows:

$P_0 = 0.5$, and
$P_1 = 0$.

Then the number of new age-0 individuals at $t + 1$ will depend on the age-specific fecundity rates (b_i) and the abundances in each age-class (N_i) as follows:

$$b_0 N_0 + b_1 N_1 = (10 \times 10) + (25 \times 5) = 225.$$

The number of age-1 individuals at $t + 1$ will depend on N_0 and the probability of their surviving to age-1 (P_0) given, for example, $P_0 = 0.5$:

$$P_0 N_0 = 0.5 \times 10 = 5.$$

The total population at $t + 1$ equals the number of new age-0 individuals (225) plus the number of individuals surviving from age 0 to age 1 (5), or 230.

Leslie (1945) noted that the above problem of determining the number of individuals alive in a subsequent time period could be more simply solved by re-casting the problem as a matrix algebra problem where the coefficients b_i and P_i are arranged in what is called a projection matrix as follows:

$$\begin{bmatrix} b_0 & b_1 \\ P_0 & 0 \end{bmatrix} = M.$$

The current age-class abundances are arranged in a column vector (age-structure vector) as follows:

$$\begin{bmatrix} N_0 \\ N_1 \end{bmatrix} = N_t.$$

The solution to the question of how many individuals there are at $t + 1$ can then be found by multiplying M by N_t as follows:

$$N_{t+1} = MN_t,$$

where the multiplication is accomplished through row and column multiplication to define the new vector of age-specific abundances as follows.

Row 1 of $M \times N_t =$ 1st element of N_{t+1}
$= (10 \times 10) + (25 \times 5) = 225.$

Row 2 of $M \times N_t =$ 2nd element of N_{t+1}
$= (0.5 \times 10) + (0 \times 5) = 5.$

In matrix notation the above may be expressed as follows.

$$\begin{bmatrix} b_0 & b_1 \\ P_0 & 0 \end{bmatrix} \times \begin{bmatrix} N_{10} \\ N_{11} \end{bmatrix} = \begin{bmatrix} N_{20} \\ N_{21} \end{bmatrix}, \text{ or}$$

$$\begin{bmatrix} 10 & 25 \\ 0.5 & 0 \end{bmatrix} \times \begin{bmatrix} 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 225 \\ 5 \end{bmatrix}.$$

In general the numbers in each age-group at any arbitrary time t are determined by the numbers in the age-groups at $t = 0$ and the projection matrix (also known as the Leslie matrix) raised to the power t

$$N_t = M^t \times N_0.$$

Consider a more complicated case given as follows:

$$\begin{bmatrix} 0 & 5 & 10 \\ 0.5 & 0 & 0 \\ 0 & 0.2 & 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix} = N_{t+1}.$$

This yields the following on repeated multiplication for the first seven time periods.

$$\begin{bmatrix} 100 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 50 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 250 \\ 0 \\ 10 \end{bmatrix} \Rightarrow \begin{bmatrix} 100 \\ 125 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 625 \\ 50 \\ 25 \end{bmatrix} \Rightarrow \begin{bmatrix} 500 \\ 312.5 \\ 10 \end{bmatrix} \Rightarrow \begin{bmatrix} 1{,}662.5 \\ 250 \\ 62.5 \end{bmatrix}.$$

Population totals of 2,756.25, 5,760, and 8,865.625 are obtained by extending the analysis to 10 time periods. Note that as the population size increases, the relative numbers of individuals in each age-class vary. Extending the analysis to 100 time periods allows computation of the finite rate of population increase (λ) as the average of the ratio of N_{t+1}/N_t or as the dominant eigenvalue of the Leslie matrix following procedures outlines in Caswell (1989). Here $\lambda = 1.752$. In turn, the instantaneous rate of population increase (r) is computed as $\log_e(\lambda) = 0.5609$.

(Box continues)

Box 13.2 *(continued)*

The value of Leslie matrix models is that parameters within the Leslie matrix may be varied to examine the implications of stress or changes in fishing policy for population abundance. For example, assume a pollutant stress reduces survival probabilities for the youngest (age-0) age-class but has no effect on older age-classes. Reduction of age-0 survival to 0.4 yields the following:

$$\begin{bmatrix} 0 & 5 & 10 \\ 0.4 & 0 & 0 \\ 0 & 0.2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix} = N_{t+1}.$$

Population estimates of N_1 through N_{10} are 100, 40, 208, 160, 448, 486.4, 1,024, 1,331.2, 2,437.12, and 3,481.6. Abundance values are plotted below in comparison to the base-case scenario. The associated λ and r values for the population under this scenario are 1.5829 and 0.4592.

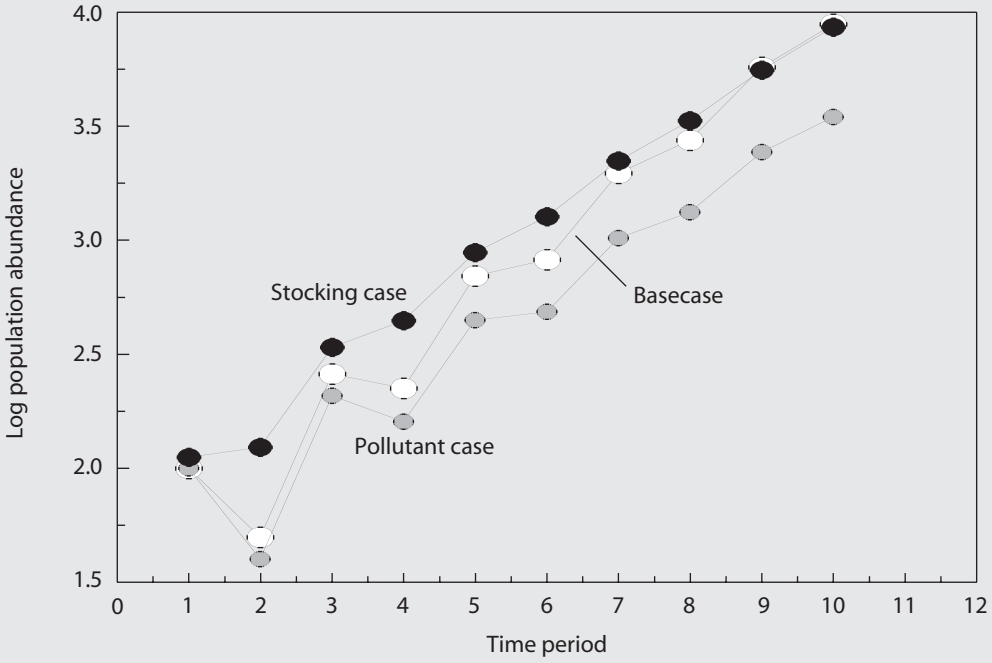


Figure Population abundance of an hypothetical population over time based on a Leslie matrix model.

stressor-induced variations in the fecundity and survivorship values used to construct the model. Model experiments are completed by repeated estimation of the model holding all elements in the projection matrix constant but one and varying the selected element to determine its effect on λ and r ($\log_e\lambda$). The proportional sensitivity of the model to changes in the matrix elements may then be measured by the elasticity of λ (ε_{ij}) as follows:

Remediative strategies may also be considered. For example, stocking may be considered as one means of offsetting age-0 losses by adding 10 age-1 and 2 age-2 individuals to the population at each time step. Note that survival probability and fecundity parameters remain the same. Under this scenario the following abundances (also plotted above) are obtained: 112, 124, 340, 445.6, 887.2, 1,271.2, 2,238.88, 3,360.16, 5,602.72, and 8,619.424.

Alternatively, implementation of catch-and-release regulations for the largest (oldest) individuals may have the effect of raising age-2 survival from 0 to 0.05. Re-computation of the period-to-period abundances and the dominant eigenvalue of the Leslie matrix here yield λ and *r* values, respectively, of 1.5875 and 0.4622. Raising the survival of age-1 individuals to 0.25 through the use of catch-and-release regulations but allowing the oldest individuals to be removed yields λ and *r* values, respectively, of 1.618 and 0.4812. Estimates of *r* resulting from a series of experiments that raise age-1 survival are plotted below in comparison to the base case to demonstrate the way in which comparative Leslie matrix model scenario results may be generated and used.

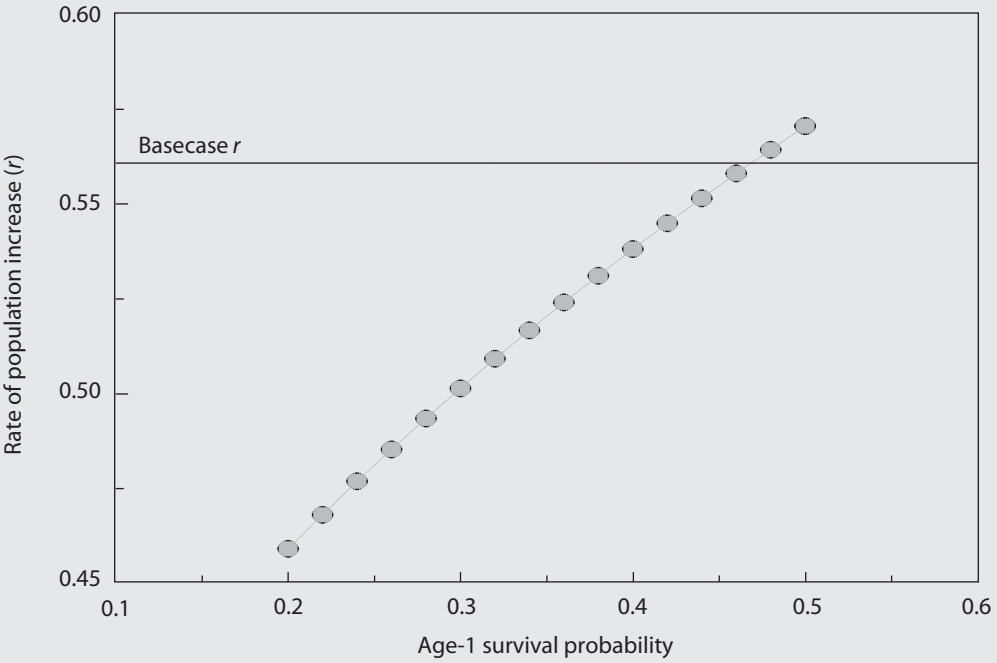


Figure Changes in the intrinsic rate of increase (*r*) with changes in age-1 survival in an hypothetical population.

$$\varepsilon_{ij} = \frac{a_{ij}\,\Delta\lambda}{\lambda\,\Delta a_{ij}}, \tag{13.10}$$

where a_{ij} is the jth element of the ith row in the projection matrix, and Δa_{ij} and $\Delta\lambda$, respectively, define the experimental change in a_{ij} and resultant change in λ. The elasticities with respect to the fecundity and survival probabilities sum to 1

(de Kroon et al. 1986), allowing the relative contributions of the matrix elements to be compared. An excellent example of the application of elasticity measures to the study of management-related issues is given in Crouse et al. (1987).

Most elements in the Leslie projection matrix can be obtained directly from life tables. Where life tables have not been previously constructed, a variety of standard fisheries-related techniques exist for deriving the required matrix elements. Catch-curve analysis may be used to derive survival estimates for age-classes fully recruited to the fishery (Robson and Chapman 1961; Chapter 6), but due attention must be paid to possible selective biases in the catch data (Ricker 1975). Horst (1977) provides an example of the use of this technique in the estimation of Leslie matrix elements for a population of cunner. Horst (1977) regressed the natural logarithm of the number of fish of age (t) collected (N_t) in a sampling program against age (t) as follows:

$$\log_e(N_t) = a + bt. \tag{13.11}$$

The slope coefficient b can be used to estimate age-specific survival as e^{-b}. Extrapolation of the regression-based adult age-specific survival estimates to young of the year, however, is not reasonable owing to the usually high, differential mortalities experienced in this age-class (Vaughan and Salia 1976). Extrapolation to juvenile age-classes may also be unwise. Estimates of age-0 and juvenile survival can be obtained by analogy by use of field data for a closely related species (e.g., Landahl et al. 1997) or by determining the values for age-0 and juvenile survival that yield an equilibrium population size.

Typically, when age-specific mortality and fecundity are known, the Leslie matrix model is used to study population trajectories over time. However, if an equilibrium population is assumed and the age-specific fecundity schedule is known, the mortality schedule can be determined instead. If all mortality rates but the age-0 mortality rate are determined independently of the Leslie model, then the model may be used to deduce the residual age-0 mortality required to yield an equilibrium population size ($\lambda = 1$) as follows (Vaughan and Salia 1976):

$$P_0 = \frac{1}{b_{i+1} + \sum_{i=1}^{K-1}\left[b_i\left(\prod_{j=1}^{i} P_j\right)\right]}, \tag{13.12}$$

where K is the number of age-classes in the population starting from 0, b_i is the average number of eggs produced per individual in the ith age-class, and P_j is the probability of survival in the jth age-class. Boreman (1997) details a conceptually similar approach based on field data whereby the survival rate from egg to spawner (S_T) is expressed as

$$S_T = \frac{N_S}{N_E}, \tag{13.13}$$

where N_E and N_S, respectively, are the total number of spawned eggs and spawners in a given year. Assuming the population is in equilibrium implies S_T will remain constant and can be partitioned into age-specific survival rates as follows:

$$S_T = S_0 \cdot S_1 \cdot S_2 \cdot \ldots \cdot S_N. \tag{13.14}$$

The age-0 survival rate, for which data are typically scarce, can then be estimated by combining equations (13.13) and (13.14) and re-arranging to obtain

$$S_0 = \frac{N_S/N_E}{S_0 \cdot S_2 \cdot \ldots \cdot S_N}, \tag{13.15}$$

Other methods for estimating the elements of the Leslie matrix model when incomplete population data exist for juvenile age-classes are discussed by Aalto and Newsome (1980), and results of computer simulations to test proposed methods are described. Although examples pertaining to lacustrine yellow perch populations are given, the mathematical complexity of the suggested approaches precludes their widespread use.

When cohort data are available, or populations in stable environments are being modeled, the instantaneous per capita rate of mortality may be used to estimate survival probabilities for the unit time step of the model (e.g., days, weeks, or years) as follows:

$$Z = [\log_e(N_{t1}) - \log_e(N_{t2})], \tag{13.16}$$

where N_{t1} and N_{t2} are the abundances of the cohort at times t_1 and t_2 or the abundances of two consecutive age-classes under the assumption the population being studied is in a stable environment. Survivorship is then defined as

$$S = e^{-Z}. \tag{13.17}$$

The use of Z implies that the decline in the abundance of a cohort over the unit time step is exponential, although the true pattern of mortality within the time step is not known. Conceptually the method of estimating survivorship is akin to that used by Horst (1977). Here, however, survivorship is allowed to vary between cohorts, and the method described above allows specific recognition of variable, age-dependent rates of mortality.

Age-specific fecundity values are best estimated directly for the species of concern from field data by means of a gravimetric (Bagenal and Braum 1978) or other suitable techniques. Fecundity, however, may be inferred in a two- or three-step iterative process based on previously estimated relationships. First a length–age relation (e.g., von Bertalanffy) is obtained. Then, a fecundity–length relationship is obtained and converted to a fecundity–age relationship with the use of the age–length relationship. Alternatively, a weight–length relationship is obtained

from which an estimated fecundity–weight relationship may be converted to a fecundity–length relationship before being converted to a fecundity–age relationship by means of the age–length relationship. Conversions from statistically robust (reasonable r^2 and regression coefficients with P-values < 0.05) average age, length, and weight relationships are permissible within the context of the Leslie matrix model because it ignores individual variability and works with age-specific averages.

Important assumptions involved in the application of Leslie matrix models to fish population bioassessment problems include (1) life history parameters (P_i and b_i) are independent of time and population density; (2) parameters are constant within age-classes; (3) spawning occurs over a short duration of time at approximately the same date each year; and (4) there is no appreciable net migration into or out of the population under study. Although the assumptions limit the applicability of the modeling framework, in many instances the detailed data required to build more complex models are not available (e.g., Crouse et al. 1987). The Leslie matrix approach does allow the effects of stress to act separately and differentially on critical population processes. Mortality and reproductive effects can be directly represented via changes in the age-specific schedules. Effects on growth can be represented indirectly via downward adjustments in the fecundity schedule or via an increment in the mortality schedule to reflect size-dependent survival. As a result of this flexibility there are numerous examples of the application of Leslie matrix models to the study of population level stress.

13.3.2.2 Examples of Leslie Matrix Model-Based Studies

In an early example of the application of Leslie matrix models to the study of possible stressor impacts on fish populations Jensen (1971b) examined the effects of increasing juvenile mortality (age-0) by 0, 5, 15, 25, 50, and 95% by means of a Leslie model for yield fitted to data on the brook trout population for Hunt Creek, Michigan (McFadden et al. 1967). The analysis showed that small increases (5%) in age-0 mortality decreased the yield of the brook trout fishery, and with 50% additional mortality the population became extinct. The analysis further suggested that significant lags between the initial action of a stressor and the appearance of the effect in yield measures used to monitor population status could be especially dangerous for sustainable fishery management.

Concerns over the possible effects of density-dependent adjustments to survival rates caused later modelers to adapt the Leslie matrix framework to include density-dependent feedbacks explicitly. For example, Vaughan (1981) included density dependence when attempting to determine the effects of juvenile mortality on the fate of yellow perch populations subjected to power plant entrainment in Lake Michigan. Similarly, DeAngelis et al. (1980) used a compensatory model that incorporated density dependence, in the form of a young-of-the-year survival term, to predict the stability and equilibrium return times for fish populations (striped bass, winter flounder, white perch, Atlantic menhaden, and cunner) following perturbations. The analysis demonstrated that the dominant eigenvalue of

the linearized Leslie matrix (e.g., λ) could be used to approximate the equilibrium return time (T_R) of the population following a perturbation as $T_R = \log_e 0.05 / \log_e \lambda$. Comparison of the return times for modeled fish populations did not correlate strongly with any single parameter of the Leslie matrix model. In general, however, T_R increased with increases in survival rates (P_i) and the number of age-classes in the population (DeAngelis et al. 1980). Therefore, the greater the relative contribution of the older age-classes to reproduction, the longer the time it takes to return to predisturbance population values. Schaaf et al. (1987) also examined return times from hypothesized pollutant impacts by using Leslie matrix models of multiple stocks of eight species of estuarine-dependent fish populations. Pollutant impacts were introduced through changes in first-year survival, and return to predisturbance abundances were tracked. Without compensation, stocks responded to a single 50% reduction in first-year survival by taking an average of 10 years to equilibrate to 88% of their predisturbance abundances.

Typical of pollutant-related impact studies is the work of Landahl et al. (1997), where Leslie matrix modeling methods were used to address the question of whether documented contaminant-related reductions in reproduction and survival rates in fish populations in Puget Sound were sufficient to affect overall fish abundance. A model for English sole was constructed using recent historical data for the investigation of contaminant effects. Laboratory testing data on the effects of contaminants on reproduction, including impaired gonadal development, reduced spawning ability, and decreased egg and larval viability, were then incorporated into the fecundity component of the model. Results suggested that declines in the fecundity component of the model equivalent to those observed in field studies were sufficient to decrease the population growth rate *(r)* substantially if density-dependent effects were weak or moderate. A compensation for loss of recruits due to contaminant effects was observed when strong density-dependent population regulation was assumed. Extensions of the simple pollutant–population models described by Johnson et al. (1998) to include specific consideration of spatial and local population grouping responses to a stressor are given in Chaumot et al. (2002, 2003) and demonstrate the level of complexity to which Leslie matrix models may be taken to mimic the complexity of actual environments.

In addition to pollutant impacts, Leslie matrix models have been used to investigate the effects of exploitation (Hayes et al. 1995) and to understand the role of compensatory mechanisms in the population dynamics of lake trout under varying stressor regimes (Ferreri and Taylor 1996). Hayes et al. (1995) used literature-derived data to model the population dynamics of largemouth bass and walleye subject to competitive fishing on the joint attainment of the management objectives of maintaining population abundance and population size structure. Although results of the analysis are case specific, the analysis does demonstrate methods for the development of an analytical framework useful for assessing the sensitivity of population abundance and size structure to management action.

Ferreri and Taylor (1996) used Leslie matrix methods to explore the role of compensatory mechanisms in the population dynamics of lake trout in lakes Michigan and Superior in the pre–sea lamprey period (prior to 1950), the post–sea lamprey, high-exploitation period (1951–1961), and a current post–sea lamprey, moderate-exploitation period (1985–1993). Comparisons were made on the basis of the finite rate of population increase (λ) as computed from Leslie matrix models incorporating stressor-regime-specific mortalities. To estimate compensatory potential, regime-specific mortality factors (other than natural mortality) were set at 0, and λ was re-computed. The compensatory scope of the population was then defined as the difference between the regime-specific maximum for λ and λ equal to 1 (a stable population). Individual growth rates (size) and age-specific fecundity rates changed in response to the different levels of lake trout abundance during each of the stressor periods. Lake trout during the sea lamprey dominant period, which experienced the lowest abundance and highest mortality levels, exhibited the fastest individual growth rates and the highest age-specific fecundities. The high rates contributed to the larger compensatory scope exhibited by lake trout during the sea lamprey dominant period (1951–1961) as compared with the pre–sea lamprey or current periods. As Ferreri and Taylor (1996) demonstrate, Leslie matrix evaluations of compensatory fish population responses to varying mortality sources can aid fisheries managers in assessing both population level consequences of stressor action and the management of productive fisheries.

13.2.2.3 Extensions to Stage-Based Models

Leslie matrix models are not appropriate for modeling all fish populations. Species that show developmental plasticity, for which age is a poor descriptor of demographic attributes (e.g., reproduction), are not suited for Leslie-based analysis. For example, when maturation schedules are weakly correlated with age and depend on individual size and growth controlled through environmental conditions, the Leslie model will yield inaccurate predictions of population growth rates. Lefkovitch (1965), however, demonstrated that the Leslie matrix model was a special case of a more general class of matrix models:

$$\mathbf{N}_{t+1} = \mathbf{A}\mathbf{N}_t, \tag{13.18}$$

where the elements (a_{ij}) of the projection matrix (**A**) are termed transition coefficients and describe how the abundance of one stage is related to the abundance of other stages. An element of the transition matrix can represent average stage-specific fecundity or the average propensity of individuals to move from one stage to the next or remain in a given stage from one time step to the next. Within the Lefkovitch matrix framework, there is no necessary relationship between age and stage. Like the Leslie matrix, the Lefkovitch assumes a constant time step, but the duration of time an individual spends in each stage-class is not necessarily the same as the time step.

In population level studies, the Lefkovitch matrix typically takes on the following form:

$$\mathbf{A} = \begin{bmatrix} P_0 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ G_1 & P_1 & 0 & \cdots & 0 & 0 \\ 0 & G_2 & P_2 & \cdots & 0 & 0 \\ 0 & 0 & G_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & G_{k-1} & P_k \end{bmatrix}, \tag{13.19}$$

where b_i, P_i, and G_i, respectively, define stage-specific fecundity, the probability of surviving and remaining in the same stage, and the probability of surviving and growing into the next stage.

The technique of mutliplying the projection matrix **A** by the population vector $\mathbf{N}_t$ described above for the Leslie matrix modeling framework is similarly used in stage-based modeling to forecast future population states (Caswell 1989). Furthermore, the dominant eigenvalue of the projection matrix yields an estimate of the finite rate of population increase λ $(= e^r)$, just as in the Leslie matrix case.

While stage-based models overcome the conceptual difficulties associated with modeling developmentally plastic species, they are not without some their own unique drawbacks. The main problem is the large number of coefficients that must be estimated to construct the model appropriately. A second is ensuring that chosen stages make clear biological sense and are not arbitrarily imposed upon the model for reasons of data convenience. There have been several comparative assessments of age- and stage-based models published in the literature, though none specific to fish populations. In general, comparative studies have favored stage-based over age-based models for organisms whose growth and reproductive responses are strongly driven by environmental conditions (Werner and Caswell 1977).

13.2.2.4 *Summary Appraisal of the Matrix-Based Approach*

The age structure detail in Leslie matrix models allows a population to be described in terms of its long-term abundance, intrinsic rate of natural increase, reproductive potential, resilience, extinction risks, or some combination of these factors (Landahl et al. 1997). Application of Leslie models to the study of stressor effects, however, requires that accurate age-specific survival and reproductive rates under various stressor scenarios are available. Generally, these rates have been derived based on limited laboratory tests and longitudinal studies of species ecology (Barnthouse et al. 1990). Questions about the validity of extrapolations from laboratory tests to field conditions exist (Power and McCarty 1997) and recommend against the use of non-field-validated, laboratory-estimated stressor response data when constructing Leslie matrix models for population bioassessment studies.

In constructing models, it is also important to remember that stressor-related impacts on survival and reproduction will be influenced by the extent to which stressor effects are mitigated by population regulating mechanisms such as emigration, immigration, or density dependence (Power 1997). Where migrations are likely to increment or decrement population numbers, age-specific mortality factors must be appropriately adjusted to reflect such effects. In addition, when recruitment processes are known to conform to standard stock–recruitment relationships (e.g.,

Beverton–Holt or Ricker; Chapter 4), due consideration of these relationships in model construction should also be given. Accordingly, Leslie matrix models without appropriate density-dependent adjustment are viewed in the ecological literature as only first approximations of how population levels may vary under a given set of conditions, appropriate only for use in studying expanding populations that have yet to approach carrying capacity (Mendelssohn 1976).

Critical to the Leslie matrix approach is the assumption of stability in the population age distribution over time. This is often unrealistic for populations subjected to random anthropogenic or environmental disturbances, which therefore display high variability in year-class strength. Furthermore, a Leslie matrix model will be inappropriate when critical survival and reproductive values do not depend on age but on physiological development that is independent of age. As noted previously, Leslie matrix models have also been criticized for failing to include considerations of density dependence or temporal changes in critical survival parameters. As a result, matrix survival coefficient constants are now often viewed as random variables or made explicitly density dependent such that the value the coefficient takes on in any one time period is functionally dependent on abundance in the period (Goodyear 1985; Manly 1990). An explicit example for fish populations of the incorporation of density-dependent parameters in the construction of matrix models is given in Van Winkle et al. (1978).

Drawing survival coefficients randomly from a distribution with a defined mean and variance further allows the Leslie matrix model to mimic the impact of varying environments on populations. Although density dependence and temporal parameter changes may be accounted for in the Leslie framework, the adaptations come at the cost of increasing model complexity. Introduction of varying parameter values limits the interpretative value of the intrinsic rate of population increase because repeated changes in parameter values imply the stable age distribution on which the parameter depends mathematically will never be reached. Accordingly, a detailed description of the population growth rate parameter distribution in terms of its mean, variance, and empirical density function are necessary to interpret probable population level effects when models are largely stochastic. Regardless of improvements, Leslie models still focus on abundance and do not produce information on other population parameters routinely used by biologists to assess the effects of environmental perturbations on populations, including rates of growth, length-frequency distributions, condition factors, and age-at-maturity (Power 1997).

13.2.3 Stock-Assessment-Based Methods

To conduct a formal stock assessment it is necessary to understand the dynamics of the population being assessed. A key objective of stock assessment procedures is the description of how the stock or population of interest has responded to fishing pressure in the past. With knowledge of the relationship between stock densities and fishing effort, it is possible to assess stock productivity and predict future sustainable stock harvesting levels (Haddon 2001). Historical interests in determining equilibrium harvest values for exploited stocks explain the traditional

dichotomous representation of total mortality (Z) in fish populations as the sum of fishing mortality (F) and natural mortality (M) (see, for example, Ricker 1975 and Chapter 6). Natural mortality is rarely estimated directly because of the difficulties associated with gathering and interpreting appropriate field data (Shepard 1988), but it may be estimated indirectly as the difference between total and fishing mortality. Conceptually, the difference between fishing-induced mortality and mortality imposed by other forms of anthropogenic stress (e.g., pollution) are often small. Both result in the permanent removal of individuals from the population. Accordingly, it is possible to adapt traditional fisheries management approaches to the problem of population bioassessment, including (1) surplus production models, (2) yield or yield-per-recruit models, and (3) stock–recruitment models. Brief descriptions of each modeling approach, the data required to complete a population level bioassessment, and the ways in which anthropogenic stressors other than fishing may be incorporated into the modeling frameworks are given below. For further theoretical details on mathematical form, estimation, and examples of applications to the study of exploitation-related issues, readers are referred to the detailed descriptions given in previous chapters of this book, Hilborn and Walters (1992), and Haddon (2001), the latter being particularly good at providing easily understood examples. A final category of aggregate mortality impact-based assessment methods developed over the years for assessing entrainment mortality impacts is also discussed in this section.

13.2.3.1 *Surplus Production Models*

Surplus production models aim at describing the dynamics of exploited populations with a minimum of biological information. The models assess the production from a stock above that required to replace losses due to natural mortality by means of the application of a logistics-based response curve. Surplus production models are denominated in biomass and require time series data on total harvest biomass and total fishing effort in standardized units per unit of time. Catch per unit of standardized effort is assumed to be proportional to fishing mortality. In the absence of fishing mortality, the changes in population biomass are a nonlinear function of existing population biomass as follows (Hilborn and Walters 1992):

$$B_{t+1} = B_t + \frac{r}{p} B_t \left[1 - \left(\frac{B_t}{K} \right)^P \right] - C_t, \tag{13.20}$$

where B_t is population biomass in period t, r is the intrinsic rate of population increase, K is carrying capacity, p is an assymetry term, and C_t is the total catch in period t. Catch is estimated recursively by estimating catch per unit effort (C/f_t) as qB_{t-1}, where q is the technical coefficient embodying fishing technology that describes the proportion of the stock taken with each unit fishing effort (E_t). Then C/f_t is substituted into the standard catch per unit effort relation to yield $C_t = C/f_t \times E_t$. Further details on fitting surplus production models to data are given in Chapter 8, Hilborn and Walters (1992), and Schnute and Richards (2002).

The effect of stress on a population can be analyzed using surplus production models through variations in q, E_t, K, or r. Stress-related mortality losses may be

treated as equivalent to biomass removals and analyzed through incrementing E_t or q. Declines in habitat suitability resulting from contaminant or physical disturbance (e.g., macrophyte removal) effects may be modeled through reductions in K, population carrying capacity. Changes in individual growth or reproduction are conceptually more difficult to assess. Stress ultimately affects one or more of the three basic processes governing fluctuations in abundance: mortality, somatic growth, and reproduction. If stress is chronic (nonlethal), individual growth-mediated fecundity or direct fecundity responses are likely to be induced in the affected population (Adams 2002). In either case, the population growth parameter (r), which equals the difference in the instantaneous rates of natality and mortality (Krebs 2001), will change. If only one of the underlying natality or mortality processes is affected, the parameter r may be altered using field-derived measures of changes in fecundity or mortality. If both natality and mortality processes are affected, changes in r must reflect the net effect of opposing changes on r, and a detailed life table assessment is likely to be required before surplus production analysis can be used to infer population level changes. Example computations using the above approach are given in Box 13.3.

Surplus production models have the advantage of requiring data routinely collected by monitoring agencies (e.g., catch and fishing effort time series) and a well-developed application and interpretation literature exists (e.g., Chapter 8). Unfortunately, surplus production models make a number of strong assumptions. For example, it is assumed that the equilibrium population and stable age structure are attained instantaneously for every level of fishing effort. Catch per unit effort is assumed to be proportional to population biomass and determined by the catachability coefficient (q). Net migration is also assumed to be zero. Haddon (2001) notes that equilibrium-based surplus production models should be avoided for detailed fisheries assessments, a view countered by Schnute and Richards (2002) who argue surplus production models are nevertheless valuable tools for understanding and communicating the inevitable limits of biomass loss for a population. Consideration of the advantages and disadvantages of surplus production models when coupled with the general availability of software packages for model estimation (Chapter 8) suggests surplus production models are probably best used as a screening tool for assessing a large number of fish populations that may be at risk from anthropogenic disturbances or as a ranking tool for assessing preliminary impact scenarios for a population known to be at risk. In either instance, additional research can then be directed toward the selection and implementation of more complex assessment methods tailored to the specifics of the case or cases of most concern.

13.2.3.2 *Yield or Yield-per-Recruit Models*

Yield-per-recruit models improve on surplus production models with a more explicit representation of the processes governing individual growth (Beverton and Holt 1957). The general form of the yield-per recruit models is (Gulland 1988)

$$\frac{dY_T}{dt} = FN_T W_T, \tag{13.21}$$

Box 13.3 Surplus Production Analysis

Consider the following data for a commercial alewife fishery used in Chapter 8 to illustrate the application of surplus production models.

Table Catch and effort data for alewife fishery described in Box 8.7.

Year	Catch (kg)
1986	90,000
1987	113,000
1988	155,860
1989	181,128
1990	198,584

The parameters initial biomass (B_0), carrying capacity (K), catchability (q), and the intrinsic rate of growth (r) required to solve the model for biomass in period t are given as $B_0 = 732{,}506$; $K = 1{,}160{,}771$; $q = 0.0001484$; and $r = 0.4049$ as determined in the example in Box 8.7. The resulting set of computations form the base case scenario for determining the possible impacts of reductions in the intrinsic rate of population increase (2% decline in r to 0.396773), with the latter scenario requiring re-computation of the predicted biomass series by means of equation (13.19). To determine further the offsetting reductions in catch required to compensate for changes in biomass caused by reducing r, a third set of calculations is performed using equation (13.19) to predict biomass. Example computations are given below for 1.1% and 1.6% reductions in the catch rate.

Table Predicted changes in biomass of alewife fishery given changes in intrinsic rate of population increase (r) and reductions in catch rate. The first row represents the initial biomass.

Year	Catch (kg)	Base case biomass	Biomass with 2% r decline	Biomass with 2% r decline, 1.6% catch decline	Biomass with 2% r decline, 1.1% catch decline
		732,506	732,506	732,506	732,506
1985	90,000	751,933	749,745	751,185	750,735
1986	113,300	745,867	741,789	744,874	743,910
1987	155,860	697,954	692,172	697,408	695,773
1988	181,128	629,503	621,922	629,646	627,235
1990	198,584	547,577	537,897	548,559	545,234

The 1.1% reduction in catch is insufficient to offset the biomass reductions caused by a decline in the intrinsic rate of populations increase, whereas a 1.6% reduction in catch eventually more than compensates for the initial biomass reductions. Note, however, that scenario predictions are increasingly likely to be in error as a result of the strong equilibrium and stable age structure assumptions made by surplus production models. Due caution in making long-term predictions of possible stressor effects, therefore, is required.

where Y_t is the yield to the fishery in period t, F is the instantaneous rate of fishing mortality, and N_t and W_t are functions describing population numbers and average individual weight at age t. Population abundance is typically assumed to decline exponentially with age as

$$N_t = Re^{-Z(t-t_c)}, \tag{13.22}$$

where t_c is the age-class at which fish are recruited to the fishery, R equals $N(t_c)$ and is the number of recruits in the exploited age-classes, and Z is the instantaneous rate of total mortality. The instantaneous rate of total mortality may be decomposed into F, the instantaneous rate of fishing mortality, and M, the instantaneous rate of natural mortality (Ricker 1975). Average individual weight at age t (W_t) is usually defined using a von Bertalanffy equation (Beverton and Holt 1957) as follows:

$$W_t = W_\infty(1 - e^{-K(t-t_0)^3}), \tag{13.23}$$

where W_∞ is asymptotic weight, K is a parameter governing the rate at which maximal weight is approached, and t_0 is the theoretical age at which weight equals 0. By substituting equations (13.22) and (13.23) into equation (13.21) and integrating over all age-classes, the following equation is obtained:

$$Y = FRW_\infty \sum_{n=0}^{3} \frac{b_n e^{-nK(t_c - t_0)}}{Z + nK}, \tag{13.24}$$

where b_n takes on the values of 1, –3, 3, –1, respectively, for $n = 0, 1, 2, 3$. Further details of the model's mathematical details are given in Ricker (1975), Hilborn and Walters (1992), and Haddon (2001).

This model is closed, assuming no net migration. The model further assumes all sources of mortality remain constant over the life of the fish and are independent of population density. Estimates of mortality, von Bertalanffy growth parameters, recruitment to the fishery, and maximum age may be obtained using the techniques discussed in Chapters 4 through 10.

The effects of stress may be incorporated by varying the parameters used to estimate the model. Acute (lethal) stressors may be represented by variations in the total mortality rate or as additions to the sum of natural and fishing mortality. Age-specific changes in mortality can be included by altering t_c, the age-class at which fish are recruited to the fishery. Chronic effects leading to changes in individual growth rates can easily be incorporated by varying K in the underlying von Bertalanffy weight equation. Representation of reproductive effects, however, is difficult and must be dealt with indirectly by decreasing recruitment. The approach has the disadvantage of having to include the effects of mortality prior to recruitment in the adjustment made to R and the addition of a complex secondary assessment to partition the actual decline in R due to reductions in fecundity and the effects of natural morality.

Results of scenarios incorporating numerous changes to model parameters may be plotted on isopleth diagrams defining the contours of equivalent yields resulting from varying combinations of selected model parameters. Isopleth diagrams are useful for understanding the joint response of yield to possible multiple stressor action (e.g., declines in both growth as presented by K and mortality as represented by Z) characteristic of cumulative impacts on fish populations (e.g., Ricker 1975). Although yield-per-recruit models provide a realistic representation of events affecting fish during the life cycle of a particular cohort, they are clearly more problematic when used to describe events acting on a succession of generations (Gulland 1988). This weakness undoubtedly restricts the use of yield-per-recruit models in population bioassessment to cases in which the consequences of single perturbations (e.g., accidental fish kills) for fish populations are of concern.

13.2.3.3 *Stock–Recruitment Models*

Stock–recruitment (SR) models describe the average relationship between the abundance of mature individuals in a population and the number of progeny recruited to the mature population in the next generation. Where detailed life history data are available, SR models may be used to describe the relationship between density or abundance at any two life stages. General mathematical expressions developed to describe available observational data for fish populations were proposed by Ricker (1954) and Beverton and Holt (1957), as follows.

$$\text{Ricker:} \quad R = \alpha S e^{-\beta S}; \text{ and} \tag{13.25}$$

$$\text{Beverton and Holt:} \quad R = \frac{S}{a + bS}, \tag{13.26}$$

where R is the number of recruits, S the size of the parent (spawning) stock, and α, β, a, and b are parameters estimated from the stock–recruitment data. Further detailed discussion of the derivation and estimation of stock–recruitment models is given in Chapter 4.

Stock–recruitment models represent changes in population numbers brought about by both density-independent and density-dependent population regulating factors (Ricker 1973). Accordingly SR models may be used to represent the effects of a wide variety stressors (e.g., acute or chronic) on population numbers. Density-independent causes of mortality (Z_i) generally include all abiotic causes for which mortality removes a set proportion of the population at all densities (e.g., physiological stressors such as temperature). Density-dependent (compensatory) mortality (Z_c) includes all biotic mortality factors for which mortality is positively correlated with population density (e.g., predation and disease). Mortality rates at the origin estimate mortality from density-independent causes. As density-independent mortality does not change with increasing population density, total mortality may be estimated simply as the sum of all mortalities at a given density. Instantaneous rate of compensatory mortality may be deduced by subtracting

density-independent mortality from the recruitment rate at the given population density (Ricker 1973) to obtain for the Ricker curve,

$$Z_c = \beta S, \tag{13.27}$$

and for the Beverton–Holt curve,

$$Z_c = \log_e(1 + bS/a). \tag{13.28}$$

If the average number of eggs laid per fish (E) is known (Ricker 1973), estimates of density-independent mortality can be computed for the Ricker curve as

$$Z_i = \log_e E - \log_e \alpha, \tag{13.29}$$

and for the Beverton–Holt curve as

$$Z_i = \log_e E - \log_e a. \tag{13.30}$$

The effects of changes in Z_i on population recruitment may be determined as discussed in Box 13.4. As with other modeling-based methods, the analysis assumes a stable age structure from which an average fecundity value can be derived for a closed population (Vaughan et al. 1984). In addition, it is assumed that all population-dependent mortality occurs prior to the selected recruitment age and that only stock size affects survivorship.

The separability of mortality into density-dependent and density-independent factors allows stress to be incorporated separately into population level bioassessments in numerous ways. The effects of mortality acting on a population prior to the selected age of recruitment can be modeled by increasing β in the Ricker curve or by incrementing b, or decrementing a, in the Beverton–Holt curve. The effects of declines in reproduction can be incorporated through reductions in E in the expression for Z_i, and the effects of multiple sources of stress can be modeled by simultaneous changes in either Z_c or Z_i.

A second approach to using SR models to assess possible population level impacts (Savidge et al. 1988) makes no attempt to separate mortality into its density-independent and density-dependent components. By use of models denominated in numbers of spawners such that parents and recruits represent discrete generations, replacement stock values are determined (e.g., Ricker 1975), and the effect of incremental mortality is introduced through proportional adjustment of Ricker or Beverton–Holt model parameters. Two cases are considered: the case in which incremental mortality occurs prior to the period of generational compensatory adjustment and the case in which incremental mortality occurs after the period of generational compensatory adjustment. Expressions for the percentage change in the equilibrium population size (PC) of a population exposed to a stressor *(m)* incremental to the stresses of natural variation in physicochemical environmental factors may be defined in terms of SR model parameters, respectively, for the Ricker and Beverton–Holt models as shown below (Savidge et al. 1988).

In case 1, mortality acts prior to the period of generational compensation.

$$\text{Ricker: PC} = 100\left\{\left[\frac{\log_e[\alpha(1-m)]}{(1-m)\log_e(\alpha)}\right]-1\right\}, \text{ and} \tag{13.31}$$

$$\text{Beverton–Holt: PC} = 100\left[\frac{-am}{(1-m)(1-\alpha)}\right]. \tag{13.32}$$

Box 13.4 Stock–Recruitment Analysis

Table Hypothetical data for a longitudinal study of a stream-dwelling population of brown trout.

Year-class	Spawning stock (*S*)	Recruits (*R*)
1985	1,034	1,443
1986	505	1,705
1987	390	1,680
1988	574	1,918
1989	1,032	1,283
1990	642	1,830
1991	803	1,718
1992	1,768	550
1993	1,630	758
1994	941	1,475
1995	1,400	1,005
1996	460	1,788
1997	230	1,393
1998	160	1,193
1999	1,700	743
2000	1,768	650
2001	551	1,770

The Ricker stock–recruitment (SR) curve may be estimated from the data by means of the linear transformation $\log_e(R/S) = \log_e\alpha - \beta S$ to obtain

$$\log_e(R/S) = 2.2173 - 0.00185S,$$

for which r^2 is 0.997 and regression coefficient *P*-values are less than 0.05. Transforming the intercept of the estimated curve to $\alpha = e^{\log_e\alpha} = 9.1823$ allows the Ricker SR model to be written in its more familiar form as

$$R = 9.1823Se^{-0.00185S}.$$

A plot of the data used to estimate the model and the estimated model is given in the plot below with the equilibrium replacement value for the population ($\log_e\alpha/\beta$) marked as point A on the one-for-one replacement line for the population.

(Box continues)

Box 13.4 *(continued)*

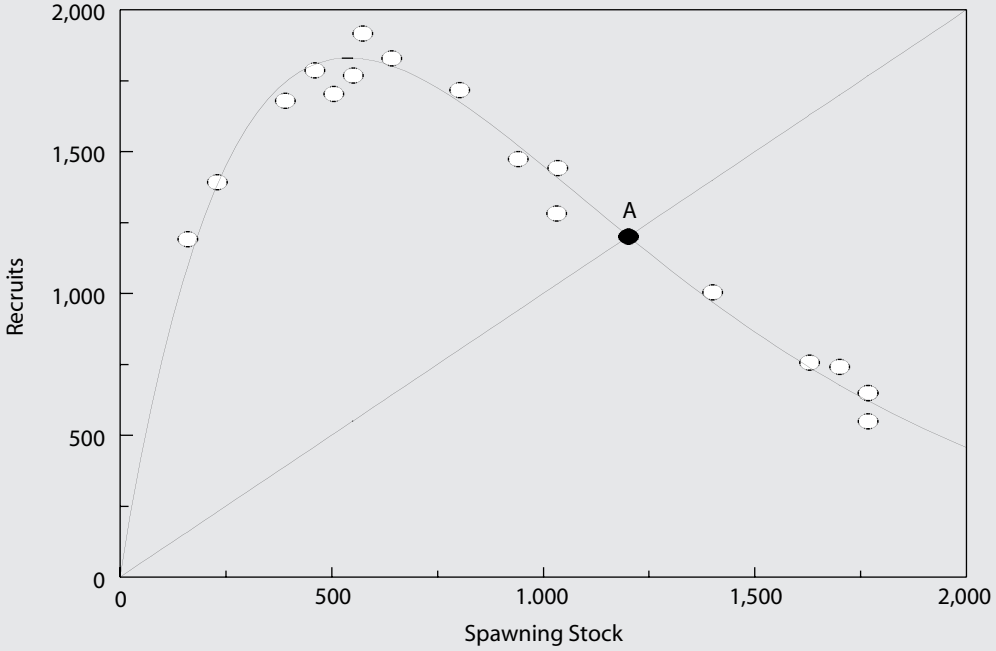


Figure Plot of data used to estimate the Ricker SR model and the estimated model with the equilibrium replacement value for the population marked as point A (at intersection with one-for-one replacement line).

As an example of how SR models might be used to study the possible effects of hydroelectric generation activities on fish populations, assume that average adult fecundity (E) equals 1,000. Then the population-independent mortality term may be defined from equation (13.28) as

$$Z_i = \log_e E - \log_e \alpha = \log_e(1{,}000) - 2.2173 = 4.6905 .$$

Now suppose that the utility wishes to predict the consequences of entrainment mortality on the population. Entrainment mortality affects only the population-independent source of mortality, Z_i. If estimates show entrainment mortality to be 10%, then Z_i would increase from 4.6908 to 5.1599. The resulting change in Z_i implies, through re-arrangement of equation (13.26), a decrease in the SR model α parameter as follows:

$$\begin{aligned} \log_e \alpha &= \log_e E - Z_i \\ \alpha &= E e^{-Zi} \\ &= 1{,}000 e^{-5.1599} \\ &= 5.7423 . \end{aligned}$$

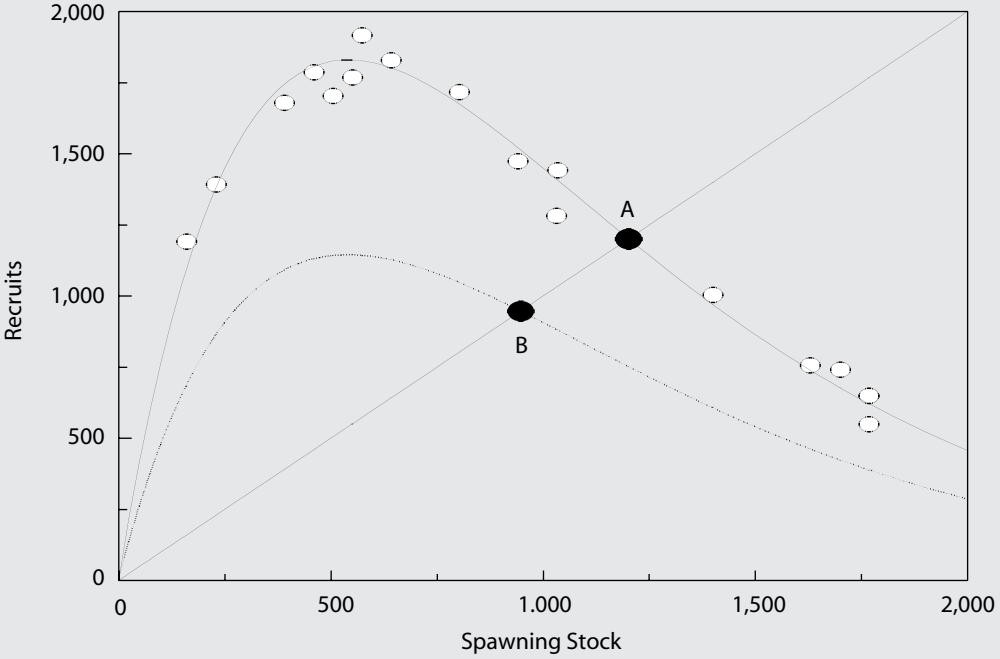


Figure Plot of data used to estimate the Ricker SR model and the estimated model with Z_i of 4.6905 (solid) and an entrainment mortality of 10% ($Z_i = 5.1599$, dashed line). Equilibrium replacement values for the population marked as points A and B (at intersection with one-for-one replacement line).

If there were no change in the population-dependent mortality factor, the resulting Ricker curve would be

$$R = 5.7423Se^{-0.00185S},$$

as represented by the dotted line in the figure above.

Inspection of the two recruitment curves indicates that recruitment has fallen for every level of the spawning stock as a result of entrainment-induced mortality. Furthermore, the equilibrium size of the population, defined by the point at which each curve cuts the 45° line, or the expression ($\log_e \alpha / \beta$), has fallen by some 21.17% from 1,198.5 to 944.8, indicating the impact of entrainment mortality is more than proportional to the increase in population-independent mortality. The result depends on the nonlinear nature of the SR model and highlights the dangers of assuming proportional population effects for even small impacts. Although the varying effects of changes in density-independent or density-dependent mortality can be incorporated into SR models, there are many factors that recruitment models cannot take into account. Effects of changing environmental factors, the potential effects of changes in population size structure, and the differential effects of a human action on varying age- or size-classes cannot be readily considered by such models. The SR relationship, therefore, cannot adequately describe the internal changes within a population that may be important early indicators of the adverse effects of human action.

In case 2, mortality acts after the period of generational compensation.

$$\text{Ricker: } PC = 100\left[\frac{\log_e(1-m)}{\log_e(\alpha)}\right], \text{ and} \tag{13.33}$$

$$\text{Beverton–Holt: } PC = 100\left[\frac{-m}{1-a}\right]. \tag{13.34}$$

In each case m is defined in terms of a proportional effect. A key assumption of the approach is that compensation operates over only a brief period early in the life cycle of the fish. Mortality added after compensation has been assumed to occur results in larger impacts on the population. Although simple and deterministic in nature, the approach yields results similar to those produced by more complicated age-structured models, except when variation in survival rates is high and the modeled population is semelparous (Savidge et al. 1988).

In summary, SR-based approaches to assessing the population level effects of stress have several advantages: mathematical simplicity, paucity of data requirements, and specific inclusion of population regulation concepts in the form of density-dependent and density-independent mortality. Nevertheless, SR-based methods are suited only for use with populations for which an adequate time series of information exists from which statistically sound SR models may be estimated. In addition, SR models cannot incorporate multiple sources of mortality or age-dependent variations in density dependence unless specific age-to-age or stage-to-stage analyses are to be conducted. Even then, interannual variations in rates will be difficult to adjust for in model estimation and are likely to introduce enough variability in the data required to estimate the SR model to preclude the model from explaining a sufficiently high proportion of observed recruitment variability.

13.2.3.4 Aggregate Assessment Methods

Aggregate assessment methods include a variety of techniques developed for estimating entrainment and fishing mortality losses. As elsewhere, the analogy between population losses resulting from fishing or entrainment and population losses resulting from the actions of other anthropogenic stressors is a good one, and the techniques may be adapted for use in population level bioassessments. Techniques include the equivalent adult method (Horst 1975), the production foregone method (Rago 1984), and the reproductive potential method (Goodyear 1988, 1993).

Equivalent adult method. Developed initially to assess the effects of power station entrainment of juveniles on eventual population abundance, the equivalent adult methods is most suited to the study of stressors having their largest impacts on juvenile age-classes (Horst 1975). The method attempts to translate losses in juvenile age-classes into the equivalent number of individuals lost in older age-classes under the assumption of the juvenile fish having been allowed to survive to reach

the older age-classes. As such the method directly accounts for the effects of natural mortality acting in the intervening period and may be used to express losses in terms of the age- or size-classes of fish that are typically the focus of commercial or recreational harvest regulations. If density-dependent mortality does not act on the population between the age-classes impacted by a stressor and the age-class at which the population level effects are assessed, population level effects may be estimated as (Horst 1975; Goodyear 1978)

$$N_a = \sum_{i=1}^{n} N_i S_i, \tag{13.35}$$

where N_a is the equivalent number of fish in age-group a, the age-class at which the population level effects are judged, N_i is the number of fish in the ith age-class that are lost, S_i is the survival rate between age-class i and a, and n is the number of age-classes between the first-impact age-class and age-class a. With additional information on population age structure and reproductive characteristics, the approach may be expanded into a Leslie matrix analysis, and the impacts on the intrinsic rate of population increase can be measured using the dominant eigenvalue. Although the approach is useful for obtaining a first approximation of potential adult population impacts of age-0 losses, significant insights into probable long-term effects of persistent losses are not obtained using the equivalent adult approach (Goodyear 1978).

Production forgone method. Losses of fish as a result of anthropogenic impacts have immediate as well as future impacts on population abundance. Fish biomass removed from an aquatic system as a result of stress reduces population biomass and the potential prey base for predators, leading to trophic cascades. One means of assessing the potential for trophic cascades is to determine what an individual fish would have produced in terms of biomass over its remaining life span. Defining P_j as the production forgone due to a stressor removing a fish of age j, the loss in biomass may be computed as (Rago 1984)

$$P_j = \sum_{i=j}^{t_{\max}} G_i B_i, \tag{13.36}$$

where $t_{\max}$ is the maximal age that can be attained by any individual in the population and G_i and B_i, respectively, are the ith age-specific instantaneous somatic growth rates and average biomasses. The equation implies that age-specific abundances (N_i) and age-specific mean weights (W_i) are known or can be estimated as follows:

$$W_i = W_{i-1} e^{Gt}, \text{ and} \tag{13.37}$$

$$N_i = N_{i-1} e^{-Zt}. \tag{13.38}$$

In either case, simple recursive exponential decline expressions may be used if other methods do not suffice. A detailed example of the application of the

production forgone method is given in Rago (1984) for a hypothetical fish population impacted by entrainment mortality.

Production forgone is argued to be more relevant for purposes of population level bioassessment than numbers lost because it includes consideration of the energy potentially transferable to other trophic levels through consumption or decomposition (Rago 1984). The key assumption of the approach is that short-term population responses to stress do not induce immediate alteration of existing survival and growth rates deducible from current population structure. Thus, the method does not require the specification of a dynamic population model, and the relative importance of larval versus juvenile and adult fish losses can be compared directly (Rago 1984) in any completed analysis. Additionally, the model requires some parameters that are difficult to estimate (e.g., juvenile age-class survivors) and is consequently restricted in its use to species with adequate pre-existing data sets. Finally, the production forgone methods is deterministic and limits users to making projections about future abundances, without allowing comment on the probabilities associated with those estimates.

Reproductive potential method. Although the equivalent adult method addresses potential abundance losses, it has been criticized for failure to provide long-term insights into the consequences of possible losses in juvenile age-classes (Goodyear 1978), and it cannot adequately deal with the nonlethal effects of chronic pollutant impacts that typically include reductions in fecundity (Adams 2002). To understand the possible consequences of stressor effects for reproductive impairment of the population as a whole, the reproductive output of an individual fish recruited to the reproductive proportion of the population may be measured as potential lifetime egg production *(P)* under conditions of optimal growth and survivorship as follows (Goodyear 1988):

$$P = \sum_{i=1}^{n} X_i R_i L_i \prod_{j=0}^{n-1} S_j, \tag{13.39}$$

where S_j is the density-independent survival probability for the jth age-class, X_i is maximum average fecundity of mature females at age i, R_i is the maximal fraction of females of age i that are mature, L_i is proportion of age-class i that is female, and n is the number of population age-classes. Multiplication of P by the total number of recruits to the reproductive age-classes yields the reproductive potential for the population. Under the conditions of stationary (constant year-to-year S_j, R_i, and L_i) and a 1:1 sex ratio among eggs, the age-0 survival rate (S_0) is given by

$$S_0 = \frac{2}{P}. \tag{13.40}$$

Stressor-induced mortality in the first year of life would affect S_0 but not P, whereas stressor-induced mortality in the older age-classes would affect P but not S_0.

Reproductive potential may also be used to compute a viability index representing the aggregate of compensatory density-dependent factors operating

throughout the life history of the study population under the assumption that the population is fluctuating about a long-term equilibrium (Goodyear 1988). The viability index *(v)* is given by

$$v = 1/PS_0, \tag{13.41}$$

where S_0 is the probability of survival from density-independent sources of mortality between egg deposition and recruitment to an age-class of concern (e.g., typically the age-class at which the population is fully recruited to the fishery or the reproductive age-classes of the population). For assessment of stressor-related effects S_0 may be broken into two component parts, survival from density-independent natural causes of mortality (S_N) and survival from density-independent anthropogenic stressor-related causes of mortality (S_S), such that

$$S_0 = S_N S_S. \tag{13.42}$$

By analogy, S_0 can be decomposed into multiple sources of mortality if the impact of multiple stressors on a single population is to be considered.

$$S_0 = S_N S_1 S_2 \ldots S_M, \tag{13.43}$$

where $S_1, S_2, \ldots, S_M$ represent survival from density-independent anthropogenic stressor-related mortality causes 1 through *M*.

From the viability index a compensation ratio (CR) may be computed. The CR is a composite measure of the changes in survival and fecundity necessary for a stressed population to attain a new equilibrium and is expressed relative to the viability index of the unstressed population as follows (Goodyear 1988):

$$\text{CR} = \frac{v_S}{v_U}, \tag{13.44}$$

where v_S and v_U are stressed and unstressed population viability indices. Typically, the CR ratio will be greater than 1 because the total survival probabilities will be greater in the unstressed population. Differences between CR and 1, therefore, will represent an index of the stressor impact. If cumulative stresses are considered, the differences between combinations of stressors will represent the marginal impact of the incremental stressor (Goodyear 1988).

13.3 VARIANCE ANALYSIS

Populations are known to vary over a wide range of densities in response to environmental fluctuations and disturbances, and there are theoretical reasons for expecting the variance of life history characteristics to increase as a function of stress (Service and Rose 1985). Accordingly, variability may provide a convenient measure of the relative degree of population level stress. Ryder (1990) reasoned

that whatever the inherent complexity of an ecological system, all systems were characterized by a normative range of variability because of intrinsic homeostatic mechanisms. Variability outside the normative range, therefore, ought to provide an easily obtainable indicator of stress. DeAngelis et al. (1990) likewise suggested that it was possible to compare within-species variability at different times—and under different conditions—and use increases in variability as evidence of population level stress.

Marshall (1978) was among the first to demonstrate the responsiveness of variability to stress with a series of laboratory experiments that examined the effects of chronic cadmium exposure on zooplankton populations. Results showed an exponential rise in variability, measured as the coefficient of variation for population abundance. Inter-individual physiological variability was subsequently suggested as a specific variability-based means of investigating stress effects (Depledge 1990). Forbes et al. (1995) examined the growth rate of gastropods to cadmium exposure and found that exposure to cadmium increased the variability in population growth rates. Results lent credibility to the use of variability as a potentially useful indicator of exposure to environmental stress. Power (1997) examined brook trout population responses to cumulative stresses within a modeling context by means of an index of stressor intensity and found broad support for the notion that variability increased as a function of stress.

To assess variation in response to stress, one must determine beforehand which endpoint to assess and which measure of variability is most appropriate. Reproductive and growth endpoints are sensitive to stress (Adams 2002) but have the disadvantage of being equally responsive to environmental variation (Wootton 1990). Summary measures like abundance and the finite rate of population increase are better insofar as they are integrative measures. Many analyses of variation in abundance choose to employ the coefficient of variation (CV) computed as

$$\mathrm{CV} = \frac{\sqrt{S^2(n)}}{\overline{X}(n)}, \tag{13.45}$$

where $s^2(n)$ and $\bar{x}(n)$, respectively, are the variance and mean of n abundance observations. Simple statistical comparison between CVs computed for the same population under differing stress regimes, or for two populations at differing sites, may be completed for samples with n_1 and n_2 each equal to 30 or more by use of a z-statistic as follows:

$$z = \frac{|CV_1^2 - CV_2|}{\sqrt{\frac{CV_1^2}{2n_1} + \frac{CV_2^2}{2n_2}}}, \tag{13.46}$$

where CV_i is the value of the ith CV computed with observations n_i. For an exact test and critical values for small n see Lohrding (1975).

Abundance-based measures of the CV, however, are sensitive to high abundance values. Furthermore, in comparisons involving samples from different locations or times, a direct comparison of computed CVs may not be valid because the coefficients would reflect the combined differences in natural and stressor-induced variability and mean abundance. Accordingly, Williamson (1984) has recommended the use of a measure of annual variability (AV) developed by Wolda (1983):

$$\mathrm{AV} = \mathrm{var}[\log_{10}(n_{t+1}) - \log_{10}(n_t)], \tag{13.47}$$

where n_{t+1} and n_t, respectively, are the population abundances at time $t+1$ and t. If the population under study is following a trend (e.g., decreasing), as many stressed populations may be doing, the AV measure is superior to the CV (Wiens 1989). In addition, the measure facilitates comparisons of different sites or species without the need for awkward corrections for the effects of density (Williamson 1984). The AV may be simply related to the standard deviation of the logarithm ($\log_{10}$) of population density (s_D) as follows (Williamson 1984):

$$\mathrm{AV} = 2s_D^2(1 - r_1), \tag{13.48}$$

where r_1 is the first serial correlation coefficient of the log-transformed abundance time series. When r_1 is not significantly different from 0, AV may be used to estimate the standard deviation of population density. When r_1 is greater than 0, AV will be smaller than when estimated using population density alone. Similarly, when r_1 is less than 0, AV will be larger than when estimated using population density alone. The significance of r_1 can be simply tested at an α of 0.05 by means of the relation (Abraham and Ledolter 1983)

$$\sqrt{n}\,|r_1| > 1.96\,, \tag{13.49}$$

where n is the number of samples in the time series from which r_1 was calculated, and r_1 is computed as (Abraham and Ledolter 1983)

$$r_1 = \frac{\sum_{t=1}^{n-1}(n_t - \overline{n})\,(n_{t+1} - \overline{n})}{\sum_{t=1}^{n}(n_t - \overline{n})^2}. \tag{13.50}$$

where, n_t is population density in the tth time period and $\overline{n}$ is the mean of the population density observations.

Whittaker (1975) proposed a coefficient of fluctuation based on sequential abundance measures (N) that may be adapted for broader use in population bioassessment as follows

$$\mathrm{antilog}\sqrt{\frac{\sum_{i=1}^{N}(\log_{10} n_i - \log_{10} n_g)^2}{t-1}}. \tag{13.51}$$

Here n_i is the ith species abundance, or other population endpoint measure, and n_g is the geometric mean of the samples:

$$n_g = \sqrt[N]{n_1 \cdot n_2 \cdot \ldots \cdot n_N}. \tag{13.52}$$

The use of logarithms implies that the measure is less sensitive than the CV to differences in absolute measures of the observations used to assess variability.

All approaches to using variability as a measure of stress suffer from the potential problem of temporal autocorrelation. Sequential measures of any endpoint are unlikely to be completely independent because of the persistence of either resident individuals or, in the case of many stress exposure problems, physiologically resistant individuals. One suggested solution is the assessment of stress in populations over periods of time in excess of generational length (Connell and Sousa 1983). Although multiperiod occupancy of a location by individuals violates the requirements of statistical independence, it does represent the biology of the system under study (Wiens 1989). Furthermore, there is comparative value in the information for population bioassessments if a population at one location is less variable than a population at another location when both are subjected to similar stress regimes.

13.4 KEY-FACTOR ANALYSIS

When studying any fish population, it is often desirable to measure the effects of all mortality factors acting on the population and to select among them to determine which factor (or factors) is most important for the observed variability in population abundances through time. Morris (1959) introduced the term key factor to describe mortality factors that contribute to variable mortality throughout the life cycle and appear to be largely responsible for observed changes in population density in successive generations. Key-factor analysis is based on the concept that it is the variation in survival rates at different life stages, and the manner in which variations in survival rates in one stage affect abundance in later stages, that determine overall population dynamics. Accordingly, appropriate use of key-factor analysis allows determination of which life stage mortality factors have been most affected by stress. The key-factor concept was subsequently developed by Varley and Gradwell in a series of papers (1960, 1963, and 1968) and Varley et al. (1973). Southwood and Henderson (2000) describe the method of key-factor analysis in detail. Elliott (1994) provides an excellent example of the application of the analysis to a fish population, and Stiling (1988) and Royama (1996) discuss the application of key-factor analysis.

The method requires data from a series of successive age- or stage-specific life tables. Consider a single generation of a population whose abundances entering age-classes or developmental stage-classes 1 to m are $N_1, N_2, \ldots, N_m$. The stage-specific survival rates (w_j) can be expressed as the number alive in stage $j + 1$ divided by the number alive in stage j, or

$$w_j = N_{j+1}/N_j. \tag{13.53}$$

Total survival to the final stage (generation survival), K, is the product of the stage-specific survival rates:

$$K = \frac{N_m}{N_1} = \prod_{j=1}^{m-1} w_j. \tag{13.54}$$

Taking the logarithms then yields

$$\log_{10}(N_m/N_1) = \log_{10}(w_1) + \log_{10}(w_2) + \ldots + \log_{10}(w_{m-1}), \tag{13.55}$$

or

$$K = k_1 + k_2 + \ldots + k_{m-1}. \tag{13.56}$$

The resulting values are typically multiplied by –1 to avoid expressing the key factors as negative values, thus $K = -\log_{10}(N_m/N_1)$ and $k_i = -\log_{10}(w_i)$. Finally, the values K and k_j are plotted against the same time axis for a series of generations and the key factor is identified by visual inspection as the stage-specific k-value whose variation most closely resembles that of K. See, for example, Elliott (1985). Caution in interpreting significance, however, is advised, as the k_i values obscure the effects of sampling errors, which may lead to misinterpretation of the data (Kuno 1971).

The graphical Varley and Gradwell method often identifies key factors, but ambiguities in interpretation can arise (Manly 1990). In particular, the visual inspection approach has been criticized on the grounds that no key factor may be immediately obvious and the relative importance of mortality factors other than that identified as the key factor are ignored (Podoler and Rogers 1975). To account for situations in which the key factor may not be obvious, or when there is a need to investigate the relative importance of changes in each mortality factor to changes in total mortality, regression of the k_i values against K and the use of the resulting regression least-squares estimates of slope has been suggested as one means of determining the relative importance of each k_i value (Podoler and Rogers 1975). The regression slope of k_j on K is given by the least-squares estimator b_j as follows:

$$b_j = \frac{\sum_{i=1}^{G} (k_{ij} - \bar{k}_j)(K_i - \bar{K})^2}{\sum_{i=1}^{G} (K_i - \bar{K})^2}, \tag{13.57}$$

where k_{ij} is the k-value for the survival in the jth stage of the ith generation; $\bar{k}_j$ is the mean k-value for survival in the jth stage; K_i is the K-value in the ith generation;

$\bar{K}$ is the mean of K for all generations; and G represents the generations of data. From the above it follows that

$$\sum_{j=1}^{m-1} b_j = 1, \tag{13.58}$$

which implies that the b_j values have an immediate use for interpreting the contribution of the jth k-value to explaining the total variation observed in K. Normalization also allows direct comparison of two or more populations for detection of differences in key factors. See, for example, Box 13.5.

Although the proposed methodology allows quantitative determination of the key factor, it is not without its drawbacks. Podoler and Rogers (1975) report a number of methodological short-comings, as follows.

1. The variables k and K are not statistically independent because errors in the estimates of the k-values will be reflected in the estimate of total mortality.
2. The test cannot be used to draw conclusions about the real contributions of each k-value to changes in population density.
3. Normal regression techniques cannot be used to estimate the level of significance associated with a particular line because the k_i values are functionally related to K.

The procedure, therefore, can be used only to select the factor that most contributes to changes in the value of total generational mortality and does not replace the need to use subsequent statistical testing procedures to establish the existence of density relationships for each mortality factor.

The various k_i values may be tested for density dependence by plotting each k_i value against the number of individuals entering the age or stage interval (N_i) on which it acts and determining the significance of the regression. If significance is found, density dependence is suspected. The method has been criticized because initial density appears on both sides of the regression equation (Elliott 1994). The problem may be overcome by first regressing the logarithm of the initial age or stage density ($\log_{10} N_i$) on which k_i acts against the logarithm of survivor density ($\log_{10} N_{i+1}$) and then regressing the logarithm of survivor density against logarithm of initial density. If both regression lines plotted on a single graph lie on the same side of unity, and have slopes significantly different from one, the relationship is concluded to be density dependent (Varley and Gradwell 1968). Southwood et al. (1989), however, have noted that regression of k-values against initial density has been shown to be a much less biased technique for detecting density dependence than initially thought. Accordingly, Elliott (1994) adopted the method for use in examination of density dependence in a stream-resident population of brown trout in the English Lake District. Kuno (1971) and Bulmer (1975), however, provide detailed statistical critiques of the problem of testing for density dependence with key factors, and Bulmer (1975) provides statistical tests that address the critical problems of measurement error and serial correlation that plague key-factor detection of density dependence.

Box 13.5 Key-Factor Analysis

Consider the data used in Box 13.1 for the 1954 cohort of brook trout from Hunt Creek, Michigan, expanded to include 1952 and 1953 (McFadden et al. 1967).

Table Data for the cohort of brook trout from Box 13.1 expanded to include 1952 and 1953.

Age-class	1952	1953	1954
Age 0	51,000	40,000	52,000
Age 1	2,694	3,162	2,118
Age 2	1,052	1,160	779
Age 3	199	183	159
Age 4	26	17	13

Defining stage-specific survival rates (w_i) as in equation (13.52), converting to $\log_{10}$ values, and summing as in equations (13.54) and (13.55) yields the following table of k_i (stage-specific survival) and K (generation survival) values. Note the values are multiplied by –1 to avoid expressing the key factors as negative values.

Table Stage-specific (k_i) and generation (K) survival for 1952–1954 cohorts of brook trout from Hunt Creek, Michigan (McFadden et al. 1967). Values are multiplied by –1 to avoid expressing the key factors as negative values.

Survival value	1952	1953	1954
k_0	1.2772	1.1021	1.3901
k_1	0.4084	0.4355	0.4344
k_2	0.7232	0.8020	0.6901
k_3	0.8839	1.0320	1.0875
K	3.2926	3.3716	3.6021

Repetition of the computations for the data given for the years 1949 to 1951 and 1955 to 1959 yields estimates of the k_i and K values given in the table below.

Table Estimates of k_i and K for an expanded brook trout data set.

Survival value	1949	1950	1951	1955	1956	1957	1958	1959
k_0	1.4801	1.4469	1.3436	1.4699	1.3008	1.3610	1.5728	1.5994
k_1	0.3246	0.3765	0.4637	0.4339	0.3956	0.3478	0.2685	0.4169
k_2	0.9017	0.9770	0.8197	0.6740	0.6497	0.6967	0.7186	0.6848
k_3	1.2231	1.1996	1.0381	0.7541	1.0998	1.2478	0.9031	1.0223
K	3.9294	4.0000	3.6651	3.3318	3.4459	3.6532	3.4630	3.7234

Combining the data in the two tables and plotting them yields the following figure.

(Box continues)

Box 13.5 *(continued)*

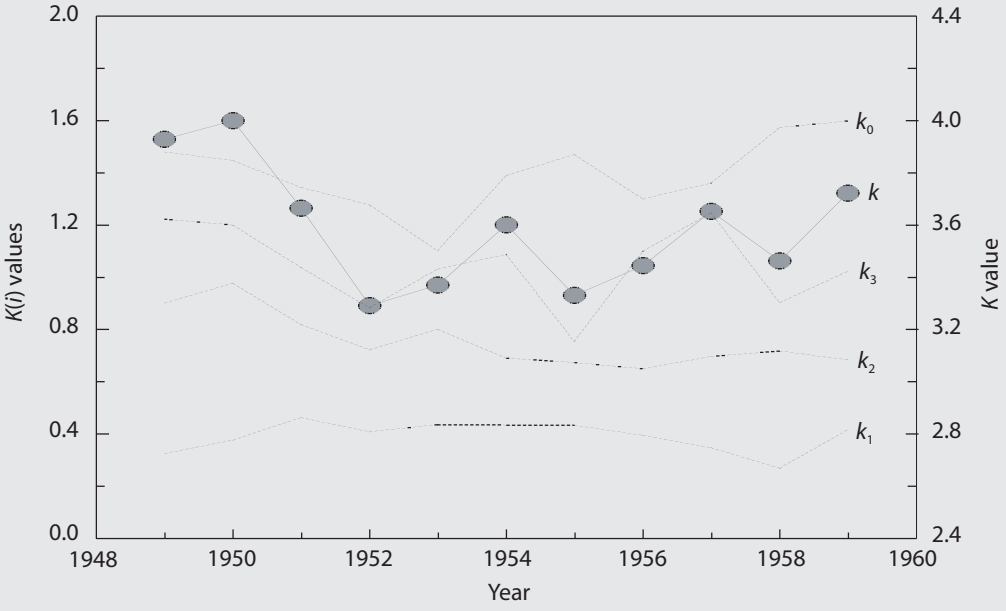


Figure Combined data for brook trout from Hunt Creek, Michigan (McFadden et al. 1967). The series of k_i values are plotted as dashed lines and the *K*-value series is plotted as a solid line with gray circles.

Inspection of the graph suggests that the variation in k_3 most closely resembles the variation in *K*. The relative importance of k_3 may be confirmed by regressing the k_i values against *K* as suggested by Podoler and Rogers (1975).

Table Regression of k_i versus *K* for combined data for brook trout.

Survival value and summation	Slope (b_j)	r^2
k_0	0.258	0.182
k_1	–0.060	0.059
k_2	0.312	0.489
k_3	0.490	0.567
$\sum_{j=0}^{3} b_j =$	1.000	

If regressions of initial on final density as $\log_{10}(N_i)$ on $\log_{10}(N_{i+1})$ and $\log_{10}(N_{i+1})$ on $\log_{10}(N_i)$ yield slopes that are significantly different from 1 and are on the same side of unity, then density dependence in the life stage *i* to *i* + 1 may be taken as demonstrated (Varley and Gradwell 1968). Significance is established using standard regression testing techniques as follows:

$$t = \frac{b_i - 1}{\mathrm{SE}(b_i)},$$

where b_i is the estimated regression slope and SE(b_i) is the standard error of the slope. Significance is determined with reference to the Student's *t*-table with $n - 2$ df. For the data used in this example, repeat application of the suggested regressions yields the results given in the table below.

Table Regressions of initial on final density as $\log_{10}(N_i)$ on $\log_{10}(N_{i+1})$ and $\log_{10}(N_{i+1})$ on $\log_{10}(N_i)$ to determine density dependence in the life stage i to $i + 1$. Significance is determined with reference to the Student's *t*-table with $n - 2$ df.

	N_i versus N_{i+1}		N_{i+1} versus N_i	
Regression	Slope	*t*-statistic	Slope	*t*-statistic
Age 0 and age 1	–0.159	–4.98	–0.310	–2.89
Age 1 and age 2	0.618	–1.72	0.745	–0.95
Age 2 and age 3	1.500	1.03	0.345	–5.90
Age 3 and age 4	1.340	1.01	0.474	–4.40

With 9 df, only regressions with *t*-statistics of an absolute value greater than 2.262 are significant. Using the criteria of Varley and Gradwell, only the age-0 to age-1 interval yields evidence of density dependence. The simpler approach applied by Elliott (1994) to the study of brown trout yields similar results when applied to the McFadden et al. (1967) brook trout data, with only the regression of k_0 on initial density yielding a significant result as shown below.

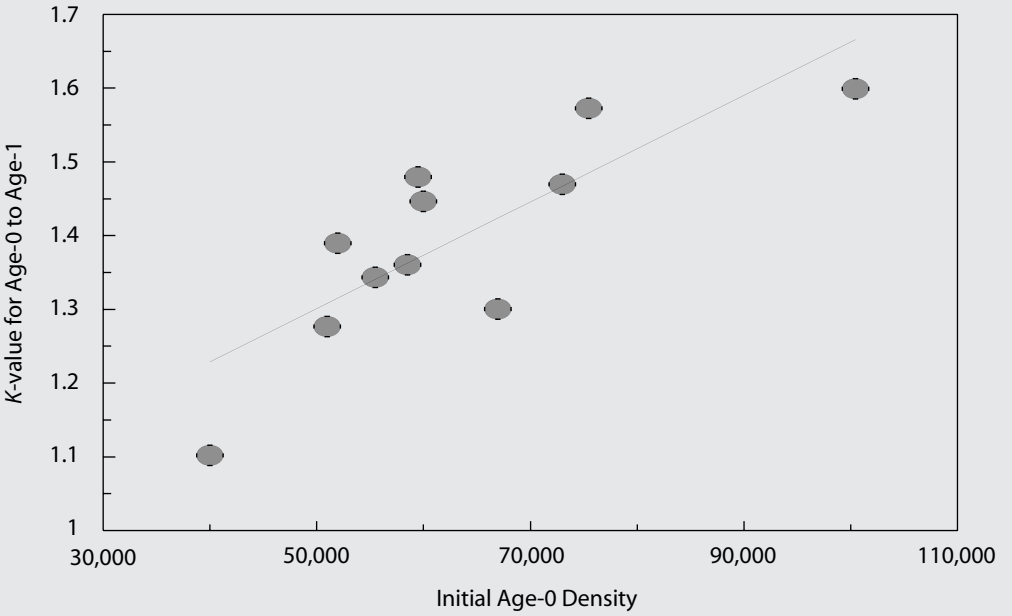


Figure Regression of k_0 on initial density of brown trout. Application of Elliott's (1994) method to the McFadden et al.'s (1967) brook trout data.

Smith (1973) argued that the graphical approach was necessarily subjective due to the problem of scale variation and suggested an analysis of variance approach, where total variation in K was systematically apportioned to the different sources of mortality in a manner similar to that proposed by Podoler and Rogers (1975). Smith (1973) also suggested that once the key factor had been determined there was further utility in ascertaining the relative importance of the remaining k-values. This involves defining the residual killing power, K_R, in ages or stages other than that identified as the key stage as follows:

$$K_R = K - k_i. \tag{13.59}$$

The remaining stage-specific k-values are regressed on K_R to determine which is the most important determinant of K_R. Repeated use of the procedure allows the age- or stage-specific k-values to be rank ordered in terms of their relative contributions to K as detailed by Smith (1973). Manly (1990), however, has noted that significance testing of the k_i values is not straightforward because sampling errors in the k_i values may induce apparent regression relationships.

None of the methods described above makes direct use of the known order in which mortality operates on the age or life stages. If mortality is highly density dependent in a given life stage, it can remove most of the variation in abundance induced by the action of variable mortality in previous stages such that density-dependent k–values may show little relationship to K, even though they are important for the determination of overall population dynamics (Manly 1990). By estimating a population model that includes density-dependent mortality, it is possible to estimate the contribution of each life stage to the variation in final stage abundance under the assumption that the k_i values are linear functions of the numbers entering the stage ($k_i = a + bN_j$). Accordingly, a key factor may be defined as the life stage that substantially increases or decreases the variation in the final stage (Manly 1990).

Statistical details of model estimation and the determination of k-value significance, along with associated computation programs, are given in Manly (1990). Although the method has great promise, it is limited in application to cases in which sufficient data exist to estimate adequately the underlying regressions required to drive the model. It is also important to note that the method assumes that k-values are linear functions of the numbers entering each stage, which may not be true. Accordingly, the validity of the model and its conclusions concerning the influence of each life stage on variation observed in the final stage will depend on the extent to which the linear model approximates any nonlinear relationship between the k-values and the numbers entering the stage (Manly 1990).

13.5 SUMMARY

There are numerous methods for assessing the possible impacts of stressors on fish populations. The basic decision in the use of population bioassessment methods is the choice between field-based and modeling-based studies. Both approaches have

their relative strengths and weaknesses and neither should be considered superior. An approach that combines the two is highly recommended, though practical for only a select few populations of high conservation concern. It is important not to regard population bioassessment as an activity separate from the routine collection and analysis of fishery data. A population bioassessment will only ever be as good as the care and attention that goes into the routine collection and analysis of population monitoring data. In that regard, population bioassessment should be looked on as the natural extension and joint consideration of the numerous methods elaborated in previous chapters of this book. The key difference between population monitoring and population bioassessment is that population bioassessment draws its conclusions from multiple lines of evidence, whether informally as in the case of field-based methods or formally as in the case of modeling techniques that explicitly combine survival, growth, and reproductive information. A summary of the approaches discussed here is given in Table 13.3.

A basic requirement of approaches to population bioassessment is the availability of adequate data. For field-based methods, the requirement can be problematic insofar as available data must be sufficient to allow investigators to partition natural variability from stressor-induced variability. Numerous modeling-based studies have made this point. For example, a Leslie-matrix-based analysis of the significance of reducing brook trout population biomass by increasing age-0 mortality by 50% indicated the first appearance of a significant change in yield statistics did not occur for at least 2–3 years (Jensen 1971b). Attempts to determine the number of years of data required to detect entrainment reductions in Hudson River white perch year-class strength concluded that at least 20 years of data would be required to detect a greater than 50% reduction in mean year-class strength (Van Winkle et al. 1981; Vaughan and Van Winkle 1982). Power and Power (1995) further demonstrated that despite stressor-induced monotonic declines in age-0 abundances, adult brook trout abundances displayed a subsidence response as predicted by Odum et al. (1979). The subsidence-induced lag between the initiation of stressor action and the ability to detect significant stressor-induced changes in an endpoint response poses a critical challenge to field-based assessment approaches, particularly if they do not include considerable detail on prior natural history and an understanding of the population regulating processes (Underwood 1989).

The field detection and measurement of stresses acting on natural populations would appear to be beset with considerable difficulties. If it were already established that a population of interest was at equilibrium, and the equilibrium level was known, then the detection of stress would simply involve determining that a known perturbation had caused the endpoint of interest (e.g., abundance) to deviate from the equilibrium value (Underwood 1989). The simple comparison, however, cannot be made because populations rarely, if ever, exist at equilibrium. Even if there are no temporal fluctuations in abundance, equilibrium cannot be established unless observations are available for a period of time exceeding the natural life expectancy of the species (Frank 1968). In cases in which the sampling

Table 13.3 Summary of reviewed approaches to assessing population level effects of stress in fishes. The basic division is between field and model approaches. Key assumptions, temporal resolution, key advantages and disadvantages, and the biological relevance of each approach are listed below. Key assumptions must always be validated if approach results are to be credible.

Approach	Key assumptions	Temporal resolution	Key advantage	Key disadvantage	Biological results relevance
			Field Based		
Before and after comparisons	Stressor addition causes differences in endpoint measures	Poor, period to period only	Quick, removes problem of selecting comparable reference sites	Retrospective, does not consider response dynamics, discounts temporal change	Low and weakly descriptive
Reference site comparisons	Stressor alone causes differences in endpoint measures	Very poor, static point-in-time samples	Quick, removes problem of considering temporal changes	Retrospective, does not consider response dynamics, discounts spatial and temporal differences	Low and weakly descriptive
Graded exposure–response	Response gradient attributable to differences in stressor intensity	Very poor, static point-in-time samples	Considers spatial aspects of effects explicitly	Weak diagnostic tool, cannot establish causation	Medium, helps describe population health, amenable to statistical testing
Sequential data sampling	Critical response variables known a priori, stability in nonmeasured variables prevails	Good to excellent depending on number and frequency of samples	Considers temporal aspects of effects explicitly	Requires long-term data and consistency of sampling effort	Medium to high, very descriptive of rates of change and trends, amenable to statistical testing

Model Based					
Life tables	Fixed mortality and natality schedules	Poor as fixed vital rate assumptions imply point-in-time analysis	Provides critical summary of current population vital statistics	Data can be hard to obtain, may confuse interannual with age-specfic variation	Good if used descriptively for a single period
Leslie matrix	Stable-age distribution reached, mortality and natality schedules fixed	Good as may be tuned to species reproductive cycles	Allows projection of future abundances based on current vital statistics	Usually deterministic, omits density-dependent effects, correction of both can be difficult	Good if projections used with caution and constancy and deterministic flaws addressed
Surplus production	Population in equilibrium, factors affecting somatic growth do not change from period to period	Good, uses age-class data to produce annual results	Provides summary index of biomass surplus to maintaining equilibrium	Does not consider possible compensation effects, deterministic	Poor, yields an aggregate population response average somatic and mortality responses
Stock recruitment	Population processes are invariant with respect to time over the period for which data available	Low if generational data used, good if age-class data used	Uses aggregate data typically available with reasonable time series	Averages all mortality factors over all age-classes between parent and recruit	Medium, yields an average response incorporating all sources of mortality
Variance analyses	Aggregate measures will not damp responses to stress	Good, requires annual data series	Yields a quick index of responses to stress	Statistical testing issues arising from auto-correlation complicate interpretation	Poor, an aggregate measure not specific to a single source of stress
Key-factor analysis	Population processes invariant with respect to time over analysis period	Good to excellent depending on data used	Quick identification of life-stage bottlenecks and density-dependent stages	Requires long-term, detailed age-class abundances	Good to excellent in that it identifies key life stages

of natural populations has been carried out over a sufficiently long period, there are few indications that populations exist at equilibrium (Connell and Sousa 1983).

Establishing the effects of stress on a population, therefore, requires more than demonstrating changes in population abundance endpoints. For stress to have had a significant effect, measured population level changes must exceed those that would normally be expected based on knowledge of inherent population variability and the routine operation of population regulating mechanisms (Underwood 1989). Unfortunately, sampling imprecision complicates attempts to gather the required data. Census approaches to estimating population parameters are typically not feasible except in cases where population abundances have declined to the point at which it is clear a population is threatened with local extinction. As a result, random sampling is used to establish a probabilistic parameter estimate that carries with it a variance and the associated problem of sampling error. Accordingly, if field-based methods are to be used, the graded-exposure–response (Adams et al. 1994) or sequential sampling (Mills 1985; Mill et al. 1987) approaches that specifically accommodate statistical tests of significance based on estimated variability are preferable to static compare-and-contrast approaches that eliminate consideration of variability on the pretense of diagnostic power.

Modeling approaches to population bioassessment are likely to help in identifying and quantifying the effects of stressor-induced changes in key population regulating processes. There is a long history of model development and use, but application of models to studying the specifics of stressor induced changes has not always been easy owing to the paucity of good, long-term data sets. Well-designed and executed monitoring programs can provide the information necessary to postulate and test stressor cause–effect relationships and, in the future, should redress this problem. Nevertheless, modeling studies will always rely on field data and cannot be viewed as an acceptable surrogate for field studies. Instead, modeling and field approaches to the study of population bioassessment must be viewed as mutually supportive means of study, requiring connected and parallel programs of study. The achievement of approach integration will require substantial interdisciplinary cooperation but should improve abilities to assess the ultimate effects of stressor action on fish populations.

13.6 REFERENCES

Aalto, S., and G. E. Newsome. 1980. Some methods of estimating the parameters of the Leslie matrix using incomplete population data. Canadian Journal of Fisheries and Aquatic Sciences 37:1140–1148.

Abraham, B., and J. Ledolter. 1983. Statistical methods for forecasting. Wiley, New York.

Adams, S. M. 1990. Status and use of biological indicators for evaluating the effects of stress on fish. Pages 1–8 *in* S. M. Adams, editor. Biological indicators of stress in fish. American Fisheries Society, Symposium 8, Bethesda, Maryland.

Adams, S. M., editor. 2002. Biological indicators of aquatic ecosystem stress. American Fisheries Society, Bethesda, Maryland.

Adams, S. M., A. M. Brown, and R. W. Goede. 1993. A quantitative health assessment index for rapid evaluation of fish condition in the field. Transactions of the American Fisheries Society 122:63–73.

Adams, S. M., K. D. Ham, and J. J. Beauchamp. 1994. Application of canonical variate analysis in the evaluation and presentation of multivariate biological response data. Environmental Toxicology and Chemistry 13:1673–1683.

Adams, S. M., and M. G. Ryon. 1994. A comparison of health assessment approaches for evaluating the effects of contaminant-related stress on fish populations. Journal of Aquatic Ecosystem Health 3:15–25.

Bagenal, T. B., and E. Braum. 1978. Eggs and early life history. Pages 166–198 *in* T. B. Bagenal, editor. IBP (International Biological Programme) Handbook 3, Methods for assessment of fish production in fresh waters, 3rd edition. Blackwell Scientific Publications, Oxford, UK.

Barnthouse, L. W. 1993. Population-level effects. Pages 247–274 *in* G. W. Suter, editor, Ecological risk assessment. Lewis Publishers, Boca Raton, Florida.

Barnthouse, L. W., G. W. Suter, and A. E. Rosen. 1990. Risks of toxic contaminants to exploited fish populations: influence of life history, data uncertainty and exploitation intensity. Environmental Toxicology and Chemistry 9:297–311.

Begon, M., J. L. Harper, and C. R. Townsend. 1990. Ecology: individuals, populations and communities. Blackwell Scientific Publications, Oxford, UK.

Berlinsky, D. L., M. C. Fabrizio, J. F. O'Brien, and J. L. Specker. 1995. Age-at-maturity estimates for Atlantic coast female striped bass. Transactions of the American Fisheries Society 124:207–215.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. Fishery Investigations, Series II, Great Britain Ministry of Agriculture, Fisheries and Food 19.

Birch, L. C. 1953. Experimental background to the study of the distribution and abundance of insects. III. The relations between innate capacity for increase and survival of different species of beetles living together on the same food. Evolution 7:136–144.

Boreman, J. 1997. Methods for comparing the impacts of pollution and fishing on fish populations. Transactions of the American Fisheries Society 126:506–513.

Bulmer, M. G. 1975. The statistical analysis of density dependence. Biometrics 31:901–911.

Cairns, V. W., P. V. Hodson, and J. O. Nriagu. 1984. Contaminant effects on fisheries. Volume 16, Advances in Environmental Science and Technology. Wiley, New York.

Caswell, H. 1989. Matrix population models: construction, analysis and interpretation. Sinauer, Sunderland, Massachusetts.

Caughley, G. 1977. Analysis of vertebrate populations. Wiley, London.

Chambers, R. C. 1993. Phenotypic variability in fish populations and its representation in individual-based models. Transactions of the American Fisheries Society 122:404–414.

Chaumot, A., S. Charles, P. Flammarion, and P. Auger. 2003. Ecotoxicology and spatial modeling in population dynamics: an illustration with brown trout. Environmental Toxicology and Chemistry 22:958–969.

Chaumot, A., S. Charles, P. Flammarion, J. Garric, and P. Auger. 2002. Using aggregation methods to assess toxicant effects on population dynamics in spatial systems. Ecological Applications 12:1771–1784.

Clark, L. R., P. W. Geier, R. D. Hughes, and R. F. Morris. 1967. The ecology of insect populations in theory and practice. Methuen, London.

Connell, J. H., and W. P. Sousa. 1983. On the evidence needed to judge ecological stability. American Naturalist 121:789–824.

Craig, J. F. 1980. Growth and production of the 1955 to 1972 cohorts of perch, *Perca fluviatilis* L., in Windermere. Journal of Animal Ecology 49:291–315.

Crouse, D. T., B. Crowder, and H. Caswell. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology 68:1412–1423.

Daniels, R. E., and J. D. Allan. 1981. Life table evaluation of chronic exposure to a pesticide. Canadian Journal of Fisheries and Aquatic Sciences 38:485–494.

DeAngelis, D. L., L. W. Barnthouse, and W. Van Winkle. 1990. A critical appraisal of population approaches in assessing fish community health. Journal of Great Lakes Research 16:576–590.

DeAngelis, D. L., L. J. Svoboda, S. W. Christensen, and D. S. Vaughan. 1980. Stability and return times of Leslie matrices with density-dependent survival: applications to fish populations. Ecological Modelling 8:149–163.

de Kroon, H., A. Plaisier, J. van Groenendael, and H. Caswell. 1986. Elasticity: the relative contribution of demographic parameters to population growth rate. Ecology 67: 1427–1431.

Depledge, M. H. 1990. New approaches in ecotoxicology: can inter-individual physiological variability be used as a tool to investigate pollution effects? Ambio 19:251–252.

Elliott, J. M. 1985. Population regulation for different life-stages of migratory trout *Salmo trutta* in a Lake District stream, 1966–83. Journal of Animal Ecology 54:617–638.

Elliott, J. M. 1994. Quantitative ecology and the brown trout. Oxford University Press, Oxford, UK.

Emlen, J. M. 1989. Animal population dynamics: identification of critical components. Ecological Modelling 44:253–273.

Evans, D. O., G. J. Warren, and V. W. Cairns. 1990. Assessment and management of fish community health in the Great Lakes: synthesis and recommendations. Journal of Great Lakes Research 16:639–669.

Ferreri, C. P., and W. W. Taylor. 1996. Compensation in individual growth rates and its influence on lake trout population dynamics in the Michigan waters of Lake Superior. Journal of Fish Biology 49:763–777.

Fogarty, M. J., M. P. Sissenwine, and E. B. Cohen. 1991. Recruitment variability and the dynamics of exploited marine populations. Trends in Ecology and Evolution 8: 241–246.

Forbes, V. E., and P. Calow. 1999. Is the per capita rate of increase a good measure of population-level effects in ecotoxicology? Environmental Toxicology and Chemistry 18:1544–1556.

Forbes, V. E., V. Møller, and M. H. Depledge. 1995. Intrapopulation variability in sublethal response to heavy metal stress in sexual and asexual gastropod populations. Functional Ecology 9:477–484.

Frank, P. W. 1968. Life histories and community stability. Ecology 49:355–357.

Goodyear, C. P. 1978. Entrainment impact estimates using the equivalent adult approach. U.S. Fish and Wildlife Service Biological Services Program FWS/OBS-78/65.

Goodyear, C. P. 1980. Compensation in fish populations. Pages 253–280 *in* C. H. Hocutt, and J. R. Stauffer, editors. Biological monitoring of fish. Lexington Books, Lexington, Massachusetts.

Goodyear, C. P. 1985. Toxic materials, fishing and environmental variation: simulated effects on striped bass population trends. Transactions of the American Fisheries Society 114:107–113.

Goodyear, C. P. 1988. Implications of power plant mortality for management of the Hudson River striped bass fishery. Pages 245–254 *in* L. W. Barnthouse, R. J. Klauda, D. S. Vaughan, and R. L. Kendall, editors. Science, law, and Hudson River power plants: a case study in environmental impact assessment. American Fisheries Society, Monograph 4, Bethesda, Maryland.

Goodyear, C. P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. Canadian Special Publication of Fisheries and Aquatic Sciences 120:67–81.

Gulland, J. A. 1988. Fish population dynamics, 2nd edition. Wiley, New York.

Haddon, M. 2001. Modeling and quantitative methods in fisheries. Chapman and Hall/CRC, Boca Raton, Florida.

Hassell, M. P. 1986. Detecting density dependence. Trends in Ecology and Evolution 4: 90–93.

Hayes, D. B., W. W. Taylor, and H. L. Schramm. 1995. Predicting the biological impact of competitive fishing. North American Journal of Fisheries Management 15:457–472.

Healey, M. C. 1978. Fecundity changes in exploited populations of lake whitefish *(Coregonus clupeaformis)* and lake trout *(Salvelinus namaycush)*. Journal of the Fisheries Research Board of Canada 35:945–950.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London.

Horst, T. J. 1975. The assessment of impact due to entrainment of ichthyoplankton. Pages 107–118 *in* S. B. Saila, editor. Fisheries and energy production: a symposium. Heath, Lexington, Massachusetts.

Horst, T. J. 1977. Use of the Leslie matrix for assessing environmental impact with an example for a fish population. Transactions of the American Fisheries Society 106: 253–257.

Jaworska, J. S., K. A. Rose, and L. W. Barnthouse. 1997. General response patterns of fish populations to stress: an evaluation using an individual-based model. Journal of Aquatic Ecosystem Stress and Recovery 6:15–31.

Jensen, A. L. 1971a. Response of brook trout *(Salvelinus fontinalis)* populations to a fishery. Journal of the Fisheries Research Board of Canada 28:458–460.

Jensen, A. L. 1971b. The effect of increased mortality on the young in a population of brook trout, a theoretical analysis. Transactions of the American Fisheries Society 100:456–459.

Johnson, L. L., and six coauthors. 1998. Assessing the effects of anthropogenic stressors on Puget Sound flatfish populations. Journal of Sea Research 39:125–137.

Karas, P., E. Neuman, and O. Sandstrom. 1991. Effects of pulp mill effluent on population dynamics of perch, *Perca fluviatilis*. Canadian Journal of Fisheries and Aquatic Sciences 48:28–34.

Krebs, C. J. 1999. Ecological methodology, 2nd edition. Addison Wesley Longman, Menlo Park, California.

Krebs, C. J. 2001. Ecology, 5th edition. Benjamin Cummings, San Francisco.

Kuno, E. 1971. Sampling error as a misleading artifact in key-factor analysis. Researches on Population Ecology 13:28–45.

Landahl, J. T., L. L. Johnson, J. E. Stein, T. K. Collier, and U. Varanasi. 1997. Approaches for determining effects of pollution on fish populations for Puget Sound. Transactions of the American Fisheries Society 126:519–535.

Lefkovitch. L. P. 1965. The study of population growth in organisms grouped by stages. Biometrics 21:1–18.

Lenski, R. E., and P. M. Service. 1982. The statistical analysis of population growth rates calculated from schedules of survivorship and fecundity. Ecology 63:655–662.

Leslie. P. H. 1945. On the use of matrices in certain population mathematics. Biometrika 33:183–212.

Leslie, P. H., J. S. Tenor, M. Vizoso, and H. Chitty. 1955. The longevity and fertility of the orkney vole, *Microtus orcadensis,* as observed in the laboratory. Proceedings of the Zoological Society of London 125:115–125.

Lohrding, R. K. 1975. A two-sample test of equality of coefficients of variation or relative errors. Journal of Statistical Computing and Simulation 4:31–36.

Maltby, L. 1999. Studying stress: the importance of organism-level responses. Ecological Applications 9:431–440.

Manly, B. F. J. 1990. Stage-structured populations: sampling analysis and simulation. Chapman and Hall, London.

Marshall, J. S. 1978. Population dynamics of *Daphnia galeata mendontae* as modified by chronic cadmium stress. Journal of the Fisheries Research Board of Canada 35:461–469.

McFadden, J. T. 1977. An argument supporting the reality of compensation of fish populations and a plea to let them exercise it. Pages 153–183 *in* W. Van Winkle, editor. Proceedings of a conference on assessing the effects of power-plant-induced mortality on fish populations. Pergamon Press, New York.

McFadden, J. T., G. R. Alexander, and D. S. Shetter. 1967. Numerical changes and population regulation in brook trout *Salvelinus fontinalis.* Journal of the Fisheries Research Board of Canada 24:1425–1459.

Mendelssohn, R. 1976. Optimization problems associated with a Leslie matrix. The American Naturalist 110:339–349.

Meyers, R. A. 2002. Recruitment: understanding density-dependence in fish populations. Pages 123–148 *in* J. B. Hart and J. D. Reynolds, editors. Handbook of fish biology and fisheries. Volume 1, Fish biology. Blackwell Scientific Publications, Oxford, UK.

Mills, K. H. 1985. Responses of lake whitefish *(Coregonus clupeaformis)* to fertilization of Lake 226, the Experimental Lakes Area. Canadian Journal of Fisheries and Aquatic Sciences 42:129–138.

Mills, K. H., and S. M. Chalanchuk, 1987. Population dynamics of lake whitefish *(Coregonus clupeaformis)* during and after the fertilization of Lake 226, the Experimental Lakes Area. Canadian Journal of Fisheries and Aquatic Sciences 44 (Supplement 1):55–63.

Mills, K. H., S. M. Chalanchuk, and D. J. Allan. 2000. Recovery of fish populations in Lake 223 from experimental acidification. Canadian Journal of Fisheries and Aquatic Sciences 57:192–204.

Mills, K. H., S. M. Chalanchuk, L. C. Mohr, and I. J. Davies. 1987. Responses of fish populations in Lake 223 to 8 years of experimental acidification. Canadian Journal of Fisheries and Aquatic Sciences 44 (Supplement 1):114–125.

Mills, K. H., B. R. McCulloch, S. M. Chalanchuk, D. J. Allan, and M. P. Stainton. 1998. Growth, size, structure, and annual survival of lake whitefish *(Coregonus clupeaformis)* during eutrophication and oligotrophication of Lake 226, the Experimental Lakes Area, Canada. Archiv fuer Hydrobiologie Special Issues and Advances in Limnology 50:151–160.

Minns, C. K. 1992. Use of models for integrated assessment of ecosystem health. Journal of Aquatic Ecosystem Health 1:109–118.

Mohr, L. C., K. H. Mills, and J. F. Klaverkamp. 1990. Survival and development of lake trout *(Salvelinus namaycush)* embryos in an acidified lake in northwestern Ontario. Canadian Journal of Fisheries and Aquatic Sciences 47:236–243.

Morris, R. F. 1959. Single factor analysis in population dynamics. Ecology 40:580–588.

Munkittrick, K. R., and D. G. Dixon. 1989. Use of white sucker *(Catostomus commersoni)* populations to assess the health of aquatic ecosystems exposed to low-level contaminant stress. Canadian Journal of Fisheries and Aquatic Sciences 46:1455–1462.

Nicholson, A. J. 1933. The balance of animal populations. Journal of Animal Ecology 2 (Supplement 1):132–178.

Nicholson, A. J. 1954. Compensatory reactions of populations to stresses and their evolutionary significance. Australian Journal of Zoology 2:1–8.

Odum, E. P., J. T. Finn, and E. H. Franz. 1979. Perturbation theory and the subsidy-stress gradient. Bioscience 29:349–352.

Pianka, E. R., and W. S. Parker. 1975. Age-specific reproductive tactics. American Naturalist 109:453–464.

Podoler, H., and D. Rodgers. 1975. A new method for the identification of key factors from life table data. Journal of Animal Ecology. 44:85–115.

Power, M. 1997. Assessing the effects of environmental stress on fish populations. Aquatic Toxicology 39:151–169.

Power, M. 2002. Assessing fish population responses to stress. Pages 379–429 *in* S. M. Adams, editor. Biological indicators of aquatic ecosystem stress. American Fisheries Society, Bethesda, Maryland.

Power, M., and L. S. McCarty. 1997. Fallacies in ecological risk assessment practices. Environmental Science and Technology 31:370A–375A.

Power, M., and S. M. McKinley. 1997. Using fish population models in hydro project evaluation. Hydro Review 16:36–40, 69–71.

Power, M., and G. Power, 1995. A modeling framework for analyzing anthropogenic stresses on brook trout *(Salvelinus fontinalis)* populations. Ecological Modelling 80:171–185.

Rago, P. J. 1984. Production forgone: an alternative method for assessing the consequences of fish entrainment and impingement losses at power plants and other water intakes. Ecological Modelling 24:79–111.

Rago, P. J., and R. M. Dorazio. 1984. Statistical inference in life-table experiments: the finite rate of increase. Canadian Journal of Fisheries and Aquatic Sciences 41:1361–1374.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada 11:559–623.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Rigler, F. H. 1982. Recognition of the possible: an advantage of empiricism in ecology. Canadian Journal of Fisheries and Aquatic Sciences 39:1323–1331.

Robson, D. S., and D. G. Chapman. 1961. Catch curves and mortality rates. Transactions of the American Fisheries Society 90:181–189.

Roff, D. A. 1992. The evolution of life histories: theory and analysis. Chapman and Hall, New York.

Royama, T. 1996. A fundamental problem in key factor analysis. Ecology 77:87–93.

Ryder, R. A. 1990. Commentary. Ecosystem health, a human perception: definition, detection, and the dichotomous key. Journal of Great Lakes Research 16:619–624.

Ryder, R. A., and C. J. Edwards. 1985. A conceptual approach for the application of biological indicators of ecosystem quality in the Great Lakes basin. Great Lakes Fishery Commission, Windsor, Ontario.

Sandstrom, O. 1994. Incomplete recovery in a coastal fish community exposed to effluent from a modernized Swedish bleached kraft mill. Canadian Journal of Fisheries and Aquatic Sciences 51:2195–2202.

Savidge, I. R., J. B. Gladden, K. P. Campbell, and J. S. Ziesenis. 1988. Development and sensitivity analysis of impact assessment equations based on stock–recruitment theory. Pages 191–202 *in* L. W. Barnthouse, R. J. Klauda, D. S. Vaughan, and R. L. Kendall, editors. Science, law, and Hudson River power plants: a case study in environmental impact assessment. American Fisheries Society, Monograph 4, Bethesda, Maryland.

Schaaf, W. E., D. S. Peters, D. S. Vaughan, L. Coston-Clements, and C. W. Krouse. 1987. Fish population responses to chronic and acute pollution: the influence of life history strategies. Estuaries 10:267–275.

Schnute, J. T., and Richards, L. 2002. Surplus production models. Pages 105–126 *in* J. B. Hart and J. D. Reynolds, editors. Handbook of fish biology and fisheries. Volume 2, Fisheries. Blackwell Scientific Publications, Oxford, UK.

Seyle, H. 1976. The stress of life. McGraw-Hill, New York.

Service, P. M., and M. R. Rose. 1985. Genetic covariation among life-history components—the effect of novel environments. Evolution 39:943–945.

Shepard, J. G. 1988. Fish stock assessments and their data requirements. Pages 35–62 *in* J. A. Gulland, editor. Fish population dynamics, 2nd edition. Wiley, New York.

Shuter, B. J. 1990. Population-level indicators of stress. Pages 145–166 *in* S. M. Adams, editor. Biological indicators of stress in fish. American Fisheries Society, Symposium 8, Bethesda, Maryland.

Shuter, B. J., and H. A. Regier. 1989. The ecology of fish and populations; dealing with interactions between levels. Pages 13–49 *in* C. D. Levings, L. B. Holtby, and M. A. Henderson, editors. Proceedings of the national workshop on effects of habitat alteration on salmonid stocks. Canadian Special Publication of Fisheries and Aquatic Sciences 105.

Smith, R. H. 1973. The analysis of intra-generation change in animal populations. Journal of Animal Ecology 42:611–622.

Southwood, T. R. E., M. P. Hassell, P. M. Reader, and D. J. Rogers. 1989. Population dynamics of the viburnum whitefly *(Aleurotrachelus jelinekii)*. Journal of Animal Ecology 58:921–942.

Southwood, T. R. E., and P. A. Henderson. 2000. Ecological methods, 3rd edition. Blackwell Scientific Publications, Oxford, UK.

Stiling, P. 1988. Density-dependent processes ad key factors in insect populations. Journal of Animal Ecology 57:581–594.

Swanson, S. M., R. Schryer, R. Shelfast, P. J. Kloepper-Sams, and J. W. Owens. 1994. Exposure of fish to biologically treated bleached-kraft mill effluent. 3. Fish habitat and population assessment. Environmental Toxicology and Chemistry 13:1497–1507.

Underwood, A. J. 1989. The analysis of stress in natural populations. Biological Journal of the Linnaean Society 36:51–78.

USEPA (U.S. Environmental Protection Agency) 1992. Framework for ecological risk assessment. USEPA/630/R-92/001, Washington, D.C.

Usher, M. B. 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. Journal of Applied Ecology 3:355–367.

Usher, M. B. 1969. A matrix model for forest management. Biometrics 25:309–315.

van den Heuvel, M. R., M. Power, M. D. MacKinnon, and D. G. Dixon. 1999. Effects of oil-sands related aquatic reclamation on yellow perch *(Perca falvescens)* II. Validation studies of chemical and biochemical indicators of exposure to oil-sands related waters. Canadian Journal of Fisheries and Aquatic Sciences 56:1226–1233.

Van Winkle, W., D. L. DeAngelis and S. R. Blum. 1978. A density-dependent function for fishing mortality rate and a method for determining elements of a Leslie matrix with density-dependent parameters. Transactions of the American Fisheries Society 107: 395–401.

Van Winkle, W., D. S. Vaughan, L. W. Barnthouse, and B. L. Kirk. 1981. An analysis of the ability to detect reductions in year-class strength of the Hudson River white perch *(Morone americana)* population. Canadian Journal of Fisheries and Aquatic Sciences 38:627–632.

Varley, G. C., and G. R. Gradwell. 1960. Key factors in population studies. Journal of Animal Ecology 29:399–401.

Varley, G. C., and G. R. Gradwell. 1963. The interpretation of insect population changes. Proceedings of the Ceylon Association for the Advancement of Science 18:142–156.

Varley, G. C., and G. R. Gradwell. 1968. Population models for the winter moth. Pages 132–42 *in* T.R.E. Southwood, editor. Insect abundance. Blackwell Scientific Publications, Oxford, UK."

Varley, G. C., G. R. Gradwell, and M. P. Hassell. 1973. Insect population ecology. Blackwell Scientific Publications, Oxford, UK.

Vaughan, D. S. 1981. An age structure model of yellow perch in western Lake Erie. Pages 189–216 *in* D. G. Chapman and V. F. Gallucci, editors. Quantitative population dynamics. International Cooperative Publishing House, Fairland, Maryland.

Vaughan, D. S., and S. B. Saila. 1976. A method for determining mortality rates using the Leslie matrix. Transactions of the American Fisheries Society 105:380–383.

Vaughan, D. S., and W. Van Winkle. 1982. Corrected analysis of the ability to detect reductions in year-class strength of the Hudson River white perch *(Morone americana)* population. Canadian Journal of Fisheries and Aquatic Sciences 39:782–785.

Vaughan, D. S., R. M. Yoshiyama, J. E. Breck, and D. L. DeAngelis. 1984. Modeling approaches for assessing the effects of stress on fish populations. Pages 259–278 *in* V. W. Cairns, P. V. Hodson, and J. O. Nriagu, editors Contaminant effects on fisheries. Volume 16, Advances in Environmental Science and Technology. Wiley, New York.

Walters, C. J., J. S. Collie, and T. Webb. 1989. Experimental designs for estimating transient responses to habitat alteration: is it practical to control for environmental interactions?

Pages 13–20 *in* C. D. Levings, L. B. Holtby, and M. A. Henderson, editors Proceedings of the national workshop on effects of habitat alteration on salmonid stocks. Canadian Special Publication of Fisheries and Aquatic Sciences 105.

Walthall, W. K., and J. D. Stark. 1997. Comparison of two population-level ecotoxicological endpoints: the intrinsic (r_m) and instantaneous (r_i) rates of increase. Environmental Toxicology and Chemistry 16:1068–1073.

Werner, P. A., and H. Caswell. 1977. Population growth rates and age versus stage-distribution models for teasel (*Dipsacus sylvestris* Huds.). Ecology 58:1103–1111.

Whittaker, R. H. 1975. Communities and ecosystems. Macmillan, New York.

Wiens, J. A. 1989. The ecology of bird communities. Volume 2, Processes and variations. Cambridge University Press, Cambridge, UK.

Williamson, M. 1984. The measurement of population variability. Ecological Entomology 9:239–241.

Wolda, H. 1983. Seasonality of Homoptera on Barro Colorado Island. Pages 319–330 *in* G. E. Leigh, A. S. Rand, and D. M. Windsor, editors. The ecology of a tropical forest. Oxford University Press, Oxford, UK.

Wootton, R. J. 1990. Ecology of teleost fishes. Chapman and Hall, London.

14 Analysis of Movement and Habitat Use from Telemetry Data

Kevin B. Rogers and Gary C. White

14.1 INTRODUCTION

Telemetry in freshwater environments began in the 1950s (Trefethen 1956; Stasko and Pincock 1977; Mitson 1978), and a fairly extensive body of literature has developed since then. Unfortunately, the literature contains very few rigorous treatments of telemetry data. The data sets generated during a telemetry project quickly become so massive that analyses could not really flourish until personal computers became widespread. Now that personal computers are pervasive, the absence of complex analyses that produce quantitative results cannot be blamed on a lack of processing power. Though fish telemetry projects continue to proliferate, the dearth of quantitative analyses used in them is perhaps due to the absence of a readily available synopsis of the various approaches that can be employed and exposure to the various software packages that have been developed to perform such analyses. Fortunately, sophisticated personal computer programs that run on a variety of platforms are available free of charge from a variety of sources (e.g., White and Garrott 1990; Kenward 1992; Larkin and Halkin 1994; Hooge et al. 2001). Our objective for this chapter is to help provide a foundation for synthesizing some of the many divergent approaches and software that can be used in the analysis of telemetry data.

While we recognize that the emphasis of this book lies with data analysis, tackling telemetry analyses without first carefully considering an adequate study design can lead to untrustworthy results. Perhaps the first question that should be asked regarding proposed telemetry-based research is whether telemetry is really necessary to answer the question of interest. Some have implemented costly and labor intensive telemetry studies because of the high-tech allure when a more mundane approach would have yielded better results and been considerably less labor intensive. For instance, course-scale movement and range extent can often be examined with conventional mark–recapture programs that allow for the tagging of thousands of individuals. However, there are certainly many situations in which telemetry is the only realistic option for addressing the questions of interest. Fine-scale movement and habitat use studies often fall into this latter category.

If one is certain that all other options have been explored, one can find details on telemetry equipment in Henderson et al. (1966), Stasko and Pincock (1977), and Winter (1996). Specific methods of transmitter attachment are discussed in Hart and Summerfelt (1975), Ross (1981), Schramm and Black (1984), and Petering and Johnson (1991). Both radio and ultrasonic transmitters are used in fisheries work, each with its own benefits and limitations. In addition to the standard transmitters that provide horizontal locations of a fish's position, many specialized transmitters exist that can provide additional information. Perhaps the most commonly used specialized transmitters are temperature sensitive (Kelso 1978; Schulz and Berg 1992). Not only is ambient temperature important for bioenergetics work (see Chapter 12), but generally, if the lake is stratified, the depth of the fish can be inferred as well if a temperature profile of the lake has been obtained. Alternatively, several authors have used pressure-sensitive transmitters that measured depth directly (Warner and Quinn 1995; Lee and Bergersen 1996; Baldwin et al. 2002). Numerous authors have explored the use of heart rate telemetry (Priede and Young 1977; Armstrong et al. 1989; Lucas et al. 1991, 1993) and tail beat frequencies (Stasko and Horrall 1976; Ross et al. 1981; Johnstone et al. 1992) to try to assess energy budgets of host fish. Because heart rate can be affected by stimuli other than exercise, electromyogram transmitters have become increasingly popular for indirectly assessing oxygen consumption (Weatherley et al. 1982) and swimming speed and activity (Demers et al. 1996; Økland et al. 1997). Because these specialized techniques generally require specialized analyses, they will not be discussed here. Interested readers should consult the literature cited for each topic.

14.2 FUNDAMENTAL CONSIDERATIONS IN TELEMETRY-BASED RESEARCH

14.2.1 Representative Samples

Paramount to any telemetry project is the underlying assumption that those animals carrying transmitters are representative of the population as a whole. Historically, this has been evaluated subjectively as in "the animals appeared normal." More rigorous treatment of this assumption is warranted, however, as the repercussions can be severe if this assumption is violated. If, for example, the implantation procedure resulted in substantial internal infections of most fish that received transmitters and infection induced extreme lethargy, one would expect these fish to move significantly less than their untelemetered counterparts. If movement and habitat use by telemetered fish is uncharacteristic, then there is little point in conducting the study, as it will be uninformative about the population of interest. Unfortunately, this dilemma is difficult to resolve and is usually complicated by the absence of controls (Doerzbacher 1980). The assumption is difficult to test because telemetry is used presumably to gather information that one cannot obtain any other way, making it difficult to compare data with untagged controls. Numerous studies have been conducted with dummy

transmitters in laboratory settings and have failed to document negative effects on growth, feeding, condition, or swimming behavior (e.g., Moore et al. 1990; Martin et al. 1995; Swanberg and Geist 1997; Brown et al. 1999; Cote et al. 1999; Cooke and Bunt 2001). The program MARK (White and Burnham 1999) can be used to compare survival rates between fish carrying transmitters and those monitored with conventional tag recovery studies, allowing estimation of the impact of the transmitter on survival of telemetered fish compared with fish tagged in a conventional manner. Now that transmitter battery life has been greatly enhanced, merely documenting survival of fish carrying transmitters over an extended period of time may lend some support to the notion that they are behaving normally. Additionally, if no differences in growth or condition can be detected in the same waters between telemetered fish and conspecifics tagged with conventional methods, then transmitters are probably not negatively affecting the host fish.

There are several things one can do in order to minimize the potential negative influence of the transmitter, including using the smallest possible transmitter that will still allow collection of the required data. Some have advocated keeping the weight of the transmitter to less than 2% of recipient's body weight (Gallepp and Magnuson 1972; Ross and McCormick 1981), although Brown et al. (1999) found that transmitters weighing up to 12% of the body weight did not affect swimming performance. The transmitters, if not internal, should be inconspicuous and unobtrusive. Finally, data acquisition should be delayed until animals have had a chance to become accustomed to the extra ballast afforded by the transmitter, so that fish behavior can become representative. Activity patterns, growth, and condition may be abnormal at least 2 weeks following surgery (Smith 1974; Manns and Whiteside 1979; Knights and Lasse 1996; Paukert et al. 2001). If resources permit, monitoring the behavior and health of captive fish carrying transmitters in a controlled setting can also be helpful.

Additionally, one must be concerned that the approach used to capture fish will not lead to an unrepresentative sample and thereby biased statistics. For instance, gill nets sample moving fish more than they do stationary ones. If all fish used in the study were acquired from gill-net sets, one might run the risk of implanting transmitters in a more mobile subpopulation, biasing movement estimates for the population as a whole. The radio-marked sample must be representative of the entire population if correct inferences from radio-marked fish are to be applied to the entire population.

Timing of the study and data collection is critical as well; fish should be observed over a time frame that is representative of the question of interest. One's inference regarding behavior patterns is limited to the seasons in which one observed the fish. If one is exploring habitat use but only locates telemetered fish in the middle of the day, then one cannot infer what habitats these same fish use at night. This notion seems trivial, but telemetry studies have often extrapolated behavior patterns outside the window of observation. To do so is simply not supported by the data.

14.2.2 Data Format

Analysis of telemetry data requires that information be stored in a readily usable digital database because rigorous treatment of telemetry data cannot be addressed without a computer. The pervasiveness of geographic information systems (GISs) has greatly facilitated the analysis of spatial telemetry data (Rogers and Bergersen 1996) for both map generation and analysis. Conventional spreadsheet programs (e.g., Microsoft's Excel or Corel's Quattro Pro) are also very capable platforms for compiling data and performing some of the more rudimentary data analyses.

Basic telemetry data are at least three-dimensional (x and y in space and z in time), so that in addition to the tag number and date and time, some metric of horizontal position must be recorded in the database. A variety of mapping systems have been used for describing position, such as the township-range land-mapping system, the latitude–longitude system, and the Universal Transverse Mercator (UTM) system. The township-range land-mapping system is difficult to work with, and because of survey errors, distances between locations cannot be computed reliably (White and Garrott 1990). The latitude–longitude system is not rectangular, making calculation of the distance between any two points less intuitive. The UTM system is the coordinate system of choice for mapping locations in inland waters because subsequent analysis is so much easier than it is in a latitude–longitude system. The UTM approach is based on the metric system, which is the universal standard for science. It provides a Cartesian coordinate system within each zone, allowing easy calculation of distances between points and greatly simplifying triangulation to locate animals. Now that global positioning systems (GPSs) are widely used in telemetry projects, obtaining UTM coordinates is trivial as well.

In the UTM system, the world is divided into 60 zones between 80°S latitude and 84°N latitude that each span 6° of longitude and are bisected by a central meridian. The polar ice caps are not included because the width of the zones becomes zero at the North and South poles. In addition, the coordinate system is not continuous because coordinates from a spherical surface cannot be plotted on a Cartesian coordinate system without breaks. Though this projection spans the globe, its power lies in mapping finer scales because error and distortion increase for regions that cover more than one zone. The UTM coordinates consist of a zone descriptor and two seven-digit numbers with units in meters. The y-coordinate increases with distance from the equator. The x-coordinate describes the distance from the central meridian. Five hundred thousand meters (3° longitude) are added to each x-coordinate within each zone so that negative x-values don't result. As one heads east within a zone, x-coordinates increase; y-coordinates increase as one heads north.

14.2.3 Study Design

There are basically three kinds of telemetry studies that can be conducted. The first are exploratory descriptive studies that are very common in the early literature.

Usually there is no attempt at formulating testable hypotheses. Some include evaluations of home range size and movement, but they are limited to learning about what an animal does but not why (Sanderson 1966; White and Garrott 1990). Descriptive studies that simply map fish locations are of limited value. The second variety of telemetry studies are correlative in nature and are becoming increasingly prevalent. These studies try to link movement or habitat use to environmental features that may be important to the well being of the fish. Although relationships can be documented, they do not necessarily imply cause and effect. Manipulative experiments with both spatial and temporal controls make up the third type of telemetry study. These are the only ones that can establish why animals do what they do and are therefore preferred.

Careful study design is critical if useful results are to be obtained. The notion that analyses are restricted to summarizing data that has already been gathered is a major misconception. Consideration of analysis goals should precede data collection to ensure that the study is worthwhile and to ensure that the appropriate data will be collected (Kenward 1992). If the parameters to be estimated or hypotheses to be tested are not defined, then an optimum strategy for data collection cannot be formulated and sample sizes cannot be determined. Sample size and power calculations (see Chapters 1 and 3) are mandatory before initiating a study because they determine whether or not proposed experiments will adequately address the questions that are posed. Either literature searches or pilot experiments can provide the background information needed to make these calculations. Simply employing the latest technology does not guarantee that quality research will follow. If funding or time constraints will prevent one from achieving the sample sizes necessary to detect a biologically significant difference, then there is no point in conducting the study. Unfortunately, cost and labor must be factored in, as transmitters are expensive and monitoring them has historically been labor intensive. Precision of estimates should be adequate to answer the questions being addressed, even if attaining adequate precision means focusing on a narrower set of questions, such as a portion of the population (e.g., one species, one sex, or a single basin). To answer one question well is far better than to address many questions poorly. Experiments with low sample sizes can lead to variable results with little power to detect differences in metrics of interest.

14.2.4 Pseudoreplication

The independence of successive observations of a given fish has always been a big concern in telemetry work (Byers et al. 1984; Swihart and Slade 1985; Alldredge and Ratti 1986; Thomas and Taylor 1990; Cresswell and Smith 1992). The closer in time two locations occur, the more likely they are to be autocorrelated. Autocorrelation has received much attention because many of the statistical techniques used in telemetry work require that observations be independent. The assumption of independence is violated on two levels when locations are taken as the sampling unit and records for all individuals are combined or pooled (Aebischer et al. 1993), as is common in fish telemetry analyses. If locations are taken to be

the sampling unit, then points close in time are serially correlated (Swihart and Slade 1985). In addition, if locations are pooled across individuals, then the natural heterogeneity found between individuals is eliminated, resulting in statistical tests that yield significant results more often than they should. Pooling data across individuals is only justified if all individuals being monitored act similarly (Aebischer et al. 1993), which is rarely the case in natural systems.

This violation of observation independence, however, is largely an artifact of how the telemetry data are analyzed. If inferences are to be made about the population of animals, the experimental units in a telemetry project are the individual animals, not the individual location estimates (Aebischer et al. 1993; Winter 1996; Otis and White 1999). Though precise estimates of an animal's movement can be achieved by intensive sampling of an individual, we are generally interested in how the population as a whole behaves. Hypothesis tests should use the variation among individuals to assess significant effects. As such, the power of a given study will then be driven by the number of fish monitored more than the number of locations obtained for each fish (Otis and White 1999). Using each location as a sampling unit while analyzing telemetry data is pseudoreplication and is the equivalent of "statistical malpractice" (Hurlbert 1984). The perceived dilemma associated with autocorrelation is mitigated when the data are analyzed correctly and experimental unit assignment is restricted to individual fish only (Kenward 1992). In fact, serial correlation between locations of an individual fish is not necessarily a bad thing if sampling is representative, as increased sampling effort (resulting in points closer in time) will better describe what an animal is actually doing (Aebischer et al. 1993). The emphasis on obtaining uncorrelated location estimates has been misguided and has allowed the more egregious violation of pseudoreplication to persist.

14.3 ESTIMATION OF TELEMETRY ERROR

Telemetry error is introduced into fish telemetry projects from two sources. The first potential introduction or error occurs when trying to determine the actual position of a telemetered fish. Unlike radio telemetry, ultrasonic telemetry can make this source of error negligible, especially in small basins. The second potential introduction occurs when converting that fish position to a pair of coordinates that can be used to plot the location on a map. The advent of GPS technology has greatly simplified this task. If fine-scale habitat work is needed, then one might consider setting up a differential GPS (Rogers and Bergersen 1996). By deploying a fixed GPS base station, one can subtract erroneous deviations in position recorded on the stationary unit from positions registered on a mobile unit to achieve added accuracy. For deeper-dwelling fishes living in lentic systems, one can maneuver the monitoring boat directly over the fish while obtaining a GPS reading (Wilkerson and Fisher 1997; Paukert and Fisher 2000). The primary concern with this approach in shallow systems is that continual harassment of the fish may actually alter the very behavior that one is hoping to monitor. In these situations, it may be advisable to maintain some distance between the observer and the

subject by employing a triangulation technique to determine fish location (Springer 1979; Lenth 1981; Nams 1989; Saltz and White 1990; White and Garrott 1990). For both approaches, it is imperative that experiments are conducted to verify the methodology and evaluate what accuracy can be expected (Box 14.1).

In addition to legitimate sources of error, simple data entry errors can dramatically alter the results of an analysis. It is imperative that the data sets are first subjected to algorithms that can identify potentially erroneous data points (White and Garrott 1990). Dates and times should increase chronologically, and movement between fish locations should be assessed to ensure that they are reasonable. Finally, the location of a fish should be reasonable as well (e.g., if a recorded observation has a fish on dry ground, then the raw data forms should be reviewed).

14.4 SPATIAL DISTRIBUTION

The first step in the analysis of telemetry data should be the creation of location maps. With the rapid expansion of GIS technology, the generation of maps is trivial and greatly facilitates subsequent analyses. Generally, the time dimension is eliminated, and fish locations are overlaid on a map of the perimeter first, then on maps of other habitat features (e.g., depth, habitat type, and temperature) to evaluate use. Many of the subsequent analyses of telemetry data will focus on fusing spatial information into one dimension, which results in a loss of information. Often an examination of the raw data can be more revealing in terms of illustrating the importance of habitat features that do not show up in conventional summary statistics. This examination can lead to a better understanding of the system as a whole. In many packages, animated graphics can reintroduce the time dimension by portraying movement through time. Although it can be difficult to do quantitative work with these graphics, they can be instructive when combined with maps of habitat features.

Of interest is whether the distribution of fish locations on the map is random. Samuel and Garton (1985) propose a Cramér–von Mises statistic to test whether the distribution of fish locations follows a bivariate uniform distribution. White and Garrott (1990; their Appendix 7) provide SAS code to perform the test. Another conceptually simple approach to evaluate randomness compares the distribution of the fish locations to an equal number of random locations placed in the same basin (Rogers 1998), which is analogous to a randomization test. One metric to use as a test statistic would be the variance of the distances from each random point to the nearest fish location. If the distribution of the fish locations is fairly random, then the distances between fish locations and random locations should not be highly variable. On the other hand, if fish locations are highly clustered, then distances from each random point will be highly variable, depending on whether the point is near or far from the cluster. The entire process of generating the variance of these distances is then iterated 1,000 times to generate a mean variance estimate that is then used as a test statistic (Figure 14.1). Replacing the fish locations with an equal number of randomly placed "pseudolocations" generates the distribution of this test statistic. This process should be iterated at

Box 14.1 Estimation of Telemetry Error

Prominent landmarks were used to assess the location of 31 largemouth bass carrying ultrasonic transmitters in two small impoundments near Denver, Colorado, over a 4.5-year period (Rogers 1998). Because these landmarks were associated with the shoreline, it was assumed that error would be a function of distance to shore. This error was quantified by locating 50 random points on each lake's surface by means of prominent landmarks and by means of a differential global positioning system (GPS) with better than 1-m accuracy. If the GPS readings reflect true position, then error associated with using prominent landmarks was taken to be the distance between each pair of locations generated for each point by the two methods.

Distance from the 50 random locations to shore was calculated with the program FishTel (available at http://wildlife.state.co.us/Research/Aquatic/Software/) and a digitized map of the lake perimeter. Distance between each pair of locations (prominent landmark and GPS) was generated in a simple spreadsheet by use of the Pythagorean relation. As expected, error generally increased with distance from shore (see figure) with fitted regressions for both lakes having positive slopes ($P = 0.002$ in Lake Ladora and $P \leq 0.001$ in lower Derby Lake). Because fish tended to congregate offshore in winter seeking deeper warmer water, this analysis suggested that greater error was incorporated in estimates of winter movement than those movement estimates generated in summer.

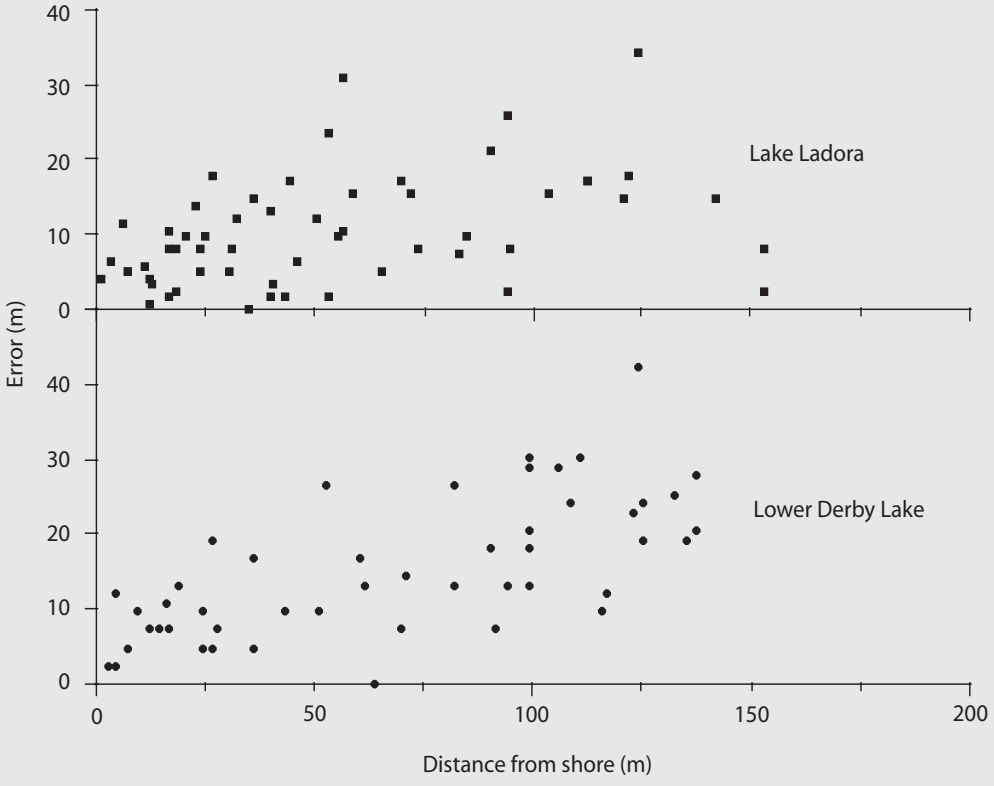


Figure Fifty random coordinates were located conventionally and with a differential GPS in each lake. The distance between these points was assumed to be a measure of telemetry error (in meters). This error was then plotted as a function of distance from shore in meters (from Rogers 1998).

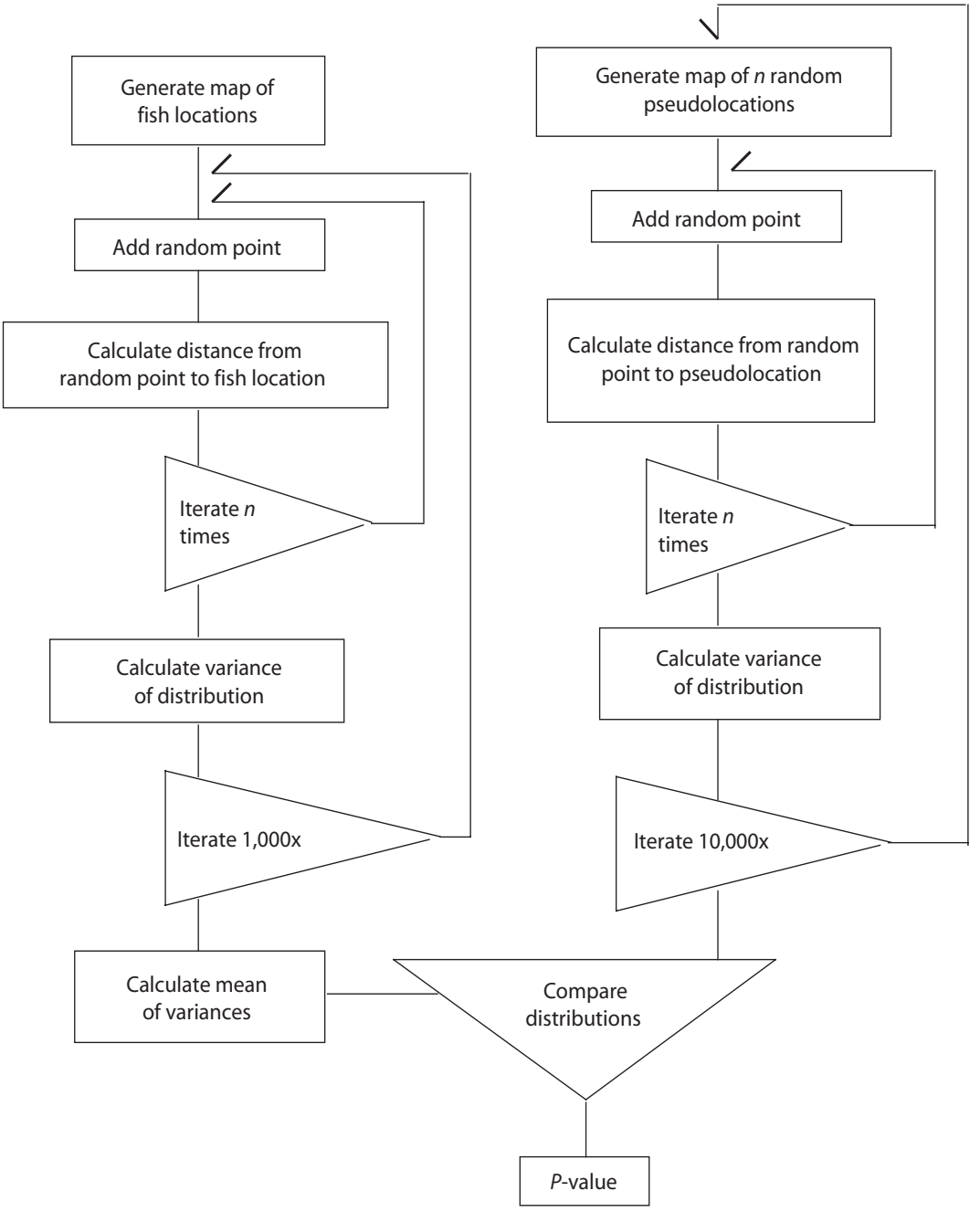


Figure 14.1 A flowchart demonstrating the process for evaluating if *n* fish locations are distributed randomly around a lake. The left arm calculates the mean of variance estimates for the each telemetered fish, whereas the right arm is used to generate the distribution of the test statistic so that a *P*-value can be assessed. If randomness of the population of fish is to be evaluated, then a grand mean can be calculated and compared to the same test statistic distribution.

least 1,000 times to generate a distribution for the expected mean variance statistic under the null hypothesis of random distribution. The actual mean variance statistic generated from the fish location data can then be compared with this distribution to determine the appropriate *P*-value (Box 14.2). Comparisons can be made between individuals, seasons, or basins, the only caveat being that the same number of fish locations must be considered in each group. If not, the group with more locations will display shorter distances on average than those groups with sparser data. A quick way to remedy this dilemma is simply to have a computer randomly drop fish locations out of each group to be compared until all groups have the same number of locations (Rogers 1998).

Usually, one is more interested in how the locations of all telemetered fish are distributed (population level). This can be achieved by averaging the mean variance for each fish across the population of telemetered fish to generate a grand mean variance that can then be compared with the sampling distribution (see figure Box 14.2) to yield a *P*-value. For this approach to be valid, one only has to ensure that all fish are located the same number of times. As mentioned previously, that goal can be easily achieved by randomly dropping observations from each data set until all fish register the same number of locations. Several alternative approaches are possible and could be substituted for this one based on the following considerations. The mean distance to the nearest fish location should be the same as the mean distance to a random location if fish locations are randomly distributed. Alternatively, the probability that the nearest location is a fish location should be 0.5 if fish locations are randomly distributed.

14.5 MOVEMENT PATTERNS

Evaluations of movement by telemetered fish are pervasive in the literature, yet it is very difficult to compare metrics across studies because observed movement is a function of how often a fish was located (Baras 1998). Estimates of fish movement are minimum estimates of displacement. Fish do not move in a straight line (Guy et al. 1994; Rogers and Bergersen 1995), so the more times a fish is located in a day, the greater total movement will appear. Better estimates of true movement can therefore be obtained if continuous tracking schedules are employed or if fish are at least monitored frequently over a 24-h period.

If one wishes to compare movement between studies or lakes, or even within a population, it is critical that subjects be located the same number of times and the time interval between locations be approximately equal. Because fish movement is usually heterogeneous among individuals (Rogers 1998), if fish are not located consistently over a given time frame, contact bias can result (Jones and Rogers 1998). This bias can dramatically affect the outcome of a study if search routines are not rigorously applied to give all telemetered fish the best chance of being located every time they are sought. If more mobile fish tend to spend more time in open water while moving, they might be located more often than are their sedentary brethren. Movement estimates for the population would therefore be biased upward. If, on the other hand, sedentary fish are easier to locate because

Box 14.2 Evaluation of Spatial Distribution

In this example, we will explore the distribution of northern pike in Lake Ladora, Colorado, in summer (Rogers 1998). Colorado's plains' reservoirs represent the southern limit of this species' range, as water temperatures frequently approach the species' tolerance limits in summer. A spring that feeds a slough of this lake ensures that water temperatures in this shallow arm are cooler than the remainder of the lake basin. As the northern pike begin to experience thermal stress, they congregate in this region despite the marginal habitat that is available (Rogers 1998).

To quantify the nonrandom nature of this spatial distribution, we will examine the locations of a northern pike (number 96) that was observed on 43 occasions over a 3-month period in Lake Ladora. A data file containing the *x*–*y* positions on a Universal Transverse Mercator (UTM) grid were read by the program FishTel, which performs the functions described in Figure 14.1. The mean variance statistic generated for this fish was 6,095 m^2. Using the spatial test statistic module of program FishTel and 43 random pseudolocations, a distribution of the test statistic was generated under a null hypothesis of random distribution (see figure).

The probability of obtaining a mean variance value of 6,095 m^2 or larger by chance was remote ($P = 0.008$). The distribution of northern pike 96 observations was therefore nonrandom.

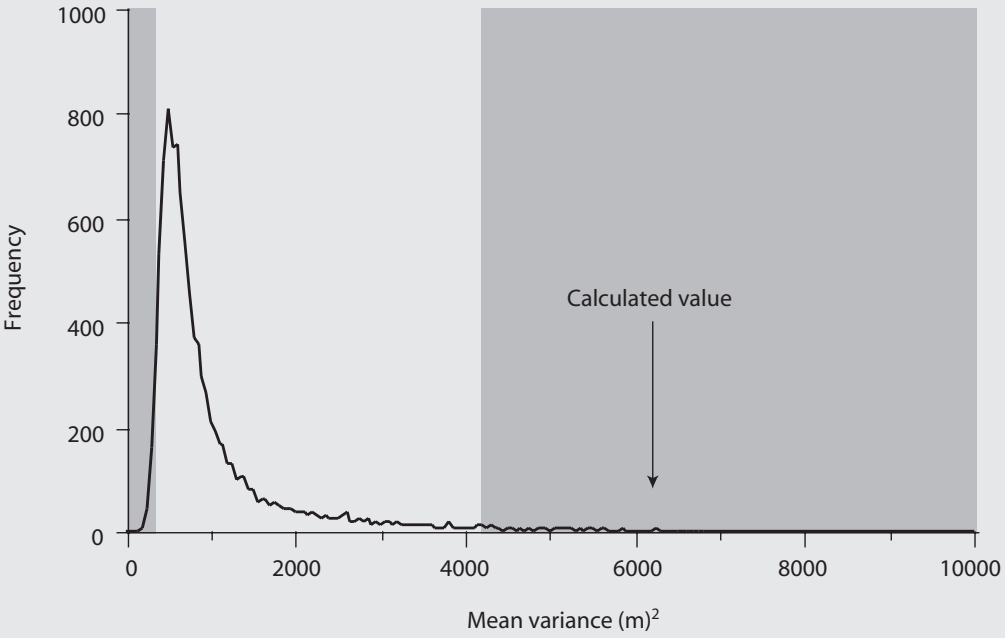


Figure The variance in distances from 43 random points to 43 randomly generated pseudolocations was calculated and iterated 10,000 times to generate a distribution of the mean of variance estimates under a null hypothesis of random distribution. The *P*-values for mean variance statistics calculated from fish locations can be assessed by determining where that value falls on this graph and evaluating what percent of the area under the curve falls to the right (or left) of the value calculated for the fish. The area under the curve in the shaded regions represents a two-tailed alpha of 0.05 (2.5% of the observations under each tail).

one already knows about where they are from the previous observation, then apparent movement may be biased downward.

Units of measure should reflect the precision that was conferred when the data were acquired. For instance, if fish were located numerous times over a 24-h period, then it might be fair to express movement as minimum displacement per hour (MDPH). If fish were located only several times per week, then movement should not be expressed in units of finer resolution than meters per day. The common practice of expressing fish movement in centimeters per second is usually absurd and implies precision that is simply not possible. Positional information cannot be accurately measured in centimeters, and reporting movement in seconds suggests that fish were followed (and locations recorded) continuously, which is rarely the case.

Quite often, as the values that describe various aspects of telemetry data, such as MDPH, increase in magnitude, the variance associated with those values increases as well. Because increasing variance conflicts with the equal variance assumption of many of the parametric statistics used in telemetry analysis, the data may first need to be transformed. A simple natural log transformation is usually all that is required (Jones and Rogers 1998; Rogers 1998) because the variance is a constant multiple of the mean. That is, the coefficient of variation is constant, and the natural log transformation creates a new variable that has constant variance. Other options exist (Ott 1988; Wilkerson and Fisher 1997). The important thing is to check to make sure that the variances have stabilized, and that they are no longer increasing with the metric of interest before the analyses are conducted.

14.5.1 Evaluation of Movement with Categorical Data

General linear models (GLMs) are very conducive to exploring variation in movement and allow for a broad spectrum of relationships to be examined (Ott 1988; Box 14.3). Heterogeneity among fish, or between seasons, years, and diel periods, can easily be evaluated in this fashion. Fish do not expend energy for swimming needlessly, so relationships between movement and other factors give some insight into which factors are associated with fish movement. As with all analyses presented here, it is important to remember that the fish carrying transmitters are the sampling units (Otis and White 1999).

14.5.2 Evaluation of Movement with Continuous Data

A variety of parameters are thought to influence movement. Because fish are poikilothermic, they tend to move less in winter when their metabolisms slow down (Casselman 1978; Cook and Bergersen 1988; Schulz and Berg 1992; Rogers 1998; Snedden et al. 1999; Bramblett and White 2001). Rogers and Bergersen (1995) documented more movement during a mild reservoir water level drawdown. Several authors have shown barometric pressure to influence fish move-

Box 14.3 Evaluation of Seasonal Movement Patterns

Variation in movement by largemouth bass in Lake Ladora, Colorado, is evaluated with a mixed-model analysis of variance (ANOVA). Seventeen largemouth bass carried transmitters during a 4.5-year study, and their movement (m/h) was evaluated on 49 consecutive full-moon 24-h tracks (Rogers 1998). Analyses were conducted with SAS version 6.

Table Portion of data set of minimum displacement per hour (m/h) by 17 largemouth bass in Lake Ladora. Fish represents unique largemouth bass; year is the last digit in 199*; seasons are given as (1) spring (March–May), (2) summer (June–August), (3) fall (September–November), and (4) winter (December–February); and minimum displacement per hour (MDPH) is given in m/h. The full data set is included in the Chapter 14 CD folder.

Fish	Year	Season	MDPH
1	4	2	18.74
1	4	3	34.55
1	4	3	32.05
1	4	3	17.76
1	4	4	0.68
1	5	4	4.83

The following SAS program is used to evaluate whether largemouth bass move differently between seasons and years (fixed effects), with the sampling unit of individual fish taken as a subject or random effect. Interactions between fish and season and fish and year are also treated as random effects. Notice that MDPH has been transformed by the natural logarithm to stabilize increasing variance found with increasing MDPH. Bonferroni's multiple-comparisons procedure was used to assess differences in MDPH between seasons.

Program

```
Options nodate ps = 40;
data LMB;
input FISH YEAR SEASON MDPH;
LNMDPH = LOG(MDPH);
proc mixed;
CLASS FISH YEAR SEASON; * These variables are categorical;
MODEL LNMDPH = SEASON YEAR; * Fixed effects in the model;
RANDOM FISH*SEASON FISH*YEAR/subjects = fish;
/* Above specifies which effects in the model are random */
lsmeans SEASON/ADJUST = BON;
run;
```

(Box continues)

Box 14.3 *(continued)*

Descriptive Output

Table Mixed model output (partial) to evaluate whether largemouth bass move differently between seasons and years. Akaike's Information Criteria (AIC), small-sample corrected AIC (AICc), and Bayesian Information Criteria (BIC) are model fit statistics, for which smaller values reflect better fit (see Box 14.4 for additional explanation of AIC).

Covariance Parameter Estimates

Covariance parameter	Subject	Estimate
Fish*Season	Fish	0.04282
Fish*Year	Fish	0.07549
Residual		0.6042

Fit Statistics

–2Residual log likelihood	414.4
AIC	420.4
AICc	420.5
BIC	422.9

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	*P* > *F*
Season	3	45	7.00	0.0006
Year	4	10	0.98	0.4629

ment (Warden and Lorio 1975; Markham et al. 1991; Guy et al. 1992; Jones and Rogers 1998). Jones and Rogers (1998) also documented a link between water clarity and movement. A simple way to explore the influence of these continuous variables is to model them against MDPH in a multiple regression framework (Box 14.4).

The Akaike's Information Criteria (AIC) model selection procedure (Burnham and Anderson 1998) provides an elegant way to resolve which variables are really important in describing movement. However, one should ensure that fish used in this sort of analysis are exposed to a biologically meaningful cycle or spectrum of values for each variable included in the model. One might not detect a significant relationship between water temperature and movement if fish were followed for only a couple of months when water temperatures were stable, and that result might have little biological relevance. Failure to detect a relationship does not imply that no relationship exists, especially in this scenario.

Least-Squares Means

Effect	Season	Estimate	SE	*df*	*t*-value	$P > \|t\|$
Season	1	2.5819	0.1642	45	15.72	<0.0001
Season	2	2.8455	0.1563	45	18.20	<0.0001
Season	3	2.7107	0.1750	45	15.49	<0.0001
Season	4	1.9358	0.1686	45	11.48	<0.0001

Differences of Least Squares Means

Effect	Season–season	Estimate	SE	*df*	*t*-value	$P > \|t\|$	Adjusted P[a]
Season	1–2	–0.2636	0.2082	45	–1.27	0.2121	1.0000
Season	1–3	–0.1288	0.2180	45	–0.59	0.5577	1.0000
Season	1–4	0.6461	0.2057	45	3.14	0.0030	0.0179
Season	2–3	0.1348	0.1974	45	0.68	0.4982	1.0000
Season	2–4	0.9096	0.2105	45	4.32	<0.0001	0.0005
Season	3–4	0.7748	0.2186	45	3.55	0.0009	0.0056

[a] Bonferroni adjustment for experimentwise error rate.

Interpretation

From this output (type 3 tests of fixed effects), it is clear that movement observed is very different between seasons ($P < 0.001$) but not from year to year ($P = 0.463$). The conservative Bonferroni multiple-comparisons test (controls experimentwise error rate) demonstrates that movement in winter (4) was significantly less than during the ice-free seasons.

14.6 HABITAT USE

Interest in evaluating the habitat used by fishes has been a cornerstone of telemetry projects. The overriding question is whether fish spend more or less time in some habitats than would be expected based on the availability of those habitats. When this disproportionate allocation of time occurs, the behavior is said to be selective. Although selection and preference are often used synonymously, selection is the process by which an animal chooses a habitat. Preference is the likelihood that a resource will be chosen if all habitats are offered up equally (Johnson 1980; Manly et al. 1993). Animals are presumed to use habitats that confer fitness, so by studying habitat use biologists can hope to assess what habitat features may be limiting. By studying where animals allocate their time, one can gain insight into how they meet their requirements for survival. Such information is useful when considering the introduction of a species and

Box 14.4 Evaluation of Environmental Effects on Movement

The potential influence of water temperature, surface elevation, barometric pressure, and change in barometric pressure are modeled against MDPH (natural log transformed) of northern pike in a Colorado reservoir. Fish in this population were monitored on consecutive full-moon 24-h tracks for at least 1 year (Rogers 1998). A mixed-model ANOVA with a random coefficients model (Littell et al. 1996) and an AIC model selection (Burnham and Anderson 1998) was used to isolate significant effects. Akaike's Information Criteria is a useful tool for selecting the model that most closely fits the theoretical distribution of the data without overparameterization (Schisler and Bergersen 1996). This metric can be thought of as the relative distance between pairs of multiple candidate models (Burnham and Anderson 1998) and allows model selection to occur in an optimization framework similar to parameter estimation. By formulating the problem of model selection across a set of candidate models, AIC provides an objective means for the selection of the best approximating model for inference (Burnham and Anderson 1998), as well as allowing the user to rank other candidate models. This minimizes the practice of data dredging and overfitting models and provides an alternative to the traditional null hypothesis testing approach. For a complete analysis of these data, additional models would likely have been considered a priori, and values of AIC would be useful in selecting the most parsimonious model supported by the data.

Table Data for northern pike being located during a 24-h track in lower Derby Lake, Colorado (partial data set). Provided are the individual fish identification (Fish ID), the water temperature (temp, °C), the maximum lake depth (depth, m) for that date, the mean barometric pressure (BP, mm), the change in barometric pressure (DBP, mm) over the 24-h period, and the mean minimum displacement (MDPH, m/h). The full data set is included in the Chapter 14 CD folder.

Fish ID	Temp	Depth	BP	DBP	MDPH
2	21	4.4	630	–1.8	31.22
2	25	5.0	630	–2.3	43.25
2	23	4.6	634	0.2	28.06
2	27	4.1	630	–1.3	23.21
2	21	3.7	633	–2.8	16.72
2	12	3.5	630	–1.3	38.52

The following SAS program evaluates whether fish movement is correlated with the environmental variables listed. Again, we must transform the MDPH to stabilize increasing variance found with increasing MDPH. As AIC values will be used to isolate the most parsimonious model, all possible models should be entertained (or at least all that make biological sense).

Program

```
OPTIONS NODATE PS = 40;
data NOP;
INFILE 'DNOPENV.TXT' FIRSTOBS = 3;
input FISH TEMP DEPTH BP DBP MDPH;
LNMDPH = LOG(MDPH);
PROC MIXED DATA = NOP;
```

```
   CLASS FISH;
   MODEL LNMDPH = TEMP DEPTH BP DBP/Solution;
   /* Prints the solution to the model */
   RANDOM INTERCEPT TEMP DEPTH BP DBP/TYPE = VC SUBJECT = FISH;
   /* TYPE = VC sets the Variance-Covariance for each subject to a variance
components type*/
   MAKE 'FitStatistics' OUT = ModelFit;
   /* Saves the AIC value into the file MODELFIT */
   DATA ModelSelection;
   LENGTH MODEL $ 18;
   SET ModelFit;
   MODEL = 'TEMP DEPTH BP DBP';
   /* The above DATA step adds the model name to the file
   Now the process is repeated for another model */

   PROC MIXED DATA = NOP;
   CLASS FISH;
   MODEL LNMDPH = TEMP/Solution;
   RANDOM INTERCEPT TEMP/TYPE = VC SUBJECT = FISH;
   MAKE 'FitStatistics' OUT = ModelFit;
   DATA ModelFit;
   LENGTH MODEL $ 18;
   SET ModelFit;
   MODEL = 'TEMP';
   PROC APPEND BASE = ModelSelection DATA = ModelFit;
   /* Append the new model's statistics to the file MODELSELECTION */

   PROC MIXED DATA = NOP;
   CLASS FISH;
   MODEL LNMDPH = TEMP DEPTH/SOLUTION;
   RANDOM INTERCEPT TEMP DEPTH/TYPE = VC SUBJECT = FISH;
   MAKE 'FitStatistics' OUT = ModelFit;
   DATA ModelFit;
   LENGTH MODEL $ 18;
   SET ModelFit;
   MODEL = 'TEMP DEPTH';
   PROC APPEND BASE = ModelSelection DATA = ModelFit;
```

Iterate last 10 lines here for all models entertained.

```
   PROC SORT DATA = ModelSelection;
   BY Value;
   /* Sort by AIC value */
   PROC PRINT DATA = MODELSELECTION;
   WHERE DESCR = 'AICC (smaller is better)';
   /* Print the summary table of sorted AICC values */
   RUN;
```

(Box continues)

Box 14.4 *(continued)*

Descriptive Output

Table Small-sample corrected Akaike's Information Criteria (AICc) model selection for dependent variable $\log_e$MDPH (LNMDPH). Smaller values of AIC, AICc, and BIC reflect better fit. Results for two best models are shown.

Selection method

Model	AICc value
TEMP	384.4
TEMP DEPTH	387.5
TEMP BP	388.3
TEMP DEPTH BP	389.7
TEMP DBP	390.6
TEMP DEPTH DBP	393.6
TEMP BP DBP	394.4
TEMP DEPTH BP DBP	395.8
DEPTH	403.4
DEPTH BP	409.1
DEPTH DBP	409.4
DBP	409.9
BP	410.4
DEPTH BP DBP	415.1
BP DBP	416.3

Covariance Parameter Estimates

Covariance parameter	Subject	Estimate
Intercept	Fish	0.1816
Temp	Fish	0
Residual		0.5929

Fit Statistics

–2Residual log likelihood	380.4
AIC	384.4
AICc	384.4
BIC	384.8

Solution for Fixed Effects

Effect	Estimate	SE	*df*	*t*-value	$P > \|t\|$
Intercept	2.4878	0.1839	8	13.53	<0.0001
Temp	0.03906	0.007112	8	5.49	0.0006

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	*P* > *F*
Temp	1	8	30.17	0.0006

Covariance Parameter Estimates

Covariance parameter	Subject	Estimate
Intercept	Fish	0.1655
Temp	Fish	0
Depth	Fish	0.003149
Residual		0.5870

Fit Statistics

–2Residual log likelihood	381.4
AIC	387.4
AICc	387.5
BIC	388.0

Solution for Fixed Effects

Effect	Estimate	SE	*df*	*t*-value	*P* > |*t*|
Intercept	3.0604	0.5955	8	5.14	0.0009
Temp	0.04392	0.008580	8	5.12	0.0009
Depth	–0.1517	0.1501	8	–1.01	0.3417

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	*P* > *F*
Temp	1	8	26.20	0.0009
Depth	1	8	1.02	0.3417

Interpretation

The best model to predict $\log_e$(MDPH) based on the AIC model selection criterion is temperature and is better by 3.1 AIC units than the second-best model (it is the absolute rather than relative difference in values that matters). Temperature appeared in the top eight models, demonstrating the importance of this variable in the model. Depth appears in four of the top eight. For the best AIC model, temperature is positively related to $\log_e$(MDPH) with a slope of 0.0391 (SE = 0.0071). Thus, northern pike moved more at warmer water temperatures. In the second best model, in addition to temperature being positively correlated with movement, depth was negatively related with a slope of –0.1517 (SE = 0.1501). Under this model, northern pike moved more in warmer water but less as water levels were reduced.

its ability to persist or potentially explaining why a species is in decline. Note, however, that just observing how fish use habitat does not allow cause and effect to be inferred. Rather, inferences from these observational studies are strictly correlational. Cause and effect can only be isolated from experiments involving manipulation of habitat.

In order to assess habitat use, one must first document what habitat is available to the fish. Often, what the biologist perceives as available and what the fish deems as available may be quite different. Biologists typically consider the contiguous wetted area as available habitat for limnetic fish, though this may not always be warranted, as restrictions in fish movement may occur based on physical barriers such as inhospitable water temperatures, excessive aquatic vegetation growth, presence of other species, or shallow waters. The advent of GPS and GIS technologies has greatly enhanced the process and accuracy of mapping habitat types and has simplified the estimation of availability.

Typically, the habitat type a fish is using is recorded either when the fish is observed in the field, or it is determined in a GIS by overlaying fish distribution maps on habitat maps (Rogers and Bergersen 1996). A broad array of approaches can be used to evaluate habitat use and resource availability at the population and individual level. Some of the more prominent methods of evaluating whether fish spend more time in some habitats than would be expected based on the availability of those habitats are discussed elsewhere (White and Garrott 1990; Alldredge and Ratti 1992; and Manly et al. 1993). Here our focus will center on methods for which individual use is known (monitoring fish carrying transmitters) and the proportion of available resource units is also known. Though similar methods exist to evaluate resource selection when resource availability is only estimated or sampled (Thomas and Taylor 1990; Manly et al. 1993), the proliferation of mapping technologies (GIS and GPS) has probably limited the need for discussion of those approaches here.

14.6.1 Chi-Square Tests

The simplest and most pervasive approach to assessing whether fish are using habitats in proportion to their availability is the use of chi-square tests. Though the Pearson statistic is more common, the log-likelihood statistic is preferred because model selection based on AIC (Burnham and Anderson 1998) can be used and because more sophisticated models using logistic regression can be built compared with the use of simple contingency tables. In practice, both often yield similar results and are asymptotically equivalent.

Researchers traditionally pooled use data and did not maintain unique identification for each animal (Neu et al. 1974). For example, Rogers (1998) followed seven adult largemouth bass over a summer in a Colorado reservoir, locating these fish, in total, 128 times (Box 14.5). Historically, these observations might simply have been analyzed as 128 independent observations, though this clearly was not the case. Locations by the same individual are correlated in time and must be treated accordingly (section 14.2.4; Otis and White 1999). Pooling may be justified if a few

Box 14.5 Evaluation of Habitat Use

This example reports habitat use by seven largemouth bass carrying transmitters during the summer of 1994 on 27-ha Lake Ladora, Colorado (Rogers 1998). Maps of individual fish locations were overlaid on a 10 × 10-m raster map of bottom type in a GIS to evaluate habitat use by individual fish and to quantify available habitat. The frequency of habitat use along with the availability of the habitat is shown in the table below.

Table The number of locations each of seven largemouth bass frequented by habitat type in Lake Ladora. Fish are identified by transmitter identification numbers, habitat types are listed in the first column, and the last column represents the number of 10 × 10 m cells of a given habitat type that were available in the lake.

Habitat type and individual total	105	132	17	2,263	2,353	285	510	Total for habitat type	Available habitat type
Silt	0	2	3	22	4	9	9	49	959
Chara	0	0	1	2	0	0	0	3	155
Pondweed	0	0	0	0	2	10	1	13	57
Milfoil	1	0	2	3	4	0	1	11	988
Coontail	9	9	3	3	0	28	0	52	503
Total for fish	10	11	9	30	10	47	11	128	2,662

The following SAS code can be used to evaluate both relevant chi-square tests.

Program

```
   data VEG;
   length Habitat $8;
   input Habitat fish1 fish2 fish3 fish4 fish5 fish6 fish7 Available;
   cards;
   Silt               0       2       3      22       4       9       9     959
   Chara              0       0       1       2       0       0       0     155
   Pondweed           0       0       0       0       2      10       1     57
   Milfoil            1       0       2       3       4       0       1     988
   Coontail           9       9       3       3       0      28       0     503
   ;
   data VEG1;

   /* The following rows read in the above data table and format it for
analysis */
   array cnt{7} fish1-fish7;
   set VEG;
   do i = 1-7;
   Fish = i;
   Count = cnt[i];
   keep Habitat Count Fish;
   output;
   end;
   proc freq;
   weight Count;
```

(Box continues)

Box 14.5 *(continued)*

```
tables Fish*Habitat/chisq;
proc transpose data = VEG out = TransposeVeg;
proc sort;
by _NAME_;
data Availablity;

/* This portion determines the amount of habitat available */
array Habitat{5} COL1-COL5;

/* The data table is transposed, so habitat is given in columns */
array Available{5} Avail1-Avail5;
retain Avail1-Avail5;
retain ChiSq TotalChiSq TotalDF TotalLocs 0;
set TransposeVeg end = Last;
if _NAME_ = 'Available' then do;
SumAvailable = sum(of COL1-COL5);
do i = 1 to dim(Available);
Available[i] = Habitat[i]/SumAvailable;
end;
end;
else do;

/* The following generates the chi-square for the first test */
ChiSq = 0;
TotalLocations = sum(of COL1-COL5);
do i = 1 to dim(Available);
if Habitat[i]>0 then
ChiSq = ChiSq+Habitat[i]*log(Habitat[i]/
(TotalLocations*Available[i]));
end;
df = Dim(Available)-1;
Prob = 1-ProbChi(ChiSq, df);
format Prob PVALUE.;
keep _NAME_ TotalLocations ChiSq df Prob;
output;

/* Chi-square for the second test */
TotalChiSq = TotalChiSQ+ChiSq;
TotalDF = TotalDF+df;
TotalLocs = TotalLocs+TotalLocations
if Last then do;
_NAME_ = 'Total';
TotalLocations = TotalLocs;
ChiSq = TotalChiSq;
df = TotalDF;
```

locations are obtained from many fish, but generally just the opposite is true, and many locations are obtained from just a few individuals. Use of habitat among fish in the same population can be highly variable. Not only can the availability of habitat to each individual vary, but there is inherent heterogeneity in use among individuals as well. This heterogeneity among individuals is eliminated if the location

```
Prob = 1-ProbChi(ChiSq, df);
output;
end;
end;
proc print;
```

Descriptive Output (Partial)

Table Statistics for fish by habitat.

Statistic	*df*	Value	Probability
Chi-square	24	93.6632	<0.0001
Likelihood ratio chi-square	24	107.1104	<0.0001
Mantel–Haenszel chi-square	1	6.5549	0.0105
Phi coefficient		0.8554	
Contingency coefficient		0.6500	
Cramer's V		0.4277	

Transmitter identification number	Chi-square	Name	Locations	*df*	Probability
105	25.473	Fish1	10	4	<0.0001
132	23.645	Fish2	11	4	<0.0001
17	2.180	Fish3	9	4	0.7026
2263	20.129	Fish4	30	4	0.0005
2353	10.373	Fish5	10	4	0.0346
285	98.854	Fish6	47	4	<0.0001
510	14.843	Fish7	11	4	0.0050
	195.498	Total	128	28	<0.0001

Interpretation

The resulting χ^2 value for the first test, χ_{L1}^2 (equation [14.1]), was 107.1 with 24 *df*. The resulting P (< 0.0001) indicates that fish are using the available habitats very differently. The value of the second test, χ_{L2}^2 (equation [14.2]), was 195.5 with 28 *df* and a P of 0.0001, also indicating that the largemouth bass were very selective in the types of habitat they used. The difference between these two chi-square tests is 88.4 with 4 *df* (P < 0.0001), which demonstrates strong selection for certain habitat types.

information is pooled across individuals. In the worst case scenario, two fish might select opposite habitats, but pooling would make the investigator think that no selection was occurring (White and Garrott 1990). With computer processing power no longer limiting, the integrity of the data should be maintained. The preferred approach is to consider the animal as the primary sampling unit, and statistical

inference should be based on use with individual fish as replicates (Manly et al. 1993; Otis and White 1999).

Several tests can be conducted in a situation for which individual fish habitat use is recorded and available habitat is known. First, we can check to see if fish are using the various habitat types in similar fashion. Following the notation presented by Manly et al. (1993), u_{ij} is the amount of habitat type i used by fish j; u_{i+} is the amount of habitat type i used by all fish; u_{+j} is the total amount of habitat units used by fish j; and u_{++} is the total number of habitat units used by all fish. The first log-likelihood test statistic (χ_{L1}^2) is

$$\chi_{L1}^2 = 2\sum_{j=1}^{n}\sum_{i=1}^{l} u_{ij}\log_e[u_{ij}/E(u_{ij})], \tag{14.1}$$

where $E(u_{ij}) = u_{i+}u_{+j}/u_{++}$. If the value is sufficiently large compared with the chi-square distribution with $(I-1)(n-1)$ df (I being the number of habitat categories and n the number of fish), then there is evidence for heterogeneity, and fish are using the habitats differently.

To examine if selection is occurring for individual habitat types by some of the fish carrying transmitters, the second log-likelihood test statistic (χ_{L2}^2) is used:

$$\chi_{L2}^2 = 2\sum_{j=1}^{n}\sum_{i=1}^{l} u_{ij}\log_e[u_{ij}/E(u_{ij})], \tag{14.2}$$

where, $E(u_{ij}) = \pi_i u_{+j}$, and π_i is the proportion of available resource units that are in category i. Selection for specific habitats is demonstrated if the chi-square is sufficiently large with $n(I-1)$ df. The difference between these two chi-squares ($I-1$ df) describes whether, on average, fish are using the various habitat types in proportion to their availability, regardless of which ones they are selecting.

Although it is recommended that the expected frequencies in a chi-square test be five or more, these tests are fairly robust to deviations from this rule (Ott 1988). Generally if 90% or more of the expected values are greater than two, there is not a serious problem (Ott 1988). However, the data shown in the example are more sparse than even this rule of thumb suggests is appropriate. Accordingly, we should be cautious in our interpretation of these results, as almost half of the expected counts are less than two. However, given the extreme significance of these three tests and their robustness to deviation, it would be difficult to maintain that the largemouth bass were not using habitats differently or displaying strong selection for certain habitat types (Manly et al. 1993).

14.6.2 Selection Ratios

Once selection has been established, we shift our attention to evaluating which types of habitats were selected. An old intuitive approach is with selection ratios (Manly et al. 1972, 1993; Hobbs and Bowden 1982). Selection is indicated with values greater than one, while avoidance of a habitat is demonstrated with ratios

less than one. Again, from Manly et al. (1993), the selection ratio for the *j*th fish and the *i*th habitat type is estimated by

$$\hat{w}_{ij} = u_{ij}/(\pi_i u_{+j}). \tag{14.3}$$

Generally, one is more interested in selection by the population as a whole, which is estimated by

$$\hat{w}_i = u_{i+}/(\pi_i u_{++}). \tag{14.4}$$

In order to generate confidence intervals (CIs) about these selection ratios, the SE can be calculated as (K. Gerow, University of Wyoming, personal communication)

$$\mathrm{SE}(\hat{w}_i) = \sqrt{\frac{n}{(n-1)(u_{++})^2}\sum_{j=1}^{n}\left(\frac{u_{ij}}{\pi_i} - \hat{w}_i(u_{+j})\right)^2}. \tag{14.5}$$

The selection ratio estimates are generated by pooling observations from all fish in the sample, but the equation takes variation in resource selection from individual fish into account (Manly et al. 1993). It is recommended that simultaneous Bonferroni CIs be constructed to ensure the probability of all intervals containing their true parameter values is $1-\alpha$ (Thomas and Taylor 1990). The intervals around each selection ratio should therefore be constructed at the $100(1-\alpha/I)\%$ level, where I is the number of intervals being constructed (one for each habitat type), such that

$$\hat{w}_i \pm z_{\alpha/2I}\mathrm{SE}(\hat{w}_i), \tag{14.6}$$

where is the *z*-score corresponding to an upper tail probability of . An example of the use of selection ratios is given in Box 14.6.

The results are summarized in Figure 14.2, along with the results that would have been obtained if we had used the historic approach of pooling all our location data and analyzed it with the traditional chi-square method (Neu et al. 1974; Byers et al. 1984; Manly et al. 1993). Although the results are similar, different conclusions are drawn, underscoring the need for conducting the analysis correctly. Under both scenarios, largemouth bass avoid milfoil, yet under the historic approach, they select strongly for pondweed. When analyzed with fish as the sampling unit, however, the sparse data prevent us from achieving the power necessary to infer selection for pondweed. The CI includes 1 (failure to demonstrate selection) at an overall $\alpha = 0.10$ (individual $\alpha = 0.02$).

14.6.3 Continuous Distribution of Availability

Occasionally the distribution of a habitat character of interest does not lend itself to categorization. Parameters such as temperature, dissolved oxygen, and depth,

Box 14.6 Application of Selection Ratios

Using the data from Box 14.5, we can use selection ratios to evaluate which habitats were selected for or against by the population of largemouth bass in Lake Ladora. Calculation of selection ratios is demonstrated based on equation (14.4) and the milfoil habitat data from Box 14.5. The proportion of available resource units that are in category *i* is given by

$$\hat{w}_{milf} = u_{milf}/(\pi_{milf} u_{++}) = 11/(0.371 \times 128) = 0.232.$$

The SE is calculated as (equation [14.5])

$$\begin{aligned} \mathrm{SE}(\hat{w}_{milf}) &= \sqrt{\frac{n}{(n-1)(u_{++})^2} \sum_{j=1}^{n} \left(\frac{u_{milf}}{\pi_{milf}} - \hat{w}_{milf}(u_{+j}) \right)^2} \\ &= \sqrt{\frac{7}{(7-1)(128)^2} \left[\left(\frac{1}{0.371} - 0.232(10) \right)^2 + \ldots + \left(\frac{1}{0.371} - 0.232(11) \right)^2 \right]} = 0.122. \end{aligned}$$

The simultaneous Bonferroni confidence intervals (CIs) are generated as (equation [14.6])

$$\hat{w}_{milf} \pm z_{\alpha/2I} \mathrm{SE}(\hat{w}_{milf}) = 0.232 \pm z_{0.10/2 \cdot 5} 0.122 = (-0.052, 0.516).$$

Because one cannot observe a negative value, the –0.052 should be replaced with 0.000. The process can obviously be facilitated with the aid of a computer, and the appropriate code is available from several sources. The plug and play application FishTel also can perform these calculations.

to name a few, are continuous in nature. One common solution is to simply break the continuum up into categories. Whereas this is a legitimate approach, this method is not as efficient as treating the variable as continuous. Further, the approach is subjective because the biologist must then decide the cutoffs for categorization. A conceptually simpler approach would be merely to measure the parameter of interest and see if that parameter changes in response to some other factor by means of a GLM.

Numerous studies have demonstrated the importance of nearshore habitats to fishes in inland temperate lakes during summer (Winter 1977; Doerzbacher 1980; Betsill et al. 1988). Others have shown that some fishes move offshore in winter, presumably in search of deeper, warmer water when ice covers the lakes (Cook and Bergersen 1988; Rogers 1998). One option for evaluating this distribution would be to divide the habitat into, say, littoral and limnetic zones and then use the chi-square approach described in section 14.6.1, but then one would have to decide what constituted those two habitat types. Rather than make spurious assumptions, one could instead simply calculate the average distance to shore for each fish during each season and use a GLM to assess if distance from shore varied between seasons (Box 14.7).

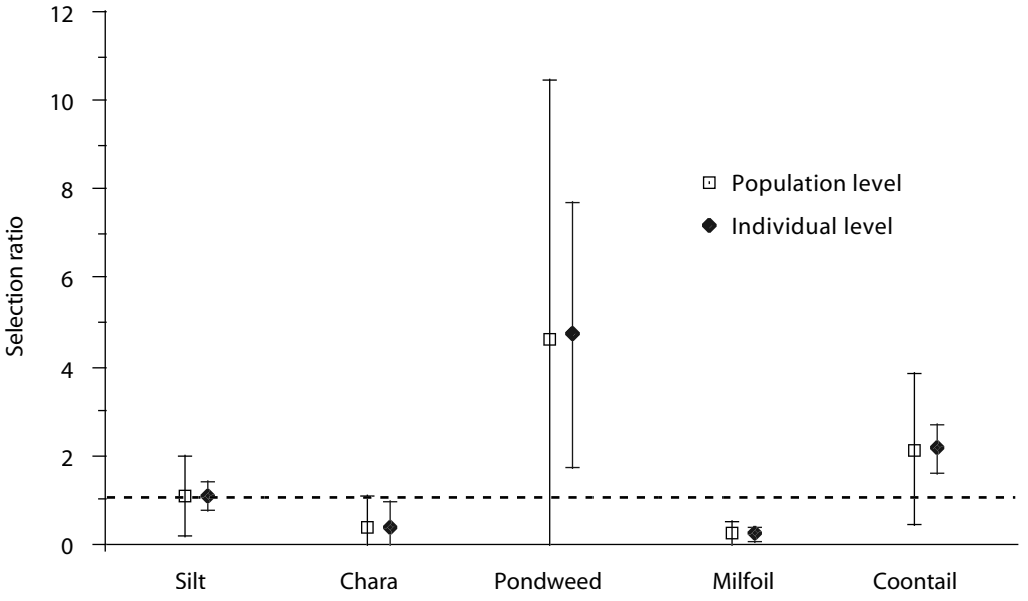


Figure 14.2 Seven adult largemouth bass were monitored during the summer of 1994 on Lake Ladora, Colorado, and their locations were plotted on a map of vegetation to assess habitat use. Selection ratios (Wi) and their associated Bonferroni-adjusted 90% confidence intervals (error bars) were used to determine if selection for (>1) or against (<1) a given habitat type was occurring. The intervals were calculated with both the traditional population level approach and the individual level approach advocated here, which uses the individual fish as sampling units.

The first two models in Box 14.7 (A and B in the SAS code) represent the traditional approach to this sort of analysis, where sources of variation are included for differences between fish and fish × season. Though legitimate, they ignore the autocorrelation of the repeated measurements (distances) taken for each fish and therefore display relatively high values of the small-sample corrected Akaike's Information Criteria (AICc) values. We would expect that distances taken close in time would have a high correlation, whereas distances taken farther apart in time would be less correlated. Model A assumes a constant residual variance across all four seasons, with no covariance between distances for a fish, whereas model B assumes a different residual variance for each season, but still with no covariance. Model E assumes a constant residual variance across seasons and that the autocorrelation of distances is constant across all of the repeated measurements for a fish. Thus, model E ignores the fact that only distances taken close in time should have a high correlation, and therefore it also gives rise to a high AICc. Only models C, D, F, and G incorporate the autocorrelation of the distances in their structure (Littell et al. 1996). Models C and D assume equal time intervals, whereas F and G do not and are therefore rewarded with the lowest AICc values. Model G allows the variance structure to vary by season as well as modeling

Box 14.7 Evaluation of Habitat Use with Continuous Variables

To quantify whether largemouth bass really do move offshore in winter in search of warmer water at greater depth, we will use the same fish observed in Box 14.3. Seventeen adult largemouth bass in Lake Ladora were monitored on monthly 24-h tracks over a 4-year study.

Table Distances (m) from shore to each fish location (partial data set). The mean distance to shore for each fish was calculated with FishTel, which determines the distance from every location registered by a fish to the perimeter of the lake. Seasons are described in Box 14.3. The full data set is included in the Chapter 14 CD folder.

Season	Fish ID	Date	Time	Distance (m)
1	2	18 March 1992	13:04	6
1	2	18 March 1992	17:20	6
1	2	18 March 1992	20:20	33
1	2	19 March 1992	08:07	19
1	2	19 March 1992	00:08	36
1	2	19 March 1992	05:32	45
1	2	16 April 1992	20:59	7
1	2	17 April 1992	00:27	5

The distance to shore is again transformed by the natural logarithm to stabilize increasing variance associated with increasing distance from shore. Using a mixed-model ANOVA where FISH and FISH*SEASON are considered random effects, a variety of models can be developed that allow for autocorrelation in the repeated observations for each fish (Littell et al. 1996).

Program

```
   proc format;
   value season 1 = 'SPRING' 2 = 'SUMMER' 3 = 'FALL' 4 = 'WINTER';
   data llmb;
   infile 'NEARSHORE.TXT' firstobs = 3;
   input SEASON FISH DATE : mmddyy8. TIME : time5. DISTANCE;
   format date date. time time. datetime datetime.;
   DateTime = DHMS(DATE,hour(TIME),minute(time),second(time));
   DateTime1 = DateTime/(60*60*24); *Convert to days;
   format season season.;
   LNDIST = log(DISTANCE);

   * MODEL A;/* Traditional approach with random effects and constant residual
variance across seasons */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL B;/* Traditional approach with different residual variance between
seasons */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
```

```
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   repeated/type = vc/* Default type */group = season;
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL C;/* Accounts for autocorrelation in distances assuming equal time
intervals */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   Repeated/subject = FISH type = ar(1);
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL D;/* Same as C while allowing different residual variance between
seasons */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   Repeated/subject = FISH type = ar(1) group = season;
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL E;/* Constant residual variance across seasons, constant
autocorrelation */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   Repeated/subject = FISH type = cs;
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL F;/* Accounts for autocorrelation in distances as a function of the
actual time interval */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   repeated/subject = fish type = sp(pow)(datetime1);
   *LSMEANS SEASON/ADJUST = BON;

   * MODEL G;/* Same as F while allowing different residual variance between
seasons */
   PROC MIXED DATA = LLMB;
   CLASS SEASON FISH;
   MODEL LNDIST = SEASON/s;
   RANDOM FISH FISH*SEASON;
   repeated/subject = fish type = sp(pow)(datetime1) group = season;
   *LSMEANS SEASON/ADJUST = BON;
   run;
```

(Box continues)

Box 14.7 *(continued)*

Descriptive Output

Table Summary of mixed ANOVA output for all models (random effects for fish). The smallest value of AICc represents the model with the best fit.

Model	AICc	*F*-value
A) Variance components	2,227.4	8.88
B) Variance components/seasons	2,220.6	8.91
C) Autoregressive	2,053.4	8.21
D) Autoregressive/seasons	1,985.2	12.24
E) Compound symmetry	2,229.4	8.88
F) Spatial power	2,063.2	10.32
G) Spatial power/seasons	1,929.8	12.47

Table Mixed ANOVA output for the spatial power/seasons model that displayed the best fit. This model describes autocorrelation in distances as a function of the actual time interval (SP[POW]) by season as well as season-specific residual variances (variance).

Covariance Parameter Estimates

Covariance parameter	Subject	Group	Estimate
Fish			0.2511
Season*Fish			0.06000
Variance	Fish	Season winter	0.6089
SP(POW)	Fish	Season winter	0.5816
Variance	Fish	Season summer	0.6653
SP(POW)	Fish	Season summer	0.000534
Variance	Fish	Season spring	0.8392
SP(POW)	Fish	Season spring	0.01965
Variance	Fish	Season fall	0.8132
SP(POW)	Fish	Season fall	0.007301

Fit Statistics

–2Residual log likelihood	1,909.6
AIC	1,929.6
AICc	1,929.8
BIC	1,937.9

Solution for Fixed Effects

Effect	Season	Estimate	SE	*df*	*t*-value	$P > \|t\|$
Intercept		3.3492	0.1824	16	18.36	<0.0001
Season	Winter	0.5809	0.2060	26	2.82	0.0091
Season	Summer	–0.5647	0.1631	26	–3.46	0.0019
Season	Spring	–0.2104	0.1794	26	–1.17	0.2516
Season	Fall	0				

Type 3 Tests of Fixed Effects

Effect	Numerator *df*	Denominator *df*	*F*-value	$P > F$
Season	3	26	12.47	<0.0001

the correlation as a function of the time interval. It is clearly the most parsimonious model, with an AICc of 1,929.8. This result demonstrates the necessity of including the time element in telemetry data, as autocorrelation is pervasive. Furthermore, this autocorrelation varies across seasons. Examination of the variance–covariance parameter estimates reveals that not only do the variances appear different between seasons, but the only strong correlation between observations close in time occurs in winter ($\hat{r} = 0.5816$), also when the smallest residual variance was estimated ($\hat{\sigma}^2 = 0.6089$). This result is expected, as we have already demonstrated that these fish move least in winter (Box 14.3). When water temperatures drop close to freezing, these poikilotherms are less likely to move far enough from a previous location to remove the autocorrelation effect. The variation among fish (0.2511) and among fish × season (0.0600) is considerably less than the residual variance within fish for each season.

The untransformed mean distance to shore for all the largemouth bass in summer was 26 m, which more than doubled in winter, with fish locations moving to 58 m from shore on average. Mean distance from shore was 39 m in spring and 41 m in fall. Largemouth bass moved offshore, presumably seeking out the deeper, warmer water found at the bottom of the ice-covered lake (Rogers 1998). This same approach can be used in a vertical plane to evaluate conveniently if fish are using various depths in proportion to their availability.

Nearly identical conclusions are obtained with this analysis if the natural logarithm transform is not applied to distance. The same autocorrelation structure, a spatial power autocorrelation function varying among seasons, was selected as the minimum AICc model. The highest autocorrelation was again estimated for winter. Because the residual variance is computed for each season, the heterogeneity of variances caused by distance is somewhat mitigated. Because the mean distance for winter was the largest, the residual variance for winter is greatest for the untransformed distances. The analysis results obtained with the untransformed distances are biologically easier to interpret, hence, in some ways preferred. However, in general, the effect of heterogeneity of variance on ANOVA results is to lower the power of the tests. We can better understand the structure of the data by performing an analysis on both the transformed and untransformed variable and examine similarity in results.

14.6.4 Alternative Approaches

Numerous approaches for analyzing resource selection have been developed over the years. Though most questions can be addressed with the methods we have already discussed, a brief summary of some of the more prominent historic approaches is presented, should the reader wish to explore other avenues.

The first approach to test resource selection rigorously was presented by Neu et al. (1974). They used chi-square analyses to examine the differences in the proportion of used versus available habitats. Analogous to the approaches presented in this chapter, chi-square tests are implemented to test the goodness-of-fit of used to available habitat considering both all habitats simultaneously and each

habitat separately. With their method, the influence of an animal's electivity is effectively weighted by the number of locations for each animal. Unfortunately, analysis is restricted to population level use data and does not allow the user to incorporate unique information for each individual fish in the analysis. As such, it assumes that locations are independent, with a lack of independence resulting in too many type I errors.

The approach developed by Marcum and Loftsgaarden (1980) is especially elegant if the availability of the habitat is only estimated. By doing a chi-square test of independence of habitat types from random locations and telemetry locations, a census of available habitat is not needed. As discussed earlier, with the proliferation of GIS and GPS available to inland fisheries professionals, conducting a census of available habitat is not difficult and does confer more power.

Johnson (1980) developed another approach that is less sensitive to availability concerns. His method ranks habitats by area so that precise estimates are not necessary and minor errors in habitat classification can be tolerated. The method does not test for habitat selection for each animal but rather uses each animal as an observation to test for a preference by the population. Unfortunately, the data from each telemetered fish is weighted equally, regardless of the number of observations recorded for each fish. This method also tends to have lower power than the chi-square approaches discussed earlier. We encourage the investigator instead simply to measure the available habitat.

Friedman's test is another rank-based approach (Alldredge and Ratti 1986, 1992). Unlike Johnson's test, this approach ranks the actual differences between proportional use and availability. Like Johnson's test, fish are compared as if sample sizes are equal for each fish, which is usually not the case. Both tests cause much higher type I error rates if animals differ in their habitat selection because animals are assumed to be blocks. This is not a problem for the chi-square goodness-of-fit tests because those are performed for each individual fish.

Lastly, logistic regression models are gaining popularity in the evaluation of habitat use, especially in wildlife journals (Hudgins et al. 1985; Hosmer and Lemeshow 1989; Agresti 1990; Mace and Waller 1996; Conner and Leopold 1998; Mysterud and Ims 1998). Logistic regression methods represent a specialized form of regression models that are designed for the analysis of categorical data, which are the most common in habitat studies. Unlike the chi-square analyses, logistic regression can also incorporate continuous habitat variables such as water temperature or depth in the analysis. Logistic regression evaluates changes in the odds of habitat use where the odds are defined as the ratio between the probability of using a habitat and the probability of not using it. In a simple study where only three habitats are available to choose from, the model equation is

$$\text{odds} = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2) = e^{\beta_0} e^{\beta_1 X_1} e^{\beta_2 X_2}. \tag{14.7}$$

When the predictor *(X)* is categorical with several categories (as would be the case when habitat is used as a predictor), one must represent that predictor by a

set of indicators (artificial variables set at 0 or 1). Given the above equation, suppose there are three habitats: A, B, and C. Then let X_1 be 1 to represent an observation in habitat A and be 0 otherwise. Observations in habitat B are given by $X_2 = 1$ and 0 otherwise. Possible values for this pair of variables is (1,0), (0,1), or (0,0), representing habitats A, B, and C, respectively. An important point is that we need only two indicators to represent three habitats. In general, the number of indicators required is one less than the number of categories. The choice of habitat that is referenced implicitly (habitat C) is arbitrary. One should simply select the habitat that makes the subsequent inferences simplest or most meaningful.

Most telemetry work is retrospective (K. Gerow, personal communication), necessitating the use of odds ratios defined by

$$\text{odds ratio} = \frac{e^{\beta_0} e^{\beta_1 (X+1)}}{e^{\beta_0} e^{\beta_1 X}} = e^{\beta_1}. \tag{14.8}$$

In logistic regression, the influence of unit changes in the predictor *(X)* is manifested in the odds ratios for each predictor. Because we are interested in the ratio of odds, one habitat category must be selected as a baseline to which other habitat categories can be compared. The selection of the baseline category is again arbitrary, and one should just select the habitat that makes explaining the subsequent data most meaningful.

Implementation is straightforward when fish are used as the sampling unit. The chosen model is fit to each individual fish, and only the parameter estimates are recorded. We are only isolating the relative odds parameters for each fish (K. Gerow, personal communication). Once accomplished, standard parametric approaches can be used to explore the distribution of the independent estimates for each parameter. Like the chi-square analyses, this approach has trouble with categories that have zero use. In fact, CIs in habitat categories with very low observed or expected numbers may be suspect, because the standard normal distribution may not accurately represent the sampling distribution of the statistic (K. Gerow, personal communication).

Although logistic regression is a powerful modeling tool for data analysis, we prefer selection ratios when selection for only habitat categories is explored. Results derived from the selection ratio approach are both more intuitive and easier to interpret. The benefits associated with using logistic regression become apparent when habitat information contains continuous metrics.

14.6.5 Conclusion

Although the above approaches are powerful tools to elucidate if fish are using various habitat types in proportion to their availability, they do not tell us if the habitats are critical to survival or reproduction (Hobbs and Hanley 1990; White and Garrott 1990). Given that heterogeneity in habitat use is pervasive within a species in the same basin, even the link between habitat use and fitness may be tenuous. We do suspect that preferences that actually decrease fitness would be

rapidly eliminated from the population through natural selection (White and Garrott 1990). Telemetry studies are generally correlational and tell you only if a habitat is preferred or not, not if it is critical. Only manipulative experiments can reveal the true importance of underutilized habitats.

14.7 HOME RANGE

Home range is defined as "that area traversed by the individual in its normal activities of food gathering, mating, and caring for young" (Burt 1943). The word normal unfortunately introduces ambiguity into the home range concept. An objective method is needed to define normal. Typically, it is defined as some probability level (95% of the locations), though selecting the appropriate probability is arbitrary as well. Evidence for the establishment of home ranges has been documented for several temperate piscivores (Lewis and Flickinger 1967; Malinin 1969; Winter 1977; Doerzbacher 1980; Ross and Winter 1981; Mesing and Wicker 1986), though the results have been questioned due to the very short duration of some of these telemetry studies (Diana et al. 1977; Diana 1980; Cook and Bergersen 1988). The time frame in which the fish are observed must be representative of the interval of interest, which is determined by the objective of the study. Despite the common practice of generating home range estimates, they rarely are related back to the original study objective. Alone, estimates of home range are of little use unless correlated with some additional parameter of importance. Even then, most approaches for determining home range are so vulnerable to criticism that their utility is questionable. Unfortunately, the most commonly used method of estimation (minimum convex polygon approach) is fraught with more problems than any other. Its persistence in the literature is presumably a tribute to its ease of calculation.

14.7.1 Minimum Convex Polygons

The oldest and most common method for evaluating home range size is the minimum-convex-polygon approach (Mohr 1947; Odum and Kuenzler 1955). In its most basic form, the locations for each fish are plotted, and the smallest convex polygon that encompasses all the locations is constructed. The area of this polygon is then an estimate of home range. This approach has several appealing attributes in that it is simple to calculate and allows for flexibility of home range shape.

As mentioned before, the original intent of the home range was to describe the area that an animal "normally" uses during a specified time frame (Burt 1943). The convex-polygon approach encompasses all points where a fish was recorded, including any rare forays. As such, the minimum-convex-polygon approach is really more of an estimate of the total range of an animal, rather than its home range. In addition, the estimate is a minimum estimate of range, because it is unlikely that fish never explored areas farther than where they were monitored. The range is also constrained to be a convex polygon with this method, which is

probably an unreasonable assumption especially in heterogeneous environments (Anderson 1982). This is especially true when studying fish in lakes with complex shorelines, as fish clearly cannot exploit the dry land that falls in their "home range." Several authors (Winter 1977; Rogers and Bergersen 1995) have removed land areas by generating convex polygons with concave intrusions to reflect the wetted minimum area of the home range. One must be objective when defining such boundaries, but as fish need water to live, this approach is probably defensible.

The most glaring problem with this approach, however, is that estimated home range size increases with increasing sample size (Jennrich and Turner 1969; Schoener 1981). The home range size is a function of the duration of the study and the number of locations used to generate the estimate (Winter 1977). To demonstrate this, daily summer locations of 13 largemouth bass in two Colorado reservoirs (Rogers 1998) were used to generate a minimum-convex-polygon estimate of home range, excluding land area that intruded on the range. In addition to calculating the convex polygon for each fish over the course of the summer, a program was written that randomly dropped 10, 30, 50, and 70% of the locations and recalculated the minimum-convex-polygon home range. Each scenario for each fish was iterated 100 times to generate a mean value (Figure 14.3). Clearly, in all scenarios, documented home range sizes would have been substantially less had fish been located fewer times during that summer. This underscores the futility in comparing home range sizes of a species across studies that use different sampling protocols.

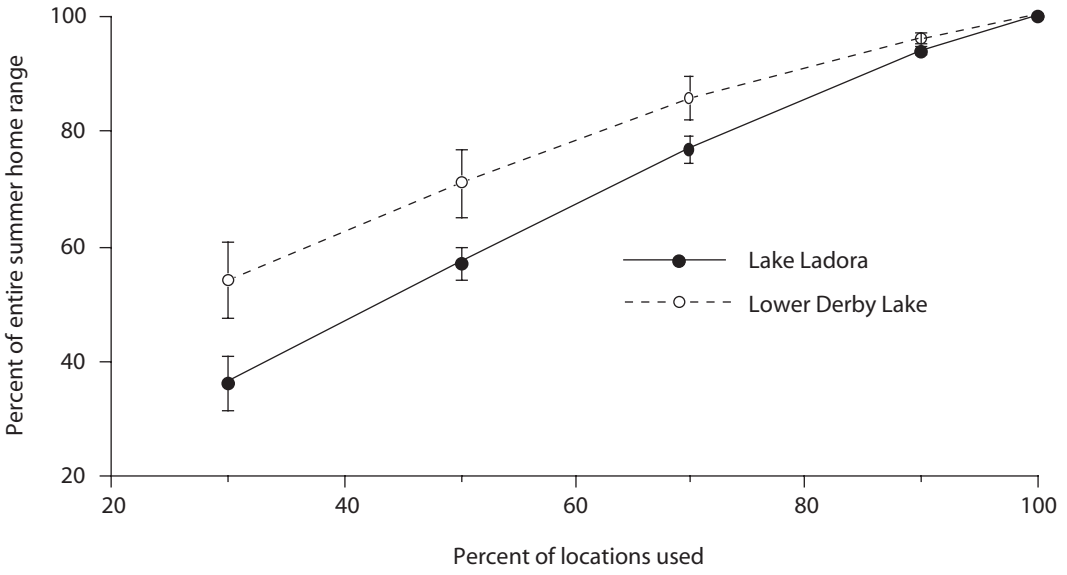


Figure 14.3 Minimum-convex-polygon home ranges were calculated for largemouth bass located in lakes Ladora and lower Derby in the summer of 1994. If the estimates are recalculated using only 30%, 50%, 70%, and 90% of those same locations, a reduction in estimated home range size is realized. Error bars represent the SE of the mean.

Odum and Kuenzler (1955) recognized this limitation and proposed a method to justify using the minimum-convex-polygon approach by plotting the size of the home range against increasing numbers of contacts used to generate the polygon. This "observation area" curve, or cumulative increase in maximum home range area with time, was drawn to determine whether home range size was stable or increasing. They defined stable as the point beyond which each additional observation produced less than 1% increase in area. If the curve did not level off, then either not enough observations were made or the animal in question did not set up a home range in the classic sense. When this approach is used, a staircased graph typically results (Figure 14.4). Although the estimated home range appears to stabilize in a number of places, it is apparent that the home range does not reach an asymptote until the locations fill the entire basin of this 27-ha lake. Terminating the study after the 20th, 40th, or 60th observation would have been made under the illusion that the home range area had stabilized, which was clearly not the case.

14.7.2 Bivariate Normal Models

According to a bivariate normal model, locations are assumed to be distributed independently; that is, fish move randomly around their home range, with their most probable location being the very center. Most models use the area of a 95% ellipse calculated around the mean location as an estimate of the animal's home range. Although 95% has traditionally been used, the number is arbitrary, and any percentage could be used with adequate justification. By not including all of the points, these methods tend to be more robust to outliers. Points close to the mean are weighted greater than those far away (Jennrich and Turner 1969), and

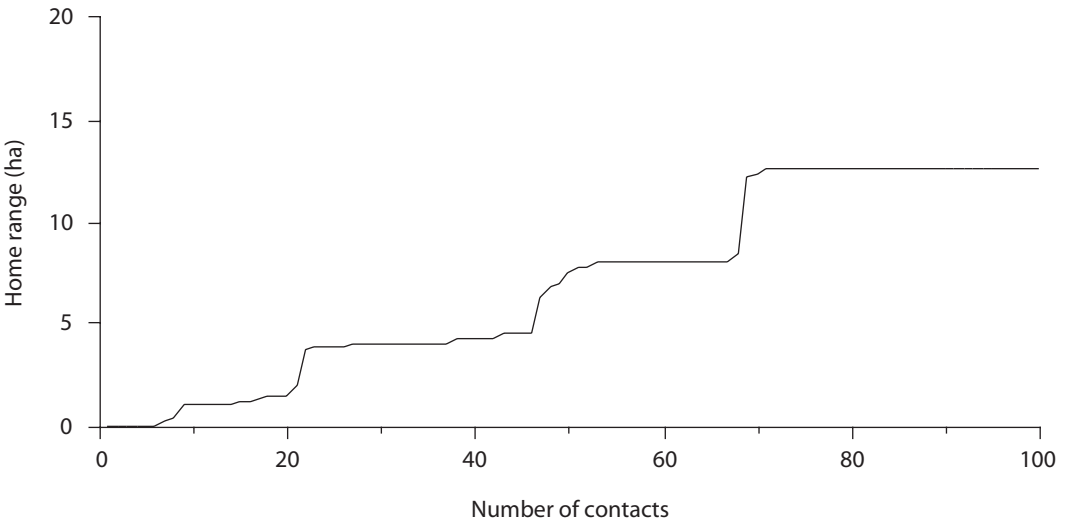


Figure 14.4 The area of the minimum convex polygon encompassing increasing numbers of consecutive contacts for the first 100 locations of largemouth bass #258 in Lake Ladora, Colorado (from Rogers 1998).

some models weight points close in time less than those far apart (Dunn and Gipson 1977). This results in home range estimates that are more consistent with the spirit of Burt's (1943) definition of the home range as the "area traversed by the individual in its normal activities of food gathering, mating and caring for young." Bivariate normal models generally do not include extreme forays that stray far from a center of activity. Numerous methods have been developed in an attempt to decouple the estimated home range size from sample size by this approach (Jennrich and Turner 1969; Dunn and Gipson 1977; Anderson 1982; Samuel and Garton 1985). The obvious benefit is that estimates of home range become more comparable between studies. Fisheries scientists have been reluctant to adopt these methods, presumably because the alternative was historically much easier to calculate. With personal computers now ubiquitous, and a plethora of software programs available (e.g., White and Garrott 1990; Kenward 1992; Larkin and Halkin 1994; Hooge et al. 2001), this should no longer be the case.

Unfortunately, bivariate normal models are still fraught with problems. Conceptually, they do not properly describe the movements of most free-living animals. Fish do not randomly bounce around their home range; they move through it with a purpose, finding food resources, shelter, and mates. Bivariate normal models assume a single center of activity and, hence, do not deal well with multiple centers of activity if a fish decides to change its movement patterns. Home ranges generated by these approaches still increase in size if a fish decides to relocate to new centers of activity on a regular basis, requiring the investigator to decide subjectively what location data to include in the analysis. Smith (1983) provides a chi-square goodness-of-fit test to evaluate whether the home range data are consistent with the assumption of bivariate normality. In addition, these models assume a bivariate normal probability distribution that may not be applicable in many biological settings. Even if range is based around one activity center, this approach is still problematic if you do not want the ellipse to overlap land that the fish cannot use. Clearly landmasses that protrude into a fish's home range would not be well represented by an elliptical home range. In these situations, estimates of home range can be highly biased (Boulanger and White 1990). Additionally, there is no reason a fish should spend most of its time in the center of its home range and little time at the periphery. In general, animal movements arise from strange sampling distributions more frequently than they arise from common, well-known distributions (Schoener 1981; Anderson 1982; Swihart and Slade 1985). Outlying locations, however, cause the ellipse to extend in the opposite direction from the outliers to compensate for their impact on the shape of the normal distribution. Although animal locations seldom fit a bivariate normal distribution, the use of the bivariate normal model for home range estimation is still worthwhile, as the incorporation of a probability model is conducive to robust estimators.

14.7.3 Other Nonparametric Approaches

Perhaps the area with the most promise in dealing with the limitations of home range analyses are the more recently developed nonparametric approaches. These

are more flexible in that they are not restricted to modeling home ranges of a particular shape. Numerous nonparametric approaches have been proposed over the years, including the minimum-convex-polygon approach (Mohr 1947) discussed earlier, as well as Fourier series estimation (Anderson 1982) and grid cell counts (Siniff and Tester 1965). With Fourier series smoothing, the location data are described by adding a finite number of sine waves of various amplitude and frequencies to generate a two-dimensional surface of the area an animal uses. The home range is then the smallest area that encompasses a given percent (e.g., 95%) of the volume of this surface (similar to the arbitrary cutoffs established with the bivariate normal models). The grid cell approach superimposes a grid on a map of the area a fish uses, and the number of locations in each cell is recorded. Although this approach makes no assumptions regarding the shape of the home range, it is very sensitive to the size of the grid cell selected for analysis and the sampling intensity (White and Garrott 1990). Recent developments have provided more robust methods for estimation and are discussed below. Software for the analysis of these methods and detailed instructions for their use can be obtained from a variety of sources (e.g., White and Garrott 1990; Kenward 1992; Larkin and Halkin 1994; Hooge et al. 2001)

14.7.3.1 Dirichlet Tessellations

When sample sizes are very large and autocorrelation is significant, Dirichlet tessellations provide a simple and robust technique for evaluating home range size (Wray et al. 1992; Hooge et al. 2001). This approach describes the spatial pattern of the locations in terms of their relative position only. The density of the locations is calculated without any assumptions about the underlying distribution of the data. The Dirichlet tessellation creates a polygon around each fish location, such that all parts of the polygon are closer to the enclosed location than any other location (Figure 14.5). Fish locations are joined by the dotted lines to form Delaunay triangles (Upton and Fingleton 1985). The perpendicular bisectors of the dotted lines then give rise to the polygons that form the tessellation (Wray et al. 1992). The home range is determined as the smallest possible area that includes 95% (or any other justifiable amount) of the polygons (location estimates). Using much smaller percentages will isolate the core areas within a home range. Areas of the home range where fish locations are concentrated then give rise to smaller polygons. The internal configuration of an animal's home range is therefore readily detected. This approach is sensitive to outliers, and the home range boundary can be difficult to establish at high inclusion percentages (Wray et al. 1992). Its use should be restricted to situations for which the location data sets are large.

14.7.3.2 Harmonic Mean

Dixon and Chapman (1980) proposed a home range estimator based on the harmonic mean of the spatial distribution of locations, using the distances from nodes on a grid to observed locations. The estimator uses the mean of the inverse distances from a node on a grid to all the locations (Seaman and Powell 1996). This mean is then inverted to generate a surface that is low where locations are most

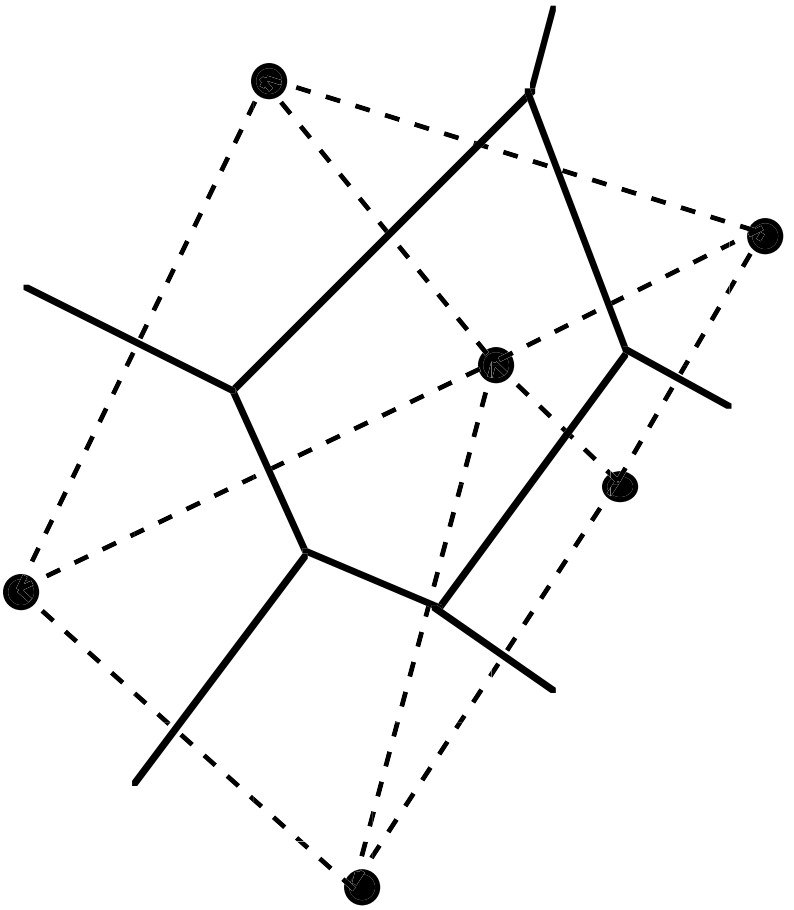


Figure 14.5 The Dirichlet tessellation is constructed from the perpendicular bisectors (solid lines) of the sides of the Delaunay triangles (dotted lines) that connect the fish location estimates. The home range is determined as the smallest possible area that includes 95% (or any other justifiable amount) of the polygons (fish location estimates).

clustered (low mean distance to observations) and high where locations are dispersed (high mean distance to other observations). The boundary of the home range can then be defined by calculating harmonic means at all locations, then at all nodes. All nodes that have harmonic mean values greater than any calculated at the locations are deemed outside the home range (Seaman and Powell 1996). The harmonic means can be converted to a frequency distribution by dividing the means at each node by the sum of the means in the home range. The area under the lowest 95% of this surface is then an estimate of home range (Seaman and Powell 1996). This estimator does not impose any particular shape on the estimated home range and can define ranges with multiple centers of activity properly. It has been shown to be less biased than the other methods presented up to this point (Boulanger and White 1990) and is useful for determining centers of

activity (Dixon and Chapman 1980). Unfortunately, the results are dependent on the origin and spacing of the grid used and the measurement units of the locations, all because a location that falls exactly on a node provides an undefined quantity. As such, the more robust kernel estimators (Worton 1987) have largely replaced this approach,

14.7.3.3 Kernel Estimators

These nonparametric density estimators represent perhaps the most intriguing of the probabilistic approaches to estimating home range (Worton 1989). They are preferred over the harmonic mean approach because they are much less sensitive to outlying locations and the choice of measurement units and the grid, and show very little bias (Worton 1995; Seaman and Powell 1996). Conceptually, the kernel method places a probability density (a mound-shaped kernel) over every location registered for a given fish that is to be used in the home range analysis. The density of the kernel is maximized directly over the fish location, but then tails off in all directions, similar to a bivariate normal function. The density is estimated at the nodes of a grid draped over this surface as the average density of all the kernels that overlap at that point. This density will be high in areas where fish are concentrated (and many kernels overlap), but low in less-frequented areas. Once this surface has been generated, contour lines can be inscribed on it that connect areas of equal density. For comparison purposes, usually the area that incorporates 95% of the utilization distribution is calculated as an estimate of home range. This value is arbitrary, however, and it is often useful to draw contour lines at multiple levels so that high-use areas can be rapidly identified.

Smoothing is critical for describing accurate home range sizes with kernel estimators (Worton 1995), and that is how the two kinds of kernel estimators are defined. In the fixed kernel estimator, the smoothing parameters are fixed over the entire surface (Worton 1989). The smoothing parameters are allowed to vary in the adaptive kernel estimator approach. Areas where the densities of fish are low receive more smoothing than do areas where fish are concentrated. Although adaptive kernel methods are thought to yield better estimates (Worton 1989; Silverman 1992), the fixed kernel approach gave the least-biased results and lowest error based on simulated data (Seaman and Powell 1996; Seaman et al. 1999) and is therefore recommended. Kernel estimates based on sparse data should be expected to overestimate true home range size, though home ranges that follow smooth unimodal distributions can be accurately described with fewer locations than can more complex distributions (Seaman and Powell 1996).

Though elegant nonparametric approaches have been developed to address some of the concerns with earlier methods of home range estimation, they still have several shortcomings. Perhaps the most problematic is that these approaches, like those that use minimum convex polygons, have no CIs around them. One cannot judge the quality of the estimate because only a point estimate is generated, without an estimate of its SE. However, SEs for the kernel estimators could be developed with a bootstrap procedure, resampling the original data with replacement to generate a distribution of calculated values (Good 1994; Edgington

1995; Manly 1997). Kernel estimators also ignore the time series nature of the data and tend to have lower precision than do parametric methods of estimation because fewer assumptions are made (White and Garrott 1990). However, the lower bias of these estimators still makes their use recommended.

14.7.4 Utility of the Home Range Concept

No one method for characterizing home range size is without flaws (Table 14.1). Many subjective decisions and assumptions must be made when generating home range estimates, limiting the biological insight they can provide (White and Garrott 1990). In particular, the sampling scheme to obtain the locations used to calculate a home range estimate must provide an unbiased picture of the animal's movements (Otis and White 1999). Many varied approaches for estimating home range size exist, but they are all vulnerable to criticism. All involve making subjective decisions that lack objective criteria, causing the resulting estimates to provide little biological insight. Home range estimates are often presented to disguise the fact that no hypotheses are being tested, providing a quantitative albeit insignificant summary of the data. Some have advocated abandoning the calculation of home range size altogether and using raw data to test hypotheses (Anderson 1982; White and Garrott 1990). This would eliminate the need to invoke spurious assumptions and biases inherent in home range estimates and confer more power to subsequent statistical tests.

Home range analyses provide an interesting exercise in data analysis but are of little interest, unless correlated with some additional parameters. Properties of the home range should have adaptive significance. For example, home range size

Table 14.1 A subjective summary of various traits associated with the minimum convex polygon (MCP), wetted MCP (MCPw), Jennrich–Turner bivariate normal (JT), Dunn–Gipson bivariate normal (DG), Fourier series (FS), grid cell count (GC), Dirichlet tessellation (DT), harmonic mean (HM), and kernel estimator (KE) methods for calculating home range. In general, only those methods that are particularly sensitive to a specific trait are included.

Trait	Method
Experiences range increases with sample size	MCP, MCPw
Experiences range increases with sampling duration	MCP, MCPw
Accounts for land intrusions	MCPw, GC, DT, HM, KE
Restricts home range to a particular shape	JT, DG
Includes arbitrary percent of locations in range	JT, DG, FS, DT, HM, KE
Assumes single center of activity	JT, DG
Accommodates multiple centers of activity	MCP, MCPw, FS, GC, DT, HM, KE
Is robust to outliers	JT, DG, DT, HM, KE
Accounts for time series nature of data	DG
Has robust statistical foundation	JT, DG
Isolates core areas	GC, DT, HM, KE
Experiences difficulty in defining boundary of home range	JT, DG, FS, DT
Depends on grid used	GC, HM

is considered an important aspect of an animal's feeding strategy and should be related to food density, metabolic needs, and the efficiency of movement, in addition to being inversely correlated with population density (Schoener 1981). Fish and Savitz (1983) attempted to use home ranges to compare trophic relationships between species in an Illinois lake. Their hypothesis was that largemouth bass would have larger home ranges than would bluegills or pumpkinseeds because the latter species were benthivores whereas largemouth bass were piscivores. The lower densities of piscine prey would necessitate that largemouth bass would have larger home ranges. Their fish were tracked on average for 43 d, and a home range was arbitrarily defined as the minimum-convex-polygon area that a fish occupied for at least 5 consecutive days. Not surprisingly, extreme variability kept them from detecting any differences in home range size between species. Usually, one can substitute raw location or movement data to test a hypothesis of interest, thereby avoiding the problems associated with home range analysis. Because movement (MDPH) and home range are highly correlated (Rogers and Bergersen 1995), Fish and Savitz (1983) may have been able to address the same relationships without having to make all of the spurious assumptions associated with analysis of home range by invoking MDPH instead. Savitz et al. (1993) used home range size to show that largemouth bass used reduced areas when given supplemental feed. Minns (1995) demonstrated that home range was correlated with fish size. Both would have drawn the same conclusions by measuring MDPH while avoiding criticisms of home range analysis.

The entire home range concept may not be as appropriate for fishes as it is for terrestrial mammals (especially those with altricial young). Burt's (1943) original home range concept was developed for mammals as the area used for foraging that surrounded a permanent home site. These areas were generally stable over long periods of time. In fact, before generating a home range estimate, it would be wise to determine if site fidelity even exists (Hooge et al. 2001). Fish seem to display more transitory ranges (Winter 1977; Cook and Bergersen 1988; Jones and Rogers 1998), perhaps due to short spawning seasons (Savitz et al. 1993) and limited or absent parental care. Changes in home range areas appear to occur with changes in prey, water temperature (Savitz et al. 1993), or body size (Minns 1995) rather than intraspecific competition. We may be stretching the original spirit of Burt's (1943) concept too far in fisheries research and should perhaps employ metrics other than home range in testing hypotheses of interest.

14.8 SUMMARY

The methods presented here are only a sampling of the varied approaches used in telemetry studies, but we hope that they will provide a foundation for customizing analyses for specific applications. With the proliferation of software programs and powerful computers to run them, many of the traditional shortcomings of telemetry work can now be addressed. Researchers can now focus on study design aspects of their work prior to the initiation of the study by ensuring that representative samples of fish from the population of interest are obtained.

Power calculations should be conducted before the initiation of field work to ensure that proposed research will be able to address the questions that are posed. Early on, one should explore the error associated with telemetry system to be used to evaluate what influence or bias it will introduce to the results. Perhaps the strongest message that should be gleaned from this chapter regards the correct treatment of the sampling unit. Because we are interested in describing how fish in a population behave, the individual fish are the sampling units and not the individual locations, as is commonly reported. Using locations as sampling units is pseudoreplication (Hurlbert 1984). The attention focused on the serial correlation of location data taken close in time is misguided now that powerful statistics packages, such as SAS, allow for the integration of autocorrelation in the data structures. Autocorrelation of location estimates is largely irrelevant when telemetry data are analyzed correctly.

After ensuring that the study design is sound, one is encouraged to explore the spatial distribution of fish locations to determine if they are distributed in a nonrandom fashion. Movement is another metric that is often explored in telemetry literature and can often be correlated with environmental attributes. Often habitat use is the primary area of interest. Though numerous approaches have been developed to characterize use, they do not reveal whether the behaviors we observe are actually a critical reflection of what is needed for the fish to reproduce and survive. Cause and effect can be isolated only by conducting manipulative experiments.

A tremendous amount of effort has been expended on developing methods to evaluate home range size. The minimum-convex-polygon approach is certainly the most common but is problematic because home range size tends to increase with sample size or sampling duration. Bivariate normal models strive to minimize those limitations by including only a portion of the locations (usually 95%). Reducing the number of locations considered makes this type of approach less sensitive to outliers, but 95% is completely arbitrary, and a few percentage points either way can have a large effect on range size. Additionally, these models unrealistically restrict the shape a home range can assume. Shortcomings of the bivariate normal models have been addressed with nonparametric approaches that allow for multiple centers of activity and accurate reflections of odd-looking home range boundaries that typically occur in nature. Many nonparametric methods also entertain only a portion of the locations recorded to decrease sensitivity to outlying locations. Though these approaches represent the most promise in home range analyses, it is difficult to establish CIs with these methods, preventing the reader from assessing the quality of the estimate. Because all approaches have some flaws, we advocate substituting alternative metrics, such as movement per hour, for comparison to parameters of interest. These other metrics generally provide the same results without invoking the myriad assumptions and subjective criteria associated with home range estimation. In fact, the whole concept of a home range may not be as appropriate for freshwater fishes as it is for small terrestrial mammals with altricial young, for which the concept was developed (Burt 1943).

14.9 A LOOK TO THE FUTURE

With improvements in technology, we expect to see increased diversity in types of telemetry applications that are implemented. We have witnessed an increase in the use of telemetry to isolate the importance of various parameters used in bioenergetics modeling in recent publications. In addition, we expect fisheries scientists to explore other applications of telemetry, such as survival rate estimation (Pollock et al. 1989; White and Garrott 1990; Bunck et al. 1995; Harmata et al. 1999; White and Burnham 1999; Skalski et al. 2001) and population estimation (Bartmann et al. 1987; White and Garrott 1990; Bowden and Kufeld 1995; White 1996; White and Shenk 2001), applications that see wide use in wildlife research but are just beginning to be deployed in inland fisheries work.

As technology advances, we also expect to see a greater emphasis on automated receiving systems. Some new innovative applications of telemetry require continuous monitoring to obtain meaningful results. This has spawned a proliferation of automated systems that either track or monitor fish 24 h/d (Hawkins et al. 1980; Pincock 1980; Armstrong et al. 1992; Lucas et al. 1992; Cooke et al. 2000, 2001; Cooke and Bunt 2001; Dieperink et al. 2001; Skalski et al. 2001). With continuous monitoring, many of the assumptions that had to be made with previous telemetry work can be validated. Not only is this approach much less labor intensive than traditional telemetry, it can yield orders of magnitude more information that personal computers can easily process. Though some of these applications are extremely specialized and will require specific types of data analyses, many will be able to expand on the approaches discussed in this chapter.

14.10 REFERENCES

Aebischer, N. J., P. A. Robertson, and R. E. Kenward. 1993. Compositional analysis of habitat use from animal radio-tracking data. Ecology 74:1313–1325.

Agresti, A. 1990. Categorical data analysis. Wiley, New York.

Alldredge, J. R., and J. T. Ratti. 1986. Comparison of some statistical techniques for analysis of resource selection. Journal of Wildlife Management 50:157–165.

Alldredge, J. R., and J. T. Ratti. 1992. Further comparison of some statistical techniques for analysis of resource selection. Journal of Wildlife Management 56:1–9.

Anderson, D. J. 1982. The home range: a new nonparametric estimation technique. Ecology 63:103–112.

Armstrong, J. D., P. M. Bagley, and I. G. Priede. 1992. Tracking deep-sea fish using ingestible transmitters and an autonomous sea-floor instrument package. Pages 378–386 *in* I. G. Priede and S. M. Swift, editors. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

Armstrong, J. D., M. C. Lucas, I. G. Priede, and L. DeVera. 1989. An acoustic telemetry system for monitoring of heart rate of pike, *Esox lucius* L., and other fish in their natural environment. Journal of Experimental Biology 143:549–552.

Baldwin, C. M., D. A. Beauchamp, and C. P. Gubala. 2002. Seasonal and diel distribution and movement of cutthroat trout from ultrasonic telemetry. Transactions of the American Fisheries Society 131:143–158.

Baras, E. 1998. Selection of optimal positioning intervals in fish tracking: an experimental study on *Barbus barbus*. Hydrobiologia 371/372:19–28.

Bartmann, R. M., G. C. White, L. H. Carpenter, and R. A. Garrott. 1987. Aerial mark–recapture estimates of confined mule deer in pinyon–juniper woodland. Journal of Wildlife Management 51:41–46.

Betsill, R. K., R. L. Noble, and W. H. Neill. 1988. Distribution and habitat selection of telemetered northern and Florida largemouth bass in 2 small Texas impoundments. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 40(1986):275–286.

Boulanger, J. G., and G. C. White. 1990. A comparison of home-range estimators using Monte Carlo simulation. Journal of Wildlife Management 54:310–315.

Bowden, D. C., and R. C. Kufeld. 1995. Generalized mark–resight population size estimation applied to Colorado moose. Journal of Wildlife Management 59:840–851.

Bramblett, R. G., and R. G. White. 2001. Habitat use and movements of pallid and shovelnose sturgeon in the Yellowstone and Missouri rivers in Montana and North Dakota. Transactions of the American Fisheries Society 130:1006–1025.

Brown, R. S., S. J. Cooke, W. G. Anderson, and R. S. McKinley. 1999. Evidence to challenge the 2% rule for biotelemetry. North American Journal of Fisheries Management 19:867–871.

Bunck, C. M., C.-L. Chen, and K. H. Pollock. 1995. Robustness of survival estimates from radio-telemetry studies with uncertain relocation of individuals. Journal of Wildlife Management 59:790–794.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York.

Burt, W. H. 1943. Territoriality and home range concepts as applied to mammals. Journal of Mammalogy 24:346–352.

Byers, C. R., R. K. Steinhorst, and P. R. Krausman. 1984. Clarification of a technique for analysis of utilization–availability data. Journal of Wildlife Management 48:1050–1053.

Casselman, J. M. 1978. Effects of environmental factors on growth, survival, activity, and exploitation of northern pike. Pages 114–128 *in* R. D. Kendall, editor. Selected coolwater fishes of North America. American Fisheries Society, Special Publication 11, Bethesda, Maryland.

Conner, L. M., and B. D. Leopold. 1998. A multivariate habits model for female bobcats: a GIS approach. Proceedings of the Southeastern Association of Wildlife and Fisheries Agencies 52(1998):232–243.

Cook, M. F., and E. P. Bergersen. 1988. Movements, habitat selection, and activity periods of northern pike in Eleven Mile Reservoir, Colorado. Transactions of the American Fisheries Society 117:495–502.

Cooke, S. J., and C. M. Bunt. 2001. Assessment of internal and external antenna configurations of radio transmitters implanted in smallmouth bass. North American Journal of Fisheries Management 21:236–241.

Cooke, S. J., C. M. Bunt, and R. S. McKinley. 2000. Winter residency of smallmouth bass in a thermal discharge canal: implications for tempering pump operation. North American Journal of Fisheries Management 20:288–295.

Cooke, S. J., C. M. Bunt, J. F. Schreer, and D. H. Wahl. 2001. Comparison of several techniques for mobility and activity estimates of smallmouth bass in lentic environments. Journal of Fish Biology 58:573–587.

Cote, D., D. A. Scruton, L. Cole, and R. S. McKinley. 1999. Swimming performance and growth rates of juvenile Atlantic cod intraperitoneally implanted with dummy acoustic transmitters. North American Journal of Fisheries Management 19:1137–1141.

Cresswell, W. J., and G. C. Smith. 1992. The effects of temporally autocorrelated data on methods of home range analysis. Pages 272–284 *in* I. G. Priede and S. M. Swift, editors. Remote telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

Demers, E., R. S. McKinley, A. H. Weatherley, and D. J. McQueen. 1996. Activity patterns of largemouth bass and smallmouth bass determined with electromyogram biotelemetry. Transactions of the American Fisheries Society 125:434–439.

Diana, J. S. 1980. Diel activity pattern and swimming speeds of northern pike *(Esox lucius)* in Lac Ste. Anne, Alberta. Canadian Journal of Fisheries and Aquatic Sciences 37: 1454–1458.

Diana, J. S., W. C. Mackay, and M. Ehrman. 1977. Movements and habitat preference of northern pike *(Esox lucius)* in Lac Ste. Anne, Alberta. Transactions of the American Fisheries Society 106:560–565.

Dieperink, C., S. Pedersen, and M. I. Pedersen. 2001. Estuarine predation on radiotagged wild and domesticated sea trout (*Salmo trutta* L.) smolts. Ecology of Freshwater Fish 10:177–183.

Dixon, K. R., and J. A. Chapman. 1980. Harmonic mean measure of animal activity areas. Ecology 6:1040–1044.

Doerzbacher, J. F. 1980. Movement and home range of largemouth bass *(Micropterus salmoides)* in relation to water quality of the Atchafalaya River basin, Louisiana. Master's thesis. Louisiana State University, Baton Rouge.

Dunn, J. E., and P. S. Gipson. 1977. Analysis of radio telemetry data in studies of home range. Biometrics 33:85–101.

Edgington, E. S. 1995. Randomization tests, 3rd edition. Marcel Dekker, New York.

Fish, P. A., and J. Savitz. 1983. Variations in home ranges of largemouth bass, yellow perch, bluegills, and pumpkinseeds in an Illinois lake. Transactions of the American Fisheries Society 112:147–153.

Gallepp, G. W., and J. J. Magnuson. 1972. Effects of negative buoyancy on the behavior of the bluegill, *Lepomis macrochirus* Rafinesque. Transactions of the American Fisheries Society 101:507–512.

Good, P. 1994. Permutation tests: a practical guide to resampling methods of testing hypotheses. Springer-Verlag, New York.

Guy, C. S., R. M. Neumann, and D. W. Willis. 1992. Movement patterns of adult black crappie, *Pomoxis nigromaculatus,* in Brant Lake, South Dakota. Journal of Freshwater Ecology 7:137–147.

Guy, C. S., D. W. Willis, and J. J. Jackson. 1994. Biotelemetry of white crappies in a South Dakota glacial lake. Transactions of the American Fisheries Society 123:63–70.

Harmata, A. R., G. J. Montopoli, B. Oakleaf, P. J. Harmata, and M. Restani. 1999. Movements and survival of bald eagles banded in the greater Yellowstone ecosystem. Journal of Wildlife Management 63:781–793.

Hart, L. G., and R. C. Summerfelt. 1975. Surgical procedures for implanting ultrasonic transmitters into flathead catfish *(Pylodictus olivaris).* Transactions of the American Fisheries Society 104:56–59.

Hawkins, A. D., G. G. Urquhart, and G. W. Smith. 1980. Ultrasonic tracking of juvenile cod by means of a large spaced hydrophone array. Pages 461–470 *in* C. J. Amlaner, Jr. and D. W. Macdonald, editors. A handbook on biotelemetry and radio tracking. Pergamon Press, Oxford, UK.

Henderson, H. F., A. D. Hasler, and G. G. Chipman. 1966. An ultrasonic transmitter for use in studies of movements of fishes. Transactions of the American Fisheries Society 95:350–356.

Hobbs, N. T., and D. C. Bowden. 1982. Confidence intervals on food preference indices. Journal of Wildlife Management 46:505–507.

Hobbs, N. T., and T. A. Hanley. 1990. Habitat evaluation: do use/availability data reflect carrying capacity? Journal of Wildlife Management 54:515–522.

Hooge, P. N., W. M. Eichenlaub, and E. K. Solomon. 2001. Using GIS to analyze animal movements in the marine environment. Pages 37–51 *in* G. H. Kruse and nine coeditors. Spatial processes and management of marine populations. University of Alaska Sea Grant College Report AK-SG-01-02, Fairbanks.

Hosmer, D. W., and S. Lemeshow. 1989. Applied logistic regression. Wiley, New York.

Hudgins, J. E., G. L. Storm, and J. S. Wakeley. 1985. Local movements and diurnal-habitat selection by male American woodcock in Pennsylvania. Journal of Wildlife Management 49:614–619.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Jennrich, R. I., and F. B. Turner. 1969. Measurement of non-circular home range. Journal of Theoretical Biology 22:227–237.

Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65–71.

Johnstone, A. D. F., M. C. Lucas, P. Boylan, and T. J. Carter. 1992. Telemetry of tail-beat frequency of Atlantic salmon (*Salmo salar* L.) during spawning. Pages 456–465 *in* I. G. Priede and S. M. Swift, editors. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

Jones, M. S., and K. B. Rogers. 1998. Palmetto bass movements and habitat use in a fluctuating Colorado irrigation reservoir. North American Journal of Fisheries Management 18:640–648.

Kelso, J. R. M. 1978. Diel rhythm in activity of walleye, *Stizostedion vitreum vitreum.* Journal of Fish Biology 12:593–599.

Kenward, R. E. 1992. Quantity versus quality: programmed collection and analysis of radio-tracking data. Pages 231–246 *in* I. G. Priede and S. M. Swift, editors. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

Knights, B. C., and B. A. Lasse. 1996. Effects of implanted transmitters on adult bluegills at two temperatures. Transactions of the American Fisheries Society 125:440–449.

Larkin, R. P., and D. Halkin. 1994. A review of software packages for estimating animal home ranges. Wildlife Society Bulletin 22:274–287.

Lee, W. C., and E. P. Bergersen. 1996. Influence of thermal and oxygen stratification on lake trout hooking mortality. North American Journal of Fisheries Management 16: 175–181.

Lenth, R. V. 1981. On finding the source of a signal. Technometrics 23:149–154.

Lewis, W. M., and S. Flickinger. 1967. Home range tendency of the largemouth bass *(Micropterus salmoides)*. Ecology 48:1020–1023.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, North Carolina.

Lucas, M. C., D. F. Johnstone, and I. G. Priede. 1993. Use of physiological telemetry as a method of estimating metabolism of fish in the natural environment. Transactions of the American Fisheries Society 122:822–833.

Lucas, M. C., I. G. Priede, J. D. Armstrong, A. N. Z. Gindy, and L. DeVera. 1991. Direct measurements of metabolism, activity and feeding behaviour of pike, *Esox lucius* L., in the wild, by the use of heart rate telemetry. Journal of Fish Biology 39:325–345.

Lucas, M. C., G. G. Urquhart, A. D. F. Johnstone, and T. J. Carter. 1992. A portable 24-hour recording system for field telemetry: first results from continuous monitoring of heart rate of adult Atlantic salmon (*Salmo salar* L.) during the spawning migration. Pages 466–473 *in* I. G. Priede and S. M. Swift, editors. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

Mace, R. D., and J. S. Waller. 1996. Grizzly bear distribution and human conflicts in Jewel Basin Hiking Area, Swan Mountains, Montana. Wildlife Society Bulletin 24:461–467.

Malinin, L. K. 1969. Uchastki obitaniya i instinkt vozvroskcheniya ryb. [Home range and homing instinct of fish.] Zoologicheskii Zhurnal 48:381–391. Translated from Russian: Fisheries Research Board of Canada Translation Series 2050, 1970, Ottawa.

Manly, B. J. F. 1997. Randomization, bootstrap, and Monte Carlo methods in biology, 2nd edition. Chapman and Hall, London.

Manly, B. F. J., L. McDonald, and D. Thomas. 1993. Resource selection by animals: statistical design and analysis for field studies. Chapman and Hall, London.

Manly, B. F. J., P. Miller, and L. M. Cook. 1972. Analysis of a selective predation experiment. American Naturalist 106:719–736.

Manns, R. E., and B. G. Whiteside. 1979. Behavioral variations associated with ultrasonic tagging of Guadalupe bass in Lake Travis, Texas. Underwater Telemetry Newsletter 9:4–9.

Marcum, C. L., and D. O. Loftsgaarden. 1980. A non-mapping technique for studying habitat preferences. Journal of Wildlife Management 44:963–968.

Markham, J. L., D. L. Johnson, and R. W. Petering. 1991. White crappie summer movements and habitat use in Delaware reservoir, Ohio. North American Journal of Fisheries Management 11:504–512.

Martin, S. W., J. A. Long, and T. N. Pearsons. 1995. Comparison of survival, gonad development, and growth between rainbow trout with and without surgically implanted dummy radio transmitters. North American Journal of Fisheries Management 15:494–498.

Mesing, C. L., and A. M. Wicker. 1986. Home range, spawning migrations, and homing of radio-tagged Florida largemouth bass in two central Florida lakes. Transactions of the American Fisheries Society 115:286–295.

Minns, C. K. 1995. Allometry of home range size in lake and river fishes. Canadian Journal of Fisheries and Aquatic Sciences 52:1499–1508.

Mitson, R. B. 1978. A review of biotelemetry techniques using acoustic tags. Pages 269–283 *in* J. E. Thorpe, editor. Rhythmic activities of fishes. Academic Press, New York.

Mohr, C. O. 1947. Table of equivalent populations of North American small mammals. American Midland Naturalist 37:223–249.

Moore, A., I. C. Russell, and E. C. E. Potter. 1990. The effects of intraperitoneally implanted dummy acoustic transmitters on the behaviour and physiology of juvenile Atlantic salmon, *Salmo salar* L. Journal of Fish Biology 37:713–721.

Mysterud, A., and R. A. Ims. 1998. Functional responses in habitat use: availability influences relative use in trade-off situations. Ecology 79:1435–1441.

Nams, V. O. 1989. Effects of radiotelemetry error on sample size and bias when testing for habitat selection. Canadian Journal of Zoology 67:1631–1636.

Neu, C. W., C. R. Byers, and J. M. Peek. 1974. A technique for analysis of utilization–availability data. Journal of Wildlife Management 38:541–545.

Odum, E. P., and E. J. Kuenzler. 1955. Measurement of territory and home range size in birds. Auk 72:128–137.

Økland, F., B. Finstad, R. S. McKinley, E. B. Thorstad, and R. K. Booth. 1997. Radiotransmitted electromyogram signals as indicators of physical activity in Atlantic salmon. Journal of Fish Biology 51:476–488.

Otis, D. L., and G. C. White. 1999. Autocorrelation of location estimates and the analysis of radiotracking data. Journal of Wildlife Management 63:1039–1044.

Ott, L. 1988. An introduction to statistical methods and data analysis, 3rd edition. PWS-Kent Publishing, Boston.

Paukert, C. P., P. J. Chvala, B. L. Heikes, and M. L. Brown. 2001. Effects of implanted transmitter size and surgery on survival, growth, and wound healing of bluegill. Transactions of the American Fisheries Society 130:975–980.

Paukert, C. P., and W. L. Fisher. 2000. Abiotic factors affecting summer distribution and movement of male paddlefish, *Polyodon spathula,* in a prairie reservoir. The Southwestern Naturalist 45:133–140.

Petering, R. W., and D. L. Johnson. 1991. Suitability of a cyanoacrylate adhesive to close incisions in black crappies used in telemetry studies. Transactions of the American Fisheries Society 120:535–537.

Pincock, D. 1980. Automation of data collection in ultrasonic biotelemetry. Pages 471–476 *in* C. J. Amlaner, Jr. and D. W. Macdonald, editors. A handbook on biotelemetry and radio tracking. Pergamon Press, Oxford, UK.

Pollock, K. H., S. R. Winterstein, and M. J. Conroy. 1989. Estimation and analysis of survival distributions for radio-tagged animals. Biometrics 45:99–109.

Priede, I. G., and A. H. Young. 1977. The ultrasonic telemetry of cardiac rhythms of wild brown trout (*Salmo trutta* L.) as an indicator of bioenergetics and behaviour. Journal of Fish Biology 10:299–319.

Rogers, K. B. 1998. Habitat use by largemouth bass and northern pike on the Rocky Mountain Arsenal, Colorado. Doctoral dissertation. Colorado State University, Fort Collins.

Rogers, K. B., and E. P. Bergersen. 1995. Effects of a fall drawdown on movement of adult northern pike and largemouth bass. North American Journal of Fisheries Management 15:596–600.

Rogers, K. B., and E. B. Bergersen. 1996. Application of geographic information systems in fisheries: habitat use by northern pike and largemouth bass. Pages 315–323 *in* L. E.

Miranda and D. R. DeVries, editors. Multidimensional approaches to reservoir fisheries management. American Fisheries Society, Symposium 16, Bethesda, Maryland.

Ross, M. J. 1981. Shielded-needle technique for surgically implanting radio-frequency transmitters in fish. Progressive Fish-Culturist 44:41–43.

Ross, M. J., and J. H. McCormick. 1981. Effects of external radio transmitters on fish. Progressive Fish-Culturist 43:67–72.

Ross, M. J., W. Watts, and A. H. Young. 1981. An ultrasonic biotelemetry system for the continuous monitoring of tail-beat rate from free-swimming fish. Journal of Fish Biology 18:479–490.

Ross, M. J., and J. D. Winter. 1981. Winter movements of four fish species near a thermal plume in northern Minnesota. Transactions of the American Fisheries Society 110: 14–18.

Saltz, D., and G. C. White. 1990. Comparison of different measures of the error in simulated radio-telemetry locations. Journal of Wildlife Management 54:169–174.

Samuel, M. D., and E. O. Garton. 1985. Home range: a weighted normal estimate and tests of underlying assumptions. Journal of Wildlife Management 49:513–519.

Sanderson, G. C. 1966. The study of mammal movements—a review. Journal of Wildlife Management 30:215–235.

Savitz, J., L. G. Bardygula, T. Harder, and K. Stuecheli. 1993. Diel and seasonal utilization of home ranges in a small lake by smallmouth bass *(Micropterus dolomieui)*. Ecology of Freshwater Fish 2:31–39.

Schisler, G. J., and E. P. Bergersen. 1996. Postrelease hooking mortality of rainbow trout caught on scented artificial baits. North American Journal of Fisheries Management 16:570–578.

Schoener, T. W. 1981. An empirically based estimate of home range. Theoretical Population Biology 20:281–325.

Schramm, H. L., and D. J. Black. 1984. Anesthesia and surgical procedures for implanting radio transmitters into grass carp. Progressive Fish-Culturist 46:185–190.

Schulz, U., and R. Berg. 1992. Movements of ultrasonically tagged brown trout (*Salmo trutta* L.) in Lake Constance. Journal of Fish Biology 40:909–917.

Seaman, D. E., J. J. Millspaugh, B. J. Kernohan, G. C. Brundige, K. J. Raedeke, and R. A. Gitzen. 1999. Effects of sample size on kernel home range estimates. Journal of Wildlife Management 63:739–747.

Seaman, D. E., and R. A. Powell. 1996. An evaluation of the accuracy of kernel density estimators for home range analysis. Ecology 77:2075–2085.

Silverman, B. W. 1992. Density estimation for statistics and data analysis. Chapman and Hall, London.

Siniff, D. B., and J. R. Tester. 1965. Computer analysis of animal movement data obtained by telemetry. BioScience 15:104–108.

Skalski, J. R., and six coauthors. 2001. Estimating in-river survival of migrating salmonid smolts using radiotelemetry. Canadian Journal of Fisheries and Aquatic Sciences 58:1987–1997.

Smith, D. C. 1974. Movement and home range of largemouth bass *(Micropterus salmoides)* tagged with ultrasonic transmitters. Master's thesis. Clemson University, Clemson, South Carolina.

Smith, W. P. 1983. A bivariate normal test for elliptical home-range models: biological implications and recommendations. Journal of Wildlife Management 47:613–619.

Snedden, G. A., W. E. Kelso, and D. A. Rutherford. 1999. Diel and seasonal patterns of spotted gar movement and habitat use in the lower Atchafalaya River basin, Louisiana. Transactions of the American Fisheries Society 128:144–154.

Springer, J. T. 1979. Some sources of bias and sampling error in radio triangulation. Journal of Wildlife Management 43:926–935.

Stasko, A. B., and R. M. Horrall. 1976. Method of counting tail-beats of free-swimming fish by ultrasonic telemetry techniques. Journal of the Fisheries Research Board of Canada 33:2596–2598.

Stasko, A. B., and D. G. Pincock. 1977. Review of underwater biotelemetry with emphasis on ultrasonic techniques. Journal of the Fisheries Research Board of Canada 34: 1261–1285.

Swanberg, T. R., and D. R. Geist. 1997. Effects of intraperitoneal transmitters on the social interaction of rainbow trout. North American Journal of Fisheries Management 17:178–181.

Swihart, R. K., and N. A. Slade. 1985. Testing for independence of observations in animal movements. Ecology 66:1176–1184.

Thomas, D. L., and E. J. Taylor. 1990. Study designs and tests for comparing resource use and availability. Journal of Wildlife Management 54:322–330.

Trefethen, P. S. 1956. Sonic equipment for tracking individual fish. U.S. Fish and Wildlife Service, Special Scientific Report—Fisheries No. 179, Washington D.C.

Upton, G., and B. Fingleton. 1985. Spatial data analysis by example: point pattern and quantitative data. Wiley, New York.

Warden, R. L., and W. J. Lorio. 1975. Movements of largemouth bass *(Micropterus salmoides)* in impounded waters as determined by underwater telemetry. Transactions of the American Fisheries Society 104:696–702.

Warner, E. J., and T. P. Quinn. 1995. Horizontal and vertical movements of telemetered rainbow trout *(Oncorhynchus mykiss)* in Lake Washington. Canadian Journal of Zoology 73:146–153.

Weatherley, A. H., S. C. Rogers, D. G. Pincock, and J. R. Patch. 1982. Oxygen consumption of active rainbow trout, *Salmo gairdneri* Richardson, derived from electromyograms obtained by radiotelemetry. Journal of Fish Biology 20:479–489.

White, G. C. 1996. NOREMARK: population estimation from mark–resighting surveys. Wildlife Society Bulletin 24:50–52.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46 (Supplement):120–138.

White, G. C., and R. A. Garrott. 1990. Analysis of wildlife radio tracking data. Academic Press, San Diego, California.

White, G. C., and T. M. Shenk. 2001. Population estimation with radio-marked animals. Pages 329–350 *in* J. J. Millspaugh and J. M. Marzluff, editors. Design and analysis of wildlife radiotelemetry studies. Academic Press, San Diego, California.

Wilkerson, M. L., and W. L. Fisher. 1997. Striped bass distribution, movements, and site fidelity in Robert S. Kerr Reservoir, Oklahoma. North American Journal of Fisheries Management 17:677–686.

Winter, J. D. 1977. Summer home range movements and habitat use by four largemouth bass in Mary Lake, Minnesota. Transactions of the American Fisheries Society 106: 323–330.

Winter, J. D. 1996. Advances in underwater biotelemetry. Pages 555–590 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Worton, B. J. 1987. A review of models of home range for animal movement. Ecological Modelling 38:277–298.

Worton, B. J. 1989. Kernel methods for estimating the utilization distribution in home-range studies. Ecology 70:164–168.

Worton, B. J. 1995. Using Monte Carlo simulation to evaluate kernel-based home range estimators. Journal of Wildlife Management 59:794–800.

Wray, S., W. J. Cresswell, and D. Rogers. 1992. Dirichlet tesselations: a new, non-parametric approach to home range analysis. Pages 247–255 *in* I. G. Priede and S. M. Swift, editors. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, London.

15 Community Indices, Parameters, and Comparisons

Thomas J. Kwak and James T. Peterson

15.1 INTRODUCTION

Understanding assemblages of fishes and how their numbers and compositions change over time and space has long been a fundamental interest of aquatic ecologists and has increasingly become recognized as an important component of fisheries science and management. Whereas much of traditional fisheries management may have focused on single-species approaches, targeting sport or commercial fishes, direct or indirect biotic interactions among fishes may strongly influence target populations. Furthermore, fisheries scientists may frequently be charged with sampling fish populations to detect changes in the aquatic environment, especially those effects related to human activities (e.g., pollution, altered hydrology, or nonnative introductions), and quantitative descriptors of the entire fish assemblage are required for this purpose.

For fish assemblage descriptors to be ecologically relevant, they must be compared over time or among assemblages, and ecologists and biomathematicians have developed procedures to that end. Many of the indices and procedures that we include in this chapter have been developed for use with other taxonomic groups (e.g., plants, invertebrates, and terrestrial animals), or even engineering applications (e.g., communications, Shannon and Weaver 1949), but are equally applicable to the study of fishes. Many have been developed more thoroughly in flowing-water habitats, but the concepts and techniques transfer well to other aquatic systems. In this chapter, we outline, review, and demonstrate quantitative measures and techniques to describe and compare fish assemblages to assist the fisheries scientist in addressing practical research and management objectives.

15.1.1 Definitions

Organisms that occur in a particular place may be classified as a community or an assemblage, and the meaning of these terms varies among ecologists (Morin 1999). The difference between the definitions of these terms lies primarily in the amount and predictability of the interaction among the coexisting organisms. The term

community implies substantial and predictable interaction and may include multiple taxonomic groups such as microflora, plants, and animals, whereas an assemblage is simply the group of species found together and imparts no ecological assumption. Another term, taxocene, is a taxonomically related set of species within a community, such as plants, mammals, or birds (Hutchinson 1978). Thus, a fish community or fish assemblage is a taxocene, yet that term is seldom used. Likewise, a guild is a subset of species that share common resources by similar modes (Root 1967), and although a guild is independent of taxonomy, it is most often a subset of a taxocene. For practical purposes in fish sampling, a fish assemblage is the sum total of the individuals collected at a single sampling location by any single technique or combination of them. For the purposes of this chapter, we employ the commonly applied term fish assemblage to describe the co-occurring fishes in a sample, but we recognize that no fish exists in isolation.

15.1.2 Advantages and Limitations to a Community Approach

The importance of studying fish assemblages, over single species, was evident to early aquatic ecologists (Forbes 1887; Shelford 1929), and today the advantages of a broader ecological approach are obvious and accepted by scientists. However, pragmatic and logistic constraints faced by fisheries scientists do not always allow a holistic perspective. Thus, each investigator must balance the benefit gained in knowledge by a community approach against the additional complexity and effort for each specific application.

The aquatic community is the optimal unit of study as it regulates the flow and storage of energy and materials in the ecosystem. If the fish component is of interest, then the entire fish assemblage is the best unit of study to elucidate the function of this group in the ecosystem. The composition of a fish assemblage is a result of an integration of zoogeography and ecology. Individual fish species vary widely in their morphology, physiology, and tolerance and response to their surroundings. A number of physical factors can limit the ecological success of fish populations, including water quantity, water quality, and physical habitat structure, which in turn set the framework in which biotic interactions occur, such as growth, reproduction, trophic dynamics, and competition (Karr et al. 1986; Fausch et al. 1988; Rabeni and Jacobson 1999). These physical factors may also be quantified by a suite of more proximate measures (e.g., nutrient concentrations, depth profiles, and physical cover) and are further influenced by more broad-scale processes over watersheds and riparian zones. Thus, if any one fish population or guild is limited by a single factor, the effects of other (nonlimiting) environmental influences may not be apparent by merely sampling that fish or subset of fishes.

Fish species of special interest may be atypical in their population dynamics and response to the environment. Some sport and commercial fishes are ubiquitous and tolerant to environmental disturbance (e.g., brown trout, channel catfish, and largemouth bass), and their relative abundance and population dynamics may

depend upon harvest. In contrast, many threatened or endangered fishes are endemic specialists that are extremely sensitive to environmental perturbation (e.g., desert fishes). A widespread, tolerant fish may show no response to habitat degradation or biotic disturbance, whereas a sensitive species may have been extirpated at the earliest signs of perturbation. Thus, single fish species of economic or political importance that are frequently emphasized in fishery surveys or biological assessments may not accurately represent environmental conditions and ecosystem health.

The utility of a community approach is clearly demonstrated in a study by Berkman and Rabeni (1987) to quantify the effects of siltation on stream fishes. They classified fish species from assemblages among sites into guilds based on habitat use, reproductive modes, and feeding behavior. Their guild analysis indicated that species with similar ecological requirements showed a common response to habitat degraded by siltation. Analyses of any single species in their research would likely have been inconclusive, yet the results at the assemblage level yielded strong scientific inference over previous qualitative and anecdotal findings. Further, results examined from a guild approach may allow stronger inference with regard to testing hypotheses about population regulation, as a similar pattern observed among populations within a guild is more conclusive evidence than are trends within a single species.

Valid reasons to forego a community approach in fisheries science also exist. Sampling, sorting, and quantifying all species of a diverse assemblage can be difficult and time consuming, and subsequent data analyses and reporting can be complex. Fish diversity is low in some aquatic ecosystems (e.g., coldwater streams and arctic lakes) and may be dominated by a single species. In such cases, population studies of ecologically important species are reasonable, regardless of practical constraints. Finally, the process of fisheries management and the funding environment for research are strongly governed by economic, sociocultural, and political forces (Krueger and Decker 1999), which may mandate tactical, single-species approaches despite their weaker scientific validity.

Other factors should be considered when contemplating community versus single-species approaches, as no clear criteria exist to guide such decisions. Most fish sampling data are inherently variable over space and time. High variance strongly limits one's ability to detect phenomena statistically, such as the impact of management actions, and can only be overcome by increasing sample size for a given sample design (Chapter 3). Fish assemblage attributes, such as species richness, are generally less variable than are density or abundance estimates for individual species (Peterson and Rabeni 1995). Hence, studies that utilize assemblage data usually require smaller sample sizes to obtain precise estimates and can ultimately be more cost effective than are single-species approaches. The reduction in sample size, however, must be balanced by the additional effort required to process each sample. Thus, any fish sampling protocol should be guided by objectives and scale, specific to the situation, but must be balanced by logistic considerations.

15.1.3 Strategies for Analysis of Community Data

The approach and techniques to employ in analysis of community level data depend upon the objectives of the study and the form and quantity of data. Objectives might be to describe or to compare assemblages. Among descriptive objectives, there may be an emphasis on assemblage structure or ecosystem integrity; comparative objectives may require grouping or ranking of assemblages. Assemblage data may be collected as catch per unit effort or absolute abundance and may be binary (presence–absence), ordinal (ranks), or quantitative (counts), with variable numbers of assemblages and replicates. In Figure 15.1, we present a flow diagram that depicts selection criteria for analytical techniques for community data and may serve as a preliminary guide to the fisheries scientist. Details of criteria, advantages, and shortcomings of each technique are detailed in the appropriate section referenced in the flow diagram. We present example SAS programs (SAS Institute 2004) for performing many of the procedures that we outline in this chapter, analyzing a large-river fish assemblage data set (Box 15.1) along with corresponding results as program output (Boxes 15.4–15.13), but other statistical software applications also perform these procedures. A nonexhaustive list of such software applications is presented in Table 15.1, describing which procedures may be performed using each application. Further, many software applications (e.g., R, SAS, or SPSS) allow more advanced statistical treatment of data depending on the computer programming skills of the user.

15.1.4 Topics Covered

The emphasis of this chapter is on structure, not function; that is, we cover quantitative descriptions of fish assemblage composition and comparison of assemblages rather than describing and understanding community processes and interactions. Assemblage structure is the numerical abundance of each species in the community, and descriptors may include totals or various subtotals of those abundances as well as estimates of biomass. The first step in describing a fish assemblage is designing a sampling protocol to meet specific data requirements, and we provide considerations and suggestions to facilitate that initial process. Once data have been gathered, the process of reducing the resulting data matrix into more meaningful and comparable indices is usually warranted, and we outline those methods in this chapter. We then describe statistical procedures to compare composition among assemblages and conclude with pragmatic suggestions on approaches and interpretation for fisheries scientists. For approaches, methods, and examples that quantify fish community processes, interactions, and forces which, in turn, structure fish assemblages, a topic not covered in this chapter, we refer the reader to Crowder (1990), Gerking (1994), and Matthews (1998); see Krebs (1998), Morin (1999), and Southwood and Henderson (2000) for outstanding general texts on methods for community ecology.

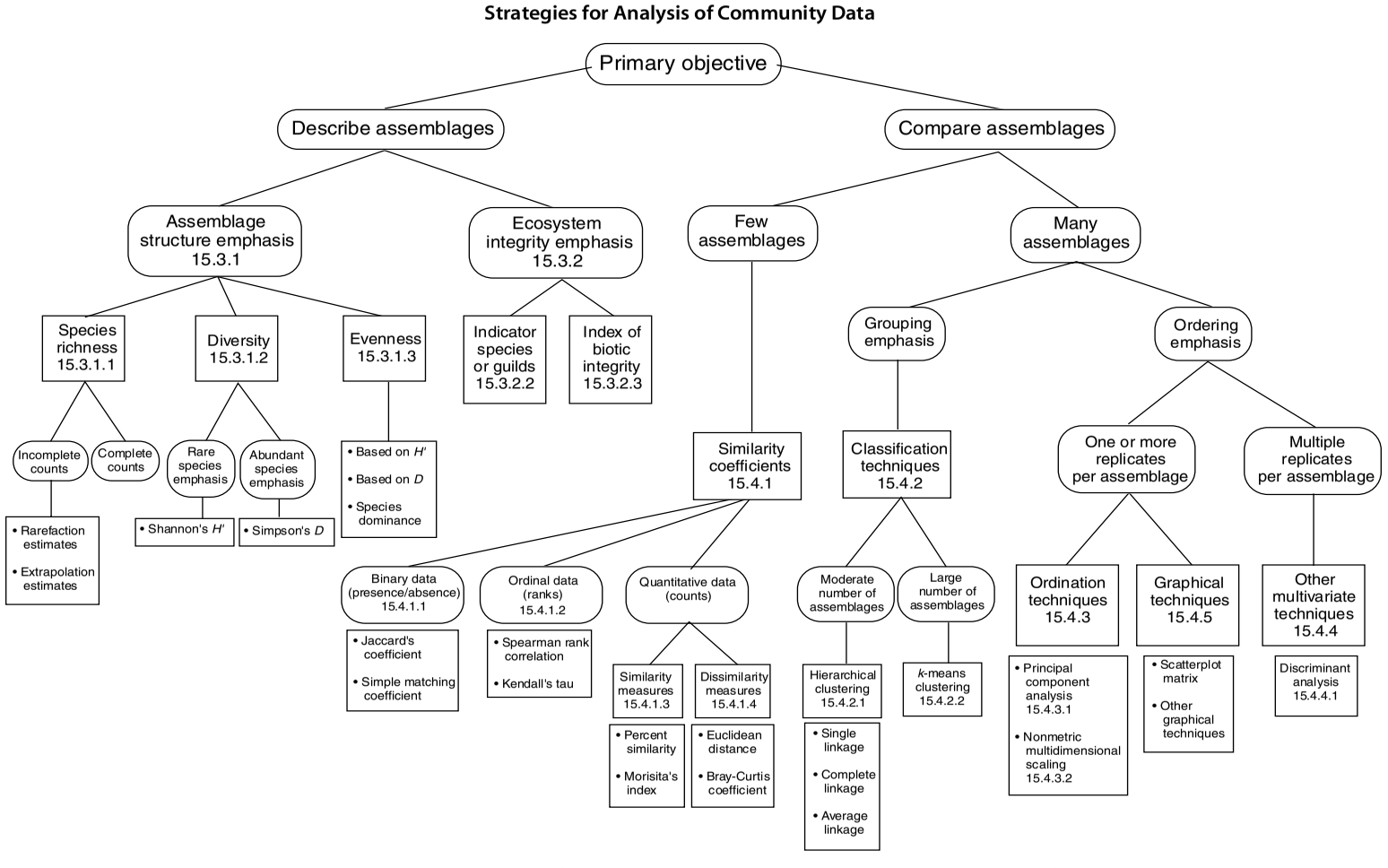


Figure 15.1 Flow diagram depicting selection criteria for techniques available to describe and compare fish assemblages based on objectives and data availability. Objects with rounded borders depict decisions related to objectives or data; those with square borders depict analytical techniques. Numbers refer to appropriate sections within this chapter.

Box 15.1 Sample Data Set and Structural Indices

During a 2-week period in 1988, a survey of the fishes of the Kankakee River, Illinois, was conducted using a boat-mounted electrofisher. Six sites (stations) were each sampled eight times with effort standardized among samples (Peterson 1989; Kwak 1993).

Table The sum of number of individuals from eight samples from each of six sites according to site and species; rare species (occurring in less than 5% of samples) are omitted.

	Station number					
Species and total	1	2	3	4	5	6
Longnose gar	6	7	0	4	26	5
Gizzard shad	164	90	6	6	432	194
Bluntnose minnow	42	33	29	3	44	35
Bullhead minnow	0	0	0	1	15	0
Common carp	13	58	10	14	36	13
Hornyhead chub	0	0	0	0	7	0
Mimic shiner	10	11	10	0	2	0
Redfin shiner	0	22	4	0	0	2
Rosyface shiner	8	89	5	15	8	35
Sand shiner	1	22	4	1	5	1
Spotfin shiner	19	24	3	2	23	8
Striped shiner	45	32	69	14	51	14
Suckermouth minnow	0	0	0	0	4	1
Black redhorse	0	0	4	0	1	0
Golden redhorse	34	0	35	36	9	55
Northern hog sucker	5	2	10	7	2	7
Shorthead redhorse	35	0	8	2	35	22
Quillback	5	2	1	4	17	14
River redhorse	2	0	2	0	0	5
Silver redhorse	8	1	1	3	14	4
Smallmouth buffalo	0	3	0	0	3	0
Brook silverside	4	10	13	9	10	6
Bluegill	2	0	0	2	1	0
Green sunfish	1	42	6	3	7	1
Largemouth bass	0	3	2	0	8	0
Longear sunfish	35	94	26	39	48	37
Orangespotted sunfish	1	0	0	8	31	3
Rock bass	30	3	15	31	27	62
Smallmouth bass	143	59	195	151	165	204
Banded darter	4	0	0	1	0	1
Blackside darter	1	9	1	0	4	0
Johnny darter	0	6	3	0	2	0
Logperch	25	0	51	42	24	7
Slenderhead darter	5	0	6	7	2	2
Total	648	622	519	405	1,063	738

(Box continues)

Box 15.1 *(continued)*

Table Indices of fish assemblage structure for six stations. Shannon's index (H') is given by equation (15.3); Simpson's D is given by equation (15.4); and species dominance is given by equation (15.7).

	Station number					
Assemblage structural index	1	2	3	4	5	6
Richness						
Species	26	22	26	24	31	25
Family	7	7	6	7	7	7
Diversity						
Shannon's H'	2.43	2.56	2.29	2.28	2.31	2.24
Simpson's $(1 - D)$	0.864	0.903	0.818	0.821	0.798	0.832
Evenness						
Based on H'	0.746	0.828	0.704	0.718	0.672	0.695
Based on $(1 - D)$	0.898	0.946	0.850	0.857	0.825	0.867
Species dominance (3 species)	0.543	0.439	0.607	0.573	0.610	0.614

Table 15.1 Nonexhaustive list of statistical software applications for analyzing community level data, according to technique. Symbols indicate that the application incorporates some (S), most (M), or all (A) of the corresponding techniques outlined in this chapter.

Statistical software application	Similarity measures	Hierarchical cluster analysis	*K*-means cluster analysis	Principal component analysis	Nonmetric multidimensional scaling	Discriminant analysis	Graphical analysis
BMDP	S	A	A	A	A	A	S
JMP	S	A	A	A		A	M
Minitab	S	A	A	A		A	S
R	S	A	A	A	A	A	M
SAS/STAT	M	A	A	A	A	A	M
S-plus	M	A	A	A	A	A	M
SPSS	S	A	A	A	A	A	S
STATA	S	A	A	A			
Statistica	S	A	A	A	A	A	S
Systat	M	A	A	A	A	A	M

15.2 SAMPLING CONSIDERATIONS AND ASSUMPTIONS

As with all fishery assessments, the analysis and interpretation of fish community indices are significantly influenced by the quality and quantity of data. Fish sampling bias can obscure relationships or, worse, suggest false relations (Bayley and Dowling 1993), and sample variance can affect the ability to detect relationships statistically (Peterson and Rabeni 1995). Standardized sampling protocols help

maintain data quality by ensuring that data are collected in a consistent manner over space and time. The influence of sampling bias and variance on single-species approaches and the importance of standardized sampling have previously been covered elsewhere (e.g., Brown and Austen 1996; Chapters 2 and 3), and most of the principles are applicable to community approaches. There are, however, several considerations unique to sampling entire fish assemblages that we consider here.

When using community indices and comparing communities, the primary assumption is that fish samples are representative of the "true" fish assemblage. That is, the number and types of species caught and their relative abundances accurately reflect those of the fish assemblage occupying the study area (e.g., lake or stream). Fish assemblages, however, are composed of species of different sizes, forms, and behaviors that can affect their vulnerability to capture by any sampling gear. Consequently, samples are influenced by these characteristic differences to varying degrees, resulting in an inaccurate representation of the fish assemblage. Similarly, fish species often use habitats that differ in size, structure, and distribution; hence, the types and allocation (relative amounts) of habitats sampled can also significantly affect the adequacy of the fish assemblage sample. It is important that fisheries scientists identify potential influences on data quality and develop sampling designs that minimize these, so that analyses of fish assemblages will be based on reliable data. Below, we identify some noteworthy influences on the quality of fish assemblage data and discuss methods to minimize their influence and to evaluate the adequacy of sampling designs.

15.2.1 Sources of Sampling Bias

Aquatic ecosystems are defined by characteristic physical, chemical, and biological attributes that may simultaneously influence sampling efficiencies and regulate fish assemblage structure. For example, water depth can affect the efficiency of many fish sampling methods and also can influence the structure of fish assemblages. Consequently, observed differences in the structure of fish assemblages collected in study areas with very different depths could be due to sampling efficiency, assemblage structure, or both. Failing to account for differences in sampling efficiency when comparing locations with different physical characteristics and species assemblages, or among samples collected with different methods, can introduce a systematic error or bias into the data, which can invalidate experiments or observational studies (Hurlbert 1984). To minimize the influence of sampling bias on fish assemblage studies, scientists should collect fishes with the most efficient method or combination of methods for which bias is known and sample under circumstances where catchability is reliable. Thus, fisheries scientists must consider those major factors affecting sampling efficiency and choose the most appropriate gear or combination of gears for their particular sampling situation. When sampling conditions are particularly challenging (e.g., large rivers and reservoirs), fisheries scientists also should consider an analysis based on more qualitative measures (e.g., species presence or rank abundance) or estimate

species abundances using mark–recapture or other methods that can account for sampling efficiency differences.

15.2.1.1 *Species and Body Size*

The efficiency of most sampling gears is influenced by the species and size of fish encountered (Hayes et al. 1996 and references therein). Body shape or morphology can influence a fish's vulnerability to capture. For example, species with cryptic coloration and reduced or absent swim bladders are often difficult to locate when stunned during electrofishing. Species-specific behaviors, such as vertical position in the water column, also affect sampling efficiency. Benthic species (e.g., darters, sculpins, and North American catfishes), particularly those using deepwater habitats, and wide-ranging pelagic species (e.g., temperate basses and herrings) are difficult to sample effectively. Body size, within and among species, is also an important factor affecting sampling efficiency. For most sampling gears, the lowest efficiencies tend to be for the extreme sizes of fish (i.e., very small and large individuals). These sampling biases often result in fish assemblage samples that overrepresent species and sizes that are most vulnerable to sampling. To minimize the influence of these biases, fisheries scientists can develop sampling efficiency models to adjust sampling data for differences in catchability. These estimates, however, require extensive gear evaluations to develop efficiency models. Excluding small fishes and species that are difficult to catch from analyses and using species presence–absence or rank abundances for the assemblage analysis (see sections 15.4.1.1 and 15.4.1.2) also can minimize this source of bias.

15.2.1.2 *Habitat Characteristics*

The physical characteristics of a sampling location can affect the efficiency of most fish sampling gears. The dimensions of a sampling location (e.g., water depth or stream width) can change the capture efficiency of a variety of gears. Sampled areas wider and deeper than the effective catch area (e.g., electrical field size or seine dimensions) can reduce capture efficiency (Bayley and Dowling 1990). In rivers, high current velocities can displace stunned fish from the electrical field before they are captured, facilitate fish escape from seines, and prohibit the use of some passive sampling gears (e.g., gill nets). Similarly, water transparency (color and turbidity) can greatly influence the application and bias of underwater observation techniques and electrofishing and passive-sampling gears. Structures within the sampling area (e.g., vegetation, woody debris, and boulders) can provide a refuge for fishes and can limit sampling efficiencies (Bayley and Dowling 1990; Rodgers et al. 1992). These biases often result in samples that overrepresent those species occupying habitats that are easier to sample and underrepresent those in habitats that impair sampling. Similar to species and size biases, the influence of habitat biases can be minimized by adjusting catch data for differences in sampling efficiency. When sampling efficiency estimates are unavailable, biologists can minimize habitat biases by grouping habitat types into strata (Chapter 3). With such designs, comparisons of fish assemblages should be restricted to similar

habitat types within each stratum. Fisheries scientists should also consider expending greater effort in habitats that are difficult to sample to ensure adequate representation of the fish assemblage in these areas.

15.2.1.3 Gear Type

In some instances, using the proper sampling gear can reduce (not eliminate) the influence of species, size, and physical habitat biases on the analysis of fish communities. Thus, selecting the proper sampling gear is one of the most critical components of a fish community study. Electrical gears are among the most efficient and widely used techniques for sampling fish assemblages in relatively shallow waters, such as small- to medium-sized streams and lake and river shorelines. Sampling in deeper waters would likely require the use of active (e.g., dredges or trawls) or passive (e.g., hoop, trap, and fyke nets) techniques designed to sample these habitats more effectively (Hubert 1996). When sampling conditions (habitats) vary considerably within a study area (e.g., large lakes, rivers, and reservoirs), no single sampling gear can adequately sample the entire fish assemblage. Hence, a multi-gear approach, in which the most effective gear(s) is used in each habitat type, can provide the most complete estimate of fish assemblage structure. However, because such estimates can be biased to varying, unknown degrees, they should not be accepted as representative of the "true" fish assemblage unless the effectiveness of each gear can be evaluated.

15.2.2 Sampling Season

Season can have a profound influence on fish assemblage structure in many freshwater ecosystems. Fishes often migrate seasonally to fulfill one or more life history requirements (e.g., spawning or juvenile rearing) and to seek refuge during severe environmental conditions (Hall 1972; Schlosser 1982; Bayley and Osborne 1993; Peterson and Rabeni 1996; Grossman et al. 1998). Thus, the structure of the fish assemblage in open systems (e.g., streams and rivers) is likely to vary among seasons where fish can freely migrate to or from study areas. In closed systems (e.g., lakes and ponds), seasonal movements can affect the assemblage structure within habitat types, thereby increasing variance (see section 15.2.3). Similarly, fish movements within a season also can alter variability of assemblage samples. Fish movement is generally greatest during the spring and fall in the northern hemisphere, which can increase sample variance. To avoid the influences of seasonal fish movement, fisheries scientists should limit comparisons of fish assemblages to similar seasons and, if possible, to seasons with the least amount of fish movement (e.g., summer for temperate, warmwater stream fishes).

Fish growth and recruitment to sampling gear also influence seasonal measures of assemblage structure. Young-of-the-year (age-0) fishes—usually the most abundant age-class—are often recruited to sampling gears by the end of their first growing season. This increases the probability of collecting less abundant and difficult to sample species (Gray 1987; Wright 1988), resulting in perceived increases in the number of species during such periods (Peterson and Rabeni 2001).

Eliminating age-0 fishes from an analysis can reduce the influence of seasonal gear recruitment but may also reduce estimates of species richness. When research objectives include analyses of age-0 fish, sampling should be conducted later in the growing season when age-0 fish are larger and more vulnerable to sampling.

15.2.3 Fish Assemblage Sampling Designs

Sampling design is an essential component of fish assemblage studies. Comparisons of fish assemblage structure require data that accurately reflect the true species composition and species' relative abundances. One means of ensuring accurate representation of the species assemblage is through effective sampling design (also see Chapter 3). Fish distribution and assemblage structure are influenced by physical habitat features and resource availability (Gorman and Karr 1978; Schlosser 1982). Thus, designs should ensure that all habitat types in a study area are properly represented. High variance among samples, another factor affecting the accuracy of fish assemblage estimates, is influenced by species-specific characteristics (e.g., behavior) and sampling conditions (e.g., gear type and habitat features) and can be overcome only by increasing sample size. Increasing sample size also improves the likelihood of detecting rare species in the assemblage. The diverse nature of freshwater ecosystems (e.g., habitat types and species) dictates that no single sampling design is best for all community level studies. Rather, the best approach will depend upon the objectives of the study, type of system being studied, and characteristics of the fish assemblage. Here, we discuss two basic approaches to sampling fish assemblages that can be modified to fit most community level studies in freshwater systems. We also strongly encourage scientists to evaluate the adequacy of their particular sampling design to ensure data quality.

15.2.3.1 Quadrat Sampling

Quadrat sampling entails dividing a study area into sample units (quadrats) and sampling a random selection of them. With this design, sample variance can be minimized by sampling greater numbers of quadrats. The number of samples required to meet study objectives can be determined with traditional statistical techniques (Chapter 3) and by analyzing species accumulation curves (section 15.3.1.1). Greater efficiency (lower variance) can often be gained by stratifying study areas according to habitat type and randomly sampling quadrats within each stratum.

In addition to sample size requirements, the size of individual quadrats should be considered prior to adopting a quadrat sampling approach. Larger sample units generally contain greater numbers of individuals, which can increase the chances of collecting an individual of another species (Connor and McCoy 1979; Angermeier and Schlosser 1989). To avoid this species–area effect, it is often preferable to maintain a consistent quadrat size among study areas or through time (when monitoring). When study areas differ substantially in size and structure, a single quadrat size could incorporate variable habitat heterogeneity among areas,

which could bias comparisons. For example, pools and riffles generally occur every five to seven stream widths in gravel-dominated streams (Leopold et al. 1964; Gordon et al. 1992). Thus, a single quadrat size based on stream length or area would incorporate a greater number of pools and riffles in smaller streams. In these instances, natural discrete morphological features (i.e., channel units; Hankin and Reeves 1988; Peterson and Rabeni 2001) could be used as sampling quadrats. These natural quadrats, however, should be sampled in proportion to their relative abundance in the study area to ensure proper representation of the fish assemblage.

15.2.3.2 Constant Ratio Sampling

Constant ratio sampling involves collecting fishes from a single sample unit, the dimensions of which are scaled relative to the size of the study area. This approach is generally used for stream studies where the size of the sample unit (stream reach length) is proportional to stream width (e.g., station length equals 35 stream widths). The size of the sample unit needed to obtain a representative sample is determined by examining the cumulative catch of species (see section 15.3.1.1) with increasing sample unit size. Thus, this design differs from quadrat sampling in that it attempts to standardize the sampling effort at a single location and time in order to obtain a representative sample of the fish assemblage. The required sample unit size, however, can vary widely among systems due to differences such as habitat characteristics, sampling efficiency, and fish abundance and assemblage structure (Lyons 1992; Angermeier and Smogor 1995; Paller 1995). Hence, no single ratio or proportion will likely be adequate for sampling fish assemblages in all freshwater systems. Additionally, data collected following a constant ratio design cannot be used to make statistical inferences about fish assemblage patterns within a study area because of the general lack of replication (i.e., only one sample collected).

15.2.4 Data Standardization

There are innumerable ways to collect and quantify fish assemblage samples. Numerical abundance for each species may be expressed as total catch, relative catch as a proportion of that of all species, catch per unit effort, catch per area, or catch per linear distance; in addition, adjusted or absolute abundance (density or biomass) may be estimated (see Ricker 1975 for examples). Furthermore, the definition of a fish assemblage for quantitative purposes may vary widely (Rahel et al. 1984; Grossman et al. 1990; Matthews 1998).

Investigators may simplify analyses and reduce variable sampling bias by defining the fish assemblage as a subset of the actual sample. This practice may exclude age-0 fish, juvenile fish, or rare species from analyses. Rare species, those found in less than 5% of the collections (Gauch 1982), are generally excluded in most community analyses because (1) it is unlikely that rare species significantly influence the dynamics of the fish community, (2) the occurrence of a rare species may be a random event unrelated to any life history requirement, and (3) many

multivariate statistical techniques are sensitive to rare species, which could distort meaningful, significant trends. However, patterns in occurrence of rare species among assemblages may provide insight into assemblage organization and form the basis of conservation strategies.

Data may also be mathematically transformed to reduce the importance of extreme values (e.g., logarithmic or square root). If data are omitted or transformed for subsequent analyses, it must be justified on some ecological, statistical, or theoretical basis, and data must be modified objectively, based on a systematic criterion. Data manipulation, such as omitting rare species or transformation, will change results of virtually all of the indices and procedures presented in this chapter and should be considered carefully.

Sampling techniques, effort, area, and resulting numerical expressions of fish abundance should be standardized within a study if possible and described in sufficient detail to allow comparisons among studies. Standardizing fish collection techniques and sites may reduce the effects of variable sampling bias associated with gear type, habitat, or time. Standardized data and manipulation will limit erroneous conclusions that may be drawn from statistical artifacts or comparison of incongruous data. However, of utmost importance in community level studies is that a precise definition of the assemblage be provided and data collection and analyses be described in sufficient detail to allow interpretation and additional analyses by others. This detail should include (1) the organisms considered to compose the assemblage (e.g., size or age criteria, rare species criteria, fin fishes, shellfish, other invertebrates, or aquatic herpetofauna); (2) sampling techniques, effort, area and boundaries, and timing; (3) data form and units; and (4) any data manipulation prior to analyses.

15.3 COMMUNITY INDICES—THEIR CHARACTERISTICS AND ESTIMATION

It is indeed appealing and useful to attempt to summarize the abundance data for multiple species in an assemblage into a single number describing assemblage structure. However, this section must begin with a word of caution. Abundant critiques and revisions in the ecological literature suggest that there is no "silver bullet" or perfect community index to serve all purposes. Essentially all of the indices presented below have received criticism for their shortcomings and misapplications (Hurlbert 1971; Washington 1984), and we urge the fisheries scientist to view these indices as relative values with variable precision, accuracy, and reliability. They are most appropriately used to compare assemblage data collected in a standardized manner within a study, rather than as broad, comparative tools among studies or over expanded scales. Index selection should consider statistical robustness, data availability, and specific study objectives. An approach utilizing several indices may prove useful to verify findings and to balance shortcomings of any single index.

Two primary approaches to quantify community structure have been developed by ecologists and applied to fish assemblages. They are the use of (1) community structural indices based directly on field samples and (2) biotic indices

based on the relative abundance of indicator organisms. Both approaches are applicable to describing fish assemblage characteristics and may be related to environmental quality, but biotic indices are especially suited to quantifying ecosystem health or ecological integrity, which may be reduced by pollution, stressful environmental conditions, or habitat degradation and destruction. As such, biotic indices do not directly represent assemblage structure. Structural indices are broad, quantitative descriptors of assemblage structure, and biotic indices are specific parameters based on a subset of indicator organisms within the assemblage, and their applications are not interchangeable.

15.3.1 Structural Indices

The relative abundance of species, other taxa, or other meaningful categorical attribute within an assemblage may be combined into a single measure that is intended to describe the state of the community. The most common of these measures is species diversity, which incorporates the number of species in an assemblage (species richness), as well as the relative abundance of those species (evenness) (Figure 15.2). Whereas such structural indices are usually calculated at the species level, it is equally appropriate to estimate them at any taxonomic or

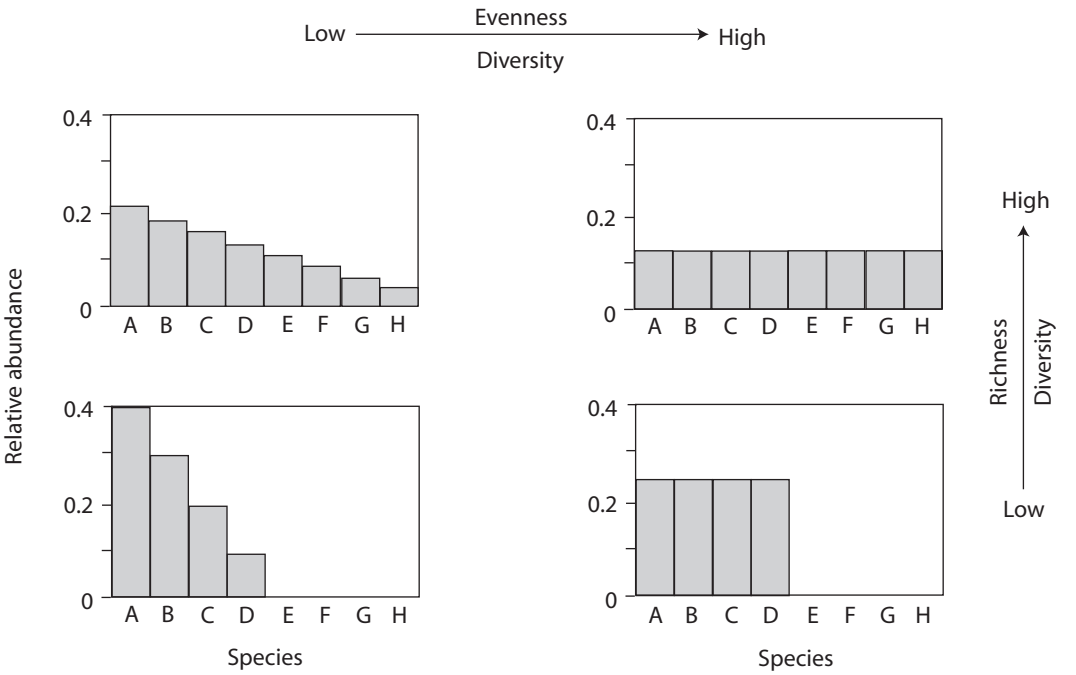


Figure 15.2 The concept of species diversity illustrated by variable relative abundance of species among assemblages. Diversity increases with increases in the number of species (richness) and equitability of distributions among species (evenness).

other hierarchical classification level (Osborne et al. 1980). The level at which to estimate structural indices is often determined by the practical ability to classify organisms. For aquatic invertebrate assemblages, which may comprise species-rich assemblages or contain unresolved taxonomic groups, these indices are often estimated at levels higher than species (e.g., genus or family). But for fish assemblages, the practical level of identification and structural index estimation is usually the species. Most assemblage indices are based on taxonomic classification, as are the examples in this chapter, but it is worth considering other biological and ecological attributes of fishes, in addition to, or in lieu of, taxonomy to describe assemblage structure. Among those that may reveal insightful, functional patterns are life history or morphological traits, habitat affinity at various spatial scales, and tolerance to environmental conditions (Bain et al. 1988; Winemiller and Rose 1992; Poff and Allan 1995; Angermeier and Winston 1999; Quinn and Kwak 2003).

15.3.1.1 Species Richness

The simplest and oldest assemblage structural index is species richness—simply a count of the number of species represented in an assemblage. Again, fish assemblage richness may also be expressed at genus, family, or other classification level. For demonstration purposes, an example data set from the Kankakee River, Illinois, is provided in Box 15.1. Among the six sites sampled, fish species richness varied from 22 to 31 species, and family richness varied from six to seven families. The variation between taxonomic levels in this example (nine species versus two families) illustrates the differing utility among levels of classification; structural indices based on very broad class levels are unlikely to provide the resolution necessary to be ecologically relevant.

While the concept of species richness appears simple, that is rarely the case. If an investigator is able to collect or count all individuals of an assemblage in a sampling area, expressing species richness is simple—the count of species present. However, when sampling fishes in aquatic environments, this is rarely the case, and scientists usually collect a sample of the assemblage rather than a complete count. Such samples are incomplete and variably biased by sampling technique and associated influences of sampling habitat, effort, area, and time, as discussed above—all of which affect the ability to sample or detect a species. This limitation is especially important when detecting rare species is a priority. In general, the larger the sample or the greater the number of samples collected, the greater the number of expected species. Consequently, it may be misleading to compare species richness among samples or sites that are based on incomplete counts with varying sample sizes, area sampled, or effort expended. But how can you estimate the number of species not detected? Several approaches to this sampling problem have been developed that are applicable to fish species richness.

Estimating species richness by rarefaction. Rarefaction is a statistical method to compare species richness among assemblage samples of different sizes (i.e., different numbers of individuals per sample). This procedure was first developed by Sanders (1968) to compare marine benthic assemblages and was later corrected for an

error by Hurlbert (1971) and Simberloff (1972). It mathematically "rarefies" a large sample of known species richness to estimate what richness would be for samples of fewer individuals. If rarefaction is performed for a number of sample sizes, a rarefaction curve can be constructed for that assemblage to serve as a tool for comparing species richness among assemblages for equal sample sizes. This process thus reveals differences in species richness among assemblages, independent of sample size.

The rarefaction algorithm assumes a hypergeometric distribution of the species–abundance relationship as

$$E(S_n) = \sum_{i=1}^{S} \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right], \tag{15.1}$$

$E(S_n)$ = expected species richness of a random subsample;
S = total number of species in the collection;
N = total number of individuals in the collection;
N_i = number of individuals of species i;
n = number of individuals in the random subsample; and
$\binom{N}{n}$ = number of combinations of n individuals that can be selected from a sample of N individuals, or $N!/n!(N-n)!$.

An example calculation of expected species richness of a smaller sample is presented in Box 15.2, and the algorithm to estimate the variance of $E(S_n)$ may be found in Heck et al. (1975). Once a large sample has been rarefied, expected species richness from smaller samples of that assemblage can be compared to samples of equal abundance from other assemblages to compare richness among assemblages, independent of sample size. Plotting rarefaction curves (see Box 15.2) of multiple assemblages on a single plot is a useful means to compare species richness. Furthermore, if fish density (number per area) of an assemblage is estimated (Chapter 8), expected species richness as a function of sampling area can also be plotted (i.e., species density curve; see Gotelli and Graves [1996] for an example).

Rarefaction can be a useful fisheries or ecological tool (see examples by Glowacki and Penczak [2000] and Quinn and Kwak [2003]). Management strategies, ecological assessment, and hypothesis testing require information on species richness and diversity, independent of sampling size, area, or effort. Fish assemblages may be compared using rarefaction over time or among locations, and associated precision may be estimated as confidence intervals. Rarefaction may also be employed in the development of monitoring programs to determine the sufficient sample size and effort to detect an acceptable proportion of species present.

Box 15.2 Estimation of Species Richness by Rarefaction

A cumulative sample of fishes from station 1 of the Kankakee River, Illinois, included 648 individuals representing 26 species (see Box 15.1). Below, we estimate the expected species richness from a sample of 100 individuals.

From equation (15.1),

$$E(S_n) = \sum_{i=1}^{S} \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right]$$

$$E(S_{100}) = \left[1 - \frac{\binom{648 - 6}{100}}{\binom{648}{100}} \right] + \left[1 - \frac{\binom{648 - 164}{100}}{\binom{648}{100}} \right] + \left[1 - \frac{\binom{648 - N_i}{100}}{\binom{648}{100}} \right] + \left[1 - \frac{\binom{648 - 5}{100}}{\binom{648}{100}} \right].$$

(longnose gar) (gizzard shad) (23 other species) (slenderhead darter)

The summation term for longnose gar is calculated as

$$\binom{648 - 6}{100} = \frac{642!}{100!(642 - 100)!} = 1.7668 \times 10^{119}$$

$$\binom{648}{100} = \frac{648!}{100!(648 - 100)!} = 4.8506 \times 10^{119}$$

$$\left[1 - \frac{1.7668 \times 10^{119}}{4.8506 \times 10^{119}} \right] = 0.6358.$$

Thus,

$E(S_{100})$ = longnose gar term + gizzard shad term + total of 23 other species terms + slenderhead darter term
= 0.6358 + 1.0 + 15.4213 + 0.5687
= 17.626 species.

Therefore, we would expect a random sample of 100 fish from station 1 to include about 18 species. We may then calculate expected species richness for a number of other smaller sample sizes by repeating the calculation above, varying n to develop a rarefaction curve. For station 1, some values of $E(S_n)$ are $E(S_{50}) = 14.040$; $E(S_{200}) = 21.110$; $E(S_{300}) = 22.940$; $E(S_{400}) = 24.107$; $E(S_{500}) = 24.975$; and $E(S_{600}) = 25.693$, resulting in the rarefaction curve below.

(Box continues)

Box 15.2 *(continued)*

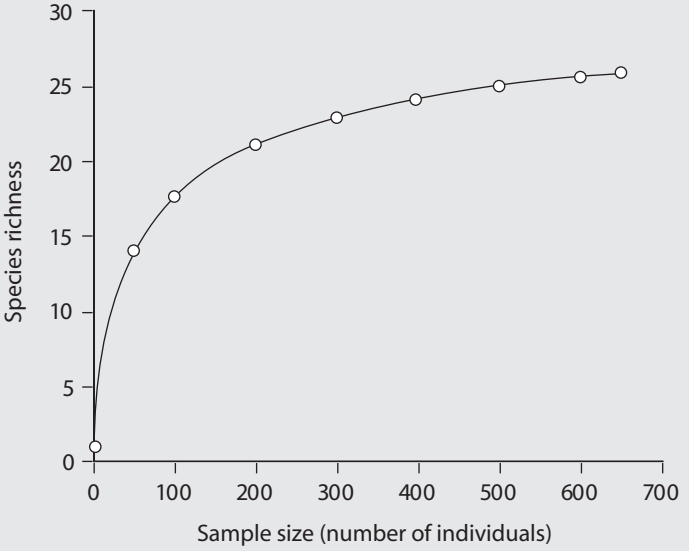


Figure Rarefaction curve to estimate expected species richness of the fish assemblage at station 1 (from Kankakee River, Illinois; Box 15.1) based on sample size. Species richness cannot be estimated for sample sizes that exceed those of the data upon which the curve has been developed.

Note that in the case where $(N - N_i)$ is less than n, then $\binom{N - N_i}{n} = 0$ by definition (no combination of a greater number of individuals can be chosen from a set of fewer individuals). In this case

$$\left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}}\right] = 1 .$$

Unfortunately, the equations above require calculations of very large numeric values (e.g., 648!, as above) that cannot be processed directly on a personal computer. Numbers greater than 170! cannot be stored in floating-point, double-precision arithmetic on a typical personal computer. However, we can perform the calculations on lower values by using the natural logarithm of the gamma function, $\Gamma(x)$ (GAMMALN function in Microsoft Excel), where the factorial of any integer (x) may be calculated as

$$x! = e^{\log_e[\Gamma(x+1)]} .$$

Equation (15.1), expressed using the gamma logarithm, is

$$E(S_n) = \sum_{i=1}^{S} [1 - e^{\{[\log_e \Gamma(N - N_i + 1)] - [\log_e \Gamma(n+1)] + [\log_e \Gamma(N - N_i - n + 1)]\} - \{[\log_e \Gamma(N+1)] - [\log_e \Gamma(n+1)] + [\log_e \Gamma(N - n + 1)]\}}] .$$

(Box continues)

Box 15.2 *(continued)*

Each summation term is expressed as a formula in Excel as

$$1 - (\text{EXP}((\text{GAMMALN}(N - N_i + 1) - (\text{GAMMALN}(n + 1) + \text{GAMMALN}(N - N_i - n + 1))) - (\text{GAMMALN}(N + 1) - (\text{GAMMALN}(n + 1) + \text{GAMMALN}(N - n + 1))))).$$

The Excel formula for the longnose gar term in the example above would be

$$1 - (\text{EXP}((\text{GAMMALN}(648 - 6 + 1) - (\text{GAMMALN}(100 + 1) + \text{GAMMALN}(648 - 6 - 100 + 1))) - (\text{GAMMALN}(648 + 1) - (\text{GAMMALN}(100 + 1) + \text{GAMMALN}(648 - 100 + 1))))).$$

These tedious computations are best carried out in a spreadsheet application or a specifically developed computer program, such as that provided by Krebs (1998).

Rarefaction has several limitations and assumptions that must be considered in its use and interpretation of results. Rarefaction curves may not be extrapolated beyond the number of individuals in the largest sample. Thus, we only address the question of undetected species for smaller samples. Rarefaction should be applied only to samples from similar habitats using similar sampling techniques. For example, the rarefaction curve developed from sampling fishes of the Kankakee River by means of a boat-mounted electrofisher (Box 15.2) should not be compared with samples from that river collected using other gears or from other water bodies, which we know support different species diversities. Rarefaction assumes a random distribution of individuals, which is rarely true for fishes, and it does not incorporate information about species identity or relative abundance among species.

Estimating species richness by extrapolation. Rarefaction estimates species richness for smaller samples of individuals, but an investigator may wish to estimate the total number of species in an assemblage (i.e., how many species remain undetected?). Careful extrapolation is required to address such questions. A simple technique to extrapolate species richness beyond the boundaries of empirical data is to develop a species accumulation curve. In this procedure, the cumulative number of species collected is plotted against increasing numbers of combined samples of equal effort or area. If samples are quadrat samples within a larger area, their sequential order on the plot should be random; if the samples are a time series from the sample site, they may be applied sequentially or randomly. Various regression techniques have been used to model the resulting relationship, but the linear regression of cumulative species as a function of the logarithm (base 10 log-linear model) generally performs well on empirical data and is simple to apply (Palmer 1990). This plot and resulting regression model may be used to estimate species richness, coinciding with larger numbers of samples (greater effort or sampling area).

We present the species accumulation curve for eight sequential fish assemblage samples from station 1 of the Kankakee River, Illinois, as an example (Figure 15.3, see Box 15.1 for summed data). After sampling the same area eight times with equal effort during a 2-week period, we collected 26 species. However, cumulative species richness increased from 18 species after our first sample to that cumulative total. Examination of the species accumulation curve, extrapolated to 100 samples, suggests that total species richness for that area is greater and that additional sampling would have increased our species count (35.6 species, Figure 15.3b). For example, if resources were available to collect twice as many samples (16), we can use the species accumulation regression function (Figure 15.3) to estimate that we could expect to collect 2–3 more species at that site (28.5 species).

The log-linear model is nonasymptotic; that is, species richness will continue to increase with samples. There are several other asymptotic models available that may be applied to species accumulation curves and parametric and nonparametric estimators to extrapolate estimates of species richness to a maximum number of species; these are reviewed by Colwell and Coddington (1994). For a single sample of an assemblage, they identified an eloquent, nonparametric estimator by Chao (1984) as the best for estimating total species richness (S_{total}) as

$$S_{total} = S_{obs} + (a^2/2b), \tag{15.2}$$

where S_{obs} = observed number of species in a sample, a = number of species represented by a single individual in the sample, and b = number of species represented by exactly two individuals in the sample.

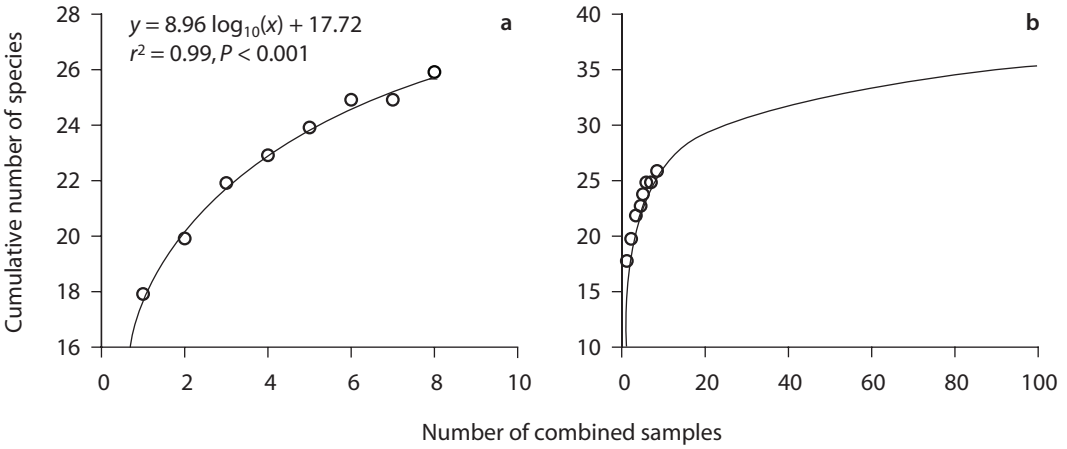


Figure 15.3 Species accumulation curve for eight sequential fish samples collected from station 1 on the Kankakee River, Illinois (Box 15.1), describing the relationship of cumulative species richness (y) and number of samples (x) fitted to a log-linear function. Left panel (a) is the curve within the range of empirical data; right panel (b) is the same relationship extrapolated to 100 samples.

Thus, for our Kankakee River fish assemblage example (station 1; Box 15.1), the total estimated species richness for that site would be 30 species ($S_{obs} = 26$, $a = 4$, $b = 2$). However, our species accumulation function for that site (Figure 15.3) suggests that about 24 samples (if $y = 30$, $x = 23.47$) would be required to detect 30 species. A variance estimator for S_{total} and application of equation (15.2) to presence–absence data are found in Chao (1984). Recently, ecologists have applied models, originally developed to estimate population size and related parameters, to estimate species richness for communities that include species with varying detection probabilities; such models require multiple samples but may offer advantages over other extrapolation techniques (Boulinier et al. 1998; Bayley and Peterson 2001).

Similar limitations and assumptions, noted for rarefaction above, apply to extrapolation techniques for estimating species richness. However, we urge extreme caution in applying any of these techniques or any statistical procedure that derives estimates by extrapolation beyond the boundaries of empirical data. The nature and form of a relationship may change at wider ranges of variables. This is dangerous territory indeed, and caution and common sense must be exercised to avoid reporting and accepting erroneous and invalid findings. Extrapolation procedures are best employed to develop hypotheses and management scenarios for testing rather than to be used as a basis for critical management decisions.

15.3.1.2 Diversity

Diversity indices combine information on the number of species in an assemblage (richness) and their relative abundance (evenness). Unfortunately, there is no correct means of assigning proportional weighting between these two components, and thus dozens of diversity indices have been developed and applied by ecologists seeking to improve on previous forms (Hurlbert 1971; Washington 1984). What, then, does a diversity index convey? A diversity index is a parameter describing assemblage structure, but any relationship to ecological function, such as productivity or stability, remains unclear. Diversity indices have been criticized for lack of biological relevance and should be considered only one of many tools available to describe assemblage structure —they are not a substitute for in-depth examination (Hurlbert 1971; Pielou 1975; Washington 1984).

We provide equations and examples of two common diversity indices applied at the species level for fishes—Shannon's H' and Simpson's D. Shannon's index of diversity (Shannon and Weaver 1949) has endured ongoing criticism yet remains widely used in biology, and it is the most widely applied diversity index in aquatic systems (Washington 1984). Despite its theoretical and ecological shortcomings, its use is probably justified as a comparative index until such time that a more suitable alternative becomes accepted. It was independently developed by Shannon and Wiener at about the same time and is often referred to as the Shannon–Wiener index or function. Shannon's index (H') is based on information theory and is defined as

$$H' = -\sum_{i=1}^{S} (p_i)(\log_e p_i), \tag{15.3}$$

where s = number of species, and p_i = proportion of the total sample represented by the ith species.

Another less common, but among the simplest, diversity (or concentration) index is Simpson's D (Simpson 1949). Simpson's index of diversity is based on the notion that diversity is inversely related to the probably that two individuals sampled at random from an assemblage will be of the same species. Thus,

$$D = \sum_{i=1}^{s} (p_i^2), \tag{15.4}$$

D = Simpson's measure of concentration,
$1 - D$ = Simpson's diversity index,
s = number of species, and
p_i = proportion of the total sample represented by the ith species.

Shannon's index is sensitive to changes in rare species in the community and is considered a type I diversity index, whereas Simpson's index is influenced to a greater extent by abundant species and is a type II index (Peet 1974; Krebs 1998). Thus, selection of a diversity index should not be an arbitrary process and may vary with specific study objectives and investigator interests. Shannon's, Simpson's, and other diversity indices have been represented by various forms that are based on the original theory proposed (Pielou 1975; Washington 1984; Krebs 1998), so investigators should report the exact algorithm used to compute the index rather than simply citing a reference. Example calculations of species diversity indices for the Kankakee River fish assemblage are presented in Box 15.3.

There are many alternative statistical and practical methods to describe diversity. Shannon's and Simpson's diversity indices are nonparametric measures that imply no assumption about the species abundance distribution of an assemblage. An alternative approach used by community ecologists to describe diversity is to use statistical sampling theory to fit distribution models to species abundance data; examples of these include the logarithmic series, lognormal, geometric series, uniform, and broken-stick distributions (Gotelli and Graves 1996; Krebs 1998). Because of the complexity and lack of theoretical or biological justification for these statistical distribution approaches, nonparametric indices, such as Shannon's and Simpson's have become more widely applied in aquatic science (Washington 1984; Krebs 1998).

15.3.1.3 Evenness

Evenness is a measure of the equitability in relative abundance among species. To report only diversity as an assemblage structural index confounds the effects of species richness and evenness; thus, it is appropriate to report richness, diversity, and evenness when describing fish assemblage structure. There are many approaches to quantifying evenness, but the most common is to express it as a proportion of estimated diversity relative to the corresponding maximum diversity for the specific number of species and sample size.

Box 15.3 Calculation of Species Diversity, Evenness, and Dominance.

Below, we estimate species diversity and evenness of the fish assemblage sample from station 1 of the Kankakee River, Illinois, which included 648 individuals representing 26 species (Box 15.1).

Shannon's Diversity Index (H')
From equation (15.3),

$$H' = -\sum_{i=1}^{S} (p_i)(\log_e p_i).$$

H' = – [(0.009)($\log_e$0.009) + (0.253)($\log_e$0.253) + (p_i)($\log_e p_i$) + (0.008)($\log_e$0.008)]
(longnose gar) (gizzard shad) (23 other species) (slenderhead darter)
= – [(–0.042) + (–0.348) + (–2.002) + (–0.039)]
= 2.431 nats/individual.

Units of expression for H' are "nats per individual" (if calculated using $\log_e$, from information theory; Pielou 1975), but it is usually reported as a unitless index value. Although use of $\log_e$ (as above) has become convention for calculating H', any logarithm base may be applied (e.g., $\log_2$ or $\log_{10}$), as resulting index values are easily converted (see Krebs 1998 for multipliers).

Evenness Based on Shannon's Index (J')
From equation (15.5), with H' = 2.431 (above) and 26 species (s),

$$J' = \frac{H'}{H'_{\max}} = \frac{H'}{\log_e s}.$$

$$J' = \frac{2.431}{\log_e 26} = \frac{2.431}{3.258} = 0.746.$$

All measures of evenness range from 0 to 1.0 and are unitless proportions.

Simpson's Diversity Index (1 – D)
From equation (15.4),

$$D = \sum_{i=1}^{S} (p_i^2).$$

D = 0.009^2 + 0.253^2 + p_i^2 + 0.008^2
(longnose gar) (gizzard shad) (23 other species) (slenderhead darter)
= 0.00008 + 0.06401 + 0.07197 + 0.00006
= 0.13612, and
$1 - D$ = 0.86388.

This result ($1 - D$) is the probability, without units, that two individuals selected randomly from this sample will be different species.

(Box continues)

Box 15.3 *(continued)*

Evenness Based on Simpson's Index (V′)
From equation (15.6), with $1 - D = 0.8638$ (above) and 26 species,

$$V' = \frac{1 - D}{(1 - D)_{\max}} = \frac{1 - D}{1 - 1/s}, \text{ and}$$

$$V' = \frac{1 - 0.1361}{1 - 1/26} = \frac{0.8639}{0.9615} = 0.8985.$$

Dominance
Based on the three most numerous species, from equation (15.7), dominance (D) is given by

$$D_3 = \sum_{i=1}^{3} p_i.$$

D_3 = 0.253 (gizzard shad) + 0.221 (smallmouth bass) + 0.069 (striped shiner)
= 0.543.

Maximum value for species dominance would be 1.0. This result is a proportion without units but may also be appropriately expressed (multiplied by 100) as a percentage.

For Shannon's diversity index (H'), the corresponding index of evenness (J') is calculated as

$$J' = \frac{H'}{H'_{\max}} = \frac{H'}{\log_e s}, \tag{15.5}$$

where $H'_{\max} = \log_e s$ = maximum possible value of Shannon's index, and s = number of species.

There is disagreement on a theoretical upper limit for Shannon's index, but in practice, it rarely exceeds 5.0 for biological assemblages (Washington 1984).

The analogous equation to calculate evenness (V') for Simpson's diversity index ($1 - D$) is

$$V' = \frac{1 - D}{(1 - D)_{\max}} = \frac{1 - D}{1 - 1/s}, \tag{15.6}$$

where $(1 - D)_{\max} = 1 - 1/s$ = maximum possible value of Simpson's index, and s = number of species.

The maximum value that Simpson's index may attain is nearly 1.0. Because many variants of Shannon's and Simpson's diversity indices have been proposed, there are an equal number of corresponding algorithms to estimate evenness associated with those measures of diversity (Pielou 1975; Washington 1984; Krebs 1998). For this reason, we suggest consistently reporting the explicit equation used to estimate evenness, as well as that for diversity. Example calculations of species evenness based on Shannon's and Simpson's diversity indices for the Kankakee River fish assemblage are presented in Box 15.3.

Another simple assemblage structural index related to evenness is species dominance, which may be expressed as the relative abundance of a subset of the most numerous species. For example, the proportion of the assemblage composed of the two or three most abundant species would, in general, be inversely related to evenness. The equation to calculate species dominance for the three most abundant species (D_3) is simply

$$D_3 = \sum_{i=1}^{3} p_i , \tag{15.7}$$

where p_i = proportion of the total sample represented by the ith species. Species dominance may be estimated for a variable number of dominant species (usually two to three). An example calculation is presented for the Kankakee River fish assemblage in Box 15.3.

15.3.2 *Biotic Integrity Indices*

The concept of using indicator organisms as descriptors of environmental quality may date back centuries, but it was not until the early twentieth century that it became formalized. The "Saprobiensystem," developed by Kolkwitz and Marsson (1908) in Europe, delineated zones of organic enrichment and classified animal species that occupy them. That early biotic index was later applied to river systems and modified (Chandler 1970), and this led to the prolific development of a variety of biotic indices for aquatic invertebrates that appears to continue without consensus (Washington 1984; Rosenburg and Resh 1993). While indicator species, such as common carp or salmonid species, have been recognized in fisheries science for decades, and other multimetric indices have been proposed (e.g., Gammon's [1976] index of well being), the development and first widespread application of a formal biotic index based on fishes is attributed to James Karr and his colleagues (Karr 1981; Karr et al. 1986).

15.3.2.1. *Rationale*

Biotic integrity of an ecosystem is the ability to support and maintain a balanced, integrated, adaptive community with assemblage characteristics and functional organization similar to a natural habitat in the region that has not been impaired by human activities (Karr et al. 1986). Systems with biotic integrity are more resistant and resilient to natural disturbances and may withstand substantial human influences. Ecological integrity integrates aspects of the chemical and physical

state of the ecosystem with the biological. Whereas aquatic systems with ecological integrity may support productive fisheries or other products and services, ecological integrity is not necessarily correlated with productivity or diversity.

Biotic indices are developed to describe or quantify ecological integrity based on known or suspected relationships between indicator organisms and their environment and may also include assemblage structural indices. Indicator organisms may be selected because they are particularly sensitive or tolerant to environmental degradation, and both types may be incorporated into a single biotic index. Effective biotic indices cannot be universal; as fauna and environmental stresses change regionally, so will suitable indicator organisms. Thus, a biotic index developed for a specific region and environmental stressors may require modification for a different fauna and environmental relationships. Unfortunately, biotic indices are often applied uncritically to systems other than those for which they were developed.

The concept and practice of biotic indices have been widely lauded and criticized on various grounds (Suter 1993; Davis 1995; Simon 1999b). In general, criticisms include a perceived lack of ecological meaning, predictability, diagnostic power, and direct application to water resource regulation. Such criticisms apply to many of the multimetric indices or multivariate techniques covered in this chapter and have been refuted by those successfully applying biotic indices. The differences between opponents and proponents are primarily philosophical and can be overcome by caution and reason in application of techniques and interpretation of results.

15.3.2.2 Indicator Species and Guilds

Fishes are especially well suited as taxa to indicate environmental quality (Karr et al. 1986; Simon 1999b). They occur in all but the most degraded waters; they can accurately reflect environmental conditions at multiple scales; life history and geographic distribution information is extensive for many species; and effective techniques are available to collect them. Finally, fishes are relatively more visible, understood, and valued by regulators, politicians, and the general public than are other aquatic organisms.

The indicator fish approach is simple and easily applied without intensive data needs or analysis. Indicator fishes or guilds may be particularly sensitive or tolerant to environmental degradation. The application is more biologically relevant when indicator guilds are used because the effect of their occurrence may imply ecological function, such as feeding or reproduction, rather than specific responses of individual species. Examples of fish guilds to be considered are those based on feeding and trophic relations (Gerking 1994), reproduction (Balon 1975), or habitat (Grossman and Freeman 1987; Bain et al. 1988). Furthermore, higher levels of fish taxa (e.g., families or genera, such as Salmonidae or darters of genera *Ammocrypta*, *Etheostoma*, and *Percina*) may be considered indicator taxa.

Disadvantages of the indicator fish approach lie primarily in its subjectivity and ecological basis. Although several lists partitioning fish guilds exist (e.g., Balon 1975; Karr et al. 1986; Halliwell et al. 1999; Simon 1999c), standard criteria for

guild delineation and selection of appropriate guilds are lacking. Another problem is that mechanisms unrelated to ecological integrity may influence occurrence or ecological success of a fish taxon or guild; these may include zoogeography, biotic interactions, or harvest (Fausch et al. 1990). Further complicating the use of indicator fishes or guilds is that responses to environmental conditions in fishes can vary with space, time, and type or degree of environmental stress, which could confound conclusions among ecosystems or years.

15.3.2.3 Index of Biotic Integrity

Since its conception and original development for wadeable, warmwater streams in the Midwestern United States, the index of biotic integrity (IBI, Karr 1981; Karr et al. 1986) has been modified, as intended by the original authors, and applied to virtually all other aquatic ecosystems, including coldwater streams, large rivers, lakes, estuaries, and highly modified habitats, and to various regions of the United States (Simon 1999a). Today, the IBI is widely applied and serves the function of a conceptual and procedural framework for assessing biological integrity based on fish assemblages rather than a prescribed, specific protocol.

The IBI was designed as a composite index to assess biological integrity of aquatic ecosystems by integrating attributes of the fish assemblage, population, and individual by means of relative abundance of species and condition of individuals in a representative sample of the assemblage. Although assemblage structural indices (section 15.3.1) utilize relative abundance data and may reflect ecological conditions in some applications, they were not conceived and designed for that function. The primary advantage of the IBI is that it was specifically developed and refined, based on ecological relationships of fishes, to describe ecological integrity and anthropogenic alterations of aquatic ecosystems.

The original IBI framework included 12 metrics that describe various aspects of fish species composition, trophic composition, abundance, and condition (Table 15.2), but metrics have been omitted, augmented, or modified in applying the IBI to other regions, habitats, and specific ecosystems, usually retaining the original ecological framework (Miller et al. 1988). Increasingly, metrics are developed systematically for a region based on metric variability and empirical relationships (Hughes et al. 1998; Angermeier et al. 2000). A number rating, or score (5, 3, or 1), based on ecological expectations is assigned to each metric, and metric scores are summed to yield a composite index score. The IBI scores may then be compared directly or ranges may be assigned to successive categorical integrity classes from very poor to excellent.

Species composition. The metrics describing species composition were intended to characterize biological integrity through measures of fish species diversity and occurrence of relatively tolerant and intolerant species. Species richness or relative abundance may be modified to reflect that of native fishes in areas affected by nonnative fishes. Occurrence and relative abundance of specific families or taxa may also vary among regions and should include species-rich groups with wide geographic distributions and include one primarily benthic taxon and one nonbenthic taxon (Karr et al. 1986). Species considered intolerant (or tolerant)

Table 15.2 Generalized fish assemblage metrics and scoring criteria for the index of biotic integrity (IBI) applied to streams (modified from Karr et al. 1986). Scores are assigned to each metric based on the sample deviation from that expected from a relatively undisturbed reference system.

	Scoring criteria		
	5	3	1
Attribute category and metric	(highest integrity)		(lowest integrity)
Species richness and composition			
Total number of fish species			
Number and identity of darter species	Expectations vary with stream size, region, and basin (see section 15.3.2.4 for discussion)		
Number and identity of sunfish species			
Number and identity of sucker species			
Number and identity of intolerant species			
Percent individuals as green sunfish	<5%	5–20%	>20%
Trophic composition			
Percent individuals as omnivores	<20%	20–45%	>45%
Percent individuals as insectivorous cyprinids	>45%	20–45%	<20%
Percent individuals as piscivores	>5%	1–5%	<1%
Fish abundance and condition			
Number of individuals sampled	Expectations vary with stream size, region, and basin (see section 15.3.2.4 for discussion)		
Percent individuals as hybrids	0	>0–1%	>1%
Percent individuals diseased or with anomalies	0–2%	>2–5%	>5%
Total IBI score (sum of 12 metrics)	60		12
Integrity class	Excellent – Good – Fair – Poor – Very Poor		

should include only 5–10% of the species that are most (or least) sensitive to human alteration of ecosystems.

Trophic composition. Trophic composition metrics are based on the premise that alterations in food resources and productivity, influenced by water and habitat quality, are reflected in the trophic structure of the fish assemblage. This extends the attributes of the IBI to other trophic levels and organisms. As habitat degrades, food resources fluctuate more, and omnivores may replace more specialized feeders. The presence of piscivores, or other top carnivores, indicates a more complex food web. For classification purposes, omnivores are defined as species that consume significant quantities of both plant and animal material, including detritus (Karr et al. 1986). In regions where insectivorous cyprinids are not common, other insectivorous fish taxa or other specialized feeder may be substituted.

Fish abundance and condition. Metrics describing fish abundance and condition were designed to incorporate population and individual level effects of environmental degradation. Obviously, the number of individuals in the sample will be dependent on effort, and thus, fish abundance must be standardized to units of catch per effort for comparison among sites (see section 15.2.4 and Chapter 7). While a metric based on fish numbers accounts for numbers of trophic links, a

metric based on fish biomass may also be incorporated to account for the magnitude of trophic transfer and energy sequestered. The metrics for hybrid individuals and for disease and anomalies are among those most difficult to apply and are frequently omitted or replaced (Miller et al. 1988). One strength of the IBI is that it is a simple field assessment, but if fish need to be preserved for later analysis, that advantage is diminished. Metrics related to nonnative species abundance or reproductive guilds have been substituted for the hybrid metric to represent a similar ecological rationale.

15.3.2.4 *Spatial Influences and Reference Systems*

The discussion above on IBI attribute categories emphasizes the practice of refining or replacing metrics for application to specific regions, and, likewise, metric expectations should be similarly scaled according to stream size and compared with least-disturbed, reference conditions. This practice will reduce the influence of confounding factors that are unrelated to human influences on fish assemblages and will improve the relationship of index scores to ecological integrity. Such considerations apply as well to comparison and interpretation of structural indices (section 15.3.1) and other methods to compare fish assemblages (section 15.4) and are discussed below.

Regional influence. The variation in fish fauna and assemblages among regions may be influenced by broad-scale factors, such as geological phenomena, river basin boundaries, historical biogeography, glaciation, and evolution (Matthews 1998). Thus, comparative use of the IBI and, in most applications, fish assemblage structural indices or multivariate techniques, should occur within a geographic region. Typical spatial frameworks for delineating regions in this context are by ecoregion, drainage basin, or other related hierarchical division (Bailey 1995; Omernick 1995; Omernick and Bailey 1997; Angermeier et al. 2000).

Stream size and longitudinal influence. Numerous studies on the distribution of stream fishes from headwater reaches downstream to large rivers suggest common patterns of change in fish assemblages along a longitudinal gradient. In most systems, species richness and diversity increase with stream size (Horwitz 1978; Vannote et al. 1980; Wiley et al. 1990), and other assemblage and population attributes, such as fish density, biomass, growth, body size, and trophic dynamics, change longitudinally (Matthews 1998). Thus, stream size or longitudinal position must be taken into account when comparing fish assemblage data among sites within and among drainage networks.

Typically, stream longitudinal position, which is generally related to channel size, is defined by stream order (Horton 1945; Strahler 1957) or watershed drainage area (in acres or hectares). Because of inconsistencies in definition and reduced precision associated with the discrete scale (consecutive integers) of stream order, drainage area is generally a more useful descriptor of stream size and position.

Fausch et al. (1984) presented a simple graphical technique to demonstrate the effect of stream longitudinal position on fish species richness and to approximate expected criteria values for IBI applications. When species richness is plotted against stream order or watershed area ($\log_{10}$ transformed), the distribution

of sites forms a right triangle, where the hypotenuse forms a positive-sloped line of maximum species richness for a river system or region. The line of maximum species richness is used in IBI practice to define "excellent" species richness (metric score = 5), which varies with stream size; similarly, sites with richness falling below the maximum expected may be rated depending on the degree of deviation below the line for a given stream order or watershed area. The line of maximum expected values can be quantitatively derived by calculating the 95th-percentile regression (Blackburn et al. 1992) rather than by visually fitting a line to perceived maximum values. This technique can be applied to other IBI metrics associated with species composition (Table 15.2) and may be less relevant to metrics associated with trophic composition or fish abundance and condition, which appear to vary less with stream position or size (Karr et al. 1986).

Examination or adjustment of other fish assemblage structural indices, in addition to IBI metrics, should be considered in most site comparisons within and among drainage networks. In the absence of data to develop such relationships for a river system or a region, site comparisons relevant to environmental quality should be conducted among stream sites of similar longitudinal position or size. Furthermore, such spatial and size effects apply to limnetic zones of lakes and reservoirs and should be considered in analogous lentic comparisons among sites.

Reference systems. Reference systems to represent undisturbed or least-disturbed ecological conditions are a critical component of any biotic or ecological assessment (National Research Council 1992; Hughes 1995; Karr and Chu 1999). Such systems are critical as a benchmark for comparison to detect and understand effects of human activities on ecosystems and to serve as a goal for ecological restoration. Therefore, selection of a reference system or definition of reference conditions is of utmost importance in biotic assessment, but no clear criteria exist for such decisions.

Variation over space and time is key to identifying reference systems or conditions. A system or location within a region with highest IBI metric scores or other appropriate biotic or physical criteria (e.g., watershed land use or riparian disturbance) may serve as reference conditions. Similarly, information from the past, recent or historical, may provide qualitative or quantitative descriptions of predisturbance conditions. Ichthyological references, graduate theses, and state and federal agency reports can be valuable sources of historical data on fish assemblages for specific regions.

Searches for information on predisturbance conditions may be difficult but perhaps the only option in regions exposed to large-scale degradation. For example, only 42 high-quality, free-flowing rivers remain in the conterminous United States, and only 2% of the rivers in those states have features sufficient to receive federal protection (Benke 1990). Further, Hynes (1970) purported that it would be extremely difficult to find any stream that has not been altered by humans and impossible to find any such river—and that assessment was made more than 30 years ago. Typically, a system or site of least disturbance must be substituted for an undisturbed reference or reference conditions from another region may be

cautiously applied. Hughes (1995) examined common, alternative, and combined approaches for determining regional reference conditions.

15.3.2.5. Biotic Integrity Indices in Practice

The IBI was an important advance in biotic assessment methods that has evolved to a concept rather than a restrictive protocol. Fisheries and aquatic scientists are free to apply their own experience, perspective, and creativity into assessing ecological integrity based on fish and invertebrate assemblages, as well as physical attributes of aquatic habitats, riparian zones, and landscapes. Development of an IBI is a rather intensive endeavor that requires sampling a substantial number of sites and careful deliberation regarding metric and criteria development and refinement. However, there is no reason that investigators undertaking more limited assessments should not use individual IBI metrics or related assemblage structural indices (e.g., richness, diversity, or evenness) singly or in aggregate as quantitative assemblage characteristics for comparison and assessment.

As with other techniques in this chapter, several restrictions apply. For any biotic index to be meaningful, it must be based on a thorough and representative sample of the fish assemblage. Thus, sampling considerations (section 15.2) are an important aspect of any assessment program. It is essential that samples reflect the fish assemblage resulting from the physical and biotic environment rather than from a sampling bias that may vary among sites. As with most indices presented in this chapter, biotic indices are relative rather than absolute values, and their application should reflect that limitation. Finally, a biotic index must have a demonstrable empirical relationship to environmental quality to be meaningful, and such steps should be incorporated into biotic index development.

15.4 METHODS TO COMPARE COMMUNITIES

Community indices are useful for summarizing and describing the structure of a fish assemblage. However, they cannot be used to compare the composition and relative abundance of species in two or more assemblages directly. For example, species richness does not take species identity into account, and the IBI does not provide an explicit means to determine how two or more assemblages differ. Several approaches have been developed by ecologists for directly comparing two or more communities, and most have been used to study fish assemblages. These techniques can be roughly categorized as (1) resemblance measures for quantifying the similarity among assemblages, (2) classification methods for grouping assemblages based on their structure, (3) ordination methods for examining the relationships among assemblages, (4) categorical data analysis methods for estimating the differences among assemblages, and (5) graphical techniques for displaying the relationships among assemblages. Each of the techniques discussed below has associated assumptions and limitations that can affect the validity of community comparisons. Consequently, ecologists often employ two or more techniques to maximize insight and validate patterns indicated by a single method

(Green and Vascotto 1978; Gauch 1982; Romesburg 1990). However, fisheries scientists should refrain from data dredging, that is, conducting several analyses until one produces results that appear to make the most sense or are statistically significant. Such an approach to community level analysis is fraught with problems and should be avoided (see Rexstad et al. 1988). Rather, fisheries scientists should carefully consider the objectives of their study and the system being examined and develop a set of questions to be addressed through their analyses.

15.4.1 Measures of Community Similarity and Their Characteristics

Fisheries biologists are often interested in quantifying the similarity among fish communities based on their species composition and abundance. Resemblance coefficients are used to measure the similarity (or dissimilarity) of two or more communities with one or more characteristics, such as species presence, abundance, density, or other community functional parameter (Gauch 1982; Romesburg 1990). Thus, resemblance coefficients are a useful and relatively flexible means for quantifying the similarity among fish assemblages. Similarity and dissimilarity, however, are generally descriptive measures rather than statistical estimates (correlation coefficients are an exception), and, as such, it is difficult to estimate the statistical significance of relationships. Such associated significance tests usually require the use of computer-intensive resampling techniques and specialized software and are of limited scientific value (e.g., Van Sickel 1997; Johnson 1999); hence, we do not recommend their general use.

There are several types of resemblance coefficients, and the use of each depends upon the characteristics of the data and study objectives. Binary coefficients are used to measure the similarity between two assemblages using only species' presence and absence. Ordinal coefficients are used when species (relative) abundances have been transformed into ranks presumably to minimize the influence of sampling bias. Quantitative coefficients require an estimate of species-specific abundance, such as density, relative abundance, and counts of the number of individuals. Although resemblance measures can differ markedly in their data requirements and calculation, the best and most useful share two desirable characteristics. First, resemblance measures should be independent of the number of individuals in a sample and the number of species in an assemblage. Second, they should increase regularly from a set minimum to a set maximum as two fish assemblages become increasingly similar (Wolda 1981). Another characteristic to consider is the sensitivity of a particular resemblance measure to size (abundance) displacements. Measures that are insensitive to size displacements consider two assemblages to be similar if their attributes (e.g., density of a species) differ by an additive or multiplicative function (Figure 15.4). Consequently, different resemblance measures can provide markedly different estimates of the similarity (or dissimilarity) between two assemblages. Fisheries scientists should be aware of these characteristics and use the resemblance measure that best meets their needs.

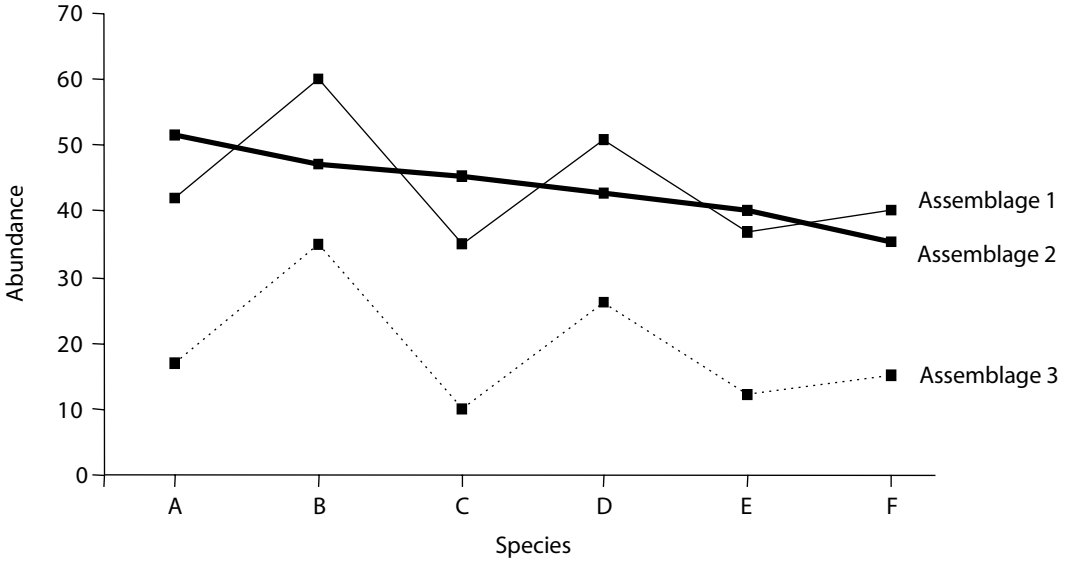


Figure 15.4 Species-specific abundances for three hypothetical fish assemblages. A resemblance measure that is insensitive to size displacement would score assemblages 1 and 3 as most similar because the species abundances differ by an additive constant (25), whereas a measure that is sensitive to size displacements would score assemblages 1 and 2 as most similar.

15.4.1.1 Coefficients for Binary (Nominal) Data

Species' presence and absence require a similarity coefficient that can be used with binary data (i.e., data that consist of two states). Binary coefficients are the simplest, but most imprecise, estimators of community similarity because they consider only species' presence or absence. Rare species and abundant species are weighted equally. Therefore, binary coefficients should be used only when species' presence are the only data available or in situations in which sampling conditions prevented the estimation of species relative abundances.

The best and most commonly used binary coefficients are the Jaccard's and simple matching coefficients, which vary from 0 to 1, with 0 indicating no species in common and 1 indicating identical species composition. Both measures are independent of the number of individuals in an assemblage (sample) and are insensitive to size displacements.

Jaccard's coefficient. The similarity, C, between a pair of assemblages j and k is calculated as

$$C_{jk} = \frac{p}{p + m}, \tag{15.8}$$

where p is the number of species that are present in both assemblages and m is the number of species present in one assemblage but not the other. An example calculation of Jaccard's coefficient is presented in Box 15.4.

Box 15.4 Calculation of Jaccard's and Simple Matching Coefficients

To calculate both Jaccard's and simple matching coefficients, first determine the number of species present and absent for both stations and the number of species occurring at one station but not another. Using the summary data for stations 1 and 2 on the Kankakee River, Illinois (Box 15.1),

number of species present at both stations is $p = 18$,
number of species absent at both stations is $a = 4$, and
number of species present at one station but not the other is $m = 12$.

Jaccard's Coefficient
From equation (15.8),

$$C_{jk} = \frac{p}{p+m}, \text{ and}$$

$$C_{1,2} = \frac{18}{18+12} = 0.60.$$

Simple Matching Coefficient
From equation (15.9),

$$C_{jk} = \frac{p+a}{p+m+a}, \text{ and}$$

$$C_{1,2} = \frac{18+4}{18+12+4} = 0.65.$$

Program
The following SAS program uses the DISTANCE macro, included in SAS/STAT software (version 6.09 or later; SAS Institute 2004), to compute the Jaccard and simple matching similarity coefficients for fish assemblages at all stations of the Kankakee River example data. Note that the path following the %INC must be changed to that of the SAS folder containing XMACRO, STDIZE, and DISTNEW macros on your computer or network. These macros are included with all versions of SAS/STAT software.

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT STATION $ LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS
SHR QLL RVR SVR SAB BKS BLG GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
LINES;
STATION1 1 1 1 0 1 0 1 0 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 1 0 1 1
STATION2 1 1 1 0 1 0 1 1 1 1 1 1 0 0 0 1 0 1 0 1 1 1 0 1 1 1 0 1 1 0 1 1 0 0
```

```
STATION3 0 1 1 0 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 0 1 1 1 0 1 1 0 1 1 1 1
STATION4 1 1 1 1 1 1 0 0 0 1 1 1 1 1 0 0 1 1 1 1 1 0 1 0 1 1 1 0 1 1 1 1 1 0 0 1 1
STATION5 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
STATION6 1 1 1 0 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 0 1 0 1 1 1 1 1 1 0 0 1 1
;
%INC '<location of SAS sample folder>/XMACRO.SAS';
%INC '<location of SAS sample folder>/STDIZE.SAS';
%INC '<location of SAS sample folder>/DISTNEW.SAS';
%DISTANCE(DATA=SPECIES, OPTIONS=NOMISS, SHAPE = SQUARE, ID = STATION,
OUT=JACCARD, METHOD = JACCARD);
PROC PRINT;
%DISTANCE(DATA=SPECIES, OPTIONS=NOMISS, SHAPE = SQUARE, ID = STATION, OUT=MATCH,
METHOD = MATCH);
PROC PRINT;
RUN;
```

Program Output

Table Jaccard's coefficient and simple matching coefficient similarity matrices for fish assemblages at all stations of the Kankakee River example data (Box 15.1).

	Station					
Station	1	2	3	4	5	6
			Jaccard Similarity Matrix			
1	1.0000	0.6000	0.7333	0.8519	0.7273	0.8214
2	0.6000	1.0000	0.7143	0.5333	0.6563	0.5667
3	0.7333	0.7143	1.0000	0.6129	0.7273	0.7000
4	0.8519	0.5333	0.6129	1.0000	0.7188	0.8148
5	0.7273	0.6563	0.7273	0.7188	1.0000	0.6471
6	0.8214	0.5667	0.7000	0.8148	0.6471	1.0000
			Simple Matching Similarity Matrix			
1	1.0000	0.6471	0.7647	0.8824	0.7353	0.8529
2	0.6471	1.0000	0.7647	0.5882	0.6765	0.6176
3	0.7647	0.7647	1.0000	0.6471	0.7353	0.7353
4	0.8824	0.5882	0.6471	1.0000	0.7353	0.8529
5	0.7353	0.6765	0.7353	0.7353	1.0000	0.6471
6	0.8529	0.6176	0.7353	0.8529	0.6471	1.0000

The Jaccard coefficient is sensitive to the direction of the coding (i.e., asymmetric); hence, presence or absence should be coded the same for all species. This measure considers only mutual presence of a species, which should minimize the influence of false absences (i.e., a species is missed) in sampling data. Therefore, Jaccard's is the preferred method for analyzing fish assemblage similarity based on species presence.

Simple matching coefficient. The similarity, C, between j and k is calculated as

$$C_{jk} = \frac{p + a}{p + m + a}, \tag{15.9}$$

where p and m are defined above and a is the number of species absent in both assemblages. An example calculation of the simple matching coefficient is presented in Box 15.4.

In contrast to Jaccard's, the simple matching coefficient uses the mutual absence of a species and should be used only when there is no potential for false absences (i.e., missed species) in the data. Use of the simple matching coefficient requires an exact definition of the species pool, which is subjective to some degree, and large species pools may inflate the number of mutual absences and, thus, similarity.

15.4.1.2 Coefficients for Ranked (Ordinal) Data

The most widely used coefficients for ranked data are also nonparametric correlation coefficients. These resemblance measures are estimated using ranks in place of actual abundance estimates, and as such, they are unaffected by nonlinear relationships between species abundances of two assemblages. Ordinal measures are not as crude as binary coefficients but are less sensitive than are quantitative coefficients (section 15.4.1.3). Hence, we recommend their use over quantitative measures only when species abundance estimates are believed to be poor due to factors such as sampling difficulties.

The most commonly used ordinal coefficients are Spearman's rank correlation and Kendall's tau (Romesburg 1990). In contrast to most resemblance measures, these coefficients have values that vary from –1 to 1, with –1 indicating different species assemblages and 1 indicating identical species composition. Both measures are strongly affected by the total number of individuals in a sample, especially when there are large numbers of species, and are insensitive to size displacements (Krebs 1998). They are most useful for comparing assemblages in low-diversity communities but should never be used as similarity measures when more than half of the abundances in one or more assemblage samples are zero (Field 1970).

Spearman's rank correlation. To estimate Spearman's rank correlation coefficient, fish abundance data are first sorted and ranked (1 = lowest) for each station. Ties (i.e., species with equal abundances) are assigned the average of the ranks. For example, two species with equal abundances in the fifth and sixth positions in the

sorted list would receive the rank 5.5. The similarity, θ, between assemblages *r* and *s* is then calculated using the Pearson product-moment correlation as

$$\theta_{rs} = \frac{\Sigma (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\Sigma (r_i - \bar{r})^2 \Sigma (s_i - \bar{s})^2}}, \tag{15.10}$$

where r_i and s_i are the ranks corresponding to species *i*, and $\bar{r}$ and $\bar{s}$ are the mean ranks of all species in assemblage *r* and *s*, respectively. An alternative formula for estimating the Spearman's rank correlation coefficient can be found in Daniel (1990). An example calculation of Spearman's rank correlation coefficient is presented in Box 15.5.

Kendall's tau. To estimate Kendall's tau coefficient for assemblages *j* and *k*, species abundances are sorted from smallest to largest for assemblage *j*. If ties are present in assemblage *j* (i.e., two or more species have the same abundance), species abundances for those species within assemblage *k* are sorted in increasing magnitude within each tied group. The similarity, τ, is then calculated as

$$\tau_{jk} = \frac{\Sigma P_i - Q_i}{n(n-1)(0.5)}, \tag{15.11}$$

where P_i and Q_i are the total number of species in assemblage *k* with higher ranks and abundances greater (P_i) and less than (Q_i) species *i* in assemblage *j*, and *n* is the total number of species.

If ties are present in assemblage *j*, corresponding species in assemblage *k* are not counted for P_i and Q_i within each tied group. The denominator in equation (15.11) is also adjusted for ties as

$$\tau_{jk} = \frac{\Sigma P_i - Q_i}{\sqrt{n(n-1)(0.5) - T_j}\sqrt{n(n-1)(0.5) - T_k}}, \tag{15.12}$$

where $T_j = (0.5)\Sigma t_j(t_j - 1)$; $T_k = (0.5)\Sigma t_k(t_k - 1)$; t_j and t_k are the number of species that are tied at a given rank in assemblage *j* and *k*, respectively; and *n* is the total number of species. An example calculation of Kendall's tau coefficient with ties is presented in Box 15.5.

15.4.1.3 *Coefficients for Quantitative Data*

The best means to characterize the similarity among fish assemblages is through the use of quantitative resemblance coefficients. In contrast to binary and ordinal measures, these coefficients use species abundance estimates and are, therefore, much more sensitive to small differences between two fish assemblages. There is a wide variety of similarity measures available for use with quantitative data. However, the two best and most widely used measures in ecology are the percent similarity

Box 15.5 Calculation of Spearman's Rank Correlation and Kendall's Tau Similarity Coefficients

Species total abundances from stations 1 and 2 of the Kankakee River, Illinois (Box 15.1), are used to illustrate the calculation of Spearman's rank and Kendall's tau coefficients.

Table Species sorted by abundance and ranked for stations 1 and 2 of Kankakee River example (Box 15.1). Also included are computations for Kendall's tau that tally the number of species at station 2 with abundances higher (P_i) or lower (Q_i) than a given species at station 1.

Species and summary statistics	Sorted abundance		Rank abundance		Number of species at station 2 with abundances	
	Station 1	Station 2	Station 1	Station 2	Higher (P_i)	Lower (Q_i)
Hornyhead chub	0	0	4.5	6.5	18	0
Suckermouth minnow	0	0	4.5	6.5	18	0
Bullhead minnow	0	0	4.5	6.5	18	0
Black redhorse	0	0	4.5	6.5	18	0
Smallmouth buffalo	0	3	4.5	17	14	11
Largemouth bass	0	3	4.5	17	14	11
Johnny darter	0	6	4.5	19	14	12
Redfin shiner	0	22	4.5	24.5	9	16
Orangespotted sunfish	1	0	10.5	6.5	15	0
Blackside darter	1	9	10.5	21	10	12
Sand shiner	1	22	10.5	24.5	8	14
Green sunfish	1	42	10.5	29	5	17
River redhorse	2	0	13.5	6.5	15	0
Bluegill	2	0	13.5	6.5	15	0
Banded darter	4	0	15.5	6.5	14	0
Brook silverside	4	10	15.5	22	9	9
Slenderhead darter	5	0	18	6.5	12	0
Quillback	5	2	18	14.5	11	4
Northern hog sucker	5	2	18	14.5	11	4
Longnose gar	6	7	20	20	9	5
Silver redhorse	8	1	21.5	13	9	3
Rosyface shiner	8	89	21.5	32	2	10
Mimic shiner	10	11	23	23	7	4
Common carp	13	58	24	30	3	7
Spotfin shiner	19	24	25	26	5	4
Logperch	25	0	26	6.5	6	0
Rock bass	30	3	27	17	5	2
Golden redhorse	34	0	28	6.5	5	0
Shorthead redhorse	35	0	29.5	6.5	4	0
Longear sunfish	35	94	29.5	34	0	4
Bluntnose minnow	42	33	31	28	2	1
Striped shiner	45	32	32	27	2	0
Smallmouth bass	143	59	33	31	1	0
Gizzard shad	164	90	34	33	0	0
Average rank	17.5	17.5				
Sum			308	150		

Spearman's Rank Correlation Coefficient

To estimate Spearman's rank, species abundances are sorted in ascending order and assigned ranks (1 = lowest) for each assemblage. Ties are assigned the average of the tied ranks. The first eight species at station 1 have abundances of 0, and their average rank is calculated as

$$\frac{1+2+3+\ldots+8}{8} = \frac{36}{8} = 4.5.$$

From equation (15.10),

$$\theta_{rs} = \frac{\sum (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\sum (r_i - \bar{r})^2 \sum (s_i - \bar{s})^2}}, \text{ and}$$

$$\theta_{1,2} = \frac{(4.5 - 17.5)(6.5 - 17.5) + (4.5 - 17.5)(6.5 - 17.5) + \ldots + (34 - 17.5)(33 - 17.5)}{\sqrt{[(4.5 - 17.5)^2 + (4.5 - 17.5)^2 + \ldots + (34 - 17.5)^2][(6.5 - 17.5)^2 + (6.5 - 17.5)^2 + \ldots + (33 - 17.5)^2]}}$$

$$= \frac{1308.250}{\sqrt{(3221.500)(3126.500)}}$$

$$= 0.412.$$

Kendall's Tau

To estimate Kendall's tau, species abundances are sorted for station 1 from lowest to highest, and because ties occur at station 1, species abundances for station 2 are also sorted within each tied group. For example, the first eight species at station 1 of the Kankakee River, hornyhead chub through redfin shiner, have abundances of 0; hence, they are also sorted by their station 2 abundances as shown in the table above.

From equation (15.12),

$$\tau_{jk} = \frac{\sum P_i - Q_i}{\sqrt{n(n-1)(0.5) - T_j}\ \sqrt{n(n-1)(0.5) - T_k}}.$$

Using the sorted abundances for station 2, P_i is estimated for each species (i) by counting the number of other species at station 2 (see table) with greater abundance. For hornyhead chub, this is any species with abundance greater than 0. However, four of these species, smallmouth buffalo through redfin shiner, have abundances greater than 0 but are not counted because their abundances are tied (all 0) at station 1. This leaves a total of 18 species with greater abundance than hornyhead chub at station 2. Then Q_i is estimated similarly by counting the number of species with lower abundance, which for hornyhead chub is none (0). The adjustment for ties in the denominator of equation (15.12), T_j and T_k, is estimated for each station by first counting the number of

(Box continues)

Box 15.5 *(continued)*

species for each tied abundance value. At station 1, eight species had 0 abundance, 4 had abundances of 1, 2 had abundances of 2, and so on. The adjustments are then estimated as

$$T_k = (0.5)\sum t_k(t_k - 1)$$

$T_1 = (0.5)[(8)(8-1) + (4)(4-1) + (2)(2-1) + (2)(2-1) + (3)(3-1) + (2)(2-1) + (2)(2-1)]$
$= 41$, and
$T_2 = (0.5)[(12)(12-1) + (2)(2-1) + (3)(3-1) + (2)(2-1)]$
$= 71$.

Thus, replacing the symbols in equation (15.12) with their corresponding values

$$\tau_{1,2} = \frac{308 - 150}{\sqrt{34(34-1)(0.5) - 41}\sqrt{34(34-1)(0.5) - 71}}$$

$$= 0.313.$$

Program

The following SAS program computes Spearman's rank and Kendall's tau similarity coefficients for fish assemblages at stations on the Kankakee River.

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT SPECIES $ STATION1 STATION2 STATION3 STATION4 STATION5 STATION6;
LINES;
Longnose_gar 6 7 0 4 26 5
(input remaining 33 species)
;
```

and Morisita's indices. Percent similarity is used on relative species abundance data, and Morisita's index is employed when data consist of the number of individuals (whole numbers) per species.

Percent similarity. To calculate the percent similarity index, species abundances in each assemblage must first be standardized to percentages by dividing the abundance of each species in a sample by the total number of fish in the sample and multiplying by 100. The similarity, P, between assemblages j and k is calculated as

$$P_{jk} = \Sigma \text{minimum}(p_{ki}, p_{ji}), \tag{15.13}$$

where p_{ji} and p_{ki} are the relative abundances of species i in assemblage j and k, respectively, and minimum indicates that the smallest of the two relative abundances is used in the summation. An example calculation of the percent similarity index is presented in Box 15.6.

```
PROC CORR NOPRINT OUTS=SPEARMAN OUTK=KENDALL;
DATA SPEARMAN; SET SPEARMAN; WHERE _TYPE_='CORR'; DROP _TYPE_;
PROC PRINT;
DATA KENDALL; SET KENDALL; WHERE _TYPE_='CORR'; DROP _TYPE_;
PROC PRINT;
RUN;
```

Program Output

Table Spearman's rank and Kendall's tau similarity matrices for fish assemblages at stations 1 and 2 of the Kankakee River data (Box 15.1).

	Station					
Station	1	2	3	4	5	6
			Jaccard Similarity Matrix			
1	1.0000	0.4122	0.7087	0.7318	0.6943	0.8540
2	0.4122	1.0000	0.4532	0.2912	0.4378	0.4087
3	0.7087	0.4532	1.0000	0.6372	0.4394	0.6565
4	0.7318	0.2912	0.6372	1.0000	0.6601	0.7953
5	0.6943	0.4378	0.4394	0.6601	1.0000	0.6989
6	0.8540	0.4087	0.6565	0.7953	0.6989	1.0000
			Kendall's Tau Similarity Matrix			
1	1.0000	0.3130	0.5448	0.5762	0.5470	0.7060
2	0.3130	1.0000	0.3609	0.2374	0.3261	0.3288
3	0.5448	0.3609	1.0000	0.4966	0.3169	0.4898
4	0.5762	0.2374	0.4966	1.0000	0.4890	0.6385
5	0.5470	0.3261	0.3169	0.4890	1.0000	0.5106
6	0.7060	0.3288	0.4898	0.6385	0.5106	1.0000

The percent similarity index, also know as the Renkonen index after its creator (Renkonen 1938), is one of the best quantitative similarity measures (Wolda 1981). It varies from 0 to 100%, with 0 indicating no species in common and 100% indicating identical species composition. It is a very robust measure that is not influenced by the number of individuals in a sample and is insensitive to size displacements.

Morisita's index. The similarity, C, between assemblages j and k is calculated following Morisita (1959) as

$$C_{jk} = \frac{2\sum X_{ij}X_{ik}}{(\lambda_j + \lambda_k)N_jN_k}, \tag{15.14}$$

where

$$\lambda_j = \frac{\sum_i [X_{ij}(X_{ij}-1)]}{N_j(N_j-1)}, \text{ and } \lambda_k = \frac{\sum_i [X_{ik}(X_{ik}-1)]}{N_k(N_k-1)}.$$

The number of individuals of species i is given by X_{ij} and X_{ik}, and N_j and N_k are the total number of individuals in assemblage j and k, respectively. Horn (1966) developed a simplified version of the index in which each λ is calculated without subtracting 1 from the total number of individuals in the assemblage. This modified version of Morisita's index is used only when abundance is expressed as a proportion, such as relative abundance and density. An example calculation of the Morisita's index is given in Box 15.6.

Morisita's index varies from 0 to 1, with 0 indicating no species in common, and 1 indicating identical species composition. Unlike other similarity coefficients, Morisita's index can be interpreted as a probability (Krebs 1998). It is not significantly influenced by the number of individuals in an assemblage sample, unless that total is very small, and is insensitive to size displacements (Wolda 1981).

15.4.1.4 Distance Measures

Distance coefficients are a special kind of quantitative resemblance measure that are used to estimate the dissimilarity between two fish assemblages. In contrast to similarity, low distance (dissimilarity) values indicate that two assemblages are more similar to one another with 0 indicating identical species composition. Interestingly, any of the similarity measures presented in this chapter can be transformed into a dissimilarity measure by multiplying its value by –1 or by subtracting from a constant corresponding to the maximum value (e.g., subtract from 100 the percent similarity index). Distance coefficients are generally used in cluster analysis (section 15.4.3) and require a quantitative measure of species-specific abundance, such as numbers of individuals, relative abundance, and density.

The most commonly used distance measures are the Euclidean and Bray–Curtis coefficients. Both measures are strongly affected by the total number of individuals and number of species in a sample and are the only resemblance measures presented in this chapter that are sensitive to size displacements.

Euclidean distance. The Euclidean distance, d, between assemblages j and k is calculated as

$$d_{jk} = \sqrt{\Sigma (X_{ij} - X_{ik})^2}, \tag{15.15}$$

where X_{ij} and X_{ik} are the abundances of species i in assemblage j and k, respectively. Euclidean distance values are strongly influenced by the number of species in an assemblage. To minimize this effect, researchers often calculate the average Euclidean distance, d', as

$$d'_{jk} = \sqrt{\frac{\Sigma (X_{ij} - X_{ik})^2}{n}}, \tag{15.16}$$

where n is the total number of species. Example calculations of Euclidean distance and average Euclidean distance are presented in Box 15.7. Euclidean distance and average Euclidean distance can both vary from 0 to infinity, with 0

Box 15.6 Calculation of Percent Similarity and Morisita's Indices

Below we calculate percent similarity and Morisita's similarity indices based on summary fish abundance data from stations 1 and 2 on the Kankakee River, Illinois (Box 15.1).

Percent Similarity

To calculate percent similarity, species abundances must be expressed as percentages. Totals of 648 and 622 fish were collected from stations 1 and 2, respectively. Thus, species-specific abundances at each station are divided by their corresponding station totals and multiplied by 100.

From equation (15.13),

$$P_{jk} = \sum \text{minimum } (p_{ki}, p_{ji}), \text{ and}$$

$P_{1,2}$ = minimum (0.926, 1.125) + minimum (25.309,14.469) + . . . + minimum (0.772, 0.000)
(longnose gar) (gizzard shad) (slenderhead darter)
= 0.926 + 14.469 +… + 0.000
= 50.185%.

Morisita's Similarity Index

From equation (15.14),

$$C_{jk} = \frac{2\sum X_{ij}X_{ik}}{(\lambda_j + \lambda_k)N_jN_k}.$$

$$\lambda_1 = \frac{6(6-1) + 164(164-1) + \ldots + 5(5-1)}{648(648-1)}$$
$$= 0.135, \text{ and}$$
$$\lambda_2 = \frac{7(7-1) + 90(90-1) + \ldots + 0(0-1)}{622(622-1)}$$
$$= 0.096.$$
$$C_{1,2} = \frac{2[(6)(7) + (164)(90) + \ldots + (5)(0)]}{(\lambda_j + \lambda_k)(648)(622)}$$
$$= \frac{63{,}236.000}{92{,}875.079}$$
$$= 0.681.$$

indicating identical assemblages and large distances indicating very different assemblage structure.

Bray–Curtis coefficient. The distance, b, between assemblage j and k is calculated following Bray and Curtis (1957) as

$$b_{jk} = \frac{\Sigma |X_{ij} - X_{ik}|}{\Sigma (X_{ij} + X_{ik})}, \tag{15.17}$$

where X_{ij} and X_{ik} are the abundance of species i in assemblage j and k, respectively. An example calculation of the Bray–Curtis coefficient is presented in Box 15.7. The Bray–Curtis coefficient varies from 0 to 1, with 0 indicating identical assemblages and 1 indicating no species in common. It also tends to be strongly influenced by abundant species and should not be used when fish assemblage samples are dominated by a few very abundant species (i.e., evenness is low, Wolda 1981).

Box 15.7 Calculation of Euclidean and Bray–Curtis Distances.

Euclidean and Bray–Curtis distances are calculated for summary fish abundance data from stations 1 and 2 of the Kankakee River, Illinois (Box 15.1).

Euclidean Distance
From equation (15.15),

$$d_{jk} = \sqrt{\sum (X_{ij} - X_{ik})^2}.$$

$$d_{1,2} = \sqrt{(6-7)^2 + (164-90)^2 + \ldots + (5-0)^2}$$
(longnose gar) + (gizzard shad) + . . . + (slenderhead darter)

$$d_{1,2} = \sqrt{31{,}488.00} = 177.449.$$

Average Euclidean Distance
From equation (15.16),

$$d'_{jk} = \sqrt{\frac{\sum (X_{ij} - X_{ik})^2}{n}}.$$

$$d'_{1,2} = \sqrt{\frac{31488.00}{34}} = 30.432.$$

Bray–Curtis coefficient
From equation (15.17),

$$b_{jk} = \frac{\sum |X_{ij} - X_{ik}|}{\sum (X_{ij} + X_{ik})}.$$

$$b_{1,2} = \frac{|6-7| + |164-90| + \ldots + |5-0|}{(6+7) + (164+90) + \ldots + (5-0)}$$

$$= \frac{630}{1{,}270} = 0.496.$$

15.4.2 Classification Techniques

Similarity coefficients are useful for examining relationships among small numbers of fish assemblages. Fisheries scientists, however, often need to compare several assemblages simultaneously for the purposes of grouping or classifying them based on their structure. Cluster analysis includes a set of techniques that can be used to examine the relationships among two or more communities and group

Program

The following SAS program uses the DISTANCE macro, included in SAS/STAT software (SAS Institute 2004), version 6.09 or later, to compute Euclidean distance for fish assemblages at all stations on the Kankakee River. Note that the path following the %INC must be changed to that of the SAS folder containing XMACRO, STDIZE, and DISTNEW macros on your computer or network.

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT SPECIES $ STATION1 STATION2 STATION3 STATION4 STATION5 STATION6;
LINES;
Longnose_gar 6 7 0 4 26 5
(input remaining 33 species)
;
PROC TRANSPOSE DATA = SPECIES OUT = SPECIES;
%INC '<location of SAS sample folder>/XMACRO.SAS';
%INC '<location of SAS sample folder>/STDIZE.SAS';
%INC '<location of SAS sample folder>/DISTNEW.SAS';
%DISTANCE(DATA=SPECIES, OPTIONS=NOMISS, SHAPE = SQUARE, ID = _NAME_,
OUT=EUCLID, METHOD = EUCLID);
PROC PRINT;
RUN;
```

Program Output

Table Euclidean distance matrix for fish assemblages at six stations on the Kankakee River (see Box 15.1).

	Station					
Station	1	2	3	4	5	6
1	0.0000	177.4486	175.2341	171.6945	275.2181	93.2202
2	177.4486	0.0000	218.1811	185.4104	380.0776	225.1000
3	175.2341	218.1811	0.0000	81.1788	435.1597	211.5349
4	171.6945	185.4104	81.1788	0.0000	435.5250	206.7970
5	275.2181	380.0776	435.1597	435.5250	0.0000	258.5672
6	93.2202	225.1000	211.5349	206.7970	258.5672	0.0000

them into classes (or clusters) based on their similarity (Romesburg 1990). Classification, however, is more of a skill than an exact science, and it requires a certain amount of ecological insight and knowledge of the systems being studied. There is no single best classification system for grouping fish assemblages or determining the exact number of groups to distinguish. Thus, a fish assemblage classification system developed for one purpose may be inappropriate for addressing another (Romesburg 1990). As with all analytical techniques, we encourage fisheries scientists to consider their objectives and planned application of their classification system carefully before attempting to develop fish assemblage classifications.

There are two basic types of cluster analyses—hierarchical and nonhierarchical—that have been used by biologists to develop polythetic classifications (i.e., classifications based on overall similarity). Among these, hierarchical methods are generally the simplest, most easy to use, and the only clustering methods considered here that are useful for examining relationships among assemblages. However, they can become cumbersome when the number of samples (i.e., assemblages) is large. Nonhierarchical methods are computationally complex and require the use of computer programs but are appropriate and useful when the number of samples is large. Below, we discuss the best, most widely used of these two clustering methods for ecological classification. For a more thorough treatment of cluster analysis and ecological classification, we recommend Romesburg (1990) and Everitt (1993).

15.4.2.1 Hierarchical Cluster Analysis

By far, the most widely used form of cluster analysis is hierarchical clustering. It is used to reveal relationships among assemblages based on resemblance measures (section 15.4.1). Hierarchical cluster analysis begins with a matrix of resemblance coefficients (see Box 15.6 for example). Pairs of assemblages are then grouped sequentially using a clustering method. Clustering begins by grouping the most similar (or least dissimilar) pair(s) of assemblages. The next most similar pair(s) is then clustered, and the process continues until all assemblages are contained in a single cluster. Results are displayed in a diagram, called a dendrogram or tree, that shows the similarities in the form of a hierarchy (hence, the name). Relationships indicated by hierarchical cluster analysis are significantly influenced by the characteristics of the resemblance measure and the clustering method. In fact, different combinations of resemblance measures and clustering methods can provide quite disparate estimates of relationships among assemblages (Figure 15.5). Because there is no truly objective means of clustering (Romesburg 1990; Krebs 1998), fisheries biologists should consider the characteristics of resemblance measures and clustering methods and choose those that make the most sense. Having previously detailed the characteristics of resemblance measures, we now describe those of clustering methods.

Single linkage. Single-linkage clustering is the simplest form of hierarchical cluster analysis. It begins with a resemblance matrix and uses the nearest-neighbor rule to define similarity or dissimilarity among clusters. For distance measures, assemblages are clustered as follows.

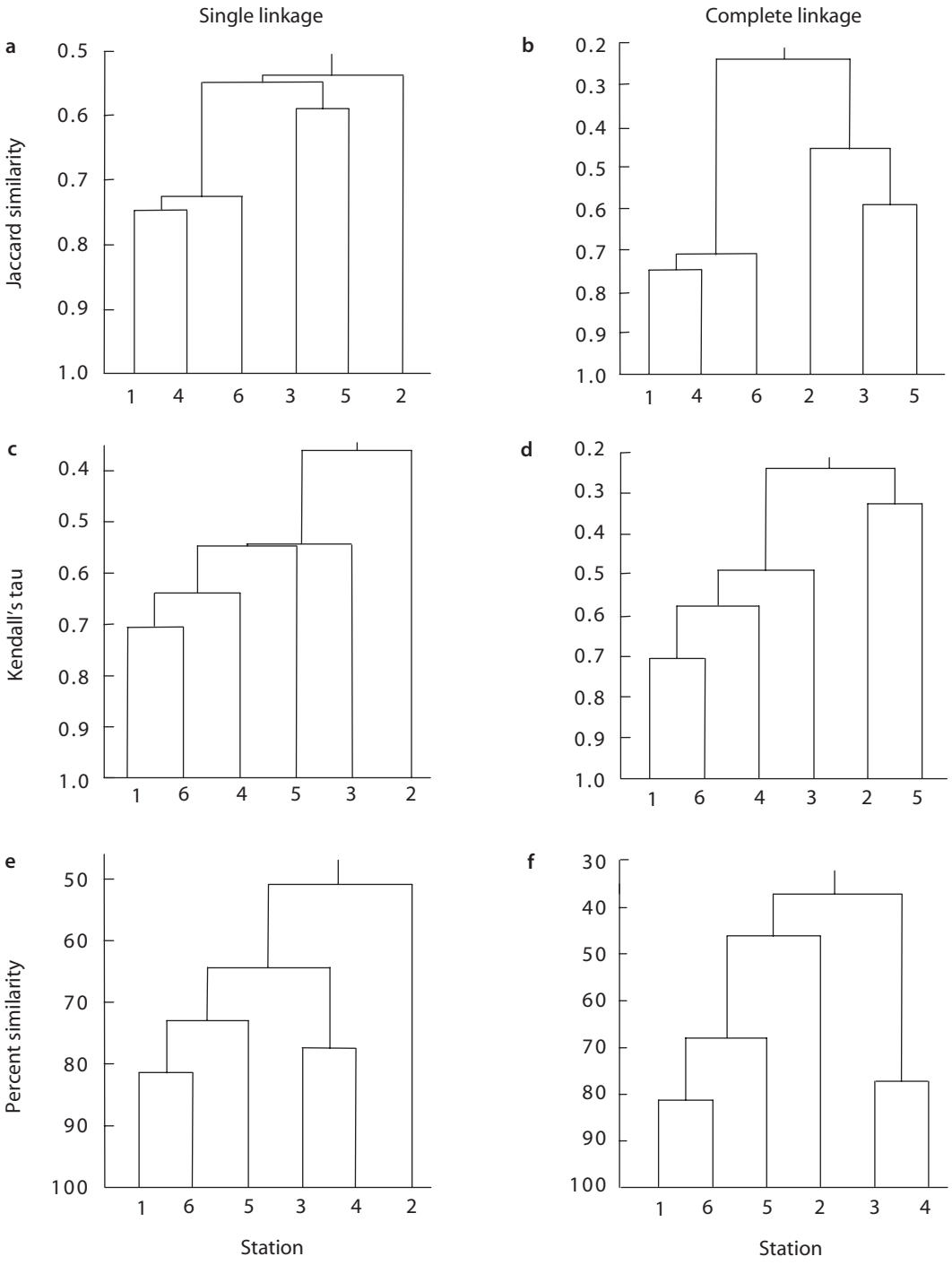


Figure 15.5 Cluster analyses of fish assemblages at six stations on the Kankakee River, Illinois (Box 15.1). Trees (a) and (b) were created using Jaccard's similarity coefficient, (c) and (d) with Kendall's tau, and (e) and (f) with the percent similarity index. Single-linkage clustering method was used for trees (a), (c), and (e) and complete-linkage for trees (b), (d), and (f).

Step 1. Find the pair of assemblages with the smallest distance (most similar) and combine into the first cluster.

Step 2. Find the second pair with the smallest distance and cluster. This pair can include two assemblages or an assemblage and the cluster created in step 1 whose nearest-neighbor distance, d, from an assemblage j is defined as

$$d_{j(m,n)} = \text{minimum } (d_{jm}, d_{jn}), \tag{15.18}$$

where (m, n) is a cluster containing assemblages m and n, and minimum indicates that the smallest distance is used in clustering.

Step 3. Find the third pair with the smallest distance and group (cluster). This pair can include two assemblages, an assemblage and a cluster, or two clusters whose nearest-neighbor distance, d, is defined as

$$d_{(j,k)(m,n)} = \text{minimum } (d_{jm}, d_{jn}, d_{km}, d_{kn}), \tag{15.19}$$

where (j, k) is the cluster containing assemblage j and k and (m, n) is defined above.

Step 3 is repeated until all of the assemblages are contained in one cluster. Single-linkage clustering is illustrated in Box 15.8.

For similarity measures, the nearest-neighbor distance is the highest similarity value, and hence, the minimum in equations (15.18) and (15.19) is replaced by a maximum.

Complete linkage. Complete-linkage clustering is also a relatively simple form of hierarchical clustering. It is very similar to single linkage except that the rules for defining distances among clusters are the exact opposite. Thus, the first step is identical to that described above for single linkage. Subsequent steps proceed as follows.

Step 2. Find the second pair with the smallest distance (most similar) and cluster. This pair can include two assemblages or an assemblage and the cluster created in step 1 whose farthest-neighbor distance, d, from an assemblage j is defined as

$$d_{j(m,n)} = \text{maximum}(d_{jm}, d_{jn}), \tag{15.20}$$

where (m, n) is a cluster containing assemblage m and n, and maximum indicates that the largest distance is used in clustering.

Step 3. Find the third pair with the smallest distance and cluster. Accordingly, the farthest-neighbor distance, d, between two clusters is defined as

$$d_{(j,k)(m,n)} = \text{maximum}(d_{jm}, d_{jn}, d_{km}, d_{kn}), \tag{15.21}$$

where (j, k) is a cluster containing assemblage j and k, and (m, n) is defined above. As above, step 3 is repeated until all of the assemblages are contained a single cluster.

Box 15.8 Clustering Procedure for Hierarchical Cluster Analysis

Below are dendrograms of fish assemblages at six stations on the Kankakee River, Illinois (Box 15.1), resulting from single- and average-linkage clustering of Euclidean distance resemblance measure of summary fish abundance data. Step labels correspond to clustering steps, detailed below.

Initial Resemblance Matrix (Euclidean Distance)

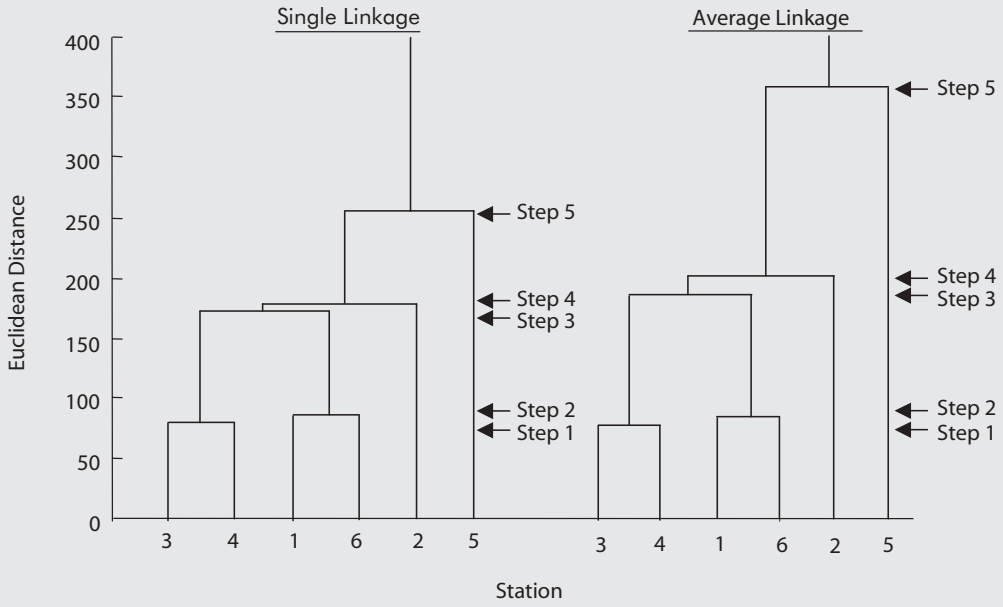


Table Euclidean distance matrix for fish assemblages at six stations of the Kankakee River data (Box 15.1).

	Station					
Station	1	2	3	4	5	6
1						
2	177.45					
3	175.23	218.18				
4	171.69	185.41	81.18			
5	275.22	380.08	435.16	435.53		
6	93.22	225.10	211.53	206.80	258.57	

Step 1

The first clustering step is to join the most similar pair of assemblages (i.e., smallest distance), which is stations 3 and 4, resulting in 1, 2, (3, 4), 5, and 6 at distance 81.18. The resemblance matrix is then

(Box continues)

Box 15.8 *(continued)*

adjusted to reflect the creation of (3, 4). For single-linkage clustering, the distance between (3, 4) and station 1 is estimated using equation (15.18), where

$$d_{j(m,n)} = \text{minimum}\,(d_{jm}, d_{jn}).$$
$$d_{1(3,4)} = \text{minimum}\,(175.23, 171.69)$$
$$= 171.69.$$

For average-linkage clustering, the distance between (3, 4) and station 1 is estimated using equation (15.22) where

$$d_{j(m,n)} = \frac{d_{jm} + d_{jn}}{N_{(m,n)}}$$
$$d_{1(3,4)} = \frac{175.23 + 171.69}{2}$$
$$= 173.46.$$

For each clustering method, compute the distances between (3,4) and the remaining stations and place into the corresponding revised matrix, shown below.

Table Second resemblance matrix based on single- and average-linkage clustering methods.

	Stations						Stations				
Station	1	2	5	6	(3,4)	Station	1	2	5	6	(3,4)
	Single-linkage distances						**Average-linkage distances**				
1						1					
2	177.45					2	177.45				
5	275.22	380.08				5	275.22	380.08			
6	93.22	225.10	258.57			6	93.22	225.10	258.57		
(3,4)	171.69	185.41	435.16	206.80		(3,4)	173.46	201.80	435.34	209.17	

Step 2

Next, join the most similar pair shown in the second resemblance matrix. For both clustering methods, this is stations 1 and 6, resulting in 2, (3,4), 5, and (1,6) at distance 93.22. As above, adjust each resemblance matrix to reflect the creation of (1,6). With single-linkage clustering, the distance between (1,6) and (3,4) is estimated using equation (15.19) and distances from the Euclidian distance matrix:

$$d_{(j,k)(m,n)} = \text{minimum}\,(d_{jm}, d_{jn}, d_{km}, d_{kn}).$$
$$d_{(3,4)(1,6)} = \text{minimum}\,(175.23, 171.69, 211.53, 206.80)$$
$$= 171.69.$$

With average-linkage clustering, the distance between (1,6) and (3,4) is estimated using equation (15.23) as

$$d_{(j,k)(m,n)} = \frac{d_{jm} + d_{jn} + d_{km} + d_{kn}}{N_{jk} + N_{mn}}$$

$$d_{(3,4)(1,6)} = \frac{175.23 + 171.69 + 211.53 + 206.80}{2 + 2}$$

$$= 191.31.$$

Compute the distances between (1,6) and the remaining stations with each clustering method, and place into the corresponding revised matrix, shown below.

Table Third resemblance matrix based on single- and average-linkage clustering methods.

	Stations					Stations			
Station	2	5	(3,4)	(1,6)	Station	2	5	(3,4)	(1,6)
	Single-linkage distances					**Average-linkage distances**			
2					2				
5	380.08				5	380.08			
(3,4)	185.41	435.16			(3,4)	201.80	435.34		
(1,6)	177.45	258.57	171.69		(1,6)	201.27	266.89	191.31	

Step 3

Join the most similar pair in the third resemblance matrix. For single-linkage distance, this is (3,4) and (1,6), resulting in 2, 5, and (1,3,4,6) at distance 171.69. With average-linkage distance, join (3,4) and (1,6) at 191.32. Compute the distances between (1,3,4,6) and stations 2 and 5 with each clustering method, and place into the corresponding revised matrix, shown below.

Table Fourth resemblance matrix based on single- and average-linkage clustering methods.

	Stations				Stations		
Station	2	5	(1,3,4,6)	Station	2	5	(1,3,4,6)
	Single-linkage distances				**Average-linkage distances**		
2				2			
5	380.08			5	380.08		
(1,3,4,6)	177.45	258.57		(1,3,4,6)	201.54	351.12	

Step 4

Join the most similar pair in the fourth resemblance matrix, 2 and (1,3,4,6), resulting in 5 and (1,2,3,4,6) at distance 177.45 for single linkage and 201.54 for complete linkage, respectively. Revise each resemblance matrix to reflect the creation of (1,2,3,4,6).

(Box continues)

Box 15.8 *(continued)*

Table Fifth resemblance matrix based on single- and average-linkage clustering methods.

	Stations			Stations	
Station	5	(1,2,3,4,6)	Station	5	(1,2,3,4,6)
	Single-linkage distances			**Average-linkage distances**	
5			5		
(1,2,3,4,6)	258.57		(1,2,3,4,6)	356.91	

Step 5

Join 5 to (1,2,3,4,6) to create one cluster (1,2,3,4,5,6) at 258.567 and 356.91 for single and average linkage, respectively.

Program

The following SAS program clusters and plots a dendrogram of fish assemblages at six stations on the Kankakee River, Illinois, with single-linkage clustering method (METHOD = SINGLE) and Euclidean distance resemblance measure.

```
OPTIONS PS = 60 LS=78;
DATA EUCLID (TYPE = DISTANCE);
INPUT NAME $ STATION1 STATION2 STATION3 STATION4 STATION5 STATION6;
LINES;
STATION1 0.000 177.449 175.234 171.694 275.218 93.220
STATION2 177.449 0.000 218.181 185.410 380.078 225.100
STATION3 175.234 218.181 0.000 81.179 435.160 211.535
STATION4 171.694 185.410 81.179 0.0000 435.525 206.797
STATION5 275.218 380.078 435.160 435.525 0.0000 258.567
STATION6 93.220 225.100 211.535 206.797 258.567 0.000
;
PROC CLUSTER DATA = EUCLID OUTTREE = SINGTREE
METHOD = SINGLE NONORM NOSQUARE NOPRINT;  ID NAME;
PROC TREE DATA = SINGTREE HORIZONTAL FC= 'C'  DIS LEVEL = 0;
RUN;
```

Program Output

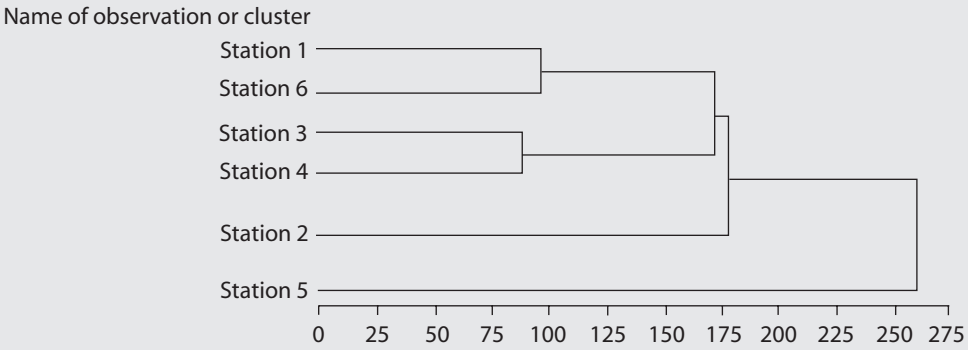


Figure Dendrogram of fish assemblages at six stations on the Kankakee River, Illinois, obtained with single-linkage clustering method and Euclidean distance resemblance measure.

For similarity measures, the farthest-neighbor distance is defined as the lowest similarity value of cluster members, and hence, the maximum in equations (15.20) and (15.21) is replaced by a minimum.

Average linkage. Average-linkage clustering, also known as unweighted pair-group with arithmetic averaging (Romesburg 1990), is the conceptual middle ground between single- and complete-linkage methods. The clustering steps are identical to those defined above. However, the distance between an assemblage, j, and a cluster (m, n) is defined as

$$d_{j(m,n)} = \frac{d_{jm} + d_{jn}}{N_{(m,n)}}, \tag{15.22}$$

and between two clusters as

$$d_{(j,k)(m,n)} = \frac{d_{jm} + d_{jn} + d_{km} + d_{kn}}{N_{jk} + N_{mn}}, \tag{15.23}$$

where N_{jk} and N_{mn} are the number of assemblages in clusters (j, k) and (m, n), respectively. Average-linkage clustering is illustrated in Box 15.8.

Dendrograms created with single-linkage clustering tend to be relatively long and narrow, whereas complete linkage tends to produce short, relatively compact trees with fewer clusters. Average linkage produces trees that are somewhat intermediate to these two extremes, and it is the most widely used clustering method. In a detailed analysis, Farris (1969) found that average linkage most faithfully represented the relationships among objects (e.g., assemblages) based on a mathematical analysis of different clustering methods. Nonetheless, we recommend that fisheries scientists create trees with two or more clustering methods and examine the fit of each. In the next section, we outline the method used to examine the adequacy of clustering methods.

Cophenetic correlation. The most appropriate means to assess the fit of each clustering method is by calculating a cophenetic correlation coefficient (Romesburg 1990). The cophenetic correlation coefficient is an unbiased measure of how well the cluster diagram represents relationships in the resemblance matrix; the largest correlation identifies the best clustering method. The cophenetic correlation coefficient is simply the Pearson product-moment correlation coefficient (equation [15.10]) between the resemblance matrix and the cophenetic matrix. The cophenetic matrix is an array of the distances among assemblages as represented in a dendrogram. It is estimated by tracing the path connecting each pair of assemblages in the dendrogram. An example calculation of the cophenetic correlation coefficient is presented in Box 15.9.

Classification. Classes are formed during hierarchical cluster analysis by cutting the tree at a specified level of similarity (Figure 15.6). Determining which level of similarity to cut the tree is highly subjective and depends upon study objectives (Romesburg 1990). For example, cutting a tree at a high level of similarity would

Box 15.9 Calculation of a Cophenetic Correlation Coefficient

Fish assemblages of the Kankakee River, Illinois (Box 15.1), were clustered with the single-linkage method (Box 15.8). Below, we illustrate calculation of the matrix cophenetic correlation coefficient. The values in the cophenetic matrix are estimated from the single-linkage dendrogram (Box 15.8) by tracing the path connecting each pair of assemblages. For example, when tracing the linkage from station 1 upward through the tree and downward to station 2, the greatest distance is 177.45. The remaining values are similarly estimated and are included in the cophenetic matrix below.

Table Resemblance matrix (Euclidean distance matrix) and cophenentic matrix from single-linkage clustering for calculation of cophenetic correlation coefficient.

	Station				
Station	1	2	3	4	5
	Resemblance Matrix				
1					
2	177.45				
3	175.23	218.18			
4	171.69	185.41	81.18		
5	275.22	380.08	435.16	435.53	
6	93.22	225.10	211.53	206.80	258.57
	Cophenetic Matrix				
1					
2	177.45				
3	171.69	177.45			
4	171.69	177.45	81.18		
5	258.57	258.57	258.57	258.57	
6	93.22	177.45	171.69	171.69	258.57

Table Side-by-side comparison of resemblance and cophenetic matrices for calculation of cophenetic correlation coefficient.

Matrix column, row	Resemblance matrix	Cophenetic matrix
1, 2	177.45	177.45
1, 3	175.23	171.69
1, 4	171.69	171.69
1, 5	275.22	258.57
1, 6	93.22	93.22
2, 3	218.18	177.45
2, 4	185.41	177.45
2, 5	380.08	258.57
2, 6	225.10	177.45
3, 4	81.18	81.18
3, 5	435.16	258.57
3, 6	211.53	171.69
4, 5	435.53	258.57
4, 6	206.80	171.69
5, 6	258.57	258.57
Average	235.36	190.92

(Box continues)

Box 15.9 *(continued)*

The resemblance matrix then is unraveled and its values are paired with the corresponding values in the cophenetic matrix as shown above. The cophenetic correlation between the resemblance (r) and cophenetic (s) matrices is calculated using equation (15.10) as

$$\theta_{rs} = \frac{\Sigma (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\Sigma (r_i - \bar{r})^2 \quad \Sigma (s_i - \bar{s})^2}}.$$

$$\theta_{rs} = \frac{(177.45 - 235.36)(177.45 - 190.92) + \ldots + (258.57 - 235.36)(258.57 - 190.92)}{\sqrt{[(177.45 - 235.36)^2 + \ldots + (258.57 - 235.36)^2][(177.45 - 190.92)^2 + \ldots + (258.57 - 190.92)^2]}}$$

$$= \frac{77{,}129.30}{\sqrt{[162{,}332.91][46{,}673.49]}}$$

$$= 0.886.$$

Cophenetic correlation coefficients would be calculated for each clustering method. The hierarchical tree (i.e., dendrogram) with the largest cophenetic correlation would be selected to infer relationships among fish assemblages.

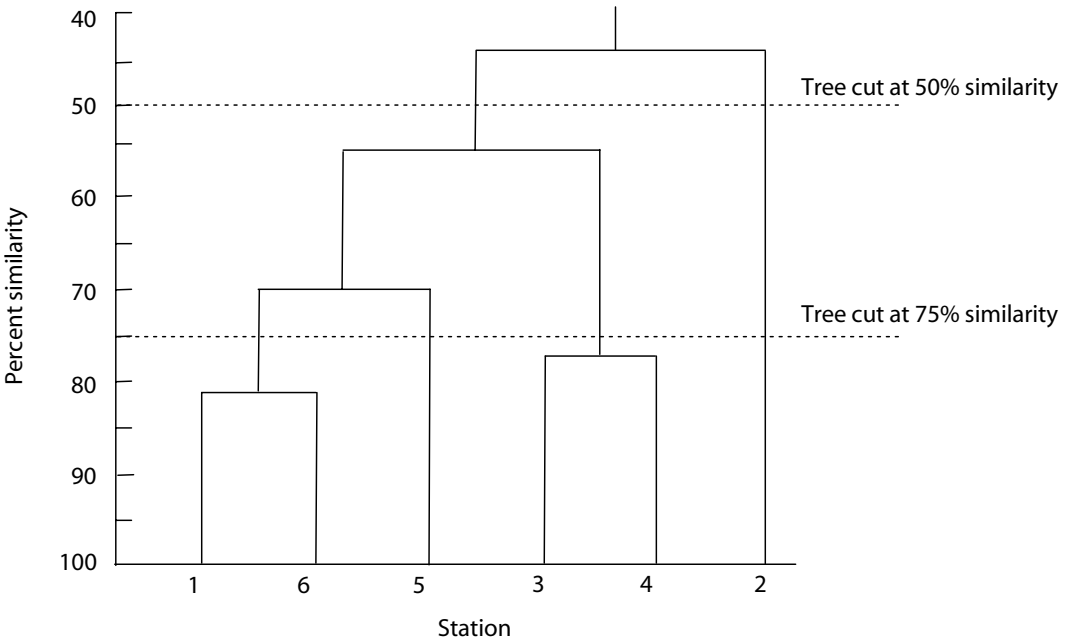


Figure 15.6 Dendrogram of fish assemblages at six stations on the Kankakee River, Illinois (Box 15.1), clustered using average linkage. Broken lines represent two cut points for classifying the fish assemblages. Four classes (1,6), 5, (3,4), and 2 are formed at 75% similarity, whereas two classes, (1, 6, 5, 3, 4) and 2, are formed at 50% similarity.

likely result in an excessively large number of groups, which would diminish one of the advantages of classification (i.e., reducing the data to a manageable size). Conversely, cutting a tree at a low level of similarity would likely introduce too much variation (heterogeneity) with each group, which could render the classification useless. Good classifications should attempt to minimize the number of groups created while simultaneously maximizing the within-group similarity.

15.4.2.2 K-Means Cluster Analysis

The most common form of nonhierarchical cluster analysis is *k*-means clustering. It is used to group assemblages into *k* clusters (i.e., *k* = number of clusters) based on their Euclidean distances. The *k*-means clustering procedures differ markedly from hierarchical clustering. It begins with a matrix of Euclidean distances and randomly assigns the assemblages to a prespecified number of clusters, *k*. Using a variance-minimizing algorithm, the assemblages are reassigned to different clusters on the basis of their similarity (distance) to other assemblages in a cluster. This process continues iteratively until the distance within each cluster is minimized and no assemblages need to be reassigned. The *k*-means clustering analysis is relatively robust to outlying data because, unlike hierarchical cluster analysis, the nature of the relationships among assemblages is unconstrained. However, the minimizing algorithm used for *k*-means clustering is inefficient when the number of samples is relatively low (Hartigan 1985). Therefore, we recommend *k*-means clustering only when the number of assemblages in the data set exceeds 30.

Similar to hierarchical cluster analysis, determining the number of clusters for a particular classification with *k*-means clustering is somewhat subjective. In fact, there are no completely unbiased methods for determining the number of clusters for any type of cluster analysis (Hartigan 1985). However for *k*-means clustering, statisticians have developed several methods that can be used to examine the relationship between within-cluster similarity and the number of clusters. This, in turn, can be used to select the optimal number of clusters (*k*) for a classification. One such method is to fit *k*-means clusters for several values of *k* and plot the overall R^2 versus the number of clusters. The overall R^2 is a measure of predictability of the fish assemblage within a cluster and is analogous to an r^2 in regression analysis. The optimal number of clusters (*k*) is considered the lowest value at which the R^2 begins to level off and reach an asymptote (Figure 15.7).

Another method for ascertaining the optimal number of clusters is the cubic clustering criterion (CCC) developed by Sarle (1983). Cubic clustering criterion is an estimator of *k*-means cluster fit. Similar to the overall R^2 method, the optimal number of clusters is determined by estimating CCC with *k*-means clustering for several values of *k* and examining a plot of *k* and CCC to find the smallest value of *k* with the lowest CCC (Figure 15.7). There are several other methods that have been developed for estimating the optimal number of clusters, all of which require specialized software and a strong background in statistical theory (e.g., Milligan and Cooper 1985; Mueller and Sawitzki 1991). For most fisheries applications, we suggest that scientists use both the R^2 and CCC to help find the optimal

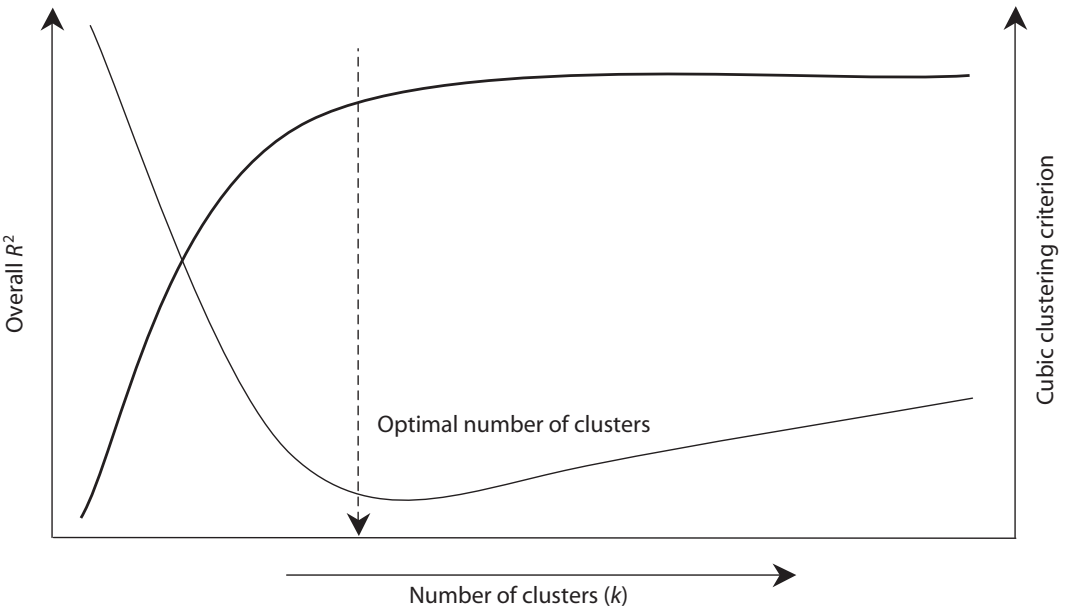


Figure 15.7 Hypothetical plot of the overall R^2 (thick line) and cubic clustering criterion (thin line) for various cluster sizes created using *k*-means clustering. Broken line represents the optimal number of clusters in the classification based on the overall R^2 and cubic clustering criterion.

number of clusters (k). An example SAS program for *k*-means clustering and the associated output are provided in Box 15.10.

15.4.3 Ordination Techniques

Ordination is the term for a variety of statistical techniques that ecologists have used to examine complex multivariate relationships among biotic communities and their environment. There are two broad classes of ordination—constrained and unconstrained. Constrained ordination is used to examine the relationship between assemblage structure and environmental gradients (e.g., stream habitat characteristics or water quality), whereas unconstrained ordination is used to examine the relationships among assemblages or species. Because we emphasize assemblage structure in this chapter, rather than function, we consider only unconstrained ordination. We refer fisheries scientists interested in learning more about constrained ordination to Ter Braak (1986) and Ter Braak and Prentice (1988).

Unconstrained ordination is used to summarize complex, multivariate relationships among assemblages and graphically display these within a small number of dimensions, usually two to three. Unlike cluster analysis, relationships are scored on a continuous scale (e.g., they need not be hierarchical or placed into discrete

Box 15.10 *K*-Means Clustering Analysis

The following SAS program performs *k*-means clustering with PROC FASTCLUS on the summary fish abundance data for six stations on the Kankakee River, Illinois (Box 15.1). Three-letter codes are used in place of species names (see Box 15.11 for key to codes). The number of clusters, $k = 3$, is specified by the MAXCLUSTERS command. Note that this example is for illustration only. The *k*-means clustering procedure should be used only when the number of assemblages (samples) exceeds 30.

Program

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT STATION $ LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR SVR
SAB BKS BLG GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
LINES;
STATION1 6 164 42 0 13 0 10 0 8 1 19 45 0 0 34 5 35 5 2 8 0 4 2 1 0 35 1 30 143 4 1 0 25 5
STATION2 7 90 33 0 58 0 11 22 89 22 24 32 0 0 0 2 0 2 0 1 3 10 0 42 3 94 0 3 59 0 9 6 0 0
STATION3 0 6 29 0 10 0 10 4 5 4 3 69 0 4 35 10 8 1 2 1 0 13 0 6 2 26 0 15 195 0 1 3 51 6
STATION4 4 6 3 1 14 0 0 0 15 1 2 14 0 0 36 7 2 4 0 3 0 9 2 3 0 39 8 31 151 1 0 0 42 7
STATION5 26 432 44 15 36 7 2 0 8 5 23 51 4 1 9 2 35 17 0 14 3 10 1 7 8 48 31 27 165 0 4 2 2 4 2
STATION6 5 194 35 0 13 0 0 2 35 1 8 14 1 0 55 7 22 14 5 4 0 6 0 1 0 37 3 62 204 1 0 0 7 2
;
PROC FASTCLUS DATA = SPECIES MAXCLUSTERS = 3 OUT = CLUSTER SHORT;
VAR LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR SVR SAB BKS BLG
GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
ID STATION;
PROC SORT; BY CLUSTER;
PROC PRINT NOOBS; VAR CLUSTER STATION DISTANCE;
RUN;
```

Program Output

Table Output for SAS' FASTCLUS procedure for *k*-means clustering of fish abundance data for six stations on the Kankakee River, Illinois (Box 15.1). The number of clusters, $k = 3$, is specified by the MAXCLUSTERS command. Settings for other commands (Replace = FULL, Radius = 0, and Maxiter = 1) are left at SAS/STAT default settings; see SAS manual for details (SAS Institute 2004). Note that *k*-means clustering procedure should be used only when the number of assemblages (samples) exceeds 30 (illustrated here with six samples). The root mean square between members within-group standard deviation is given by RMS SD. Maximum distance from seed to observation is the greatest difference between a random-number seed to an observation in that cluster. Criterion based on final seeds = 14.8912. Overall R^2 is a measure of predictability of the fish assemblage within a cluster. Further explanation of table values follows.

Cluster Summary

Cluster	RMS frequency	SD	Maximum distance from seed to observation	Nearest cluster	Distance between cluster centroids
1	1		0	3	319.3
2	2	22.4843	92.7052	3	129.7
3	3	20.3094	125.7	2	129.7

Statistics for Variables

Variable	Total STD	Within STD	R^2	$R^2/(1 - R^2)$
Longnose gar (Other 32 species)	9.14330	2.89636	0.939793	15.609272
Slenderhead darter	2.60768	3.32499	0.024510	0.025126
Overall	31.24607	21.05932	0.727448	2.669025
Pseudo *F*-statistic	4.00[a]			
Approximate expected overall R^2	0.7274[a]			
Cubic clustering criterion[b]				

Cluster Membership

Cluster	Station	Distance
1	5	0.000
2	2	92.705
2	4	92.705
3	1	61.506
3	3	125.706
3	6	91.995

[a] The two values are invalid for correlated variables.
[b] None calculated because of small sample size.

The output summary indicates that three clusters were specified (MAXCLUSTERS = 3). The remaining variables in the summary are additional clustering options; see SAS manual for details. In the first part of the table, distance between cluster centroids is used to examine the overall relationship among cluster members. Clusters 2 and 3 have the smallest distances, and hence, their members are more similar than those in cluster 1. The statistics for variables portion of the table is used to examine the change in R^2 with the number of clusters (k). The overall R^2 (0.727) and cubic clustering criterion (none was calculated because of small sample size) would be used to estimate the optimal number of clusters for the classification (Figure 15.7). The final portion of the table contains cluster membership and members' Euclidean distances from the cluster centroid (i.e., the cluster mean), which can be used to examine similarity among cluster members. For example, cluster 3 contains stations 1, 3, and 6. Among these, stations 1 and 6 are the most similar because they have the smallest distances (61 and 92).

clusters), and hence, ordination is a better technique for examining relationships among assemblages. Unconstrained ordination is generally not used for classification but can be useful for verifying those relationships indicated by cluster analysis and for suggesting alternative classifications.

There are several unconstrained ordination methods, and each has its advantages and limitations. The most commonly used techniques are principal component analysis (PCA) and nonmetric multidimensional scaling (NMDS). Of the two, PCA produces the most detailed and quantifiable measures of the relationships among assemblages (Cliff 1987). However, in an extensive evaluation of ecological ordination methods, Minchin (1987) found that PCA performed very poorly when relationships were nonlinear, whereas NMDS was the most robust technique. Further, he suggested that NMDS was the most appropriate ordination method for ecological applications. Relationships among fish assemblages are likely to be nonlinear to varying degrees. Hence, the relationships indicated by PCA could be biased. To maximize ecological insight and minimize the potential for bias, we recommend that fisheries scientists use both techniques and compare their results.

15.4.3.1 Principal Component Analysis

Principal component analysis reduces species (relative) abundances into linear combinations (i.e., principal components) that are uncorrelated with each other (Stevens 1992). It emphasizes the variation among assemblages, rather than similarities, and assumes that approximately linear relationships exist among fish assemblages. Prior to PCA, species abundances are standardized by estimating the correlation or covariance between species. Fisheries scientists should always standardize using correlation, which is the default for most statistical software. Based on the correlation matrix, PCA begins by finding the linear combination of species-specific abundances that accounts for the greatest amount of variation among samples. This linear combination is called the first principal component, which is simply a linear regression using (standardized) species abundances as predictors of a principal component score. The collection of the corresponding linear regression coefficients (one coefficient per species) is called the eigenvector of the first principal component, and the amount of variance explained is estimated as the eigenvalue.

Following the estimation of the first principal component, PCA finds a second linear combination (regression) of species abundances that accounts for the largest amount of the remaining variance (i.e., after the variance attributable to the first component is removed) and is pairwise uncorrelated with the first component. This is the second principal component, which also has an eigenvector and eigenvalue. The third principal component then is derived to be uncorrelated to the first two components and accounts for the third largest amount of variance. The process of creating uncorrelated linear combinations (principal components) is continued with each component accounting for the remaining variation until none remains. Thus, PCA attempts to summarize the pattern of variation among assemblages with a smaller number of components (compared to the number of

species) that accounts for most of the variance in the original data set. For most applications, this can be accomplished with fewer than five principal components.

The relationships among assemblages are determined by examining plots of principal component scores, with similar assemblages being located close together and dissimilar ones farther apart. For example, fish assemblages in stations 1 and 6 on the Kankakee River are similar to one another but are very different from those at stations 2 and 5 (Figure 15.8). Principal component scores are computed

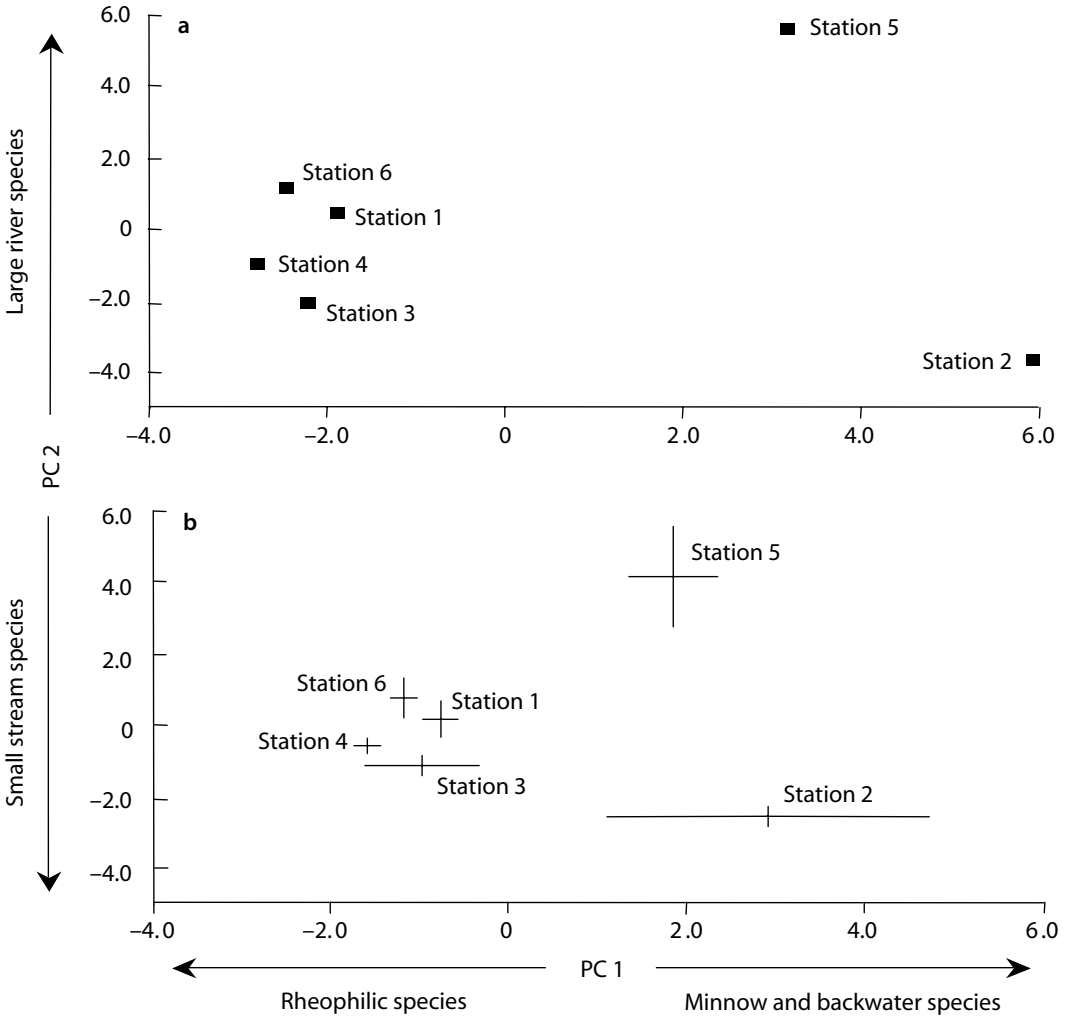


Figure 15.8 Plots of the first two principal components (PCs) of fish assemblages at six stations on Kankakee River, Illinois. The upper plot (a) was created using the PC scores from an analysis with summary data in Box 15.1. The bottom figure (b) was created using the PC scores from an analysis using eight sequential fish samples collected at each station (see Box 15.1 for summed data). The crossbars in panel (b) represent the mean and standard errors of sample scores for each station and can be used to examine the degree of overlap in assemblage structure. Principal component axes were interpreted using component loadings (Box 15.11).

for each assemblage (or sample) using the principal component eigenvectors (i.e., linear regression coefficients), and hence, each assemblage has a score for each principal component.

Each principal component is interpreted for ecological meaning by examining the principal component loadings, which are the Pearson's correlation coefficients between the component scores and species-specific abundances. These are also referred to as factor loadings or simply factors. Large loadings (in absolute value) are interpreted as having the greatest influence on that component. For example, if three species loaded high on a particular component, the component would be interpreted in terms of some characteristic these species have in common, such as taxonomy, habitat use, or environmental tolerances. Identifying the loadings of components to use for interpretation, however, is somewhat subjective. Because loadings are Pearson correlations, many fisheries scientists interpret only loadings that are statistically significant. These statistical tests are significantly influenced by sample sizes (Cliff 1987). For example, a species with a 0.20 loading would share only 4% of the variance with the principal component but could be statistically significant when sample sizes are large. Rather, a good general rule-of-thumb is to use species with loadings greater than |0.4| for interpretation when there are small numbers of species in the analysis (<20) and |0.6| for more species-rich assemblages (Stevens 1992). An example SAS program and output for PCA are provided in Box 15.11.

Theoretically, the maximum number of principal components in an analysis of fish assemblages is equal to the number of variables (species) analyzed. Thus, researchers must choose how many components to include in their analysis. There is no perfect criterion for selecting the number of components to retain. However, several techniques have been developed to assist the analyst. The most widely used of these is called the scree test (Cattell 1966). With this method, eigenvalues are plotted against their corresponding principal component number (i.e., first component = 1, second = 2, etc.). Eigenvalues generally decrease rapidly (steep slope) with increasing component number and then level off (shallow slope). The transition between these two slopes is known as the break point (Figure 15.9), and the components with eigenvalues above the breakpoint are retained for subsequent analysis. Another widely used approach is to retain all of the components with eigenvalues greater than 1 (Kaiser 1960). This method virtually ensures that important components will be retained but also tends to include additional components that have little explanatory value (Hakstian et al. 1982). A third criterion is based on practical considerations. There are two difficulties associated with using large numbers of principal components for examining the relationships among assemblages. First, three dimensions are difficult to display in a single figure and are substantially more difficult to comprehend than two. Second, four or more dimensions render ordination practically useless as a method of understanding complex relationships. Thus, we recommend that only two components be retained in fish assemblage analyses if these account for at least 70% of the variance; otherwise three components should be retained. However, if three components account for less than one half the variation, we suggest refraining from using PCA and consider NMDS as a more parsimonious alternative.

Box 15.11 Principal Component Analysis

The following SAS program performs principal component analysis (PCA) on the summary fish abundance data from six stations of the Kankakee River, Illinois (Box 15.1). Component loadings are estimated for the first three principal components, and scores are plotted for the first two components.

Program

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT STATION $ LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR SVR
SAB BKS BLG GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
LINES;
STATION1 6 164 42 0 13 0 10 0 8 1 19 45 0 0 34 5 35 5 2 8 0 4 2 1 0 35 1 30 143 4 1 0 25 5
STATION2 7 90 33 0 58 0 11 22 89 22 24 32 0 0 0 2 0 2 0 1 3 10 0 42 3 94 0 3 59 0 9 6 0 0
STATION3 0 6 29 0 10 0 10 4 5 4 3 69 0 4 35 10 8 1 2 1 0 13 0 6 2 26 0 15 195 0 1 3 51 6
STATION4 4 6 3 1 14 0 0 0 15 1 2 14 0 0 36 7 2 4 0 3 0 9 2 3 0 39 8 31 151 1 0 0 42 7
STATION5 26 432 44 15 36 7 2 0 8 5 23 51 4 1 9 2 35 17 0 14 3 10 1 7 8 48 31 27 165 0 4 2 2 4 2
STATION6 5 194 35 0 13 0 0 2 35 1 8 14 1 0 55 7 22 14 5 4 0 6 0 1 0 37 3 62 204 1 0 0 7 2
;
PROC PRINCOMP DATA=SPECIES OUT = PRIN;
VAR LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR SVR SAB BKS BLG
GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;

PROC CORR NOPRINT OUTP = LOADING NOSIMPLE NOPROB;
VAR PRIN1 PRIN2 PRIN3;
WITH LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR SVR SAB BKS
BLG GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
TITLE '*** COMPONENT LOADINGS ***';
PROC PRINT DATA = LOADING NOOBS; WHERE _NAME_ NE '';

PROC PLOT DATA=PRIN; PLOT PRIN2 * PRIN1 = '*' $ STATION /;
TITLE '*** PRINCIPAL COMPONENT SCORES ***';
RUN; TITLE; QUIT;
```

Program Output

Table Partial output of SAS program for principal component (PC) analysis. See "Component Loadings" at end of table for species abbreviations. Component loadings are Pearson correlation coefficients.

Simple Statistics (Partial)

	LOG	GZS	BLM	BUM	CAP
Mean	8.000000000	148.6666667	31.00000000	2.666666667	24.00000000
SD	9.143303561	159.1963149	14.81890684	6.055300708	19.17289754

Correlation Matrix (Partial)

	LOG	GZS	BLM	BUM	CAP	HOC	MIS	RDS	RYS
LOG	1.0000	0.9211	0.4812	0.9609	0.4438	0.9644	–0.3062	–0.1519	–0.1120
GZS	0.9211	1.0000	0.6948	0.8522	0.2727	0.8719	–0.3171	–0.2574	–0.1588
BLM	0.4812	0.6948	1.0000	0.3722	0.2583	0.4298	0.3477	0.0687	0.0333

(Box continues)

Box 15.11 *(continued)*

Eigenvalues of the Correlation Matrix

PC	Eigenvalue	Difference	Proportion	Cumulative
1	13.4751939	3.2555850	0.3963	0.3963
2	10.2196089	5.5254938	0.3006	0.6969
3	4.6941151	1.7841269	0.1381	0.8350
4	2.9099883	0.2088945	0.0856	0.9206
5	2.7010937	2.7010937	0.0794	1.0000
6	0.0000000	0.0000000	0.0000	1.0000
7	0.0000000	0.0000000	0.0000	1.0000
(Remaining 27 PCs)				

Eigenvectors

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
LOG	0.148222	0.259093	0.030061	–0.029427	–0.064273	–0.019554	–0.064777
GZS	0.104941	0.279880	–0.065667	0.058686	0.087773	–0.009625	–0.043083
BLM	0.104793	0.139913	–0.073891	0.061684	0.477267	–0.031307	–0.012889
(Remaining species and 27 PCs)							

Component Loadings

SpeciesSpecies	Abbreviation	PC1	PC2	PC3
Longnose gar	LOG	0.54410	0.82827	0.06513
Gizzard shad	GZS	0.38522	0.89472	–0.14227
Bluntnose minnow	BLM	0.38468	0.44727	–0.16009
Bullhead minnow	BUM	0.40982	0.84812	0.30129
Common carp	CAP	0.98347	–0.11637	–0.08999
Hornyhead chub	HOC	0.42955	0.84878	0.29467
Mimic shiner	MIS	0.32226	–0.53181	0.20040
Redfin shiner	RDS	0.74136	–0.64720	–0.10607
Rosyface shiner	RYS	0.68056	–0.54750	–0.41512
Sand shiner	SAS	0.87490	–0.48009	–0.02615
Spotfin shiner	SFS	0.81830	0.28838	–0.28268
Striped shiner	STS	0.12277	0.11097	0.73543

15.4.3.2 Multidimensional Scaling

Nonmetric multidimensional scaling models the relationships among two or more assemblages in a specified number of dimensions based on their similarity or dissimilarity (Kruskal and Wish 1984). It is a robust ordination technique that does not require an assumption of normality or linearity of relationships among assemblages and can usually fit a model in fewer dimensions than can PCA. Nonmetric

Component Loadings (*continued*)

SpeciesSpecies	Abbreviation	PC1	PC2	PC3
Suckermouth minnow	SUM	0.35570	0.89406	0.15593
Black redhorse	BLR	–0.18744	–0.11634	0.87537
Golden redhorse	GOR	–0.92506	0.00075	–0.23981
Northern hog sucker	NHS	–0.84410	–0.32879	0.33545
Shorthead redhorse	SHR	–0.09627	0.83134	–0.22237
Quillback	QLL	0.08542	0.87966	–0.25527
River redhorse	RVR	–0.55565	0.01944	–0.41357
Silver redhorse	SVR	0.17510	0.94825	–0.02850
Smallmouth buffalo	SAB	0.96851	0.21751	0.08016
Brook silverside	BKS	0.28808	–0.23178	0.86445
Bluegill	BLG	–0.32942	0.30074	–0.14319
Green sunfish	GSF	0.84888	–0.52367	–0.05485
Largemouth bass	LMB	0.68457	0.57695	0.43733
Longear sunfish	LOS	0.89030	–0.34000	–0.26148
Orangespotted sunfish	OSF	0.31102	0.86433	0.25246
Rock bass	ROB	–0.59897	0.43907	–0.52590
Smallmouth bass	SMB	–0.74860	0.43763	0.23112
Banded darter	BAD	–0.44120	0.08443	–0.52494
Blackside darter	BLD	0.97080	–0.23332	–0.01678
Johnny darter	JOD	0.80895	–0.46298	0.30348
Logperch	LOP	–0.70992	–0.34408	0.55491
Slenderhead darter	SLD	–0.69835	0.16420	0.50953

The first two principal components accounted for 69.7% of the variation among fish assemblages and were retained for analysis. The first component accounted for 39.6% of the variation and loaded heavily and positively on minnows (RDS, RYS, SAS, SFS) and backwater species (CAP, SAB, GSF, LMB, LOS) and negatively on rheophilic species (GOR, NHS, SMB, LOP, SLD). The second component accounted for 30.1% of the variation and loaded heavily and positively on large-river species (LOG, GZS, BUM, HOC, SUM, SHR, QLL, SVR, OSF) and negatively on small stream-dwelling species (RDS). The plot of the first two principal components (Figure 15.8a) indicates that assemblages at stations 2 and 5 had larger numbers of minnow and backwater species, whereas stations 1, 3, 4 and 6 contained greater numbers of rheophilic species. Station 5 also included greater numbers of large-river species, and station 2 had greater numbers of small stream-dwelling species.

multidimensional scaling begins with a matrix of resemblance coefficients (section 15.4.1) and uses an iterative procedure and algorithm to find the set of coordinates for each assemblage that, when plotted, most closely approximates the relationships indicated by the resemblance matrix. For example, similar assemblages should be located closer together in an NMDS plot and dissimilar assemblages farther apart. The number of coordinates depends upon the specified number of dimensions (e.g., two dimensions = two coordinates). The degree of

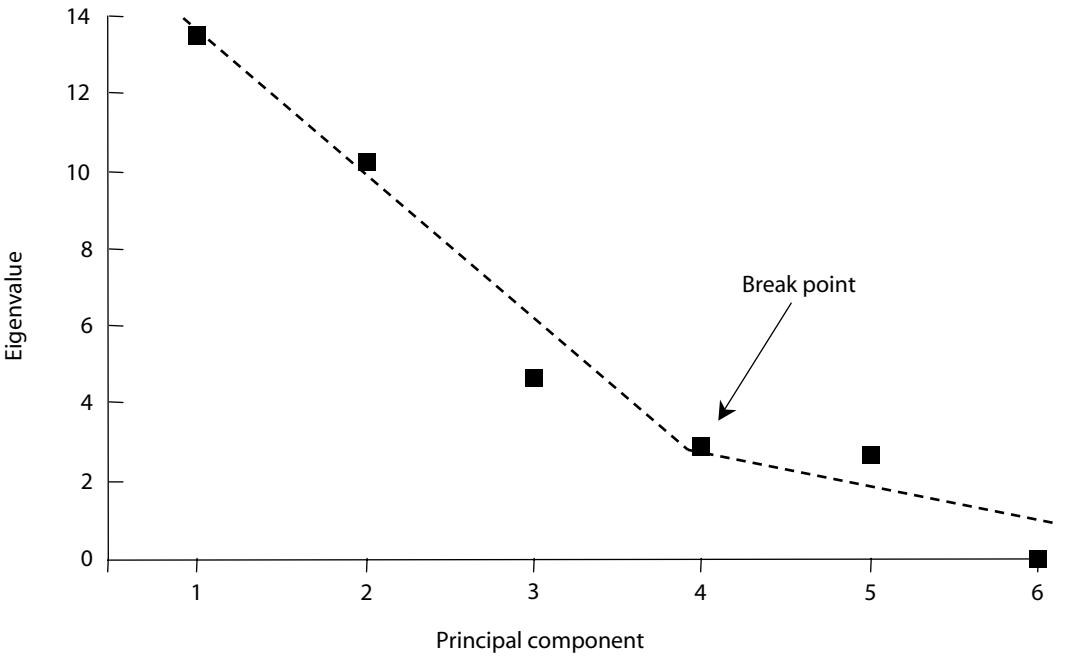


Figure 15.9 Scree plot of principal component eigenvalues from the analysis of Kankakee River fish assemblages (Boxes 15.1 and 15.11). Broken line is shown for illustration, and arrow indicates break point. Based on the scree test, principal components 1–3 would be retained for analysis.

correspondence between the resemblance coefficients and an NMDS plot is measured as stress, with lower stress values indicating better fit. Thus, NMDS iteratively finds the best-fitting coordinates by minimizing stress.

There are several algorithms that can be used to fit NMDS coordinates (Green and Rao 1972). The calculation and theory behind each are technical and beyond the scope of this chapter. However, there are two practical methods fisheries scientists can use to examine the fit of an NMDS model and determine the optimal number of dimensions. The first is the final estimate of stress or other related measure (e.g., badness-of-fit in SAS Institute 2004) that is estimated during the final iteration. Low values indicate better fit, and NMDS models with stress greater than 0.15 should be considered suspect (Kruskal and Wish 1984). The second method is inspection of a Shepard diagram, which is a scatterplot of the estimated similarity (or dissimilarity) among assemblages as shown in the NMDS plot versus the observed (actual) similarities (Shepard 1963). A Shepard diagram for a good-fitting model should resemble a smooth curve or a straight line (Box 15.12). Diagrams that resemble a stair-step or L-shaped function indicate a poor-fitting model, due to an incorrect specification of number of dimensions or the use of an incorrect algorithm. An example SAS program and output for NMDS are provided in Box 15.12.

Box 15.12 Nonmetric Multidimensional Scaling

The following SAS program performs nonmetric multidimensional scaling (NMDS) analysis on the summary fish abundance data from six stations of the Kankakee River, Illinois (Box 15.1). The resemblance measure is percent similarity, the scaling algorithm is monotonic (LEVEL = ORDINAL), and two dimensions are specified (DIM = 2). For more options, consult the SAS manual (SAS Institute 2004).

Program

```
OPTIONS PS = 60 LS=78;
DATA PSI;
INPUT STATION $ STATION1 STATION2 STATION3 STATION4 STATION5 STATION6;
LINES;
STATION1 100.0 50.8 62.6 56.7 73.0 81.4
STATION2 50.8 100.0 38.0 37.3 49.6 46.5
STATION3 62.6 38.0 100.0 77.5 44.5 59.2
STATION4 56.7 37.3 77.5 100.0 43.7 64.3
STATION5 73.0 49.6 44.5 43.7 100.0 68.1
STATION6 81.4 46.5 59.2 64.3 68.1 100.0
;
PROC MDS DATA = PSI DIM=2 SHAPE=SQUARE LEVEL=ORDINAL OUT=MDSOUT OUTRES=RESID NONORM;
ID STATION;
DATA MDSOUT; SET MDSOUT; WHERE _TYPE_ = 'CONFIG';
PROC PLOT DATA=MDSOUT; PLOT DIM2 * DIM1 = '*' $ STATION /;
PROC PLOT DATA=RESID; PLOT FITDATA * FITDIST = '*' /;
TITLE "SHEPARD DIAGRAM";
RUN;
TITLE;
QUIT;
```

Program Output

Table Nonmetric multidimensional scaling analysis on the summary fish abundance data from six stations of the Kankakee River, Illinois (Box 15.1). The resemblance measure is percent similarity, the scaling algorithm is monotonic (LEVEL = ORDINAL), and two dimensions are specified (DIM = 2). Other command settings are SAS defaults (Shape = SQUARE, Condition = MATRIX, Coef = IDENTITY, Formula = 1, Fit = 1, Mconverge = 0.01, Gconverge = 0.01, Maxiter = 100, Over = 2, and Ridge = 0.00010). The convergence criteria were satisfied.

Iteration	Badness of-fit type	Criterion	Change in criterion	Convergence measures	
				Monotone	Gradient
0	Initial	0.1655			
1	Monotone	0.0237	0.1418	0.1608	0.7467
2	Gau-New	0.0153	0.008313		
3	Monotone	0.0101	0.005246	0.0115	0.7041
4	Gau-New	0.0101	0.0000406		
5	Monotone	0.003382	0.006670	0.009464	0.5672
6	Gau-New	0.002792	0.000590		0.0221
7	Gau-New	0.002792	6.8241×10^{-7}		0.000167

(Box continues)

Box 15.12 *(continued)*

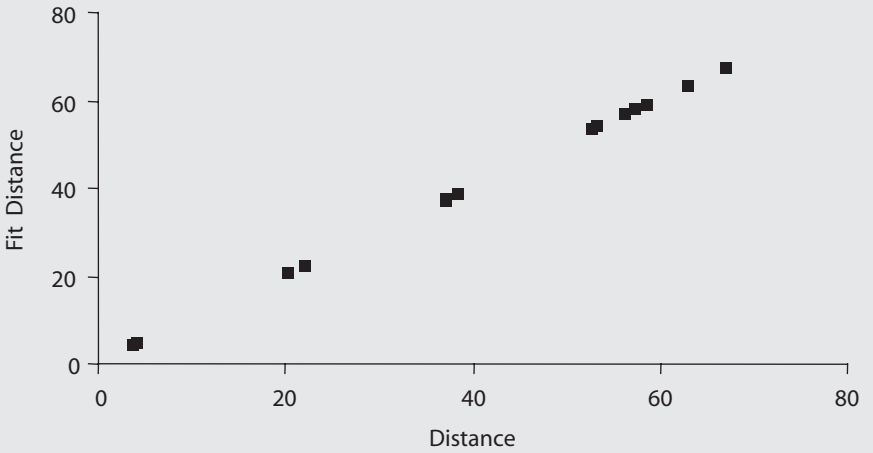


Figure Shepard diagram, which is a scatterplot of the estimated similarity (or dissimilarity) among assemblages as shown in the NMDS plot versus the observed (actual) similarities (Shepard 1963).

The badness-of-fit criteria (stress) declined smoothly among iterations and converged after seven iterations at 0.002792. The Shepard diagram displayed a smooth, linear relationship, which indicated a good fit for the two-dimensional NMDS model. The two-dimensional NMDS plot indicated that assemblages at stations 1 and 6 were most similar, as were those at stations 3 and 4 (Figure 15.10c). Fish assemblages at stations 2 and 5, however, differed from one another and from those at the other four stations. These relationships are also consistent with those indicated by PCA (Figure 15.8).

Similar to PCA, relationships among assemblages are determined by examining NMDS plots, and similar assemblages are located proximally and dissimilar assemblages distally. These relationships, however, are influenced by the characteristics of the resemblance measure, which is similar to hierarchical cluster analysis. For example, NMDS with Jaccard's similarity index suggests that fish assemblages at all stations on the Kankakee River are very different from one another, whereas NMDS with Kendall's tau suggests that assemblages at stations 1, 4, and 6 are most similar, and the percent similarity index suggests that the assemblage at station 4 is most similar to station 3 (Figure 15.10). Unlike PCA, NMDS has no formal or quantifiable means to interpret the various dimensions (i.e., no loadings). Thus, it requires insight and further examination of assemblage characteristics to determine the underlying pattern (gradient) associated with each dimension.

Nonmetric multidimensional scaling models can be fit with one to n dimensions, where n is the number of assemblages (samples). The optimal number of dimensions can be determined, similar to PCA, by examining a scree plot of stress

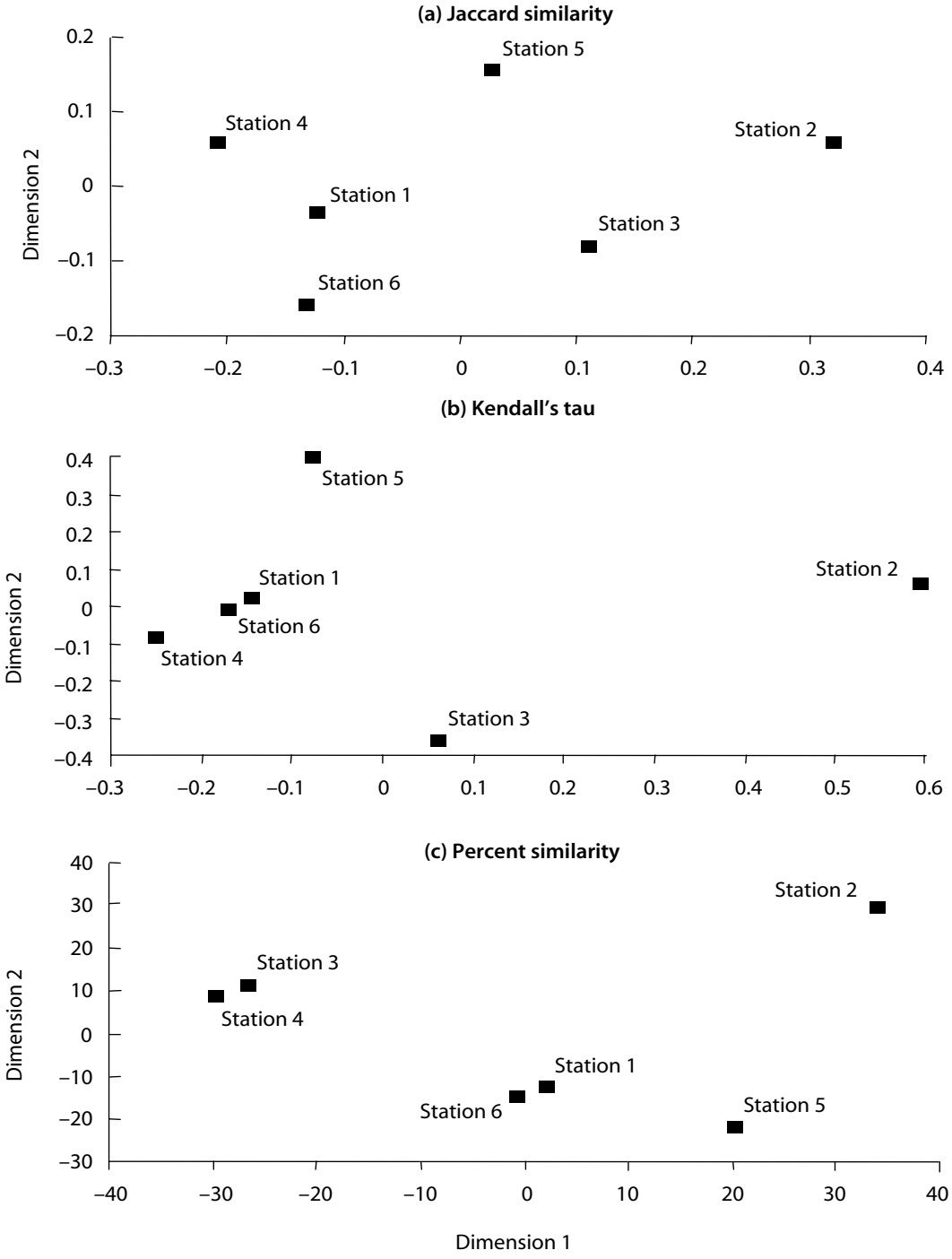


Figure 15.10 First two dimensions for nonmetric multidimensional scaling analysis of fish assemblages at six stations on the Kankakee River, Illinois (Box 15.1), with (a) Jaccard's similarity, (b) Kendall's tau, and (c) percent similarity resemblance matrices. Similar assemblages are located proximally and dissimilar assemblages are more distant.

against the corresponding number of dimensions or by examining a Shepard diagram (Green and Rao 1972). However, as with PCA, we suggest that fisheries scientists limit the maximum number of dimensions to three due to practical considerations, such as interpretability.

15.4.4 Other Multivariate Techniques—Discriminant Analysis

Fisheries scientists often need to determine how two or more fish assemblages differ and if assemblage composition is predictable. In these instances, assemblages are treated as discrete response categories (e.g., assemblages A, B, and C) with species (relative) abundances as their characteristics (i.e., predictors). Categorical data analysis is the generic term for a variety of statistical techniques for analyzing data with categorical responses (Agresti 1990). It can be used to find the species (or combination thereof) that best characterize an assemblage and can also be used to examine predictability of assemblage structure (Peterson and Rabeni 2001). In contrast to the other techniques in this chapter, categorical data analysis can be employed only when multiple samples (replicates) are collected from each assemblage. Hence, the quality of categorical data analysis is significantly influenced by sample size. To ensure reliable results, fisheries scientists should use categorical data analysis only when each assemblage has at least 20 samples (Cliff 1987).

There are several categorical data analysis techniques, the most widely used of which is discriminant analysis. Discriminant analysis is a linear statistical technique, requiring assumptions of normality and constant variance. It is relatively robust to minor violations of these assumptions (Stevens 1992) and should be appropriate for most practical applications in fisheries. However, biologists should consider alternative techniques when data are severely nonnormal and variances are heterogeneous. Fisheries scientists interested in alternative techniques may consult Agresti (1990) for logistic methods, Hand (1982) for nonparametric methods, and Breiman et al. (1984) for tree-based methods.

Discriminant analysis is a multivariate statistical technique, related to PCA, that reduces species (relative) abundances into linear combinations (i.e., discriminant functions) that are pairwise uncorrelated (Lachenbruch 1975; Klecka 1980). Discriminant analysis begins by finding the linear combination of species-specific abundances that accounts for the greatest differences among assemblages; hence, it is the best discriminator for separating (characterizing) the assemblages. This is in contrast to PCA, which simply accounts for the greatest amount of variation in the data. The linear combination is called the first discriminant function, which is a linear regression with species abundances as predictors. The amount of variance among assemblages that is explained by the discriminant function is estimated as the eigenvalue.

Similar to PCA, discriminant analysis then finds a second discriminant function that accounts for the largest amount of the remaining differences among assemblages and is pairwise uncorrelated with the first function. This is the second discriminant function, which is the second best discriminator for separating

the assemblages. The process of fitting discriminant functions then continues with consecutive discriminant functions representing smaller and smaller differences among assemblages. The maximum number of discriminant functions is determined by number of assemblages and the number of species analyzed. If the number of assemblages (k) is fewer than the number of species, the maximum number of discriminant functions is $k - 1$; otherwise, it is equal to the number of species. The statistical significance of each discriminant function can be determined with a residual testing procedure (Stevens 1992). Statistically significant functions are generally retained for interpretation.

Discriminant functions can be interpreted by two methods. The first is to examine the standardized discriminant function coefficients, which are estimated by multiplying each raw coefficient by the standard deviation of the corresponding species abundance. The second is to examine the discriminant function–variable correlations, which are analogous to PCA loadings. For both methods, the species with the larger coefficients and correlations (absolute value) are considered to have the greatest influence on the function, but they occasionally provide conflicting results. For example, a species can have a high standardized coefficient and a low correlation (and vice versa) for the same function. This generally occurs when some species abundances are strongly correlated. Discriminant function coefficients are partial regression coefficients. That is, they are estimated after the effects of the other species have been removed. Hence, they tend to be influenced by intercorrelations among species. The discriminant function–variable correlation, however, is a more direct estimate of the relationship between a species and the function and is generally more stable when sample sizes are smaller (Stevens 1992). Therefore, we recommend use of discriminant function–variable correlations to interpret discriminant functions and use of the standardized coefficients to determine which variables are redundant (i.e., correlated).

Fish assemblage characteristics are interpreted by examining plots of discriminant scores. These scores are computed for each replicate sample by means of the discriminant function coefficients and are usually averaged for each assemblage. Discriminant function scores are generally plotted in two dimensions. The separation among assemblages along a discriminant function axis corresponds to the degree to which they differ on a particular function.

Assemblage predictability. Discriminant analysis can also be used to classify samples into one or more groups (e.g., assemblages) based upon species composition and abundances, and it can be used, via a V-fold cross-validation procedure, to assess the accuracy of assemblage structure classifications (Peterson and Rabeni 2001). In this procedure, samples are randomly placed into V groups, the samples from one group are excluded, and a model is fit with the data in the remaining $V - 1$ groups. The excluded group's samples are then classified using the discriminant model. This procedure is repeated for each group, and the proportion of misclassifications, among groups, is used to assess the predictability of the assemblage structure. A special case of cross-validation occurs when V equals the total sample size, which is called "leave-one-out" cross-validation (Lachenbruch 1975). Although cross-validation is a useful technique for examining the accuracy of as-

semblage structure classifications, it is noteworthy that high classification errors can also result from poorly fitting models due to factors such as nonnormal data. Hence, fisheries scientists should consider examining the error rate for various models to ensure that misclassification errors are the result of unpredictable fish assemblage structure rather than a poor-fitting model.

Discriminant analysis example. We demonstrate discriminant analysis by use of our example data set from the Kankakee River, Illinois, where six sites were each sampled eight times (see Box 15.1 for summed data). In total, 3,995 individuals and 34 species were collected during the survey. The 48 samples (6 sites × 8 samples) were used to assess differences among the assemblages at each station and to determine if assemblage structure was predictable. In Box 15.13, we present the SAS program used to perform the discriminant analysis of the fish assemblages and the associated output.

Discriminant analysis of those fish assemblages indicated that the first three of five functions were statistically significant and accounted for 93.3% of variance among assemblages. The minimal amount of variance explained by the remaining two functions suggested that they were redundant; consequently, they were dropped from the analysis. The first function discriminated among assemblages based on the abundance of longnose gar, bullhead minnow, largemouth bass, blackside darter, and golden redhorse and accounted for 53.2% of the variance. The second function accounted for 32.3% of the variation and discriminated among assemblages based on the abundance of logperch. The third function discriminated among assemblages based on the abundance of redfin shiner, gizzard shad, shorthead redhorse, silver redhorse, and smallmouth bass and accounted for 7.9% of the variation among assemblages.

Biplots indicated that assemblages at stations 2 and 5 differed from the others with higher densities of longnose gar, bullhead minnow, largemouth bass, and blackside darter and lower densities of golden redhorse (Figure 15.11). Assemblages at stations 3 and 4 also tended to have higher densities of logperch than did those at stations 1 and 6, whereas the assemblage at station 2 could be differentiated from that of station 5 by having higher densities of redfin shiner and lower densities of gizzard shad, shorthead redhorse, silver redhorse, and smallmouth bass. As expected, the relationships among assemblages indicated by the discriminant function biplots were virtually identical to those suggested by the principal component plots (Figure 15.8). This finding reflects the similarity in procedures used to calculate discriminant functions and principal components.

The leave-one-out cross-validation procedure indicated a poor overall classification error rate of 50%, which was lower than would be expected by random (83.3%). The greatest assemblage predictability was for stations 2 and 5, with classification error rates of 25%. The high classification error rates for assemblages at stations 1 and 4 (75%) suggested that their assemblages were relatively unpredictable. The fish assemblage samples from stations 3 and 4 were most often misclassified as one another, which suggested that they were the most similar assemblages.

Box 15.13 Discriminant Analysis

Program

```
OPTIONS PS = 60 LS=78;
DATA SPECIES;
INPUT STATION $ LOG GZS BLM BUM CAP HOC MIS RDS RYS SAS SFS STS SUM BLR GOR NHS SHR QLL RVR
SVR SAB BKS BLG GSF LMB LOS OSF ROB SMB BAD BLD JOD LOP SLD;
LINES;
(input data lines)
;
PROC DISCRIM DATA= SAMPLES NOCLASSIFY SHORT CANONICAL OUT = SCORE OUTSTAT= STATS;
CLASS STATION;
DATA STDCOEF; SET STATS; WHERE _TYPE_ = 'SCORE';
PROC TRANSPOSE DATA = STDCOEF OUT = STDCOEF; ID _NAME_;
PROC PRINT NOOBS;
TITLE '** STANDARDIZED CANONICAL DISCRIM FUNCTION COEFFICIENTS **';
DATA CORR; SET STATS; WHERE _TYPE_ = 'STRUCTUR';
PROC TRANSPOSE DATA = CORR OUT = CORR; ID _NAME_;
PROC PRINT NOOBS;
TITLE '** VARIABLE CANONICAL DISCRIM FUNCTION CORRELATIONS **';
PROC MEANS DATA = SCORE NOPRINT; BY STATION; VAR CAN1 CAN2 CAN3 CAN4 CAN5;
OUTPUT OUT = CANMEANS MEAN = CAN1 CAN2 CAN3 CAN4 CAN5;
PROC PRINT NOOBS;
TITLE '** MEAN CANONICAL SCORES **';
RUN;
TITLE;
QUIT;
```

Program Output

Table Eigenvalues for canonical discriminant functions based on fish abundance data from six stations of the Kankakee River, Illinois (Box 15.1). Eigenvalues estimate the amount of variance among assemblages that is explained by the discriminant function.

Function	Eigenvalue	Difference	Proportion	Cumulative
1	71.9663	28.2795	0.5315	0.5315
2	43.6868	33.0355	0.3227	0.8542
3	10.6513	5.3225	0.0787	0.9328
4	5.3288	1.5621	0.0394	0.9722
5	3.7667	0.0278	1.0000	

Table Residual test of discriminant functions testing the null hypothesis that the canonical discriminant functions in the current row and all that follow are 0.

Function	Likelihood ratio	Approximate *F*-value	Numerator *df*	Denominator *df*	*P* > *F*
1	0.00000087	4.60	170	49.838	<0.0001
2	0.00006367	3.32	132	42.446	<0.0001
3	0.00284502	2.14	96	33.825	0.0068
4	0.03314817	1.74	62	24	0.0672
5	0.20978877	1.63	30	13	0.1756

(Box continues)

Box 15.13 *(continued)*

Table Standardized canonical discriminant function coefficients (Can1–5), which are estimated by multiplying each raw coefficient by the standard deviation of the corresponding species abundance. Fish species abbreviations are given in Box 15.11.

Species	Can1	Can2	Can3	Can4	Can5
BAD	2.22457	–0.48972	0.64003	–0.79062	0.08280
BLG	–1.68944	–0.30942	–1.21861	1.09008	–0.81268
BLD	–0.98030	–0.68695	0.07823	0.57650	–0.92412
BLR	–1.17705	–0.17038	–0.79393	–0.72426	0.29653
BLM	–3.33386	–2.79803	–0.39095	–1.58555	0.49815
BKS	–1.49200	–1.21403	–0.76243	–0.05284	0.78281
BUM	–5.06659	2.34851	0.23292	1.26939	0.22113
CAP	–0.19464	–0.26258	0.75233	–0.29933	0.40622
GOR	2.07015	–1.73841	–0.17166	1.57722	–0.99228
GSF	–0.03456	–1.95688	–1.17854	–0.52994	–1.59839
GZS	3.41728	–2.95221	1.35361	–1.60174	–1.54082
HOC	1.53056	–1.61176	2.94328	–0.66275	0.21363
NHS	–4.99973	2.02565	–0.84181	–0.05209	1.14016
JOD	3.16295	–0.00467	0.71130	–0.66375	–0.43669
LOG	–0.85862	–1.16071	0.60182	–0.65685	0.30545
LMB	–4.06244	4.20083	–2.07917	0.50675	1.50131
LOP	1.89960	5.17398	0.23476	–0.36722	–1.63614
LOS	2.59106	–0.68042	–1.03481	1.71630	–0.57696
MIS	1.08908	1.69851	0.55843	1.01839	–2.34229
SHR	2.20270	–0.64959	0.60190	–0.76786	0.69662
OSF	–1.05600	0.03937	–0.54739	2.12078	–1.37153
QLL	–2.04450	0.26525	0.15850	1.02008	0.58319
RDS	1.05416	2.87170	–0.73884	–0.99056	1.55117
RVR	–2.45861	1.14824	–0.26552	0.00837	0.90398
ROB	–0.46791	–1.95681	–0.83695	–0.67286	0.56299
RYS	1.55350	–3.03223	–0.76696	2.59213	–1.64421
SAS	–4.88904	1.58556	0.67048	–1.08944	2.55940
SVR	–1.30226	0.15896	0.09574	–0.37498	–0.22452
SLD	0.45848	–2.07767	0.53864	–0.47588	–0.45735
SMB	4.38130	2.71061	2.16573	–0.02066	1.69134
SAB	–1.18708	–1.36536	–1.25310	0.41939	–0.07079
SFS	–2.65351	2.37492	0.55013	–0.41617	–0.50139
STS	1.88926	1.12735	0.66837	0.01674	0.07120
SUM	2.62921	0.93052	0.48263	0.23275	0.03670

Table Variable–canonical discriminant function correlations (Can1–5), or the discriminant function–variable correlations, which are analogous to PCA loadings. Fish species abbreviations are given in Box 15.11.

Species	Can1	Can2	Can3	Can4	Can5
BAD	0.23811	–0.14468	0.13118	–0.05207	–0.21766
BLG	0.11125	0.01246	0.14516	0.16810	–0.31712
BLD	–0.44303	–0.04790	–0.20957	–0.18865	–0.13599
BLR	0.00311	0.34322	0.03631	–0.14784	0.18479
BLM	–0.12122	–0.21289	0.30191	–0.33172	–0.02541
BKS	–0.13964	0.05261	–0.08348	0.04821	0.07028
BUM	–0.55716	0.14021	0.50691	0.21321	–0.00930
CAP	–0.32469	–0.00431	0.11495	–0.00423	–0.03166
GOR	0.61771	0.00697	0.26700	0.15737	0.17672
GSF	–0.39040	0.03887	–0.31092	–0.12480	–0.06229
GZS	–0.31260	–0.33744	0.45179	–0.11546	0.02747
HOC	–0.40317	0.08760	0.36608	0.09808	0.01048
NHS	0.28916	0.18871	–0.04219	0.04614	0.13469
JOD	–0.30252	0.14858	–0.26393	–0.22838	0.03947
LOG	–0.46128	–0.10398	0.40494	0.13456	–0.11298
LMB	–0.46566	0.16627	0.15313	–0.04932	0.04056
LOP	0.34406	0.54093	0.27623	0.12823	–0.09737
LOS	–0.10547	–0.04831	0.03102	0.09769	–0.13537
MIS	–0.06203	0.06283	–0.14418	–0.38620	–0.20545
SHR	0.14154	–0.16356	0.54160	–0.20291	0.00342
OSF	–0.42115	0.07686	0.46911	0.37731	–0.00603
QLL	–0.10668	–0.30767	0.40928	0.16711	0.24778
RDS	–0.22552	–0.07220	–0.46580	–0.24083	0.02768
RVR	0.24577	–0.12735	0.09575	–0.15821	0.24201
ROB	0.17531	–0.16398	0.45425	0.20168	0.20621
RYS	–0.11243	–0.30703	–0.25619	–0.09533	0.05399
SAS	–0.38583	–0.01101	–0.28355	–0.16861	–0.05566
SVR	–0.18831	–0.07547	0.52481	0.06790	–0.16480
SLD	0.36209	0.27481	0.06532	0.12263	–0.09039
SMB	0.34025	0.18790	0.41150	0.18409	0.23553
SAB	–0.40926	–0.03277	–0.03778	–0.04076	–0.06411
SFS	–0.33468	–0.20200	0.20987	–0.19623	–0.14918
STS	–0.07438	0.10649	0.13131	–0.18815	–0.16235
SUM	–0.34175	–0.01094	0.35967	0.10455	0.12402

(Box continues)

Box 15.13 *(continued)*

Table Mean canonical scores (Can1–5). Canonical scores are computed for each replicate sample by means of the discriminant function coefficients and are averaged for each station assemblage.

Station	Can1	Can2	Can3	Can4	Can5
1	7.78886	–3.07963	2.56286	–2.12270	–2.64169
2	–8.98018	–3.55705	–5.12470	–1.34585	–0.77444
3	4.46916	10.80398	–0.87505	–1.87546	1.61638
4	5.14086	2.54201	–1.64292	4.16695	–1.14496
5	–13.00349	2.19170	4.37312	0.79536	0.06922
6	4.58479	–8.90100	0.70670	0.38170	2.87549

Table Cross-validation summary to assess the accuracy of assemblage structure classification.

Station and summary	Station 1	2	3	4	5	6	Total
	Number of Observations and Percent Classified into Station						
1	2	0	1	4	0	1	8
	25.00	0.00	12.50	50.00	0.00	12.50	100.00
2	0	6	1	0	0	1	8
	0.00	75.00	12.50	0.00	0.00	12.50	100.00
3	0	1	3	4	0	0	8
	0.00	12.50	37.50	50.00	0.00	0.00	100.00
4	3	0	2	2	0	1	8
	37.50	0.00	25.00	25.00	0.00	12.50	100.00
5	0	2	0	0	6	0	8
	0.00	25.00	0.00	0.00	75.00	0.00	100.00
6	2	0	0	1	0	5	8
	25.00	0.00	0.00	12.50	0.00	62.50	100.00
Total observations	7	9	7	11	6	8	48
	Error Count Estimates for Station						
Rate	0.7500	0.2500	0.6250	0.7500	0.2500	0.3750	0.5000

15.4.5 Graphical Techniques

There are a number of graphical techniques available with statistical computer applications to plot multivariate data, which are useful to describe and compare visually fish assemblage compositions. Such graphical techniques are helpful in examining broad trends among samples, detecting relationships among assemblages, and identifying outlier data. A convenient and simple graphical technique of this type is the scatterplot matrix. This graphical matrix is a series of two-dimensional biplots of the species abundances comparing two assemblages or samples,

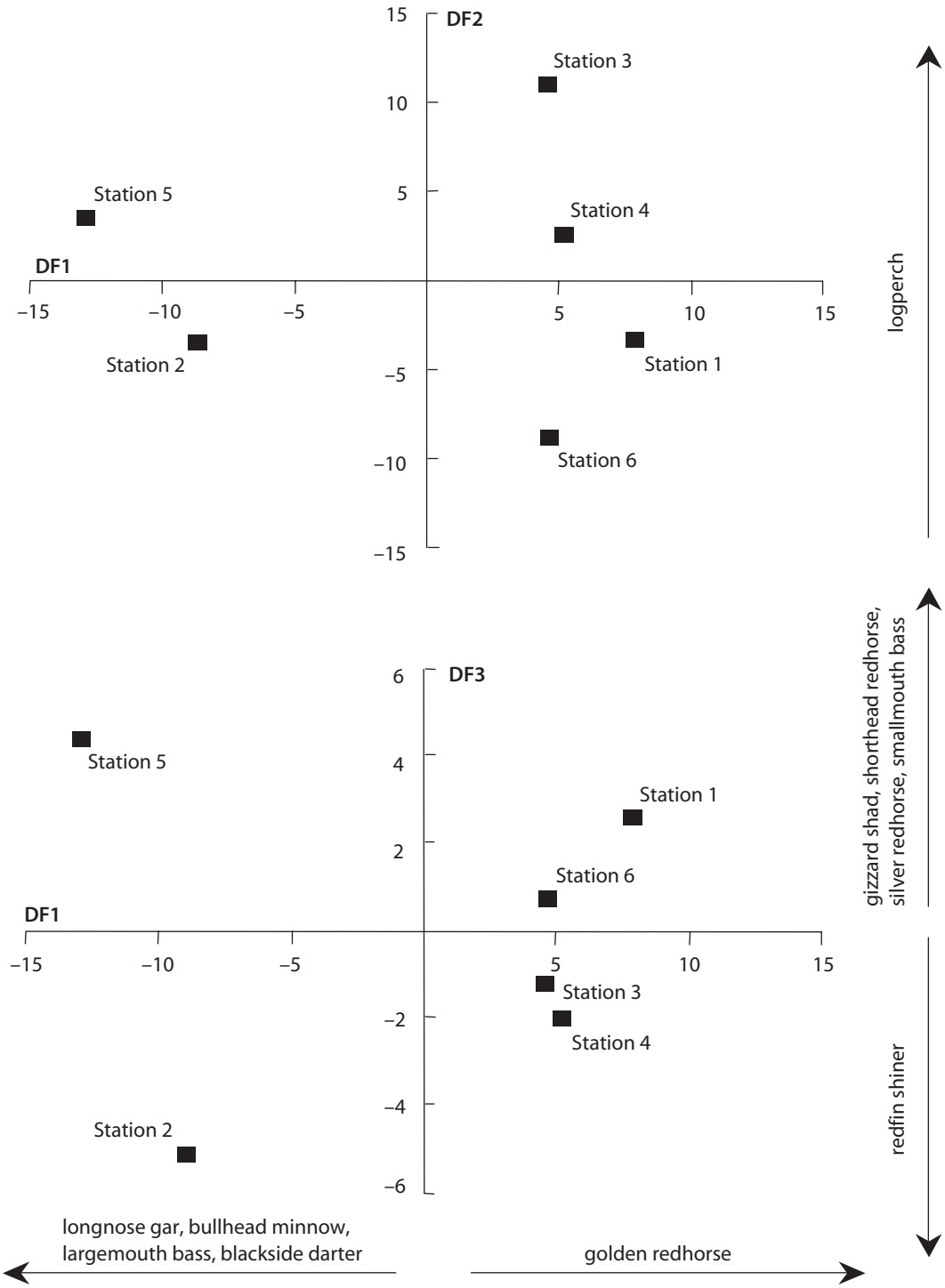


Figure 15.11 Discriminant function (DF) biplots for three functions of fish assemblages at six stations on the Kankakee River, Illinois (Boxes 15.1 and 15.13). Function interpretation and direction of positive influence (arrows) are shown at right and the bottom.

where each point on the plot represents the number of individuals (or density) of a species. When these biplots are fit with regression lines or ellipses and arranged in a matrix, trends in relationships among assemblages may be revealed and outlier species or assemblages can be identified.

Presented in Figure 15.12 are scatterplot matrices, including 95% bivariate normal density ellipses fit to each plot, that compare the six fish assemblages from the Kankakee River, Illinois example data (Box 15.1). Assemblage similarity is interpreted by examining the shape of the ellipse. Similar assemblages tend to have more linear, elongate ellipses, whereas a circular ellipse depicts dissimilar assemblages. An exploratory scatterplot matrix (Figure 15.12a) to identify similarities among assemblages can suggest a reordering of assemblage samples to more clearly reveal groups of similar assemblages (Figure 15.12b). This visual presentation suggests that assemblages at stations 1, 5, and 6 share a similar composition of fish abundances, as do the assemblages of stations 3 and 4, and that the station 2 assemblage is distinct from the others. This finding is virtually the same as that derived by multidimensional scaling, based on a percent similarity resemblance matrix (Figure 15.10c). Examination of individual points (one for each species) within the scatterplot matrix also indicates a number of outlier species that can explain differences among assemblages and may warrant additional analysis and interpretation.

Other graphical techniques that may be applicable to revealing attributes of, and relationships among, fish assemblages are Chernoff faces, star plots, sun-ray plots, and Andrews' plots (Johnson 1998). These techniques all share the properties that trends are relative, rather than absolute, and their detection depends upon the discerning eye and interpretive ability of the fisheries scientist; still, these techniques may reveal findings that could remain undisclosed by other more quantitative procedures.

15.5 SUMMARY

Several common themes emerge from this chapter that characterize a general approach and provide guidance toward the description and comparison of fish assemblages. (1) Usually, more than one quantitative approach or technique is available to describe or compare fish assemblages. (2) The most appropriate approach depends on scientific objectives and data form, quality, and quantity. (3) Comparison of results from more than one approach may be useful to elucidate trends and overcome bias or artifacts of any single technique. (4) Analytical techniques should be selected prior to analyses, based on objectives and application, rather than posthoc conformity to expectations. (5) Most results related to fish assemblages are relative values that are meaningful in a comparative context rather than in an absolute sense. In such applications, absolute probabilities (*P*-values) and statistical significance (α-levels), to which many fisheries scientists are accustomed, are less applicable, and reliance upon them may confuse interpretation. This observation, however, does not excuse the scientist from a quantitative approach; on the contrary, intensive data description, exploration, and comparison

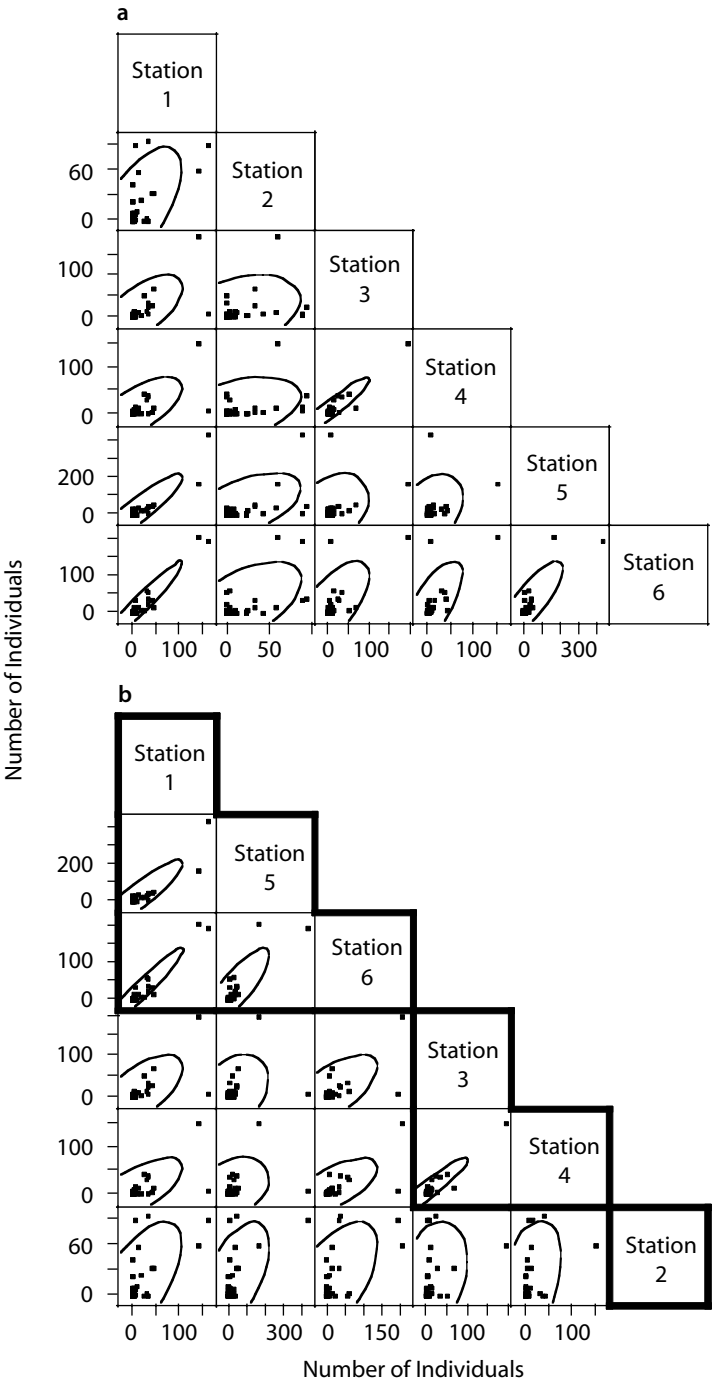


Figure 15.12 Scatterplot matrices of fish assemblages at six stations on the Kankakee River, Illinois (Box 15.1), with stations arranged by number (a) and according to similarity of assemblages (b). Curved-line enclosures represent the 95% bivariate normal density ellipse. Station plots enclosed by thick-lined boxes (b) depict similar fish assemblages.

are required to study fish at the community level. (6) Results of quantitative techniques are only valid if associated assumptions are not violated to a substantial degree. (7) The quality of results depends on quality of data. Discussion of data quality, logistic constraints, sampling bias and efficiency, analytical limitations, and other sampling and analytical concerns should not be avoided. Finally, the complexity of analyses at the community level precludes any strict protocol and allows for development of novel approaches that are limited only by the knowledge and creativity of the scientist; such quantitative methods and our understanding of them are likely to improve further in time.

15.6 REFERENCES

Agresti, A. 1990. Categorical data analysis. Wiley, New York.

Angermeier, P. L., and I. J. Schlosser. 1989. Species–area relationships for stream fish. Ecology 70:1450–1462.

Angermeier, P. L., and R. A. Smogor. 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distribution. Canadian Journal of Fisheries and Aquatic Sciences 52:936–949.

Angermeier, P. L., R. A. Smogor, and J. R. Stauffer. 2000. Regional frameworks and candidate metrics for assessing biotic integrity in mid-Atlantic highland streams. Transactions of the American Fisheries Society 129:962–981.

Angermeier, P. L., and M. R. Winston. 1999. Characterizing fish community diversity across Virginia landscapes: prerequisite for conservation. Ecological Applications 9:335–349.

Bailey, R. G. 1995. Description of the ecoregions of the United States, 2nd edition. U.S. Department of Agriculture, Forest Service, Miscellaneous Publication 1391, Washington, D.C.

Bain, M. B., J. T. Finn, and H. E. Booke. 1988. Streamflow regulation and fish community structure. Ecology 69:382–392.

Balon, E. K. 1975. Reproductive guilds of fishes: a proposal and definition. Journal of the Fisheries Research Board of Canada 32:821–864.

Bayley, P. B., and D. C. Dowling. 1990. Gear efficiency calibrations for stream and river sampling. Illinois Natural History Survey, Aquatic Ecology Technical Report 90/08, Champaign.

Bayley, P. B., and D. C. Dowling. 1993. The effects of habitat in biasing fish abundance and species richness estimates when using various sampling methods in streams. Polskie Archiwum Hydrobiologii 40:5–14.

Bayley, P. B., and L. L. Osborne. 1993. Natural rehabilitation of stream fish populations in an Illinois catchment. Freshwater Biology 29:295–300.

Bayley, P. B., and J. T. Peterson. 2001. Species presence for zero observations: an approach and an application to estimate probability of occurrence of fish species and species richness. Transactions of the American Fisheries Society 130:620–633.

Benke, A. C. 1990. A perspective on America's vanishing streams. Journal of the North American Benthological Society 9:77–88.

Berkman, H. E., and C. F. Rabeni. 1987. Effect of siltation on stream fishes. Environmental Biology of Fishes 18:285–294.

Blackburn, T. M., J. H. Lawton, and J. N. Perry. 1992. A method for estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. Oikos 65:107–112.

Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998. Estimating species richness: the importance of heterogeneity in species detectability. Ecology 79:1018–1028.

Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:326–349.

Breiman, L., J. H. Friedman, R. Olshen, and C. J. Stone. 1984. Classification and regression trees. Wadsworth International, Belmont, California.

Brown, M. L., and D. J. Austen. 1996. Data management and statistical techniques. Pages 17–62 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Cattell, R. B. 1966. The meaning and strategic use of factor analysis. Pages 174–243 *in* R. B. Cattell, editor. Handbook of multivariate experimental psychology. Rand McNally, Chicago.

Chandler, J. R. 1970. A biological approach to water quality management. Water Pollution Control 69:415–421.

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. Scandinavian Journal of Statistics 11:265–270.

Cliff, N. 1987. Analyzing multivariate data. Harcourt Brace Jovanovich, New York.

Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London B 345:101–118.

Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species area relationship. American Naturalist 113:791–833.

Crowder, L. B. 1990. Community ecology. Pages 609–632 *in* C. B. Schreck and P. B. Moyle, editors. Methods for fish biology. American Fisheries Society, Bethesda, Maryland.

Daniel, W. W. 1990. Applied nonparametric statistics, 2nd edition. PWS-Kent, Boston.

Davis, W. S. 1995. Biological assessment and criteria: building on the past. Pages 15–29 *in* W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Everitt, B. S. 1993. Cluster analysis, 3rd edition. Halsted Press, New York.

Farris, J. S. 1969. On the cophenetic correlation coefficient. Systematic Zoology 18: 279–285.

Fausch, K. D., C. L. Hawkes, and M. G. Parsons. 1988. Models that predict standing crop of stream fish from habitat variables: 1950–85. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, General Technical Report PNW-GTR-213, Portland, Oregon.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society 113:39–55.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. Pages 123–144 *in* S. M. Adams, editor. Biological indicators of stress in fish. American Fisheries Society, Symposium 8, Bethesda, Maryland.

Field, J. G. 1970. The use of numerical methods to determine benthic distribution patterns from dredgings in False Bay. Transactions of the Royal Society of South Africa 39:183–200.

Forbes, S. A. 1887. The lake as a microcosm. Bulletin of the Peoria Scientific Association 77–87. Reprinted, Bulletin of the Illinois State Natural History Survey 15(1925): 537–550.

Gammon, J. R. 1976. The fish populations of the middle 340 km of the Wabash River. Purdue University, Water Resources Research Center Technical Report 86, Lafayette, Indiana.

Gauch, H. G. 1982. Multivariate analysis in community ecology. Cambridge University Press, New York.

Gerking, S. D. 1994. Feeding ecology of fish. Academic Press, San Diego, California.

Glowacki, L., and T. Penczak. 2000. Impoundment impact on fish in the Warta River: species richness and sample size in the rarefaction method. Journal of Fish Biology 57:99–108.

Gordon, N. D., T. A. McMahon, and B. L. Finlayson. 1992. Stream hydrology: an introduction for ecologists. Wiley, West Sussex, UK.

Gorman, O. T., and J. R. Karr. 1978. Habitat structure and stream fish communities. Ecology 59:507–515.

Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution Press, Washington, D.C.

Gray, J. S. 1987. Species-abundance patterns. Pages 53–67 *in* J. H. R. Gee and P. S. Giller, editors. Organization of communities past and present. Blackwell Scientific Publications, Oxford, UK.

Green, P. E., and V. R. Rao. 1972. Applied multidimensional scaling. Holt, Rinehart, and Winson, New York.

Green, R. H., and G. L. Vascotto. 1978. A method for the analysis of environmental factors controlling patterns of species composition in aquatic communities. Water Research (Great Britain) 12:583–590.

Grossman, G. D., J. F. Dowd, and M. Crawford. 1990. Assemblage stability in stream fishes: a review. Environmental Management 14:661–671.

Grossman, G. D., and M. C. Freeman. 1987. Microhabitat use in a stream fish assemblage. Journal of Zoology 212:151–176.

Grossman, G. D., R. E. Ratajczak, M. Crawford, and M. C. Freeman. 1998. Assemblage organization in stream fishes: effects of environmental variation and interspecific interactions. Ecological Monographs 68:395–420.

Hakstian, A. R., W. T. Rogers, and R. B. Cattell. 1982. The behavior of number-of-factors rules with simulated data. Multivariate Behavioral Research 17:193–219.

Hall, C. A. 1972. Migration and metabolism in a temperate stream ecosystem. Ecology 53:586–604.

Halliwell, D. B., R. W. Langdon, R. A. Daniels, J. P. Kurtenbach, and R. A. Jacobson. 1999. Classification of freshwater fish species of the northeastern United States for use in the development of indices of biological integrity, with regional applications. Pages 301–337 *in* T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Hand, D. J. 1982. Kernel discriminant analysis. Research Studies Press, New York.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Hartigan, J. A. 1985. Statistical theory in clustering. Journal of Classification 2:63–76.

Hayes, D. B., C. P. Ferreri, and W. W. Taylor. 1996. Active fish capture methods. Pages 193–220 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Heck, K. L., Jr., G. van Belle, and D. Simberloff. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology 56:1459–1461.

Horn, H. S. 1966. Measurement of overlap in comparative ecological studies. American Naturalist 100:419–429.

Horton, R. E. 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Bulletin of the Geological Society of America 56:275–370.

Horwitz, R. J. 1978. Temporal variability patterns and the distributional patterns of stream fishes. Ecological Monographs 48:307–321.

Hubert, W. A. 1996. Passive capture techniques. Pages 157–192 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31–47 *in* W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55:1618–1631.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Hutchinson, G. E. 1978. An introduction to population ecology. Yale University Press, New Haven, Connecticut.

Hynes, H. B. N. 1970. The ecology of running waters. Liverpool University Press, Liverpool, UK.

Johnson, D. E. 1998. Applied multivariate methods for data analysts. Duxbury Press, Pacific Grove, California.

Johnson, D. H. 1999. The insignificance of statistical significance testing. Journal of Wildlife Management 63:763–772.

Kaiser, H. F. 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurement 20:141–151.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries (Bethesda) 6(6):21–27.

Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters. Island Press, Washington, D.C.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey Special Publication 5, Champaign.

Klecka, W. R. 1980. Discriminant analysis. Sage Publications, Beverly Hills, California.

Kolkwitz, R., and M. Marsson. 1908. Ökologie der pflanzlichen Saprobien. Berichte der Deutshcen Botanischen Gesellsschaft 26a:505–519. (Translated 1967. Ecology of plant saprobia. Pages 47–52 *in* L. E. Keup, W. M. Ingram, and K. M. Mackenthum, editors. Biology of water pollution. U.S. Department of Interior, Federal Water Pollution Control Administration, Washington, D.C.)

Krebs, C. J. 1998. Ecological methodology, 2nd edition. Benjamin/Cummings, Menlo Park, California.

Krueger, C. C., and D. J. Decker. 1999. The process of fisheries management. Pages 31–59 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Kruskal, J. B., and M. Wish. 1984. Multidimensional scaling. Sage Publications, London.

Kwak, T. J. 1993. The Kankakee River: a case study and management recommendations for a stream diverse in habitat, fauna, and human values. Pages 123–141 *in* L. W. Hesse, C. B. Stalnaker, N. G. Benson, and J. R. Zuboy, editors. Restoration planning for the rivers of the Mississippi River ecosystem. U.S. National Biological Survey Biological Report 19, Washington, D.C.

Lachenbruch, P. A. 1975. Discriminant analysis. Collier Macmillan, New York.

Leopold, L. B., M. G. Wolman, and J. P. Miller. 1964. Fluvial processes in geomorphology. Freeman, San Francisco.

Lyons, J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. North American Journal of Fisheries Management 12: 198–203.

Matthews, W. J. 1998. Patterns in freshwater fish ecology. Chapman and Hall, New York.

Miller, D. L., and 13 coauthors. 1988. Regional applications of an index of biotic integrity for use in water resource management. Fisheries (Bethesda) 13(5):12–20.

Milligan, G. W., and M. C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–179.

Minchin, P. R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69:89–107.

Morin, P. J. 1999. Community ecology. Blackwell Scientific Publications, Malden, Massachusetts.

Morisita, M. 1959. Measuring of interspecific association and similarity between communities. Memoirs of the Faculty of Science of Kyushu University Series E 3:65–80.

Mueller, D. W., and G. Sawitzki. 1991. Excess mass estimates and tests for multimodality. Journal of the American Statistical Association 86:738–746.

National Research Council. 1992. Restoration of aquatic ecosystems: science, technology, and public policy. National Academy Press, Washington, D.C.

Omernick, J. M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49–62 *in* W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Omernick, J. M., and R. G. Bailey. 1997. Distinguishing between watersheds and ecoregions. Journal of the American Water Resources Association 33:935–949.

Osborne, L. L., R. W. Davies, and K. L. Linton. 1980. Use of hierarchical diversity indices in lotic community analysis. Journal of Applied Ecology 17:567–580.

Paller, M. H. 1995. Relationships among number of fish species sampled, reach length surveyed, and sampling effort in South Carolina coastal plain streams. North American Journal of Fisheries Management 15:110–120.

Palmer, M. W. 1990. The estimation of species richness by extrapolation. Ecology 71:1195–1198.

Peet, R. K. 1974. The measurement of species diversity. Annual Review of Ecology and Systematics 5:285–307.

Peterson, J. T. 1989. Kankakee River fishes of the Braidwood Station Aquatic Monitoring Area, August 1988. Illinois Natural History Survey, Aquatic Biology Technical Report 89/1, Champaign.

Peterson, J. T., and C. F. Rabeni. 1995. Optimizing sampling effort for sampling warmwater stream fish communities. North American Journal of Fisheries Management 15: 528–541.

Peterson, J. T., and C. F. Rabeni. 1996. Natural thermal refugia for temperate warmwater stream fishes. North American Journal of Fisheries Management 16:738–746.

Peterson, J. T., and C. F. Rabeni. 2001. The relation of fish assemblages to channel units in an Ozark stream. Transactions of the American Fisheries Society 130:911–926.

Pielou, E. C. 1975. Ecological diversity. Wiley, New York.

Poff, N. L., and J. D. Allan. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. Ecology 76:606–627.

Quinn, J. W., and T. J. Kwak. 2003. Fish assemblage changes in an Ozark river after impoundment: a long-term perspective. Transactions of the American Fisheries Society 132:110–119.

Rabeni, C. F., and R. B. Jacobson. 1999. Warmwater streams. Pages 505–528 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Rahel, F. J., J. D. Lyons, and P. A. Cochran. 1984. Stochastic or deterministic regulation of assemblage structure? It may depend on how the assemblage is defined. American Naturalist 124:583–589.

Renkonen, O. 1938. Statistisch-ökologische Unter-suchungen über die terrestrische Kaferwelt der finnischen Bruchmoore. Annales Zoologici Societatis Zoologicae-Botanicae Fennica Vanamo 6:1–231.

Rexstad, E. A., D. R. Miller, C. R. Flather, E. M. Anderson, J. W. Hupp, and D. R. Anderson. 1988. Questionable multivariate statistical inference in wildlife habitat and community studies. Journal of Wildlife Management 52:794–798.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Rodgers, J. D., M. F. Solazzi, S. L. Johnson, and M. A. Buckman. 1992. Comparison of three techniques to estimate juvenile coho salmon populations in small streams. North American Journal of Fisheries Management 12:79–86.

Romesburg, H. C. 1990. Cluster analysis for researchers. Krieger Publishing, Malabar, Florida.

Root, R. B. 1967. The niche exploitation pattern of the blue-gray gnatcatcher. Ecological Monographs 37:317–350.

Rosenburg, D. M., and V. H. Resh, editors. 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Sanders, H. L. 1968. Marine benthic diversity: a comparative study. American Naturalist 102:243–282.

Sarle, W. S. 1983. Cubic clustering criterion. SAS Institute, SAS Technical Report A-108, Cary, North Carolina.

SAS Institute. 2004. SAS/STAT 9 user's guide. SAS Institute, Cary, North Carolina. (Also available as SAS OnlineDoc 9.1.3.)

Schlosser, I. J. 1982. Fish community structure and function along two habitat gradients in a headwater stream. Ecological Monographs 52:395–414.

Shannon, C. E., and W. Weaver. 1949. The mathematical theory of communication. University of Illinois Press, Urbana.

Shelford, V. E. 1929. Laboratory and field ecology: the responses of animals as indicators of correct working methods. Williams and Wilkins, Baltimore, Maryland.

Shepard, R. N. 1963. Analysis of proximities as a study of information processing in man. Human Factors 5:33–48.

Simberloff, D. S. 1972. Properties of the rarefaction diversity measurement. American Naturalist 106:414–418.

Simon, T. P., editor. 1999a. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Simon, T. P. 1999b. Introduction: biological integrity and use of ecological health concepts for application to water resource characterization. Pages 3–16 *in* T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Simon, T. P. 1999c. Assessment of Balon's reproductive guilds with application to midwestern North American freshwater fishes. Pages 97–121 *in* T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Simpson, E .H. 1949. Measurement of diversity. Nature (London) 163:688.

Southwood, T. R. E., and P. A. Henderson. 2000. Ecological methods, 3rd edition. Blackwell Scientific Publications, Oxford, UK.

Stevens, J. 1992. Applied multivariate statistics for the social sciences, 2nd edition. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. Transactions of the American Geophysical Union 38:913–920.

Suter, G. W., II. 1993. A critique of ecosystem health concepts and indexes. Environmental Toxicology and Chemistry 12:1533–1539.

Ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179.

Ter Braak, C. J. F., and I. C. Prentice. 1988. A theory of gradient analysis. Advances in Ecological Research 18:271–313.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130–137.

Van Sickel, J. 1997. Using mean similarity dendrograms to evaluate classifications. Journal of Agricultural, Biological, and Environmental Statistics 2:370–388.

Washington, H. G. 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. Water Research (Great Britain) 18:653–694.

Wiley, M. J., L. L. Osborne, and R. W. Larimore. 1990. Longitudinal structure of an agricultural prairie river system and its relationship to current stream ecosystem theory. Canadian Journal of Fisheries and Aquatic Sciences 47:373–384.

Winemiller, K. O., and K. A. Rose. 1992. Patterns in life history diversification in North American fishes: implications for population regulation. Canadian Journal of Fisheries and Aquatic Sciences 49:2196–2218.

Wolda, H. 1981. Similarity indices, sample size, and diversity. Oecologia 50:296–302.

Wright, S. J. 1988. Patterns of abundance and the form of the species-area relation. American Naturalist 131:401–411.

16 Predator–Prey Interactions

David A. Beauchamp, David H. Wahl, and Brett M. Johnson

16.1 INTRODUCTION

This chapter focuses on the analysis and interpretation of predator–prey interactions among vertebrates and invertebrates in freshwater systems associated with fisheries. All aquatic species are subject to predation during some phase of life (Mittlebach and Persson 1998). Predator–prey interactions play a major role in determining the structure and function of aquatic communities (Brooks and Dodson 1965; Carpenter et al. 1985; Kerfoot and Sih 1987; Northcote 1988) by influencing parameters such as survival, size structure, growth, behavior, and distribution, as well as biodiversity and water quality of these systems. These interactions are mediated by the physical–chemical environment (Kitchell 1979; Crowder et al. 1981; Coutant 1985; Gregory 1994) and habitat characteristics (Cooper and Crowder 1979; Wiley et al. 1984; Walters and Juanes 1993; Sogard 1994), which, in turn, are affected by human-induced alterations to the environment (Coutant et al. 1979; Jenkins 1979; Sandheinrich and Atchison 1990; Mesa 1994; Mesa et al. 1994).

Predation can regulate the dynamics of prey populations directly by reducing recruitment and survival (Miller et al. 1988; Luecke et al. 1990a; Tonn et al. 1992) or indirectly by altering prey behavior (Eggers 1978; Stein 1979; Clark and Levy 1988; Lima and Dill 1990), distribution, habitat choice, foraging, or growth (Dill and Fraser 1984; Werner and Gilliam 1984; Wurtsbaugh and Li 1985; Clark and Levy 1988; Jakobsen et al. 1988; Ibrahim and Huntingford 1989; Fraser and Gilliam 1992; Milinski 1993; Sogard 1994) or by altering competition and predator–prey interactions (e.g., Paine 1980; Werner et al. 1983; Mittlebach 1986, 1988; Persson 1991; Persson et al. 2000). Humans are very efficient aquatic predators: fishing can have large direct and indirect effects by selectively removing piscine predators, thereby altering food web structure and ecosystem function (He and Kitchell 1990; Schindler et al. 1998; Gislason and Sinclair 2000; Link and Garrison 2002). Prey can also influence predators as prey quantity and quality affect feeding rates, growth, and reproductive success of predators. Temporal and spatial changes in prey availability and vulnerability may influence movement and distribution patterns of predators.

Understanding the role of predation is important for successfully managing both self-sustaining and artificially enhanced fisheries, protecting water quality, conserving sensitive species, and maintaining the ecological integrity of aquatic communities. Because larval and juvenile fishes serve as both predators and prey, recruitment success is often related to the size-dependent ability to forage on available prey (Mittlebach 1981), temporal and spatial patterns in food supply, and interactions with predators and competitors (Dettmers et al. 1996; Mittlebach and Persson 1998; Persson et al. 2000). Introductions of nonnative predators and prey have resulted in extirpations, reduced biodiversity, and significant changes in community structure (Brooks and Dodson 1965; Zaret and Paine 1973; Nesler and Bergersen 1991; Schindler et al. 1998). Managers should consider whether host waters can absorb the additional consumption demand by stocked fishes or if predation losses will undermine the goals of a stocking program. Overfishing predators or enhancing prey populations can disrupt predation rates and result in depensatory declines in predatory fish populations (Johnson and Goettl 1999; Walters and Kitchell 2001; Post et al. 2002). Trophic interactions can also affect water quality. When linkages among multiple trophic levels are strong, changes in abundance at one level can result in a trophic cascade through some or all trophic levels from apex predators to primary producers, thus affecting water transparency (Carpenter et al. 1985; Carpenter and Kitchell 1988; Luecke et al. 1990b; Brett and Goldman 1996). However, the magnitude and effects of predation vary under different environmental and ecological conditions (Abbey and MacKay 1991; Persson et al. 1991; Beauchamp et al. 1999).

In this chapter, we outline methods for analyzing predator–prey interactions at multiple levels: from mechanistic to holistic studies and from examining behavior of individual predators or prey to quantifying predator–prey interactions and determining their effects on the structure and function of freshwater populations, communities, and ecosystems (Table 16.1). Topics are organized into sections on field, experimental, and modeling approaches with regard to spatial, temporal, and size relationships, which commonly underlie predator–prey interactions and are important for minimizing interpolation–extrapolation error in the analysis. Other authors provide descriptive, empirical, and theoretical reviews of predator–prey interactions and the effects of predation on inland aquatic systems (Stroud and Clepper 1979; Zaret 1980; Kerfoot and Sih 1987; Carpenter 1988; Carpenter and Kitchell 1993).

16.2 PREDATOR–PREY QUESTIONS: OBJECTIVES AND APPROACHES

The complexity of predator–prey interactions often requires a complementary, multi-pronged approach that uses some combination of field observation and measurement, experimentation or management manipulation (e.g., stocking, introductions, harvest, or removal), and modeling. The design and analysis of a predator–prey study should be guided by specific information needs, and the study objectives and approaches should follow logically from these questions. Common questions in freshwater fisheries vary in scale and complexity from the behavior

Table 16.1 A spectrum of perspectives for studying fish predation varying from the mechanistic, organismal level (top of table) to the more holistic, ecosystem level (bottom of table). Selecting the appropriate approaches and analytical tools will depend on the research question and the ecological phenomena of interest. Note that uncertainty in measurements may be lower at the mechanistic level but transferability to dynamics in nature may be questionable. Studying more holistic phenomena may rely on integrating results from a variety of measurements and approaches, with a concomitant increase in uncertainty as errors may be compounded. Abbreviations are as follows: analysis of variance (ANOVA); multivariate analysis of variance (MANOVA); available prey to predator ratios (AP/P); forage species biomass to consumer biomass ratio (F/C); and proportional stock density (PSD)

Research question	Approaches	Analytical tools	Examples
Predator and prey behavior	Laboratory trials	ANOVA, MANOVA, multiple regression, path analysis, and ethograms	Wahl and Stein 1988; Hambright 1991; Christensen 1996; Einfalt and Wahl 1997
Predator preferences (e.g., species or size)	Laboratory trials	Electivity indices Diet composition	Hambright 1991; Einfalt and Wahl 1997
	Field measurements	Electivity indices Stable isotope analysis Optimal foraging	Pyke 1984; Mittlebach and Osenberg 1994; Vander Zanden et al. 2000
Predation rate versus prey abundance or availability	Laboratory trials	Functional responses—linear, nonlinear, and multivariate regression	Koski and Johnson 2002
"Balance" between prey and predator populations	Field measurements	Indices: AP/P, F/C, and PSD	Swingle 1950; Anderson and Weithman 1978; Ploskey and Jenkins 1982
Consumption in terms of prey biomass, nutrients, dollars, or contaminant fluxes	Field-based estimates	Fullness-gastric evacuation methods	Elliott and Persson 1978; Ney 1990
	Simulation modeling	Production-based models Bioenergetics models	Ney 1990 Kitchell et al. 1977; Ney 1993
	Stable isotopes or tracers	Mass-balance models	Trudel et al. 2000
In situ foraging behavior	Field observations and simulation modeling	Spatially explicit models	Brandt et al. 1992; Stockwell and Johnson 1997; Luo et al. 2001
		Bioenergetics models	Stewart et al. 1981; Baldwin et al. 2000; Burke and Rice 2002
		Visual foraging models	Wright and O'Brien 1984; Beauchamp et al. 1999
Ecosystem structure and function: for example, the role of fishing, trophic cascades, key stones, and cultivation effects	Natural and planned ecosystem "experiments" and simulation modeling	Bioenergetics models	Kitchell 1992; Schindler et al. 1998
		Multispecies virtual population analysis	Pope 1991
		Ecopath/Ecosim	Walters and Kitchell 2001; Walters et al. 1997

and interactions of individual predators or prey to community or ecosystem level dynamics (Table 16.1). Does predation occur? How much predation occurs? Can predation regulate specific prey populations? What factors contribute to the timing, duration, and magnitude of predation? Can the effects of predation be minimized or maximized? More specific questions may follow from the primary questions above. Which predators (species and size) eat which prey (again, by species and size); when is predation occurring (timing and duration of predation), and where is it occurring (is it basinwide or in specific or isolated habitats)? What are the appropriate spatial, temporal, and body size scales that relate to specific predator–prey interactions? What characteristics of predators, prey, or their habitat foster or inhibit predation? Do predators feed selectively on specific prey or feed opportunistically on a variety of prey in proportion to their abundance? Can prey supply support a desirable size structure and abundance of predators?

The nature of each question and desired level of resolution determine the types of studies (e.g., field sampling, experiments, and modeling) and corresponding analyses that should be employed. Effective studies require tight integration of sampling and experimental design tailored to the study objectives with explicit definition of independent and response variables, sampling or experimental units, and appropriate measurement units (e.g., numbers, biomass, nutrients, dollars, or contaminant loading). The desired level of resolution should be specified: presence versus absence of predation; a quantified estimate of predation (in biomass or numbers of prey consumed); predation translated into a mortality rate for prey; predator–prey "balance"; or the strength of different factors contributing to predator–prey interactions. Here, we briefly discuss common questions related to predator–prey interactions and introduce some alternative approaches to address them.

Is there evidence of predation (presence versus absence)? This question can be addressed by laboratory experiments (e.g., will a predator eat this prey under experimental conditions?), an approach that is particularly useful in situations with novel predators or prey, such as invasions (e.g., Nesler and Bergersen 1991; Moyle and Light 1996a, b). Presence of predation may also be addressed by sampling stomach contents of potential predators at times and locations where predation most likely occurs (will the predator eat this prey under natural conditions?); however, stomach contents provide a short-term snapshot of feeding by some individuals and may not be representative of feeding by the predator population over longer periods. Alternatively, stable isotopes (Peterson and Fry 1987) and growth provide broader-scale integration of the feeding history of a consumer but with less temporal resolution. A predation signal may be inferred from stable isotope analysis (e.g., $\delta^{15}N$), especially in relatively simple aquatic communities where the prey of interest could produce a unique isotopic signature (Vander Zanden et al. 2000; Johnson et al. 2002).

Under what conditions will predation occur? Experiments can examine specific factors that influence predation such as predator–prey size relationships (Juanes 1994), prey density (Peterman and Gatto 1978; Koski and Johnson 2002), habitat elements (Savino and Stein 1982), alternative prey, and environmental conditions; however, spatial scale, oversimplification, or experimental artifacts may constrain

the generality of these results (Huston 1999). Field data on diet and distribution can identify potentially important influences on predation if stratified by size-class of predators, at appropriate temporal and spatial scales, and accompanied by data on ambient environmental conditions. Such comprehensive studies, however, are logistically challenging. Inferences may suffer from small sample sizes in many temporal–spatial–size cells. Experimental and field data can be combined to construct foraging models (e.g., functional response models [Peterman and Gatto 1978; Eby et al. 1995], encounter rate models [Gerritsen and Strickler 1977; Beauchamp et al. 1999]; spatially explicit growth models (Brandt et al. 1992); or optimal foraging models [Mittlebach and Osenberg 1994]).

How much predation occurs? Estimates of per capita consumption rates in the wild can be computed from diel gut fullness and evacuation rates (Eggers 1977; Elliott and Persson 1978) or by using mass-balance and energy budget models (Forseth et al. 1992; Hanson et al. 1997; Trudel et al. 2000; Chapter 12). These estimates require considerable investment in the collection of data on temporal and size-specific diet composition, the thermal experience of the consumers, and for bioenergetics models, incremental growth by the consumer. Estimates of the size structure and relative or absolute abundance of predators are required to expand estimates of individual consumption to population level predation rates.

What fraction of the prey population is lost to predation (predatory impact)? This analysis compares quantitative estimates of predation losses to estimates of prey abundance. When dealing with size-structured populations over extended periods, these comparisons would be stratified by time interval and by size-classes of predators and prey. Spatial stratification might also be required if predator–prey interactions differ significantly among locations. If the predation period is long in comparison to the growth rate or reproductive cycle of prey, then predation rates should be compared with the production rate of prey rather than just to prey biomass.

Can prey populations support the desired growth and production of predators? If predator–prey relationships can be quantified, are appropriately sized prey abundant enough, at times and locations where predators can encounter and successfully consume them, to satisfy target or observed growth rates for a given density of predators? This requires the ability to translate prey biomass, size structure, abundance, and distribution into the fraction of the prey population that can be consumed through functional response curves (Holling 1966; Peterman and Gatto 1978), encounter rate models (Gerritsen and Strickler 1977), and capture success models. Alternatively, simple empirical relationships such as proportional stock densities (PSD) of predators and prey (Gabelhouse 1984), available prey to predator ratios (AP:P; Jenkins and Morias 1978), mass-balance models (production: biomass; Ney 1990), and production-conversion efficiency relationships (Eck and Brown 1985) provide holistic estimates of the prey supply or balance required to support predator populations. Ecopath and Ecosim models (Christensen and Pauly 1993, 1994, 2001) provide estimates of biomass transferred among species or functional groups at each trophic level in an ecosystem. Most of these approaches estimate trophic rates on an annual time step for whole populations or feeding

guilds but require considerably less data than interactions that can be modeled at finer temporal, spatial, taxonomic, or ontogenetic scales with bioenergetics models or field-based estimates.

16.3 CONCEPTUAL FRAMEWORK FOR ADDRESSING PREDATOR–PREY INTERACTIONS

Predator–prey interactions occur only under conditions that allow detection and successful capture of prey. It is important to recognize and define the segments of predator and prey populations included in a particular type of analysis. The predation sequence (search, encounter–detection, attack, and capture) provides a useful framework for organizing questions about predator–prey interactions and identifying the most appropriate or feasible methods (laboratory or field experiments, field measurements, or modeling) for addressing these questions (Holling 1966; Box 16.1).

Whether dealing with individuals or larger segments of predator and prey populations, the predation sequence is related to the abundance, availability, and vulnerability of prey (Box 16.2). A prey population exists at some abundance in an aquatic system, but only a fraction of that abundance may be available to predators due to incomplete temporal and spatial overlap between prey and predators. Where, when, and how predators search will determine what fraction of the prey population is available for encounter. Encounters depend on how prevailing environmental and habitat conditions affect detection of prey that overlap with predators in time and space (e.g., Beauchamp et al. 1999). Of the available fraction, a smaller proportion of prey may be vulnerable to consumption due to predator avoidance behavior (e.g., refuging), size constraints, and evasion capabilities that reduce capture success once prey are encountered. Given an encounter, prey vulnerability to capture depends on morphological and behavioral characteristics of both the prey and predator, and the probability of an attack is influenced by how inclusion of that prey in the diet would affect the rate of net energy gain to the predator.

Diet composition patterns and prey electivities from field samples subsume the combined effects of the prey abundance–availability–vulnerability hierarchy. In contrast, laboratory experiments can explore how individual factors contribute to variability in predator–prey interactions under different conditions. Models provide a conceptual framework for incorporating field data, experimental results, and theory into a mechanistic simulation of predator–prey responses to different conditions.

Predator–prey investigations vary from studies on the behavior of individual predators or prey to community and ecosystem level responses, patterns, and dynamics, which may require one or a combination of field, experimental, or modeling approaches (Table 16.1). As analyses progress from individual to systemwide responses, the potential for propagation of error increases tremendously. Therefore, identification of key factors that contribute to the variability in predator–prey responses and appropriate incorporation of these factors into experimental

Box 16.1 Elements of the Predation Sequence

Behavioral elements of the predation sequence (Holling 1966) are described below (see figure), followed by a listing of common or feasible approaches (lab or field experiments, field measurements, or modeling) for studying these topics.

Prey Search

Search. Aquatic predator–prey interactions occur in three-dimensional space and can be complex due to highly mobile predators and prey utilizing different habitats or locations over a variety of spatial and temporal scales. Predators employ either active or stationary (sit-and-wait) foraging modes (Norberg 1977; Bell 1990; O'Brien et al. 1990). The temporal distribution and movement patterns of predators and their prey reflect these foraging strategies.

Temporal and spatial overlap of predators and prey can be inferred from distribution patterns in catch per unit effort (*C*/*f*) data when sampling each with comparable methods and by assuming similar catchability among sizes, species, times, or locations (see Chapters 3, 7, and 8). Overlap may be determined more directly with biotelemetry (Chapter 14), hydroacoustics, and active net sampling in pelagic environments (Brandt 1996). At the microhabitat scale, experimental (e.g., Eklov and Hamrin 1989; Eklov and Diehl 1994; Christensen 1996), video, or acoustic surveillance methods (e.g., Boisclair 1992; Collins et al. 1991; Collins and Hinch 1993) can be effective for determining search patterns in natural habitats or experimental arenas.

The search strategy (i.e., periodicity, location, search mode, or search image) may be specialized for a specific habitat, prey type, or combination and based on the mobility of predator or prey, patch dynamics, antipredator responses, such as schooling or shoaling, diel migration, or sheltering patterns.

Approaches for studying this element are lab and field experiments, modeling, and field measurements.

Encounter. Predator–prey encounters are a function of the sensory mechanisms used to detect food or threats (e.g., visual [Wright and O'Brien 1984; Henderson and Northcote 1985; Hughes and Dill 1990; Breck 1993; Beauchamp et al. 1999], chemical [Atema 1980], or pressure and tactile [Gerritsen and Strickler 1977; Janssen 1997] detection fields). Therefore the number of encounters, at least in open-water environments, depends on the temporal distribution patterns of prey and predators and the area or volume searched times the density of prey contained within or passing through that search volume.

Approaches for studying this element are field measurements with modeling and lab experiments.

Predator Response to Encountered Prey

Optimal foraging rules (Pyke 1984; Stephens and Krebs 1986) for maximizing net energy gain per unit time provide a useful framework for predicting the predator's response after encountering a specific prey organism. Prey selectivity indices provide empirically derived measures of the proportional contributions of various prey to the diet in relation to the proportions of these prey that were available in natural or experimental environments; however, field estimates of the overall abundance or density of prey has generally been used as a crude substitute for the amount of prey actually available to, or encountered by, the predators.

(Box continues)

Box 16.1 *(continued)*

Orient or reject. Upon encounter, predators either orient toward prey and continue the predation sequence or reject the prey and resume searching or other activity.

The approach for studying this element is lab experiments.

Follow. The predator maneuvers toward prey to maintain close proximity and seek opportunity to attack. The follow segment can range from being quite prolonged to an abrupt transition from orienting to pursuing and striking prey.

The approach for studying this element is lab experiments.

Pursue. Predator accelerates toward prey to close distance in preparation for striking.

The approach for studying this element is lab experiments.

Strike. Predator attempts to grasp, injure, or stun prey. A strike can result in a hit (capture) or miss. If missed, search resumes for another prey.

The approach for studying this element is lab experiments.

Capture. Once captured, a predator could either ingest the prey or expel it and resume searching.

The approach for studying this element is lab experiments.

Ingest. The probability of successfully capturing and consuming prey given an attack is dependent on morphological and behavioral traits of both prey and predator. The net energy gained by ingesting each prey depends on its mass, energy density (J/g), and the time and energy required to capture and handle that prey item.

The approach for studying this element is lab experiments.

Digest. Digestion rate will determine the maximum volume of food that can be consumed per unit time and can depend on temperature and prey size and type.

Approaches for studying this element are lab experiments and models.

designs, stratification schemes in field sampling, and architecture and scaling of models is critically important for minimizing estimation error.

16.4 STUDY DESIGN AND ANALYSIS OF FIELD DATA

Field observations might provide the first evidence of predation. Common objectives of a field data collection program include recording the timing, duration, and spatial extent of predator–prey interactions to bound interactions in space and time; describing the size structure of the predator and prey populations (Chapter 9) and size relationships of predator–prey interactions; estimating the abundance of predators and prey (Chapters 7 and 8); and collecting data on food habits (Chapter 11) of potential predators to determine the existence or relative magnitude of predation.

Resume search. During the last step in the predation sequence, predators either continue searching for prey or initiate other activity such as vigilance, refuge-seeking, territorial defense, thermoregulation, migration, courtship, or spawning.

Approaches for studying this element are lab experiments and field measurements (telemetry).

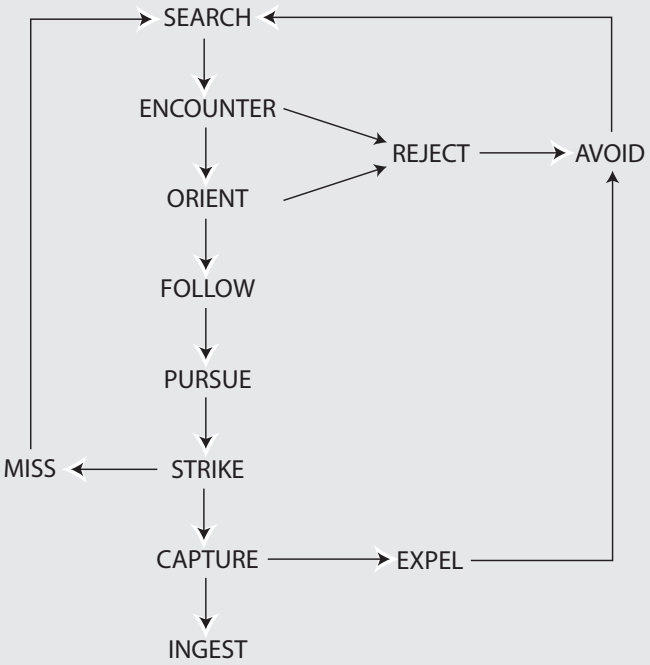


Figure Behavioral pathways from search through ingestion for a fish predator foraging on prey.

Because predation generally varies over time, with increasing body size (especially piscivory), and across some spatial dimension, these factors should be integrated explicitly into the sampling design. A stratified sampling design (Chapter 3) could be used to determine the temporal, spatial, and size-related dimensions of the predator–prey interactions. For instance, to evaluate predation on forage fishes in a lake or reservoir, the objectives of the sampling design would include (1) determining the seasonal distribution of predators (e.g., from catch per unit effort [C/f] data) among habitats and regions of the water body (e.g., benthic, pelagic, and littoral) and in relation to their prey; (2) determining the size structure of the predator and prey populations over the size range of interest; (3) determining the proportional weight contribution of forage fishes and other prey to the diet of different sizes of predators during each season; (4) and perhaps determining the abundance of predators and the abundance or availability of

Box 16.2 Hierarchical Measures of Prey Accessibility to Predators

Abundance
Interactions may not involve all members of the predator and prey populations. Consequently, abundance may not be the most ecologically relevant metric in predator–prey interactions. Population estimates or indices of relative abundance can be generated using standard methods, including mark–recapture, area or volume swept, depletion estimates, virtual population analysis (see Ricker 1975; Van Den Avyle 1993; Chapters 7–9), and hydroacoustics (Brandt 1996).

Availability
Temporal and spatial overlap allows some probability of encounter between predators and prey. Segregation may result from stage-specific distribution patterns related to environmental requirements or behavioral modification (e.g., foraging mode or predator avoidance). Ambient environmental conditions or habitat characteristics can mediate predator–prey encounters and the nature of their interactions. Data on seasonal or diel movement and distribution patterns, the effects of environmental conditions on physiology, detection capabilities of predators and prey, behavioral plasticity of predators and prey, and habitat complexity, gained through methods of capture, observation, hydroacoustics, and telemetry, give insight on availability.

Vulnerability
Only a fraction of the available prey is actually vulnerable to predators. Many factors reduce the probability of prey capture given an encounter, such as size–gape relationships, behavioral avoidance, evasion, vigilance, effects of habitat characteristics, temperature, turbidity, light, prey density, and predation rates or handling times as functions of predator–prey size relationships. Statistical analyses primarily involve analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), and linear and nonlinear regressions.

prey. Generally, samples should be collected in a spatially representative manner; whether this requires formal stratification by region, habitat type, or depth will depend on characteristics of the water body, the species involved, and the specific study question. Fish should be captured with methods adapted to collect a representative sample of the population's size structure within each season and habitat type. The sampling design should allocate all data and specimen collection methods among representative spatial units, and this procedure should be repeated seasonally or at the temporal scale that is relevant to the study question. Species residing primarily in littoral or lotic habitats can be collected with active-capture methods (e.g., beach seining, electrofishing, pop nets, angling, or toxicants) or passive-capture methods (gill nets, trap nets, traps, or set lines) appropriate for the habitat characteristics and behavior and size range of the target species. Pelagic species can also be collected via active capture (surface, midwater, or bottom trawling; purse seining; or trolling) or passive capture (gill netting, set lines, or traps). This core sampling effort supplies the data for analyzing spatial–temporal distribution patterns and population size structure and provides some or all of the specimens needed for diet analysis or for determining age and growth. Additional collections might be required to satisfy sample size requirements for size-specific, seasonal diet analysis, but data from these supplementary samples would

not necessarily be included in the analysis of size structure or spatial–temporal distribution patterns unless the additional sampling effort associated with these samples could be accounted for in calculations of C/f in an unbiased manner.

16.4.1 Distribution, Size Structure, and Abundance of Predator and Prey Populations

The distribution, size structure, and abundance of populations define boundaries on predator–prey interactions. Similar data collection methods are often employed for all three types of analyses. Distribution patterns dictate whether overlap in time and space provides the opportunity for predators and prey to interact and thus defines the availability of prey to predators. The size structure of the prey population determines the fraction of prey that is vulnerable to size-selective predators. The abundance and size structure of predators determine the magnitude of predation losses imposed on prey populations and how predation losses may be distributed among age- or size-classes of prey. If prey abundance is known, then population level predation rates can be converted into predation-specific mortality rates, and the relative importance of predation to the overall mortality of prey can be evaluated (e.g., Jones et al. 1993; Cartwright et al. 1998; Baldwin et al. 2000). Abundance estimates of predators or prey commonly represent the largest source of uncertainty in most analyses of population-level consumption.

Temporal distribution patterns vary considerably among life stages and species. These patterns must be identified and incorporated into the design of any population or community assessment program. Abundance estimates should explicitly define the target sizes or ages of the population and any restrictions imposed on the estimation procedure by sampling limitations. Different capture methods impose biases due to size or species selectivity and differential effectiveness among habitats (Chapters 3 and 9). Sampling with complementary methods can relieve some of these biases (Figure 16.1), and correction factors can be used in the analysis phase (e.g., size selectivity in gill nets: Rudstam et al. 1984; Van Den Avyle 1993). Although catch rates will not be directly comparable among sampling methods, the spatial–temporal patterns of maximum densities of each species could be inferred from the location and timing of peak C/f within each sampling method. It is important to acknowledge possible sampling biases and evaluate their potential effects on analyses either verbally or more formally through sensitivity analysis.

Population assessment techniques must be tailored for specific taxa and habitats. Population estimates from combined hydroacoustic–midwater trawl surveys have become routine for pelagic freshwater species like juvenile sockeye salmon and kokanee (Burczynski and Johnson 1986; Parkinson et al. 1994; Beauchamp et al. 1997), trouts (Stables and Thomas 1992; Yule 2000), ciscoes and whitefishes (Brandt et al. 1991; Luecke and Wurtsbaugh 1993), smelts (Burczynski et al. 1987; Appenzeller and Leggett 1995), and shads (Schael et al. 1995). Population estimates require intensive effort for littoral species (cove-rotenone treatments, depletion estimates, mark–recapture estimates, and relative abundance expansions) and river and stream fishes (e.g., depletion or mark–recapture estimators). Population

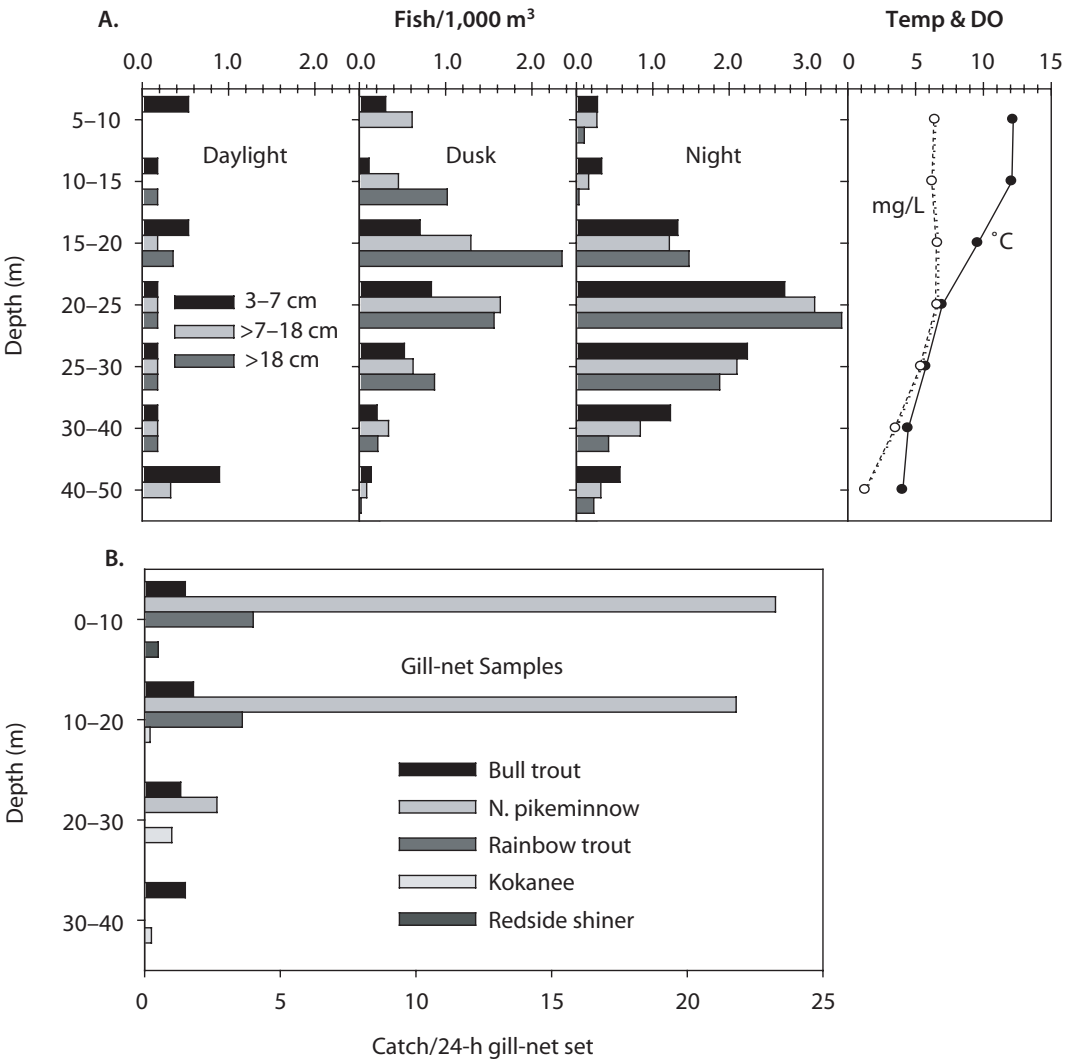


Figure 16.1 Depth distribution and relative abundance of potential predators and prey in Alturas Lake, Idaho (Beauchamp et al. 1997). Data were obtained with complementary sampling methods. (A) Hydroacoustics data provided volumetric densities for three size-classes of fish by depth and diel period. In right-most panel are temperature (temp, °C) and dissolved oxygen (DO, mg/L). Fish occupying the bottom and perimeter slope zones of the lake during daylight could not be detected until they dispersed into limnetic regions during dusk and night. Midwater trawling indicated that the pelagic targets at all depths were predominantly kokanee. Older age-classes of kokanee (in and below the thermocline) and redside shiners (above the thermocline) were vulnerable to sinking gill nets but not to trawls. (B) Depth-specific catches obtained with sinking gill nets indicated that high densities of potentially piscivorous northern pikeminnow (N. pikeminnow) and rainbow trout were primarily distributed in the upper 20 m along the slope zone, whereas predatory bull trout used all depths of the slope zone but at lower densities. Better information on movement and distribution patterns of the piscivores would require ultrasonic telemetry or more intensive hydroacoustic sampling and analysis.

assessments for deepwater demersal species or pelagic piscivore populations in freshwater systems may require a wide range of methods to estimate relative or absolute abundance of various species (e.g., hydroacoustics, volume- or area-swept approaches, distance sampling techniques, cove-rotenone treatments, depletion estimates, mark–recapture estimate, virtual population analysis, relative abundance from *C/f*, and others; Van Den Avyle 1993; Chapter 8).

16.4.2 Food Habits of Predators

A comprehensive treatment of food habits analysis is provided in Chapter 11, so only those aspects particularly relevant to analysis of predator–prey interactions are presented here. Predator–prey studies generally focus on particular prey species and perhaps some important alternative prey (e.g., a species that might buffer the effects of predation on the focal prey species). Beyond these species, additional prey items may be treated collectively as just "other food" for simplicity. Consequently, many analyses can focus on just the prey species of interest when examining the effects of different factors (independent variables) on the response variable (e.g., prey size or proportion of the focal prey in the diet). Some notable exceptions to this simplification scheme include electivity indices and optimal foraging models (Mittlebach and Osenberg 1994) because the full suite of prey species or groups is required for computations using these methods.

16.4.2.1 Explanatory Variables for Diet Analysis

Changes in feeding behavior and diet composition of predators can often be explained by factors such as body size, time (diel period, season, or year), and space (habitats, depths, regions, or geographic areas) and may vary in response to changes in availability or vulnerability of prey. At a minimum, predators should be sampled to detect seasonal changes in the proportion of the focal prey in the diet for each size-class of consumer. Spatial factors may influence diet composition of predators, with differences found between pools and riffles or nearshore and offshore zones (Beauchamp 1990; Beauchamp et al. 1992; Schindler and Scheuerell 2002), among depths (Beauchamp 1994; Stockwell and Johnson 1997), between macrophytes and open water areas (Savino and Stein 1982; Mittlebach 1984, 1988; Werner and Hall 1988; Persson and Eklov 1994), and, during prey migrations, in areas near tributaries or dams versus other zones (Poe et al. 1991; Rieman et al. 1991; Winemiller and Jepsen 1998). Conversely, the combined effects of predator movement and prey dispersal can homogenize diets from different vertical or horizontal regions of the basin (Cartwright et al. 1998; Baldwin et al. 2000, 2002).

16.4.2.2 Categorization of Continuous Data

Although diet data should always be initially examined in raw form (e.g., body length or sampling date), response variables often show similar patterns across a range of values for the explanatory variables. In these circumstances, different ranges of the explanatory variable could be grouped into categories to facilitate further analysis. Common examples include grouping continuous explanatory

variables, like predator body lengths, into discrete size categories or pooling sample collection dates into broader, ecologically relevant periods. However, pooling continuous variables into categories might not always be appropriate. The challenge is to minimize the number of categories without pooling important sources of variability. Post hoc exploration of the data might be necessary to determine the appropriate number of categories and their boundaries. For instance, size categories of predators could be determined by inspecting scattergrams of predator lengths (continuous data) and the percentages of key prey in the diet by weight to identify the size-classes corresponding with no predation, increasing predation, and one or more plateaus of high(er) predation that indicate seasonal or ontogenetic shifts in diet composition (Figure 16.2).

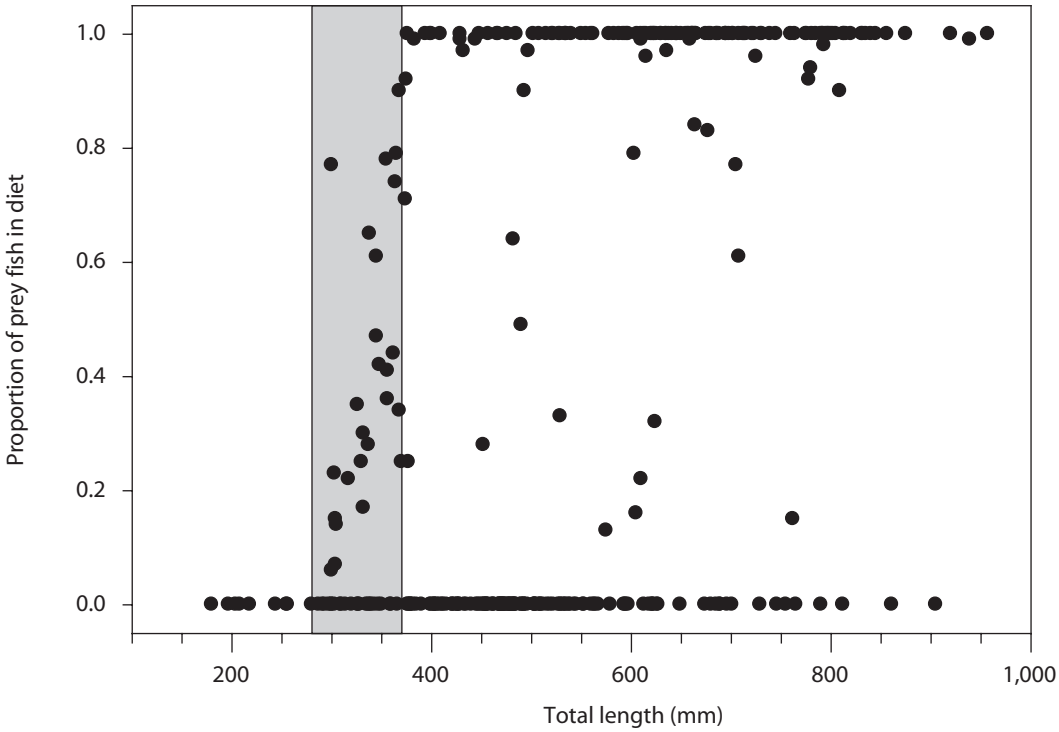


Figure 16.2 A scattergram of the proportion (by weight) of fish prey in stomachs versus predator length. The scattergram of 327 nonempty stomachs of lake trout, sampled across all seasons in Flathead Lake, Montana (D. Beauchamp, unpublished data), was used to stratify length data into size categories that minimized diet variation within size-classes by isolating most of the variability in a transition size-class (300–375 mm in total length, shaded area). Lake trout greater than 375 mm were considered fully piscivorous, with fish prey contributing up to 100% of the stomach contents. Note that fish prey were absent in some stomachs from all sizes of predators.

16.4.2.3 Determination of Diet Composition

The stomach contents of aquatic predators may contain prey that vary over one to three orders of magnitude in length and six orders of magnitude in mass (e.g., zooplankton to prey fishes); therefore, the variables chosen to characterize a predator–prey interaction will depend upon study objectives. Mean weight proportions (termed MW_i in Chapter 11; see Table 11.1 for descriptions of common diet indices) of prey in the diet are often the most useful response variable in predator–prey interaction studies. Weight proportion reflects the relative importance of each prey species to the energy budget of the consumer and can be used to estimate the loss of prey (in biomass or numbers) from predation by using bioenergetics models or other estimators of consumption (Ney 1990; Bowen 1996; Chapter 11). Counts or numerical percentages of prey in the diet are useful when evaluating prey encounter or consumption rates, as in optimal foraging models (e.g., Mittlebach and Osenberg 1994), and contribute to calculations of prey electivity. Frequency of occurrence O_i (the percentage of predators in the sample that contain a specific prey species) provides useful supplementary information about the fraction of predators in a sample that contain specific prey but should rarely be used as the primary response variable when analyzing predator–prey interactions.

When estimating the weight contribution of various prey to the diet, data should be recorded as dry weight, blotted-dry wet weight, or volumetric proportions for individual stomach samples. Samples should be collected within a sampling design that provides spatially representative samples of each size-class of consumer for each ecologically relevant period of the study. Each nonempty stomach is considered an individual sampling unit wherein the observed prey proportions are associated with the size of the consumer, the location and time of capture, and concurrent physical, chemical, and biotic conditions. For each prey category, the weight proportion, W_i, from each stomach is averaged with all other nonempty stomachs in the sample (MW_i) within each size-class of predator during each period of interest (Box 16.3). This analytical approach gives equal weight to the dietary proportions from each nonempty stomach, regardless of the level of stomach fullness. This minimizes the influence of rare stomachs that contain large quantities of a particular prey because each stomach can represent only 100% of one stomach out of N samples and attempts to reflect the average diet composition of all consumers within the same size × time cell. As with other methods, this approach has some limitations but produces less volatile estimates of diet composition.

Important underlying assumptions can be evaluated directly through pilot or supplemental studies. For instance, one can account for the assumption of similar digestion rates for all major prey taxa and test directly for diel differences in prey composition. Differential digestion among prey taxa becomes a greater concern if temperatures enter the warmer portion of a consumer's thermal range and if the prey of interest are small, soft-bodied organisms (e.g., larval fishes), which digest rapidly. This concern can be minimized by preliminary diel sampling to identify the timing of peak stomach fullness (e.g., dawn, midday, before dusk,

Box 16.3 Analysis of Diet Composition of Lake Trout by Season and Size-Class

The objective of this diet study was to estimate the proportional weight contribution of each prey type in the diet of lake trout by season and size-class in Flathead Lake (D. Beauchamp, unpublished data). To simplify the example, prey categories are reduced to just fish and invertebrates. Spatially representative samples of each size-class of lake trout were collected during winter, spring, summer, and fall from overnight sets of sinking experimental variable-mesh gill nets. Each season, two sampling locations were selected randomly in each of five regions and four depth intervals of the lake. Each lake trout stomach was considered a sampling unit, and its diet was associated with the total body length, weight, location, depth, date, and season of capture recorded on the same row. Lengths were categorized into length-classes based in part on visual inspection of the scattergram of length versus fish proportions in Figure 16.2. This abbreviated data set illustrates a useful format for analyzing diet composition data in spreadsheets or statistical packages. Other columns can be added for additional information (e.g., habitat, gear type, and diel period). Diet data are recorded first as the mass (g) of each prey category measured directly from the stomach contents (in columns Fish and Invert). Each prey category is then converted to a proportion of the total prey found in each stomach (in columns FishP and InvertP), computed as the weight of each prey category divided by the total weight of all prey from that stomach. Most analyses are conducted on prey proportions (e.g., diet composition by length-class and season). Note that prey mass and proportions are left blank for empty stomachs, as in fish 14.

Table Abbreviated data set of diet composition (fish versus invertebrate) of lake trout. The blotted wet mass of each prey category is measured for individual fish stomachs (fish number) and each prey category is then converted to a proportion of the total prey found in each stomach (in columns FishP and InvertP). Season abbreviations throughout this chapter are spring (spr), summer (sum), autumn (aut), and winter (win).

Fish number	Date	Season	Total length (mm)	Length-class (mm)	Fish weight (g)	Prey weight (g)			Diet proportions	
						Fish	Invert	Total	FishP	IInvertP
1	6/8/98	Spr	374	301–375	400	1.08	0.45	1.53	0.71	0.29
2	6/8/98	Spr	453	376–500	730	0.00	1.67	1.67	0.00	1.00
3	6/9/98	Spr	301	301–375	180	0.10	0.03	0.13	0.77	0.23
4	6/9/98	Spr	403	376–500	440	0.00	0.16	0.16	0.00	1.00
5	6/9/98	Spr	622	501–625	1,990	0.00	0.10	0.10	0.00	1.00
6	6/9/98	Spr	813	626–1,000	4,830	6.67	0.00	6.67	1.00	0.00
7	6/9/98	Spr	479	376–500	830	0.00	12.59	12.59	0.00	1.00
8	6/9/98	Spr	615	501–625	1,910	0.91	0.04	0.95	0.96	0.04
9	6/9/98	Spr	675	626–1,000	3,020	0.08	0.00	0.08	1.00	0.00
10	6/9/98	Spr	664	626–1,000	2,310	0.67	0.13	0.80	0.84	0.16
11	6/9/98	Spr	705	626–1,000	3,090	9.38	2.81	12.19	0.77	0.23
12	6/9/98	Spr	745	626–1,000	3,245	5.74	0.00	5.74	1.00	0.00
13	6/9/98	Spr	575	501–625	1,400	0.86	5.78	6.64	0.13	0.87
14	6/9/98	Spr	330	301–375	285					
15	6/15/98	Spr	293	100–300	210	0.00	0.10	0.10	0.00	1.00
16	6/15/98	Spr	287	100–300	195	0.00	0.28	0.28	0.00	1.00
17	6/15/98	Spr	332	301–375	280	0.14	0.33	0.47	0.30	0.70
18	6/15/98	Spr	379	376–500	445	0.00	0.33	0.33	0.00	1.00
19	6/15/98	Spr	539	501–625	995	0.00	0.44	0.44	0.00	1.00
20	6/15/98	Spr	845	626–1,000	4,540	7.15	0.00	7.15	1.00	0.00
…										
609	8/24/01	Sum	435	376–500		0.00	0.50	0.50	0.00	1.00

The data above represent a segment of a larger data set for predators during winter and spring. The sample sizes, mean proportions by weight (termed MW_i in Table 11.1), and 2SE for fish and invertebrates consumed by each length-class × season combination were summarized below in a Microsoft Excel spreadsheet by means of the "pivot table" analysis tool; the same summarization could also be produced using analogous tools in other spreadsheet software, PROC MEANS in SAS, or "multi-dimensional pivot tables" in SPSS.

Table Summary of lake trout diet composition data. The mean proportion by weight (*N* = number of nonempty stomachs) and 2SE for both food categories are given by length-class × season.

		Diet proportions			
Length-class (mm) and season	*N*	Fish	2SE	Inverts	2SE
100–300					
Winter	0				
Spring	8	0.00	0.00	1.00	0.00
Summer	15	0.01	0.06	0.99	0.06
Fall	12	0.00	0.00	1.00	0.00
301–375					
Winter	5	0.00	0.00	1.00	0.00
Spring	23	0.37	0.21	0.63	0.21
Summer	10	0.11	0.21	0.89	0.21
Fall	28	0.15	0.14	0.85	0.14
376–500					
Winter	25	0.21	0.17	0.79	0.17
Spring	67	0.20	0.10	0.80	0.10
Summer	20	0.10	0.14	0.90	0.14
Fall	27	0.30	0.18	0.70	0.18
501–625					
Winter	34	0.47	0.17	0.53	0.17
Spring	42	0.37	0.15	0.63	0.15
Summer	19	0.17	0.18	0.83	0.18
Fall	26	0.35	0.19	0.65	0.19
626–1,000					
Winter	26	0.83	0.15	0.17	0.15
Spring	77	0.79	0.09	0.21	0.09
Summer	13	0.74	0.25	0.26	0.25
Fall	20	0.90	0.14	0.10	0.14

after dusk, or night) and the variability in diet composition among diel periods (Figure 16.3). Net retrieval or active sampling should then be scheduled to maximize the number of samples captured during or slightly after the period of peak stomach fullness. If diet composition varies dramatically between the period with peak fullness and other periods, then stomach samples may be needed from two to three times per day to portray diet composition (or to determine feeding chronology and estimate daily consumption, see section 16.4.2.5); alternatively, prey reconstruction methods (Swenson and Smith 1973; Diana 1979) might be required to obtain good diet composition estimates.

16.4.2.4 Statistical Analysis of Diet Composition

Descriptive statistics for dietary responses by predators can be summarized effectively in spreadsheets (e.g., pivot tables in Microsoft Excel or analogous data summary tables in other spreadsheet programs, such as PROC MEANS in SAS [SAS

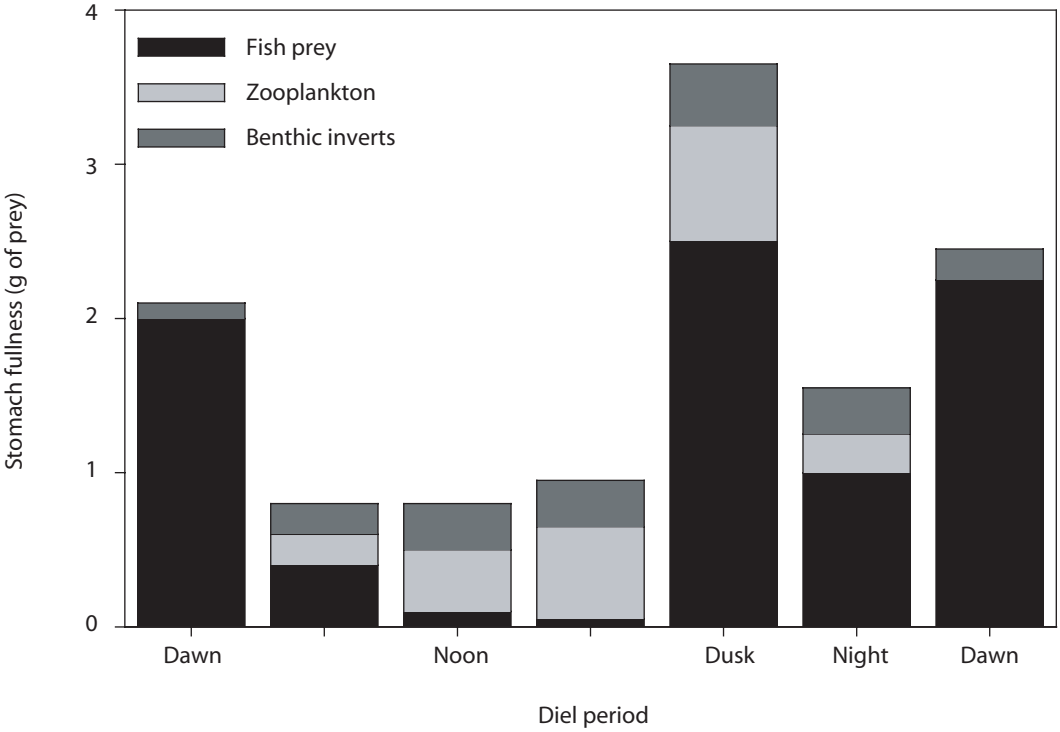


Figure 16.3 Diet composition and stomach fullness can change dramatically over a 24-h period. In this example (D. Beauchamp, unpublished data from Flathead Lake), the dusk samples captured peak stomach fullness for all three prey categories and provided the most representative diet samples for the predator. Fish prey would have been severely underrepresented if stomach contents were sampled only during mid-day, whereas zooplankton would have been seriously underrepresented in dawn samples. This potential diel variability should be considered when designing a sampling program and analyzing diet samples of predators with regard to the limitations of sampling methods and study design.

Institute 1990] or multidimensional pivot tables in SPSS [SPSS 1999]) to produce a table of averages, standard deviations, maximum, minimum, or median values, and sample sizes for response variables for any combination of independent variables or factors (e.g., Box 16.3). Careful design and formatting of the data set and careful selection of response and predictor variables can greatly facilitate these analyses.

Because proportions of key prey species in diet composition generally vary in response to more than one predictor variable, statistical methods should be capable of simultaneously examining the effects of several factors on the average proportion of key prey species in the diet of predators. Changes in proportions of key prey with size of the predator can be examined by regression, whereas the effects of seasons, size-classes of predators, and habitat groups can be analyzed effectively using various analysis of variance (ANOVA) techniques (see Chapter 11). The ANOVA techniques are relatively robust to moderate deviations from normality when sample sizes are adequate because of the central tendencies of the data. For instance, in diets of lake trout from several western lakes, mean proportions of the major prey types tended to stabilize at sample sizes of 7–15 nonempty stomachs per season × size-class cell (D. Beauchamp, unpublished data). The diet proportions of focal prey can also be transformed (e.g., by square root or arcsine transformations for weight proportions) to achieve or approach normality in the distribution of means (or slope coefficients for regressions). Alternatively, various ANOVA techniques can be applied to rank-transformed weight proportions for each prey species of interest. The single-factor version of this nonparametric method is the Kruskal–Wallis test, but this test becomes awkward and ineffective when more than one factor is involved. Multifactor ANOVAs on rank-transformed data offer less statistical power than do parametric ANOVAs of similar design but can be helpful for determining the relative importance of the main effects of different predictor variables, such as size-class, time, habitat type, or region, on the proportion of focal prey in the diet of predators. Although rank-transformed ANOVAs identify significant main effects, and levels of these effects can be compared with multiple- range tests (Conover and Iman 1981), a major disadvantage of rank-transformed ANOVAs is that significant interaction terms cannot be interpreted as in a parametric ANOVA (Hora and Conover 1984; Thompson 1991). If significant interactions exist, the relative importance of prey in response to different levels of a factor must be examined separately within each level of the other factor involved in the interaction. Despite these limitations, rank-transformed ANOVAs provide a useful framework for organizing and prioritizing how results are reported.

16.4.2.5 Field-Based Estimates of Consumption

A daily consumption rate for the average predator can be estimated from field data by using methods appropriate to the feeding chronology of the consumer, such as stomach fullness–gut evacuation rate methods and prey reconstruction (see and Ney 1990 for reviews). The most common approach involves serial sampling of stomach contents over a 24-h period. A temperature-dependent stomach

evacuation rate, R, is applied to the mean mass of food in the stomach, S, over the 24-h period to obtain a daily consumption estimate, C_d, for the average predator.

$$C_d = 24 \cdot S \cdot R, \tag{16.1}$$

where the stomach evacuation rate, R (h^{-1}), is the proportion of food digested per hour. It is obtained as the slope of the proportion of food (W_t/W_0) remaining t hours after feeding and is generally expressed as a decaying exponential function of time t (Eggers 1977; Elliott and Persson 1978):

$$W_t/W_0 = a \cdot e^{-Rt}. \tag{16.2}$$

The intercept, a, should theoretically equal 1.0, and the slope R(h^{-1}) represents the evacuation rate. The evacuation rate generally increases with increasing temperature and can also vary considerably among prey types or sizes (He and Wurtsbaugh 1993). Daily ration size should be calculated separately for different size-classes of predators or if diet composition varies considerably among size-classes. Daily consumption can be partitioned among prey types based on the proportional contribution by weight of each prey category to the diet. Multiplying daily consumption estimates of the focal prey by the number of predators in each appropriate size-class and the duration of the interaction in days generates the biomass of prey consumed by each size-class of predator over a given period.

For top predators, adequate sample sizes may be difficult to achieve for each size × time cell. Field-generated estimates of daily consumption reflect the ambient environmental conditions immediately preceding and during the period of sample collection, and estimates can vary considerably from day to day (Smagula and Adelman 1982) and seasonally (Cochran and Adelman 1982). Consequently, a field-generated estimate of consumption may have limited generality to broader periods of interest longer than 1 d.

16.4.2.6 Special Considerations for Analyzing Short-Term Acute Predation

The temporal scale of predator–prey interactions can be extremely important when attempting to quantify consumption in response to large pulses of prey (e.g., recruitment pulses, stocking, or migration of prey). Sampling should be scheduled to measure the magnitude and duration of the predatory response accurately. In systems where we have examined predation on stocked juvenile walleye, pikes, and trouts and salmons, the response has been immediate, severe, and of relatively short duration (e.g., 2 d to 2 months: Wahl et al. 1995; Cartwright et al. 1998; Baldwin et al. 2000).

Samples collected prior to stocking provide baseline data on the diet composition of predators. Stomach samples should be collected for at least three consecutive days during and after stocking, then once every 2–3 d, followed by a progressively lower frequency of sampling (Figure 16.4). The reduction in sampling frequency can be adjusted by evaluating how quickly the proportional

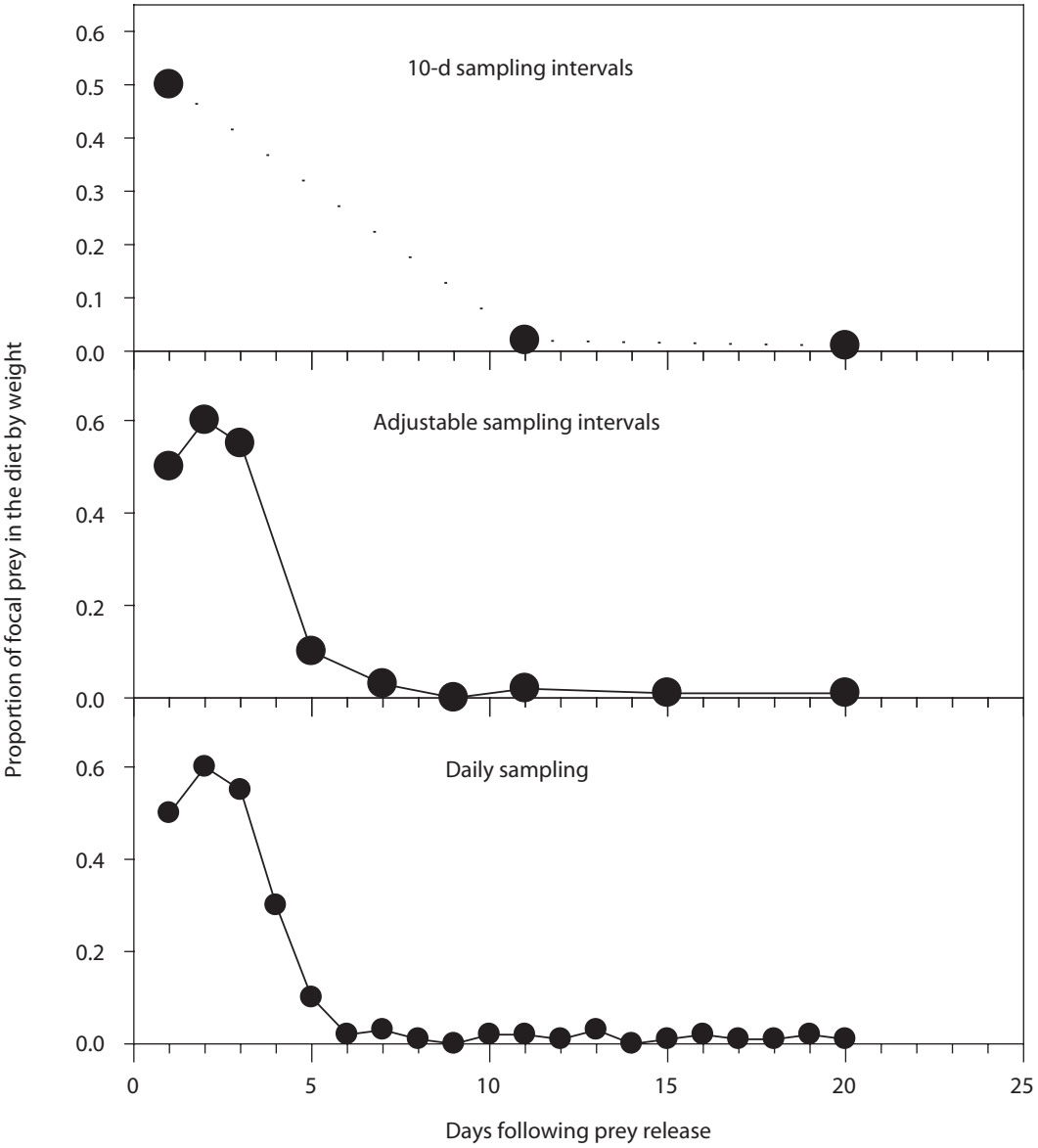


Figure 16.4 Sampling frequency can affect the amount of interpolation error in estimates of predation when acute periods of interaction are potentially short lived. In this example, a 10-d sampling interval (upper panel) missed the rapid decline of focal prey in the diet and results in an overestimate of the duration and magnitude of the predation response. The adjustable sampling frequency scheme (middle panel) captured the initial decline by sampling on three consecutive days before sampling effort was progressively reduced as the proportion of prey in the diet stabilized through time. The adjustable sampling frequency design minimized interpolation error and required less than half the effort as that of a daily sampling regimen (lower panel).

contribution of the focal prey, by weight or volume of the stomach contents, changes in the diet between sampling dates. If the prey disappears quickly from the diet (e.g., within a couple days), then sampling can be reduced to measure monthly or seasonal changes as appropriate. If the diet composition changes measurably, but at a more moderate rate, then a relatively frequent sampling schedule should be maintained (e.g., sample once every 3–7 d) until the prey of interest disappears or stabilizes at a relatively constant percentage of the diet through time. The rate of change of focal prey in the diet can be calculated by regressing its weight percentage in the diet against the number of days after the pulse of prey appeared.

The spatial extent of predator–prey interactions during acute predation periods is important as well because we need to know what fraction of the predator population is involved in order to estimate the total magnitude of predation losses. If prey are concentrated in a localized area (e.g., stocking locations or location where larval or juvenile fishes are migrating into lakes or reservoirs from streams or spawning areas), predators may concentrate in response to an influx of prey. Alternatively, predation may involve only those predators that happen to occupy that region. Studies should be designed to detect whether the relative density of predators (e.g., measured by C/f) or their dietary response (measured by the proportional weight contribution of focal prey in the diet) changes between areas of high prey concentrations and other regions. Sampling along a distance gradient from the point of prey entry as well as before and after prey entry (a BACI design, that is, before–after-control-impact design) addresses this issue.

16.5 ANALYSIS OF DATA FROM BEHAVIORAL EXPERIMENTS AND RELATED MODEL DEVELOPMENT

Several aspects of predator–prey interactions of fishes can be assessed under controlled experimental conditions to help understand foraging ecology in natural systems and answer fisheries management questions. A number of systems have been used for these types of experiments, varying from aquaria to large pools and tanks; these should be scaled to the size of the fish of interest. Experimental results can be used to construct models that simulate processes that underlie predator–prey interactions. Several techniques can be used to record behavior, including direct observation in real time (with entry directly to a portable computer) and videography.

16.5.1 Predation Sequence

The generalized predation sequence described in section 16.3 can be documented with several nonoverlapping categories (Wahl and Stein 1988; Box 16.1): inactive (resting and motionless), search (moving, not orienting to prey), observe (motionless but oriented to an individual prey), follow (moving and orienting to prey), pursue (following at burst speed), attack (striking at prey), and capture (grasping prey). Because these behaviors may be correlated, multivariate analysis of variance

(MANOVA) is the preferred analysis approach for these data (Box 16.4). Behavioral probabilities can be calculated at each step in the sequence (e.g., orient–follow–pursue or strike–capture). Handling time can be defined as the time from prey capture until search is resumed (in the vein of Werner and Hall 1974; Mayer and Wahl 1997) or as the time required to ingest prey (Holling 1966).

Important differences exist among fish functional groups in components of the predation sequence. Because planktivores ingest many small prey with relatively low mobilities, they have higher capture success (usually >80%; Confer and Blades 1975; Mayer and Wahl 1997) than do piscivores (usually <70%; Wahl and Stein 1988; Einfalt and Wahl 1997). Similarly, planktivorous prey are small and relatively defenseless, so prey are pursued, captured, and ingested within a second or less (Mittelbach 1981; Koski and Johnson 2002) whereas handling times for individual prey by piscivores are considerably higher (3–6 min for pikes, Wahl and Stein 1988; 2–10 min for walleye, Einfalt and Wahl 1997). An interesting exception is found for bluegill foraging on an exotic cladoceran *(Daphnia lumholtzi)*. Because of the large helmet and tail of this cladoceran, bluegill face foraging constraints more similar to piscivores (Kolar and Wahl 1998). In general, because of the high capture success and short handling times of planktivores, search time is of much greater importance to a particulate feeder than is capture success (Juanes 1994) or handling time. For benthivores, or drift-feeders (e.g., trouts and salmons in streams), feeding territories are common in many species, suggesting that prey are patchily distributed and search costs are high; thus consumers might shift costs from search to territory defense.

16.5.2 Optimal Foraging

Animals forage by selecting prey that maximize the rate of energy intake while minimizing costs associated with searching and handling prey (Stephens and Krebs 1986). Optimal foraging models are often employed to predict diet selection for given characteristics of the predator and prey and can successfully predict resource use (Mittelbach 1981; Mittelbach and Osenberg 1994). To increase fitness, predators choose prey that minimize energy spent on search, capture, and handling while maximizing energy intake. Energy return (J/min) can be calculated as

$$\frac{E_n}{T} = \frac{\sum_{i=1}^{n} \lambda_i E_i - C_s}{1 + \sum_{i=1}^{n} \lambda_i T_{hi} P_i}, \tag{16.3}$$

for each prey type i out of n available prey types. The parameter λ_i is the number of encounters with prey i during a feeding trial (sensu Charnov 1976); E_i is the expected energy gain (J) per individual prey item; P_i is the probability of a capture occurring after an encounter; C_s is the energy cost of searching; and T_{hi} is the handling time for each prey type i. Prey types are added sequentially until E/T is maximized (Pyke 1984). In this and many formulations of optimal foraging

Box 16.4 Multivariate Analysis of Variance for Predator–Prey Behavioral Data

The multivariate analysis of variance (MANOVA) is appropriate for cases in which several dependent variables have been measured in a single experiment. This technique is more appropriate than performing multiple univariate tests, which can increase the probability of a type I error (Scheiner 1993). As a result, this test is generally appropriate for analysis of multiple behaviors in the predation sequence (Box 16.1). Below we use a SAS program to perform a MANOVA in an example concerning the effects of predator size and prey density on behaviors in the predation sequence.

In our hypothetical example, we examine the effect of walleye length, prey density, and their interaction on individual walleye feeding behaviors in aquaria experiments (similar to T. Galarowicz, Central Michigan University, and D. Wahl, unpublished data). Foraging was examined for three different sizes of walleye (20, 50, and 100 mm) feeding at three different densities (1, 10, and 25 bluegill per m^3) of optimal-sized bluegill prey. Five replications at each size and density combination were performed with five individual walleye. The total number of searches, orientations, follows, pursuits, strikes, and captures in the predation sequence (Box 16.1) were recorded for each trial.

We first perform a MANOVA. If significant multivariate effects are found (known as a protected ANOVA), we then examine the univariate responses (ANOVAs) for significant effects. The ANOVAs that are not significant are deleted, and the remaining ANOVAs are run again. In the MANOVA, a test for significant differences among groups is based on eigenvectors (linear combinations of all dependent variables) and eigenvalues of the matrix (the amount of variation explained by eigenvectors). Pillai's trace is the measure most robust to violations of assumptions and is the most commonly used statistic.

One consideration when using MANOVA is that power decreases as the number of response variables increases, which can lead to type II error. It is also possible to use a univariate approach and a Bonferroni adjustment to alpha. In addition, MANOVA can be used only when all subjects have been measured for all response variables; it is assumed that multivariate error effects are normal and covariances equal among groups (Scheiner 1993). Other assumptions are the same as for ANOVA.

Below, we examine the effect of three distinct classes of walleye length (20, 50, and 100 mm) and prey density (1, 10, 25 per m^3), and the interaction of predator length and prey density on the total number of searches, orientations, follows, pursuits, strikes, and captures in the predation sequence. Each combination of predator size and prey density was replicated five times using different walleye. The total number of searches, orientations, follows, pursuits, strikes, and captures were recorded for each trial.

Program

```
   data fishprey;
   input length density search orient follow pursue strike capture;
   cards;
   (input data);
   proc glm;
   class length density;
   model search orient follow pursue strike capture=length density
length*density;
   manova h=length density length*density /printe printh;
   run;
```

Output and Interpretation

Table Results of MANOVA examining the total number of searches, orientations, follows, pursuits, strikes, and captures in the predation sequence for three different size-classes of walleye (20, 50, and 100 mm) feeding at three different densities (1, 10, and 25 bluegill per m^3) of optimal-sized bluegill prey.

Source	Pillai's trace	*F* (df)	*P*
Length	1.01	5.42 (12, 64)	<0.0001
Density	1.07	6.26 (12, 64)	0.0001
Interaction	0.67	1.15 (24, 136)	0.31

The MANOVA test shows that foraging behaviors of walleye were significantly affected by both walleye length and bluegill density but not by their interaction. Because the interaction was not significant, the analysis should be run again without the interaction term.

Table Individual ANOVAs for predation sequence behaviors with walleye length and bluegill density as independent variables. The *F*-statistic has 2, 36 df.

	ANOVA			
	Walleye length		Bluegill density	
Behavior	*F*	*P*	*F*	*P*
Searches	13.67	0.0001	65.47	0.0001
Orientations	12.41	0.001	57.73	0.0001
Follows	5.32	0.009	61.95	0.0001
Pursuits	17.87	0.0001	93.80	0.0001
Strikes	1.99	0.15	27.51	0.0001
Captures	2.29	0.11	28.43	0.0001

Individual ANOVAs show walleye length affected the number of each behavior in the predation sequence from search through pursuit but did not affect the number of strikes and captures. Larger walleye initiated more searches, orientations, follows, and pursuits than did smaller walleye but attacked a lower proportion of the prey. As a result, the number of captures was similar across size-classes of walleye. Density of bluegill prey significantly affected all components of the predation sequence. At higher prey densities, walleye increased the number of all individual behaviors.

models, the true costs associated with components of the predation sequence, from orientation through capture, are often not accounted for explicitly in the models. These costs are extremely difficult to determine in laboratory experiments but could be important if predation costs differ significantly among prey. Other constraints on fitness (predation, reproduction, and habitat) may result in a fish not foraging optimally, and several assumptions of these models may not be met (e.g., prey encountered sequentially rather than simultaneously). Nevertheless, these models have been usefully applied to understanding several aspects of predator–prey interactions, particularly under experimental conditions.

16.5.3 Size and Species Selection

Size-specific attributes of the predator, prey density, and characteristics of the prey can all influence predator–prey interactions. Changes in prey type may be associated with an increase in predator size. Prey size can constrain selection for a variety of prey types, including zooplankton (Bremigan and Stein 1994; Mayer and Wahl 1997) and fish (Hambright 1991; Juanes 1994; Christensen 1996; Lundvall et al. 1999). Prey encounter rates increase with size-related changes in swimming speeds (Gerritsen and Strickler 1977) and size-dependent reactive distances to invertebrates (Breck and Gitter 1983) and larval fishes. In contrast, the effect of prey size on the reaction distance of piscivorous fishes to larger (postlarval) fish prey appears to differ among taxa: sunfishes exhibited increasing reaction distance with prey size (Howick and O'Brien 1983), whereas lake trout showed no effect of reaction distance to prey size (Vogel and Beauchamp 1999). Many fish undergo ontogenetic diet shifts and change their diet as they grow. These shifts are associated with changes in ability to capture more energetically beneficial prey types successfully and are often correlated with discrete periods of growth that may occur at critical periods in the life history of fishes (Buijse and Houthuijzen 1992; Stahl and Stein 1994). Several electivity indices are available that can be used to assess changes in prey preferences by predators (see Chapter 11).

It is often assumed that fish predators actively choose the sizes and species of their prey. This concept is central to optimal foraging models, which have been used to account for the influence of prey species and sizes on diet composition of predators (Werner 1974; O'Brien et al. 1976; Werner and Mittelbach 1981). However, a review of laboratory studies with controlled prey densities showed a variety of piscivorous fish chose the smallest prey available, and many of the prey consumed were smaller than optimal (Juanes 1994), suggesting that prey capture was a passive process (Sih and Moore 1990) rather than predators actively choosing prey. Under this scenario, predators attack all sizes of prey as encountered, but differential vulnerabilities lead to smaller, more vulnerable prey being consumed in higher proportions (Juanes 1994). Although prey handling times and associated optimal foraging models have been used to explain planktivore foraging successfully, these models have been criticized as unrealistic for piscivores foraging on mobile prey. These prey require measurable energy demands involved

with stalking and capturing that can vary considerably with size of prey (Sih and Moore 1990). Foraging behavior of piscivores, such as pikes (Wahl and Stein 1988), largemouth bass (Savino and Stein 1982; Hambright 1991), and walleye (Einfalt and Wahl 1997) have been examined extensively. Differences in cost–benefit relationships (handling time/prey dry mass) have successfully predicted selection for different prey species, but piscivores sometimes select larger prey (see Figure 16.5, an example for walleye) than predicted by an optimal foraging model, at least in confined experimental arenas. For piscivores, prey encounter rates may be more important than handling times in determining prey size selection, and different decision rules may be necessary to evaluate foraging decisions (Breck 1993).

Several studies have focused on prey morphology and behavior as factors determining species selection (e.g., Wahl and Stein 1988; Einfalt and Wahl 1997). Fusiform or soft-rayed prey are generally preferred over sunfishes, and several mechanisms have been proposed to account for these preferences. Prey differ behaviorally in response thresholds (Webb 1986) and escape tactics (Savino and Stein 1982; Wahl and Stein 1988; Einfalt and Wahl 1997). These prey-specific behaviors can also interact with habitat complexity to determine foraging success of predators (Savino and Stein 1982; Savino et al. 1992). In addition, body morphology can influence capture success and handling time (Webb 1986; Wahl and Stein 1988; Hambright 1991) with predators choosing larger sizes of fusiform minnows than deep-bodied sunfishes.

16.5.4 Functional Response Curves

Prey density is an important factor determining foraging success and can influence search time and behavior of predators (see Box 16.4 for analysis of prey density and predator lengths on feeding behaviors). Functional response curves are used to describe the relationship between prey abundance and the number of prey eaten per predator (Holling 1959; Figure 16.6) and can be a useful, experimentally based approach for modeling the link between ambient prey availability and feeding rates. The type I functional response curve describes a linear increase in the number of prey consumed with prey density but is often unrealistic as the predator eventually becomes satiated. The type II curve is often more realistic and can be defined as

$$N_{\text{eaten}} = a \cdot T_i N/(1 + a \cdot T_h \cdot N), \tag{16.4}$$

where a is the attack coefficient (attacks/s), T_i is the total time available for foraging (seconds), T_h is the handling time per prey, and N is the prey density. Nonlinear least-squares regression (SAS procedure NLIN) can be used to estimate parameters for these equations (Juliano 1993; Box 16.5). The type III equation may be appropriate when the predator switches the prey of interest as that type becomes more common or the predator learns to hunt more effectively for that prey.

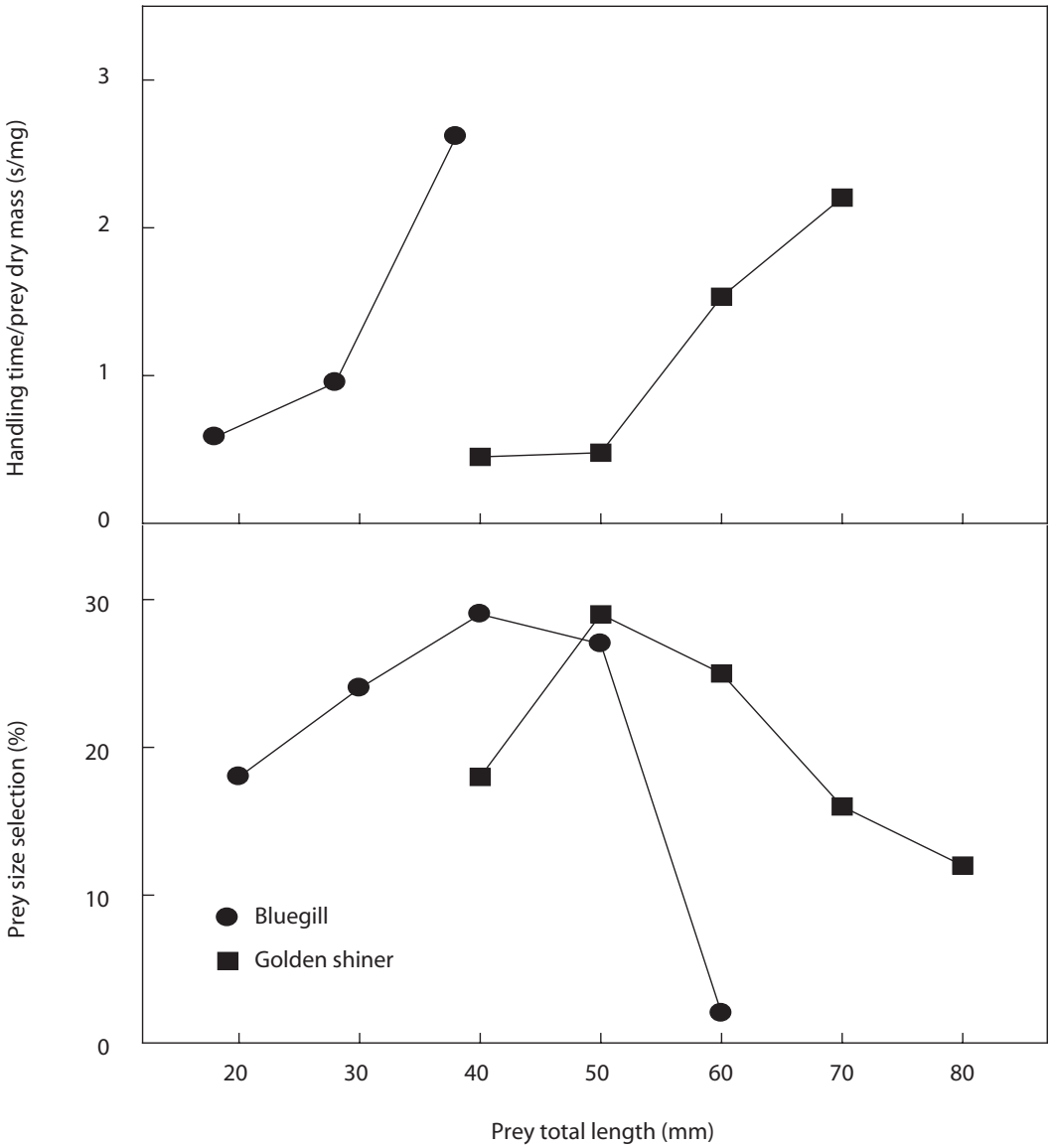


Figure 16.5 Example of cost–benefit relationships (handling time/prey dry mass) used to predict prey size and species selection for 150-mm age-0 walleye feeding on different sizes of bluegill and golden shiner prey (from Einfalt and Wahl 1997). Handling times were recorded for individual walleye in aquaria ($N = 10$) fed individual prey that varied in 4-mm size increments. To determine size preference, predators were fed five prey, one from each size-class. Minimum handling time/dry mass values indicate optimal values and suggest differences between prey species (13–20% of predator length for bluegill and 27–33% for golden shiner). Walleye have a more difficult time handling deep-bodied prey with spines such as bluegill compared with fusiform soft-rayed prey such as golden shiner. An optimal foraging construct successfully predicted differences in prey size selection with walleye choosing smaller bluegill than golden shiner; however, prey preference was for larger individuals of both species than predicted.

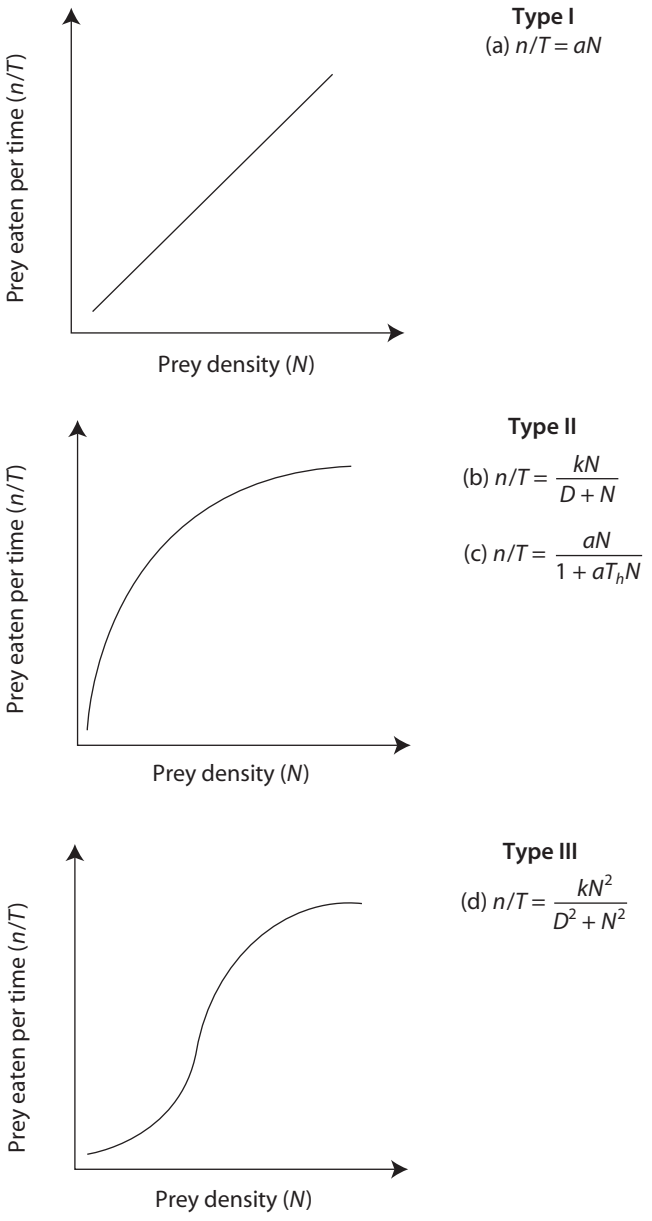


Figure 16.6 Functional response curves describing the relationship between prey abundance and the number of prey eaten per predator per unit of time. The terms of the type I equation (a) are n = number of prey, T = duration of foraging period, a = the predator's capture efficiency, and N = prey density. Some of these same terms apply to type II and type III equations. Two formulations of the type II response are shown; equation (b) gives the relationship in the form of Michaelis-Menten kinetics, where k is the asymptote (maximum feeding rate) and D is the half-saturation coefficient (prey density at which feeding rate is half of maximum). The same coefficients are used in the type III functional response (equation [d]). Equation (c) is a type II response specified in terms of foraging parameters where T_h = handling time.

Box 16.5 Functional Response Curves

Functional response curves quantify predator feeding rates as a function of prey density. The following example demonstrates how a functional response curve was determined for kokanee feeding on zooplankton.

The functional response, or the relationship between predator feeding rate and the density of its prey, is a fundamental framework for studying predator–prey interactions (Begon et al. 1996; Gotelli 1998). Three basic curves are commonly found, depending on the mechanics of the predation process (Figure 16.6). Functional responses are useful for predicting how changes in prey populations may affect fish feeding rates. They are also essential components of models that seek to predict fish foraging behavior and growth rates in the wild (e.g., Stockwell and Johnson 1997; Stockwell et al. 1999).

Laboratory experiments to estimate the parameters of a functional response involve trials in which predator consumption rate is measured under a range of prey densities. Pilot experiments can be used to estimate the variance in consumption rate among trials. That variance can then be used in a power analysis to estimate the number of trials needed to estimate the parameters of the functional response to within the desired level of precision (Chapter 3). Prey density in trials should cover the range of densities that are expected in the wild, and investigators should evaluate whether the experimental setup adequately mimics predator–prey conditions in the wild. The manner in which trials are conducted may affect fish behavior. For example, for some species intraspecific aggression in the confines of a laboratory arena may require the investigator to conduct trials with individual fish. If the same individual fish is to be used in multiple trials then the investigator should evaluate bias from possible carryover effects such as learning during previous trials.

Table Zooplankton (*Daphnia*) density, number of *Daphnia* consumed, trial duration, and kokanee consumption rate from a subset of laboratory trials by Koski and Johnson (2002).

Zooplankton density (*Daphnia*/L)	*Daphnia* consumed	Trial duration (min)	Consumption rate (*Daphnia*/min)
3.0	176	9.41	18.7
4.0	234	9.57	24.5
5.0	267	9.57	27.9
7.5	259	8.30	31.2
8.3	324	9.57	33.9
9.9	364	9.82	37.1
10.9	370	9.61	38.5
12.5	345	9.44	36.5
13.1	356	9.57	37.2
15.6	368	9.44	38.9
17.0	406	9.97	40.8
23.4	416	9.32	44.7
26.0	392	9.90	39.6

The following SAS program was used to fit a nonlinear regression by least squares and obtain the parameters of a type-2 functional response. The model's intercept was forced through the origin on theoretical grounds. Note that nonlinear regression models are more difficult to fit than are linear

ones. They also require the user to input starting parameter values; reasonable starting values are essential to insure rapid and accurate fitting of parameters. Starting values can be determined from a visual inspection of plots showing feeding rate as a function of prey density. The parameter *k* (maximum consumption rate) can be estimated from a plot if the data cover a sufficiently broad range of prey densities such that an asymptote is apparent. Given an estimate of *k* one can then estimate *D* (prey density at half of maximum consumption rate, *k*/2).

Program

```
   data typeIIfr;
   input N C;
   /* k = maximum consumption rate
   D = density at which consumption rate is 1/2 max
   N = prey density
   C = consumption rate (n/T)*/
   proc nlin;
   model C=(k*N)/(b1+N);
   parameters k = 50, D = 5; /* Input starting values for estimates of k and D*/
   output out=b p=yhat u95m=u95m l95m=195m;    /* Prints observed, predicted, and
95% confidence intervals around the mean*/
   proc print;
   run;
   proc gplot data=b;
   plot C*N yhat*N /overlay; /* Plots predicted and observed values to examine lack
of fit*/
   symbol1 value=plus;
   symbol2 interpol=join;
   run;
```

Output

Table Results of SAS program used to fit a nonlinear regression by least squares to obtain the parameters of a type-2 functional response for kokanee feeding on *Daphnia*. Parameters are *k* (maximum consumption rate) and *D* (prey density at half of maximum consumption rate, *k*/2).

		Regression			
Source	*df*	Sum of squares	Mean square	*F*-value	Approximate $P > F$
Regression	2	16121.7	8060.4	2564.13	<0.0001
Residual	11	34.5785	3.1435		
Uncorrected total	13	16155.3			
Corrected total	112	627.5			

	Parameter Estimates			
Parameter	Estimate	Approximate SE	Approximate 95%	confidence limits
b_0 (*k*)	50.0521	1.8421	45.9977	54.1065
b_1 (*D*)	4.1969	0.5419	3.0041	5.3897

(Box continues)

Box 16.5 *(continued)*

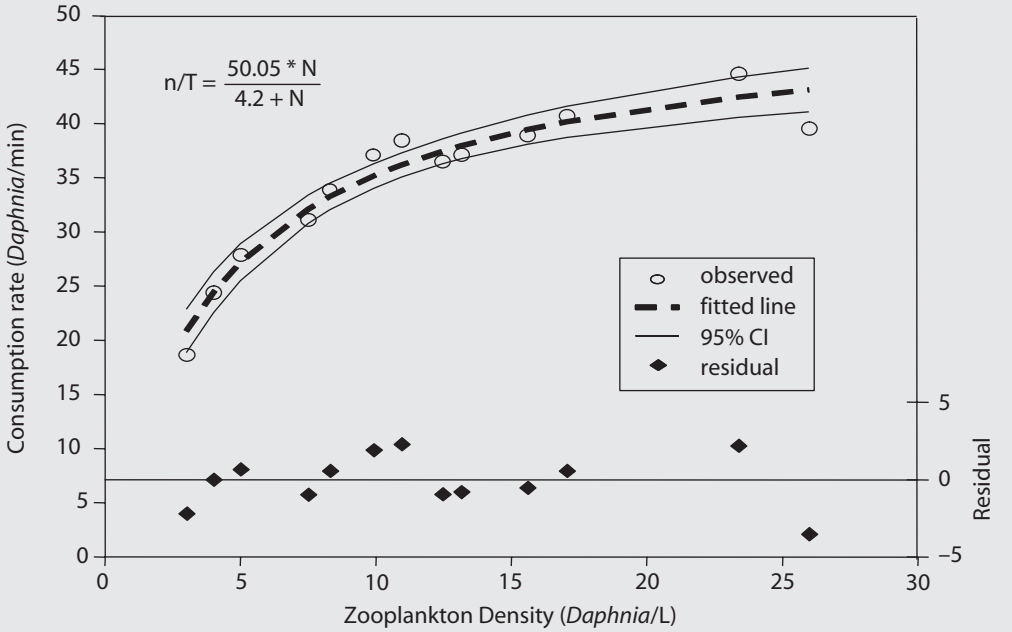


Figure Observed (circles) and predicted (dashed line) consumption rate (*n/T, Daphnia* per minute) as a function of zooplankton density (*N, Daphnia* per liter), and the upper and lower 95% confidence limits (CI, solid lines) on the regression line fit by nonlinear regression. Residuals from the regression (diamonds) are also plotted against zooplankton density.

Interpretation

Kokanee feeding rate increased rapidly as zooplankton density was increased, but the increase in their feeding rate began to slow at higher zooplankton densities, suggesting that handling time was beginning to limit consumption rate. However, feeding trials at zooplankton densities greater than the 26 *Daphnia*/L shown here (which undoubtedly occur in the wild) would be necessary to describe fully the functional response for this planktivore. Goodness-of-fit indicators suggest a good fit of the type-II model to the data: the regression explained a large fraction of the total variation in the data, there was no pattern in the residual plot (see figure), and the 95% confidence limits on the regression fit were not large. Akaike's Information Criterion (AIC; Burnham and Anderson 2002) could also be used to evaluate the appropriateness of the type-II model versus alternative models (e.g., a type I or a type III).

16.5.5 Application of Experimental Results to Management

Experimental studies of predator–prey interactions can be useful in helping guide a variety of fisheries management decisions. As an example, survival and growth can be highly variable for stocked sport fishes, and stocking strategies can be guided by information from predator–prey experiments. Losses to predation have been quantified for several species of stocked sport fishes, including pike (Stein et al.

1981; Wahl and Stein 1989), walleye (Santucci and Wahl 1993), channel catfish (Santucci et al. 1994), and Pacific salmons and trouts (*Oncorhynchus* spp.; Cartwright et al. 1998; Baldwin et al. 2000). Information on the effects of prey availability and predation on growth and survival of stocked fishes can be used to make decisions about when, where, and at what size to stock fishes (Box 16.6). Similarly, understanding of predator–prey interactions under experimental conditions have been usefully applied to predicting effects of exotic species (i.e., zebra mussels and round gobies) and have helped to guide and develop management options. These are just a few of the many ways in which these types of studies can be useful in management situations.

16.6 MODELS OF PREDATION RATES IN NATURAL SYSTEMS

Modeling approaches span a continuum of empirically through mechanistically based constructs. The experimental data section above (section 16.5) introduced models that were developed directly from controlled experiments on prey selection, optimal foraging, and functional response models, whereas this section will focus on models used to quantify predation in natural systems. Estimating the amount of prey consumed by predators is a common but challenging goal in predator–prey investigations. Predation rates can be calculated directly from field-based consumption estimates (e.g., Eggers 1977; Elliott and Persson 1978) or indirectly using empirical relationships (e.g., production:biomass ratios and predator–prey biomass ratios) or simulation models that are often supported by field measurements or laboratory experiments. Each method has strengths and weaknesses and differs considerably with regard to data requirements, validity of assumptions, and the degree of resolution along various dimensions of interest (Ney 1990). The most suitable approach will depend on the level of resolution required (e.g., annual averages versus finer temporal scales; whole population or trophic level averages versus size- or age-structured processes or individual predators; or whole-basin averages versus finer spatial scales) and on the type and quality of data or sampling resources available. Three mass- or energy-balance approaches (production-based estimates, bioenergetics models, and EcoPath with EcoSim model) and an encounter rate model are presented below. These methods should be viewed as complementary or sequential approaches, based on the primary question of interest, existing information, and available resources.

16.6.1 Production-Based Estimates of Consumption

Simple approaches such as production-based models provide first-order approximations of annual predation rates and can be computed quickly if common population dynamics data are available (Ney 1990). These estimates require knowledge of annual production by cohorts of predators (i.e., abundance and biomass at the beginning and end of a year) and food conversion efficiency. In these models,

$$P_t = G_t \cdot B_t, \tag{16.5}$$

Box 16.6 Case Study of Predation and Stocking of a Sport Fish

Stocking or introductions of sport and forage fishes are common management actions. Sport fish species are stocked to maintain fisheries in waters where habitat degradation and overexploitation have reduced existing populations or to establish new populations in waters such as ponds and reservoirs. Fish may be stocked at catchable size for immediate harvest or as juveniles in a put-grow-and-take stocking strategy. In Midwestern reservoirs, smaller northern pike are more vulnerable to predation by largemouth bass than are larger individuals (Wahl and Stein 1989), so size at stocking influences predatory mortality (see flow diagram below). Predation mortality on stocked northern pike declined from 30% for 145-mm fish to about 2% for 205-mm fish (panel **[A]**, left side flow diagram), as indicated by increased *C/f* (panel **[B]**).

Prey preference and prey demographics also influence both survival and growth of stocked piscivores. In laboratory experiments, northern pike captured herrings, carps, and minnows more successfully than they captured sunfishes (Wahl and Stein 1988; panel **[C]**), and they preferred gizzard shad over bluegill in the field. Morphology (body depth and spines) and antipredatory behavior unique to each prey species (bluegill are more evasive) contribute to differential vulnerability. As a result, northern pike grow more slowly with sunfish prey than with herrings (panel **[D]**). In addition to prey preference, availability of appropriately sized prey will also influence growth of stocked sport fishes (Madenjian et al. 1991; Santucci and Wahl 1993; Wahl and Stein 1993; Johnson et al. 1996). Access to the appropriate forage base is more important for smaller stocked fishes than for larger ones. As diet breadth increases (owing to increased gape sizes), predator dependence on specific sizes and types of prey declines (Stahl and Stein 1994). In addition to growth, prey characteristics influence survival of stocked sport fishes. Survival of northern pike is lower in sunfish communities than in those with herring or carp and minnow prey (Wahl and Stein 1988). Reduced survival may relate directly to reduced capture ability or indirectly to reduced growth and the resulting increased vulnerability to predation or disease.

Timing of stocking can influence predatory losses as well. Cool water temperatures in autumn reduce food consumption rates by largemouth bass, thus reducing the potential for predatory impact by warmwater predators. In contrast, as cooler fall temperatures reduce thermal stratification in coldwater lakes and reservoirs, spatial overlap can increase between coldwater predators like trouts and their prey (Baldwin et al. 2002), as most of the water column approaches optimal growth temperatures for the predators, and predation rates increase (Baldwin et al. 2000).

These and similar results for other stocked sport fish species suggest that stocking should be pursued within an ecological framework that integrates the relative importance of predator–prey interactions across all life stages (Wahl et al. 1995). This framework provides a guide for making management decisions concerning species, sizes, and timing of fish introductions into systems with specific characteristics. Similar approaches can assess potential negative consequences of stocking on community structure and function. In this way, the use of stocking as a management tool can be optimized.

where production P represents the amount of mass or energy accrued by a population or cohort per unit time (including growth of individuals that die during that interval), G_t is the instantaneous growth of the average individual during time t (e.g., $t = 1$ year), and B_t is the mean biomass of the population or cohort during the time interval. The instantaneous growth of the average individual, G_t, is given by

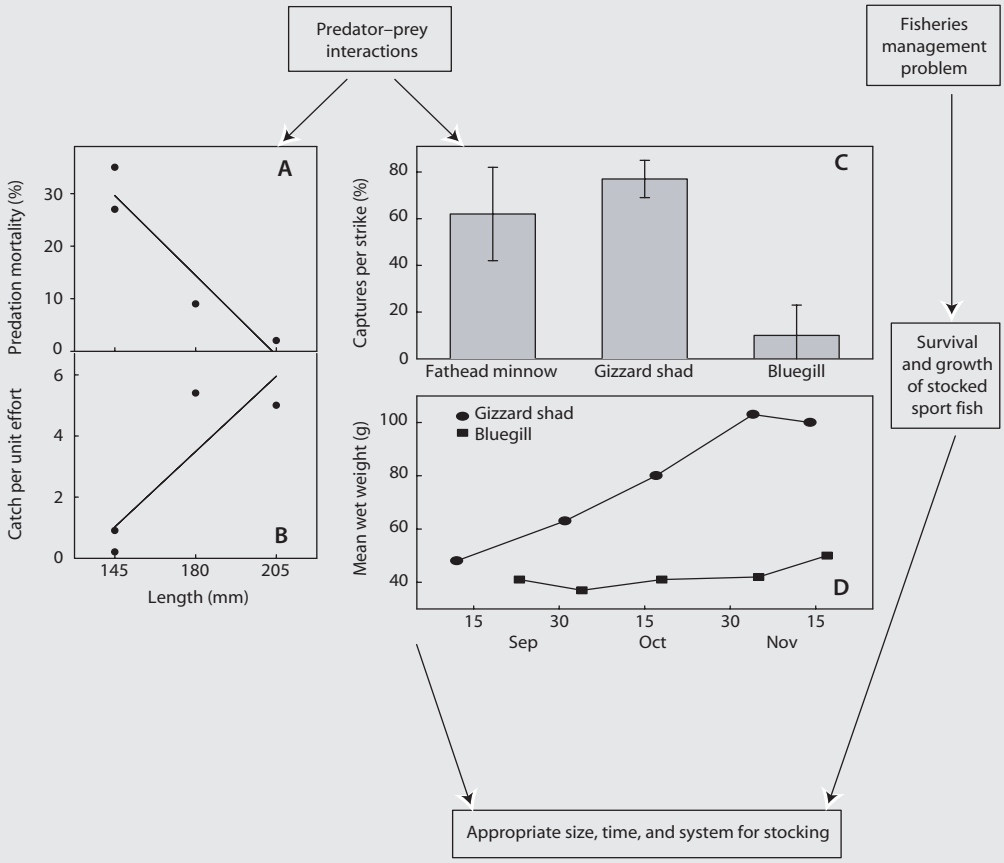


Figure Diagram of how experimental and field results can be used to address a management problem. Predation mortality for northern pike stocked into Midwestern reservoirs was higher for smaller than larger northern pike (**A**, lab experiments and field trials). This differential predation mortality was also reflected by higher catch per unit effort of the larger size-groups of northern pike stocked in reservoirs **(B)**. Juvenile northern pike exhibited higher capture success rates on soft-finned fathead minnow and gizzard shad compared with spiny-rayed bluegill in lab experiments **(C)**, and this was reflected in higher growth rates for northern pike in reservoirs where the predominant prey were gizzard shad rather than bluegill **(D).** When combined with information about the environmental conditions and community structure of the host waters, these type of results enable informed decisions about when, where, and at what size to stock sport fishes.

$$G_t = \log_e(W_t/W_0)/t\,, \tag{16.6}$$

where W_0 and W_t are the mean weights of individuals at the beginning and end of the interval, respectively. Food conversion efficiency, CE, is the amount of growth ($W_t - W_0$) by an individual that resulted from consuming a known quantity of food, C' (e.g., from long-term controlled growth experiments using natural prey):

$$\mathrm{CE} = (W_t - W_0)/C'. \tag{16.7}$$

The estimated consumption C_t over time t would then be

$$C_t = P_t/\mathrm{CE}. \tag{16.8}$$

If conversion efficiency was assumed to remain at a constant 10% for all age-classes and years, then total annual consumption of all prey species by the predator could be easily calculated in a speadsheet based on the relationships above (Box 16.7).

In reality, estimates of food conversion efficiency are rare for fish feeding in natural environments, as such estimates depend on estimates of individual consumption and growth over specified time intervals. Moreover, conversion efficiencies can change dramatically with body size, temperature, and food quality, thus potentially causing major errors in consumption estimates. Knowledge of the population dynamics of predators in some freshwater systems may also be insufficient to estimate annual production. If the predation events of interest vary considerably among size-classes of predators, through time, or among habitats, then very serious interpolation or extrapolation errors may result if the sampling design and subsequent analysis do not account for these important sources of variability. Nonetheless, such estimates could provide valuable "first-cut" evaluations for questions such as, Can predation potentially account for enough mortality to regulate specific prey populations?

If such an analysis suggested that predation removed a very small fraction of prey (e.g., mortality < 1%), then the direct effects of predation could be dismissed as a major source of mortality, even though the production-based estimator of consumption potentially erred several-fold from actual mortality. In this case, the uncertainty around the predation estimate did not exceed a threshold value at which predation losses were considered a serious source of mortality. However, if the consumption estimates suggested that predation represented, say, greater than 30% of the prey population, and the associated uncertainty resulted in a three- to fivefold difference in the predation estimate, then fisheries scientists might conclude that (1) predation could be a significant source of mortality, and (2) additional study might be required to reduce the greatest sources of uncertainty.

Ney (1990) proposed a variation of the conventional production-based approach, where consumption, C_t, represented constant multipliers of predator production, *P,* and biomass *B:*

$$C_t = 2 \cdot P + 3 \cdot B. \tag{16.9}$$

This approach implies that annual maintenance costs require the equivalent of three times the predator's body mass in food per year, and that after these metabolic demands are met, 50% of the remaining energy can be converted into growth while the remaining 50% is lost as waste. Consumption estimates from this method

Box 16.7 Production-Based Estimate of Consumption

Presented here is an hypothetical example of a production-based estimate of consumption of a focal species, prey A, by each age-class of a predator population. Fisheries scientists wish to calculate a quick, first-order estimate of predation on prey A and want to know which age- or size-classes consume the most biomass of prey A.

Table Consumption of prey A by age-class of predator. The mean body weight, W; the instantaneous growth of the average individual, G_t; abundance, N_t; cohort biomass B_t; and annual average production, P_t, of prey A in the diet varies among ages (t). Annual production, P_t, is divided by conversion efficiency (CE = 10% in this example) to estimate annual consumption, C_t, per cohort.

Predator age (t)	W(g)	G_t	N_t	B_t	P_t	CE	Total consumption C_t (g)	Percent prey A	Consumption prey A
0	0.3	3.9120	1,000	300	9,389	0.1	93,890	0%	-
1	15	2.6856	300	4,500	68,079	0.1	680,790	5%	34,040
2	220	0.7156	210	46,200	40,200	0.1	402,000	25%	100,500
3	450	0.2187	147	66,150	13,534	0.1	135,340	40%	54,136
4	560	0.0690	103	57,624	3,479	0.1	34,790	75%	26,093
5	600		72	43,218		0.1			

In this example, overall annual consumption was greatest for age-1 predators (680,790 g/year); however, since the percentage of prey A in the diet increased with the age of the predator, age-2 predators consumed the greatest annual biomass of prey A (100,500 g/year) based on this production-based estimate of annual consumption.

Note that conversion efficiency for piscivores can vary widely (e.g., CE = 5% to 30%) depending on variation in thermal conditions and energetic value of the primary prey. A first-order approximation of CE for multi-cohort piscivore populations could be 20% (e.g., Lane et al. 1979) and is consistent with a generalized annual energy budget for carnivorous fish of 60% of the consumed energy going to metabolism, 20% to waste, and 20% to growth (Brett and Groves 1979).

The data above are hypothetical, so no variability for the estimators is available. Because P_t is a product ($P_t = G_t \cdot B_t$), and C_t and CE are ratio estimates ($C_t = P_t/\text{CE}$; $\text{CE} = G/C'$), error estimates for products and ratios would be needed to compute error around the estimates for consumption using this method. Note that C' used in calculating CE would be derived independently from a different data set (e.g., consumption in a controlled growth experiment or an average from other studies) than the consumption C_t we are trying to estimate here.

$$\text{Var}(\hat{P}) = \bar{G}^2 \cdot \text{var}(B) + \bar{B}^2 \cdot \text{var}(G) - \text{var}(B) \cdot \text{var}(G).$$

$$\text{Var}(\hat{\text{CE}}) = \text{CE}^2 \left[\frac{\text{var}(G)}{G^2} + \frac{\text{var}(C')}{C'^2} - \frac{2\text{cov}(G, C')}{P \cdot C'} \right].$$

$$\text{Var}(\hat{C}) = C^2 \left[\frac{\text{var}(P)}{P^2} + \frac{\text{var}(\text{CE})}{\text{CE}^2} - \frac{2\text{cov}(P, \text{CE})}{P \cdot \text{CE}} \right].$$

compared favorably to independent field-derived estimates or bioenergetics model simulations, but the production-based and bioenergetic estimates of consumption diverged for older cohorts of predators (Ney 1993).

16.6.2 Bioenergetics Modeling

Bioenergetics modeling, particularly the Wisconsin bioenergetics model (Hewett and Johnson 1987, 1992; Hanson et al. 1997), has become a frequently used tool for estimating consumption or predation rates (e.g., Stewart et al. 1981; Hartman and Margraf 1992; Jones et al. 1993; LaBar 1993; Beauchamp et al. 1995; Baldwin et al. 2000; Johnson and Martinez 2000) or growth potential (e.g., Brandt et al. 1992) for many of the major freshwater fishes and some predatory invertebrates like mysid shrimp. Bioenergetics models are based on an energy balance equation:

$$\begin{aligned}\text{Consumption} = {}& \\ \text{Metabolism (standard and active plus specific dynamic action [SDA])} + {}& \\ \text{Waste (excretion and egestion)} + \text{Growth},&\end{aligned}$$

where maximum daily consumption and metabolism are modeled as species-specific functions of body mass and temperature (see Chapter 12 for more detail). The models are most commonly used to estimate the consumption required to satisfy growth observed or targeted over a specified time interval (Kitchell et al. 1977). These models are data intensive, requiring many species-specific parameters and extensive inputs from field data, but they offer the flexibility to address trophic responses at high temporal, spatial, and size-structured resolution if adequately supported by directed field sampling (Ney 1990, 1993; Brandt and Hartman 1993; Hansen et al. 1993).

Consumption estimates are constrained by growth increments observed for each age-class or growth cohort over specified time intervals, and seasonal consumption can be partitioned into predation rates on different prey categories by incorporating seasonal diet information from field sampling. An important advantage of the Wisconsin bioenergetics model is that it operates on a daily time step and can account for temporal changes in predator size, diet composition, temperature, and prey quality (energy density). Therefore, important short-term interactions like acute predation events can be simulated effectively. The input data requirements for this model can be demanding, but most data (size at age, size structure, diet, distribution, and temperature) are often recorded by fisheries scientists during routine population monitoring surveys. Data from routine assessments may not be collected at the frequency required to address specific predator–prey issues, but minor modifications to sampling and data recording protocols can satisfy the input requirements for these models. When compared with independent estimates of consumption, the Wisconsin bioenergetics model has performed well for a variety of salmons and trouts (Beauchamp et al. 1989; Brodeur et al. 1992; Ruggerone and Rogers 1992; Cartwright et al. 1998; Madenjian and O'Connor 1999) and largemouth bass (Rice and Cochran 1984) but differed

significantly from field-based estimates for pike (Wahl and Stein 1991) and perches (Boisclair and Leggett 1989a, 1989b, 1989c, 1989 d; but see Hewett et al. 1991 and Boisclair and Leggett 1991). A more thorough description of the Wisconsin bioenergetics model's construction, testing, refinement, and comparisons with alternative approaches (e.g., field-based stomach fullness and gut evacuation rate methods) are presented elsewhere (Ney 1990, 1993; Hansen et al. 1993; Chapter 12). We will explain how bioenergetics models can be used to quantify predator–prey interactions at multiple trophic levels and emphasize considerations for study design and analysis that minimize interpolation or extrapolation error.

Trophic interactions can be quantified by estimating the biomass or numbers and sizes of prey consumed by different predators in a food web. Consumption rates by individuals from each species or life stage can be estimated using a bioenergetics model, given field estimates of (1) incremental growth for each age-class of consumer; (2) the temporal diet composition of each age-class over the period of interest; (3) the average daily temperatures experienced by consumers (termed "thermal experience"); and (4) the energy density of the consumer and prey. Predation rates by individuals from each age- or size-class can be expanded to population level consumption rates if the mortality rates, abundance, and size structure of the consumer's population is known. As with other methods, the major challenge of this modeling approach is to minimize interpolation or extrapolation errors caused by inappropriately pooling input data across size-classes or life stages, by lumping dynamic periods of feeding, distribution, or growth with static periods, or by careless expansions from individual predation rates to population level impacts on prey populations.

16.6.2.1 Growth Inputs

Annual growth.—The accuracy of the consumption estimates are dependent on accurate estimates of growth by the predator because the Wisconsin bioenergetics model calculates the amount of food required to achieve the changes in the body weight of predators over specified time intervals. Annual growth increments should be considered the longest acceptable growth period used for fitting consumption. However, when modeling acute predation over relatively short periods, fitting consumption to an annual growth increment could produce considerable error in the estimated predation rate (discussed in following section on seasonal growth). Routine monitoring data often provide estimates of annual growth, either by tracking the modal lengths of each age-class through time or by back-calculating length at age from otoliths, scales, or other appropriate bony parts (Summerfelt and Hall 1987; Ricker 1992; Chapter 5). These length-at-age estimates are converted to weight-at-age estimates using length–weight regressions and provide a first approximation of annual weight change by different age-classes in the consumer population. Body mass at a given length can vary tremendously, so direct measurements of weight at age, when available, are preferred over the back-calculation method. Unfortunately, weights are recorded less frequently than are lengths. The accuracy and precision of these estimates will depend entirely on inherent variability of the data and the adequacy of the sample sizes for each age-class.

Accuracy and precision of annual growth increments can be improved tremendously by generating size and growth estimates from a relatively short, consistent sampling period just after annulus formation.

For older, slow-growing age-classes of long-lived species (e.g., lake trout), most of the annual growth is elaborated as seasonal gain and loss of gonadal tissue rather than somatic growth and will not be adequately reflected by annual growth increments. Therefore, the change in gonadal weight may provide a reasonable minimum estimate of both seasonal and annual growth. The Wisconsin bioenergetics model allows the user to specify the date and amount of gonadal loss for mature fishes. The question then becomes whether adults spawn every year and whether the sex ratio of spawners is highly skewed toward one gender or the other. Accurate measurement of growth becomes progressively harder with the older age-classes of long-lived or slow-growing species. Therefore, consider whether these older cohorts still represent an ecologically significant component of the predator population, based on their relative abundance, diet composition, and consumption rate. If the older cohorts impose minimal predation compared with more abundant younger cohorts, then the older cohorts may not require the additional effort needed to address minor sources of uncertainty or error.

Tracking weight change from tagged individuals can provide direct estimates of growth over a variety of time intervals; however, recapture and accurate measurement of the tagged fish in sufficient numbers at strategic times may be too limited to provide reliable estimates of growth. Tagged fish might also grow slower than untagged fish, thus potentially underestimating growth. Nonetheless, mark–recapture studies are commonly employed for population abundance estimates, and these studies can serve as the primary source of growth estimation or as supplementary information to help interpret age and growth patterns inferred from other methods.

When consumption is fitted to annual growth increments in bioenergetics models, changes in daily consumption estimates will be driven primarily by the temperature regime used in the model and, secondarily, by any large seasonal changes in energy density for a large fraction of the diet. Consequently, if the true seasonal growth pattern differs significantly from the temperature-driven growth trajectory produced by the model, short-term consumption estimates (e.g., acute predation periods) will be biased in the same direction that the simulated growth deviates from the true seasonal growth trajectories.

Seasonal Growth. More accurate estimates of consumption are generated when bioenergetics models are fit to multiple growth intervals per year instead of annual growth increments (Rice and Cochran 1984; Beauchamp et al. 1989; Wahl and Stein 1991). Aquatic organisms rarely grow at a constant rate throughout the year because temporal changes in food supply, temperature, and other environmental conditions operate independently or in concert to produce seasonal growth patterns. Because the Wisconsin bioenergetics model estimates consumption to satisfy observed growth over specified time intervals, it is important to allocate growth rates as accurately as possible to the ecologically significant periods of the year. For example, predation may be concentrated over a short period (e.g., a

week or month) when the predators' growth is either much higher or lower than the average annual growth rate. Consequently, errors in estimated predation would depend on how much the actual seasonal growth of the predators deviated from the average growth rate over the year (Figure 16.7).

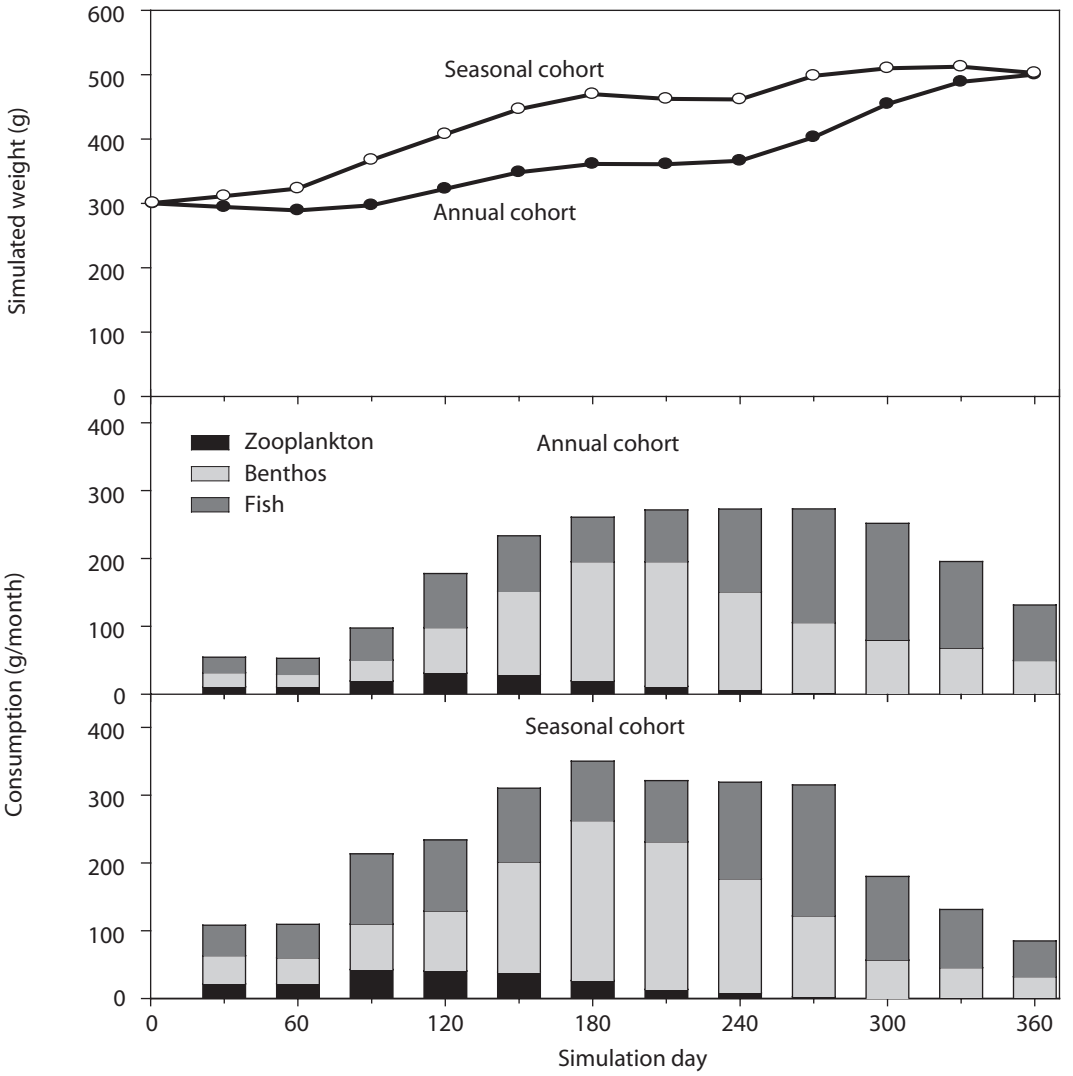


Figure 16.7 Comparison of monthly growth and consumption from bioenergetic simulations based on a single annual growth cohort versus four seasonal growth cohorts when both simulations grow from 300 g to 500 g in 365 d. Mean body weight was consistently higher each month in the seasonal-cohort simulation. The different growth trajectories resulted in different monthly consumption rates. Total consumption was estimated as 2,273 g/year for the annual-cohort and 2,677 g/year for the seasonal-cohort simulation; consumption of benthos was 20% higher per year from the seasonal-cohort estimate and fish consumption 11% higher in the seasonal-cohort simulation.

Ideally, the mean body weight of each age-class would be sampled at the beginning and end of ecologically significant periods of the year by tracking changes in mean weight for each age-class or by tracking length–frequency modes and converting to body weights. Unless age-classes can be unambiguously separated into distinct size modes, age analysis will be needed to reconcile length modes into age-specific growth. Sample size requirements and logistical or political constraints often preclude the sampling intensity needed to measure growth directly over each of the periods of interest.

Alternative methods can be employed for allocating growth among periods of the year. Short-term growth can be back-calculated from circuli spacing in scales, otoliths, or other hard parts. Seasonal changes in relative weight or condition factor for groups of age-classes (e.g., ages grouped as juveniles, subadults, and adults) may improve temporal accuracy in growth with considerably less data than would be required to estimate seasonal growth of each age-class directly. Physiological measures of short-term growth may provide additional resolution in some instances (e.g., RNA:DNA ratios or insulin-like growth factors like IGF-1 [Beckman et al. 1998]), but these methods have shown mixed results among species or are still experimental. In general, consider whether seasonal growth will vary significantly for reasons other than temperature. If so, then it will be important to allocate weight changes for each age-class of the predator into multiple growth stanzas (termed "cohorts" in the Wisconsin bioenergetics model) within year-classes to estimate appropriate consumption rates during ecologically significant periods.

16.6.2.2 Diet Inputs

For many species, diet composition changes seasonally and with increasing body size, and this variability must be captured in order to model predation rates appropriately. Consumers should be segregated into different feeding guilds, based upon differences in diet composition and trophic position. Predators can be grouped into size categories based on the statistical and graphical approaches described in Chapter 11 and section 16.4.2. Diet information for any size-class of consumer is entered as an input file into the Wisconsin bioenergetics model as the proportional contribution of each prey category in the diet by wet weight (or volume) for different dates through the period of interest. A diet input file can be constructed to contain both periods of constant diet proportions and times of rapid dietary change. This is particularly useful for limiting certain diet patterns to discrete periods (e.g., heavy predation over 1–2 months) rather than allowing certain prey items to remain in the diet inappropriately over longer intervals. When diet compositions differ between sampling dates, the Wisconsin bioenergetics model will automatically interpolate the diet proportions for every day between the dates actually entered into the diet input file. To keep diet proportions constant through a time interval, enter the same diet proportions on the first and last day of that interval (Box 16.8).

Box 16.8 Diet Data for the Wisconsin Bioenergetics Model

An example of a diet input file used in the Wisconsin bioenergetics model is given here. Diet data were entered for every sampling date (or the median date of a sampling period can be used). When diet proportions change between sampling dates (among days 1, 91, 181, and between days 271 and 365), the model linearly interpolates a daily change in proportions of each prey type as indicated in the graph. The diet proportions remain constant over periods when the same proportions are entered at the start and end of that period, as between days 181 and 271. A combination of constant and interpolated diet proportions can be used to minimize error associated with applying diet proportions over too long a period in model simulations.

Table Hypothetical diet input file for Wisconsin Bioenergetic model.

Day	Zooplankton	Benthos	Fish
1	0.05	0.60	0.35
91	0.20	0.30	0.50
181	0.05	0.30	0.65
271	0.05	0.30	0.65
365	0.00	0.60	0.40

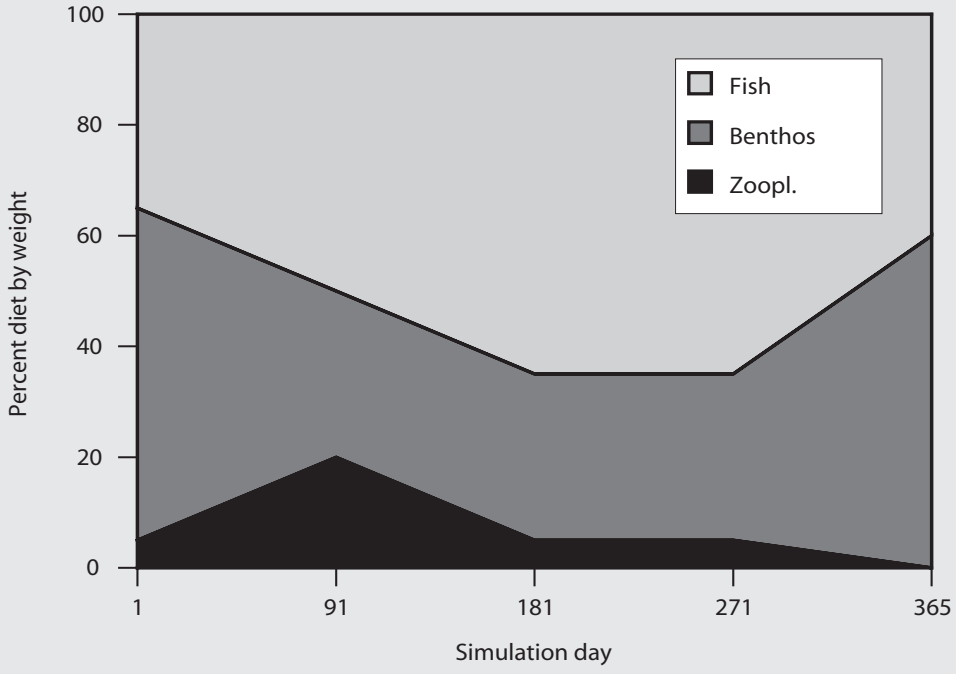


Figure Resulting model simulation of the daily change in diet from hypothetical data set (zoopl. represents zooplankton).

16.6.2.3 Thermal Experience

The thermal experience of consumers can be determined several ways based on field data or knowledge of their behavior and distribution patterns. In waters that do not thermally stratify, average daily temperature recordings from temperature loggers may be sufficient for estimating thermal experience unless organisms concentrate in thermal microhabitats (e.g., salmons and trouts congregating in groundwater intrusions). Warmwater species can often be assumed to occupy the warmest temperatures available during thermal stratification (e.g., littoral or epilimnetic regions of lakes during summer and the deepest water available during winter) and would be confined to the ambient temperature during isothermal conditions.

For pelagic and demersal species, or species with variable movement and distribution patterns, reconstructing thermal experience is more involved because the combination of vertical distribution and movement patterns with concurrent vertical temperature profiles determine the average daily thermal experience of these organisms. If vertical distribution information is available (e.g., diel hydroacoustic data for depth-specific densities of planktivores or depth-specific C/f data from gill nets), a weighted mean thermal experience for different species or size-classes can be computed for each sampling date by first multiplying the proportion of the total catch (for that species or size-class) in each depth interval by the mean temperature within that depth interval and then summing these products over all depth intervals. This sum represents the weighted average thermal exposure for the average individual of that species or size-class in the population (Box 16.9). When temporal depth distribution data or temperature profiles are unavailable, a common approach has been to assume "behavioral thermoregulation," which means that fish will seek out temperatures closest to their physiological optimum temperature for growth. This may ignore other important behavioral or physiological constraints like foraging opportunities, predator avoidance, or tolerance for hypoxia or other suboptimal environmental conditions.

16.6.2.4 Energy Density of Predators and Prey

The energy density (in terms of calories or joules per gram body weight [cal/g or J/g]) of prey will determine how much prey biomass must be consumed for a predator to obtain any given amount of energy. For example, a fish would need to consume at least 1.5 times more biomass of invertebrates with energy densities of 3,000 J/g than of fish prey with 4,500 J/g to acquire the same amount of energy. Moreover, invertebrates generally contain a relatively large fraction of indigestible material in their exoskeleton (averaging 10–17% of their body weight across many taxa and 25% for crayfishes compared with an average of 3% indigestible material in fishes. Of the energy ingested, waste losses are subtracted and metabolic costs (standard and active metabolism plus SDA) are paid before any energy is allocated for growth. The remaining energy is divided by the energy density of the consumer to convert energy into new consumer biomass. So if a predator's energy density was 6,000 J/g, and 4,000 J of energy remained after all waste and metabolic costs were accounted for and removed by the model, that remaining energy would be converted into 4,000 J/(6,000 J/g) = 0.67 g of new growth for the predator.

Box 16.9 Computation of Average Thermal Experience

Given here are two examples of computation of average thermal experience for cohorts, with and without vertical migration, occupying a range of depths under thermally stratified conditions.

Table Average thermal experience in the absence of diel vertical migration.

Depth (m)	Temperature (°C)	Fish density or *C/f*	Proportion of total fish	Proportion allocation × temperature
0	20	0	0.00	0.0
2	20	3	0.06	1.2
4	20	9	0.18	3.6
6	20	12	0.24	4.8
8	18	10	0.20	3.6
10	16	6	0.12	1.9
12	14	4	0.08	1.1
14	12	3	0.06	0.7
16	10	1	0.02	0.2
18	9	0	0.00	0.0
20	9	1	0.02	0.2
22	9	0	0.00	0.0
24	9	1	0.02	0.2
Total		50	1.00	17.5

The weighted average thermal experience on this sampling date in the absence of vertical migration was 17.5°C.

Table Average thermal experience when population undergoes diel vertical migration during periods of thermal stratification.

		Day (14 h)			Night (10 h)		
Depth (m)	Temperature (°C)	Fish density (*C/f*)	Proportion of total fish	Proportion allocation × temperature	Fish density (*C/f*)	Proportion of total fish	Proportion allocation × temperature
0	20	0	0.00	0.0	5	0.07	1.3
2	20	0	0.00	0.0	12	0.16	3.2
4	20	0	0.00	0.0	25	0.33	6.7
6	20	0	0.00	0.0	17	0.23	4.5
8	18	0	0.00	0.0	12	0.16	2.9
10	16	0	0.00	0.0	4	0.05	0.9
12	14	3	0.06	0.8	0	0.00	0.0
14	12	0	0.00	0.0	0	0.00	0.0
16	10	1	0.02	0.2	0	0.00	0.0
18	9	5	0.10	0.9	0	0.00	0.0
20	9	18	0.36	3.2	0	0.00	0.0
22	9	16	0.32	2.9	0	0.00	0.0
24	9	6	0.12	1.1	0	0.00	0.0
26	9	1	0.02	0.2	0	0.00	0.0
Totals		50	1.00	9.3	75	1.00	19.5

(Box continues)

Box 16.9 *(continued)*

To determine the weighted average thermal experience with vertical migration, the mean day thermal experience is multiplied by the hours of daylight and added to the mean night thermal experience multiplied by hours of night. The sum is devided by 24 h:

$$[(9.3°C \times 14\ h) + (19.5°C \times 10\ h)]/24\ h = 13.5°C.$$

The same computations would be repeated for other sampling dates. These values would be entered into the temperature input file, and the model would interpolate daily temperatures between the dates when thermal experience data were entered.

Energy density varies considerably among organisms and can change seasonally or with increasing body size. In the Wisconsin bioenergetics model, energy densities are provided as default values in the parameter set for each of the 40 species or life stages provided in the existing model (Hanson et al. 1997). The consumer's energy is often held constant in the model for most species; however, for members of the Family Salmonidae, energy density increases linearly with increases in body mass up to a threshold weight, then remains relatively constant or increases more gradually thereafter. Although strongly recommended, energy densities of predators and prey have rarely been measured in conjunction with a bioenergetic analysis of trophic interactions (but see Luecke and Brandt 1993; Rand et al. 1994; Hartman and Brandt 1995; Bryan et al. 1996). Energy densities of prey are generally taken from the literature (e.g., Cummins and Wuycheck 1971; Hanson et al. 1997).

16.6.2.5 Bioenergetic Simulations of Predation

The Wisconsin bioenergetics model can estimate the biomass of each prey type consumed daily by individuals from each cohort of predators. The Wisconsin bioenergetics model reports individual or population level consumption as either a daily rate (g/d) or as cumulative consumption to date (g/period) from the start of the simulation period. Total consumption over a set period can be estimated by summing daily consumption rates within that time interval. Individual consumption from each age-class or size-class can be expanded to population level consumption over any ecologically relevant time steps of daily or greater time intervals by including the initial abundance and mortality of each cohort as inputs to the model (Figure 16.8). In an alternative approach, individual consumption rates are estimated from the model, then model output for individuals is transferred to a spreadsheet, and individual consumption is multiplied at each time step by the corresponding abundance of predator cohort(s). Predator abundance would occupy a column in the spreadsheet alongside columns for the individual consumption estimates at each time step. Each cell in the column would contain the standard formula for computing predator abundance at time t (N_t):

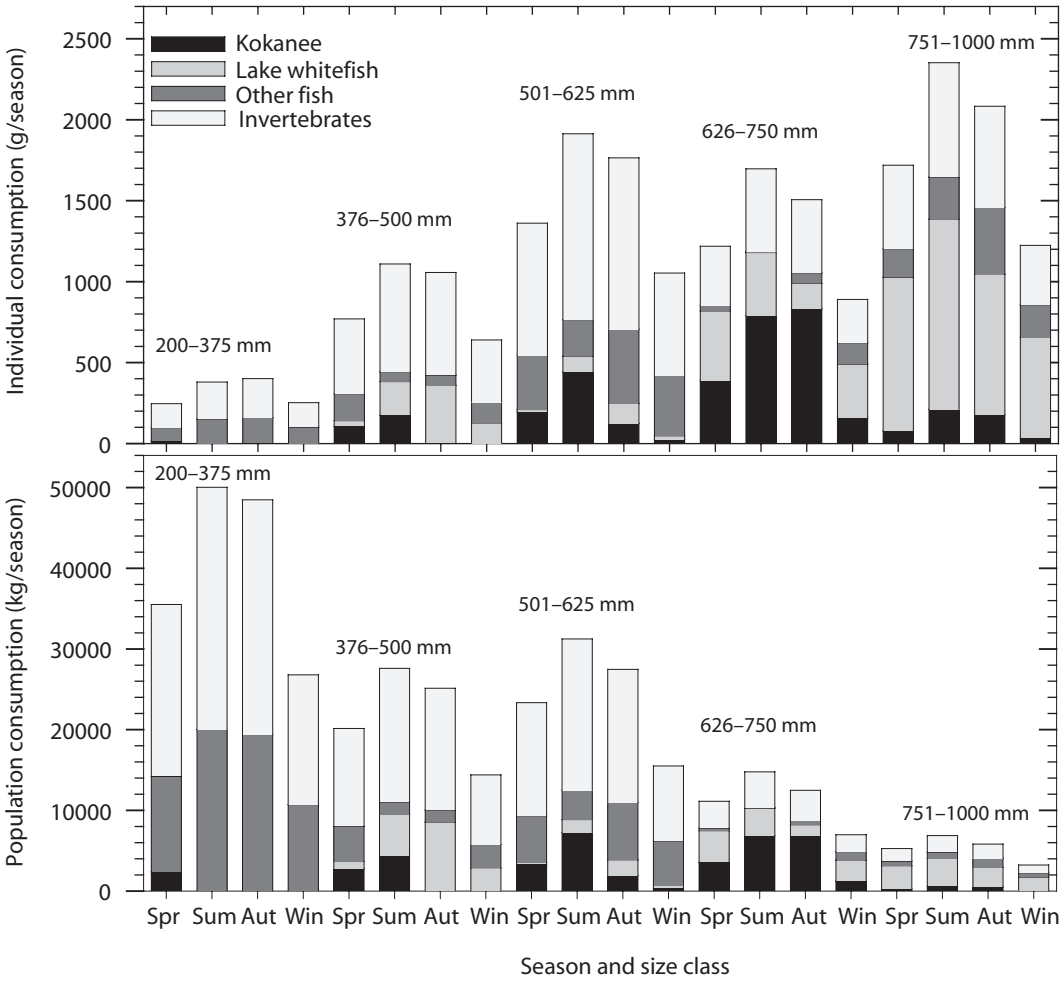


Figure 16.8 Seasonal consumption of major prey species by the average individual (upper panel) and population (lower panel) of each size-class of lake trout in Flathead Lake, Montana. Even though the largest lake trout eat more kokanee on an individual basis, the intermediate and smaller size-classes are much more important predators at the population level. Predation on kokanee varies considerably among seasons. Kokanee outgrew the smaller predators during spring or summer, but predation by the larger predators persisted through autumn and winter.

$$N_t = N_0 \cdot e^{-Zt}, \tag{16.10}$$

where t would relate to the different days in the simulation whereas the initial predator abundance, N_0, and instantaneous mortality, Z, would each refer to a fixed cell or cells where new values could be entered. Predator abundance and population level consumption would then change at every time step throughout the simulation period. The advantage of this approach is that we can rapidly explore the effects of different population dynamics scenarios for predators (e.g.,

changes in abundance or mortality rates) on predation rates of prey without rerunning the Wisconsin bioenergetics model for every scenario. Additional complexity could be added by incorporating stage-specific instantaneous mortality rates $Z = (M + F)$ to account for differences in natural mortality, M, or fishing mortality, F, among life stages. These additions would enable simulation of different management scenarios to evaluate the effects of various size and harvest limits (Luecke et al. 1994), stocking rates (Stewart et al. 1981; Jones et al. 1993), or interannual variability in survival and recruitment of predators.

The biomass of prey eaten can be converted to a numerical estimate of predation by dividing the biomass of each prey category consumed by the mean weight of an individual prey item. If prey size varies through time or varies among size-classes of predators, then the conversions from biomass to numbers of prey consumed should be computed separately for each combination of time and predator size-class. Large estimation errors can arise during conversions from prey biomass to numerical estimates of prey consumption, particularly if predation persists for more than a month when the prey (e.g., juvenile prey fishes) are growing rapidly, because the body mass of young fish can increase many-fold over relatively short periods (e.g., Cyterski et al. 2003). For example, lake trout in Flathead Lake, Montana, consumed kokanee averaging 145 mm in total length (TL) (25 g) in June and 215 mm in TL (93 g) in August, representing a 48% increase in prey length but a nearly fourfold increase in body mass over just a 2-month interval. Under these circumstances, numerical predation estimates could err considerably by careless averaging of consumption rates or prey sizes through time or across predator size-classes.

Although predation mortality is generally assumed to decline with increasing prey size, the magnitude of that decline has rarely been quantified in natural systems. Since per capita predation rates can be quantified for every size of predator, and predator–prey size relationships can be developed either experimentally (see Box 16.6, Figure 16.5) or empirically (Figure 16.9), the relationship between prey size and potential predation pressure can be formalized by including the abundance and size structure of both predators and prey. For instance, lake trout consume prey fish up to 50% of their own body length (Figure 16.9). Therefore, the number of lake trout, N_{pred}, capable of eating prey of any given length, L_{prey}, is the sum of all predators of length, L, equal to or greater than twice L_{prey} in the total population of predators, N:

$$N_{\text{pred}} = N \cdot \left(\sum_{L=2L_{\text{prey}}}^{K} n_L \Big/ \sum_{L=L_{\min}}^{K} n_L \right), \tag{16.11}$$

where n_L is the count of lake trout in length bin L of a length–frequency histogram from a representative sample of lake trout; K is the maximum length of lake trout in the population; and $L_{\min}$ is the length of the smallest lake trout included in the population estimate N (e.g., N = abundance of lake trout $\geq L_{\min}$, where $L_{\min}$ = 200 mm in TL). By quantifying how the maximum or optimum size of consumable prey increases with predator size, and using the abundance and size distribution of predators, we can examine how incremental changes in prey size affect

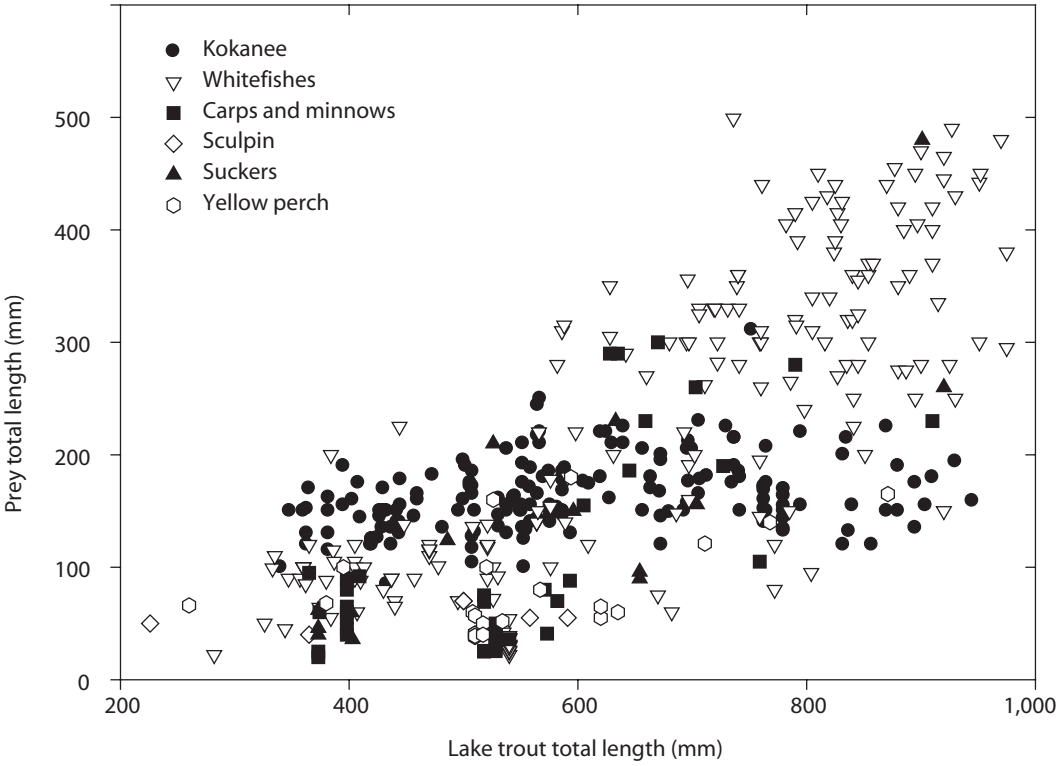


Figure 16.9 Relationship of lengths of piscivorous lake trout to the lengths of prey fishes by species in Flathead Lake, Montana. The predator–prey length relationship showed a classic wedge-shaped distribution for all fish prey collectively, but the pattern varied among prey species. Multiple age-classes of whitefishes were available year-round over a broad range of sizes, whereas most kokanee were available for only a few months after stocking and thus offered a relatively narrow size range to lake trout. This type of graph displays the size range and relative frequency of different fishes eaten by different sizes of predators. We can identify the size at which specific prey entered the diet, and size relationships can be quantified (e.g., mean, median, and lower and upper bounds; see Box 11.3).

predation potential in terms of (1) N_{pred}, the abundance of predators capable of eating a particular size of prey (Figure 16.10, upper panel); (2) the biomass and corresponding number of prey that could be consumed per unit time (e.g., kg/month or prey/month) by all predators greater than or equal to the smallest predator capable of eating a specified size of prey (Figure 16.10, lower panel); and (3) given the growth rate of prey , the period over which prey are vulnerable to predation (Figure 16.11). Larger predators can consume more biomass and larger-bodied prey, but predator abundance declines with increasing size and age due to the cumulative effects of natural and fishing mortality. These analyses can be used to evaluate trade-offs in stocking size versus changes in abundance, size

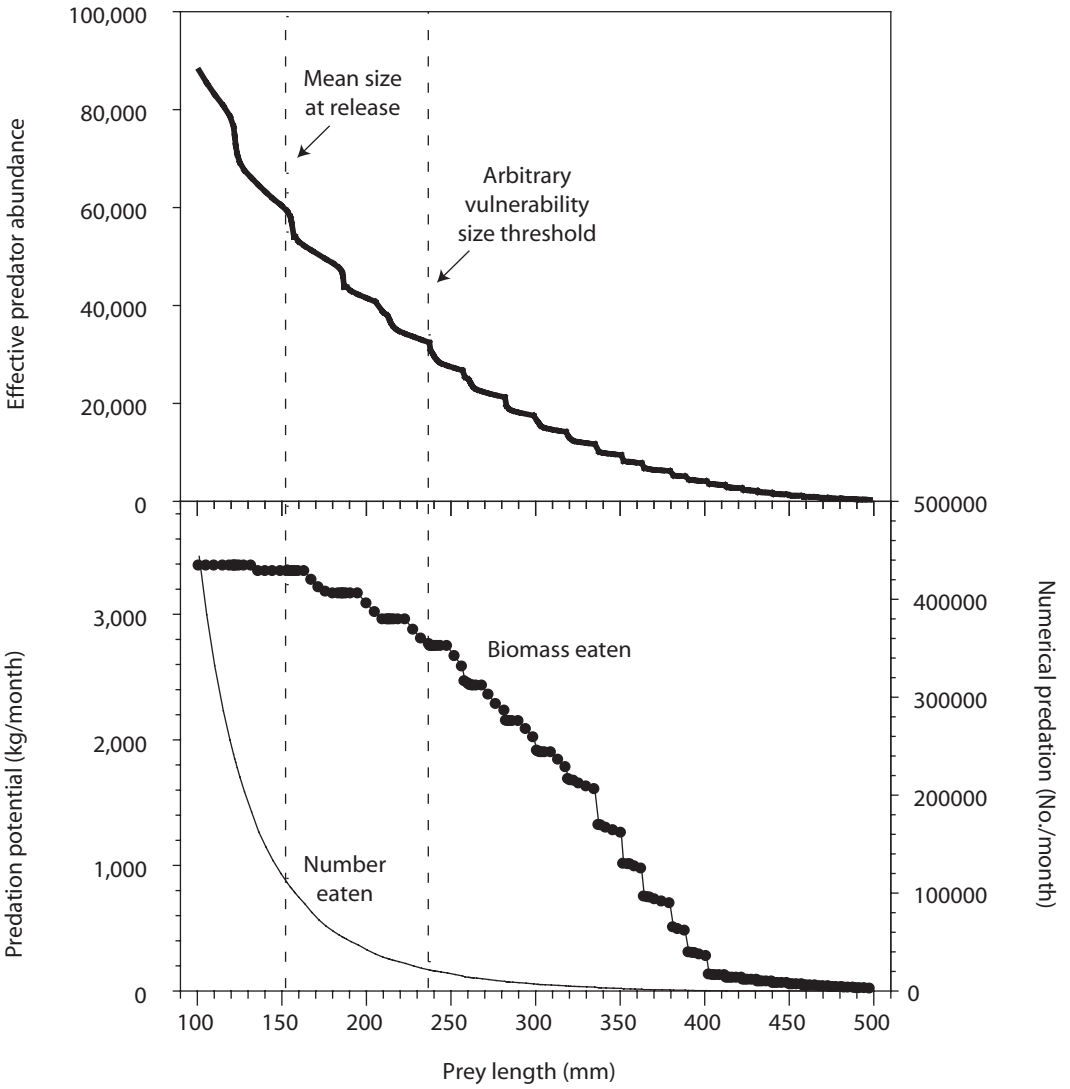


Figure 16.10 Relationships are graphed for lake trout in Flathead Lake, Montana, to demonstrate how predation declines as kokanee grow or as size at release increases. Predation rates declined as prey grew because fewer predators were large enough to capture them successfully. This reduces the total biomass of prey consumed. Also, as prey body mass increased, fewer prey were needed to satisfy the consumption demand of predators. In lower panel, the arbitrary size vulnerability threshold represents an 80% reduction in the number of kokanee lost to predation by lake trout.

structure, or survival of predators to regulate target prey populations (e.g., Box 16.10), or the feasibility of a population restoration program (Box 16.11).

16.6.2.6 Size-Structured Relative Predation Rates

When abundance estimates for the predator populations are lacking, a useful way to present model simulations is to report consumption demand in terms of

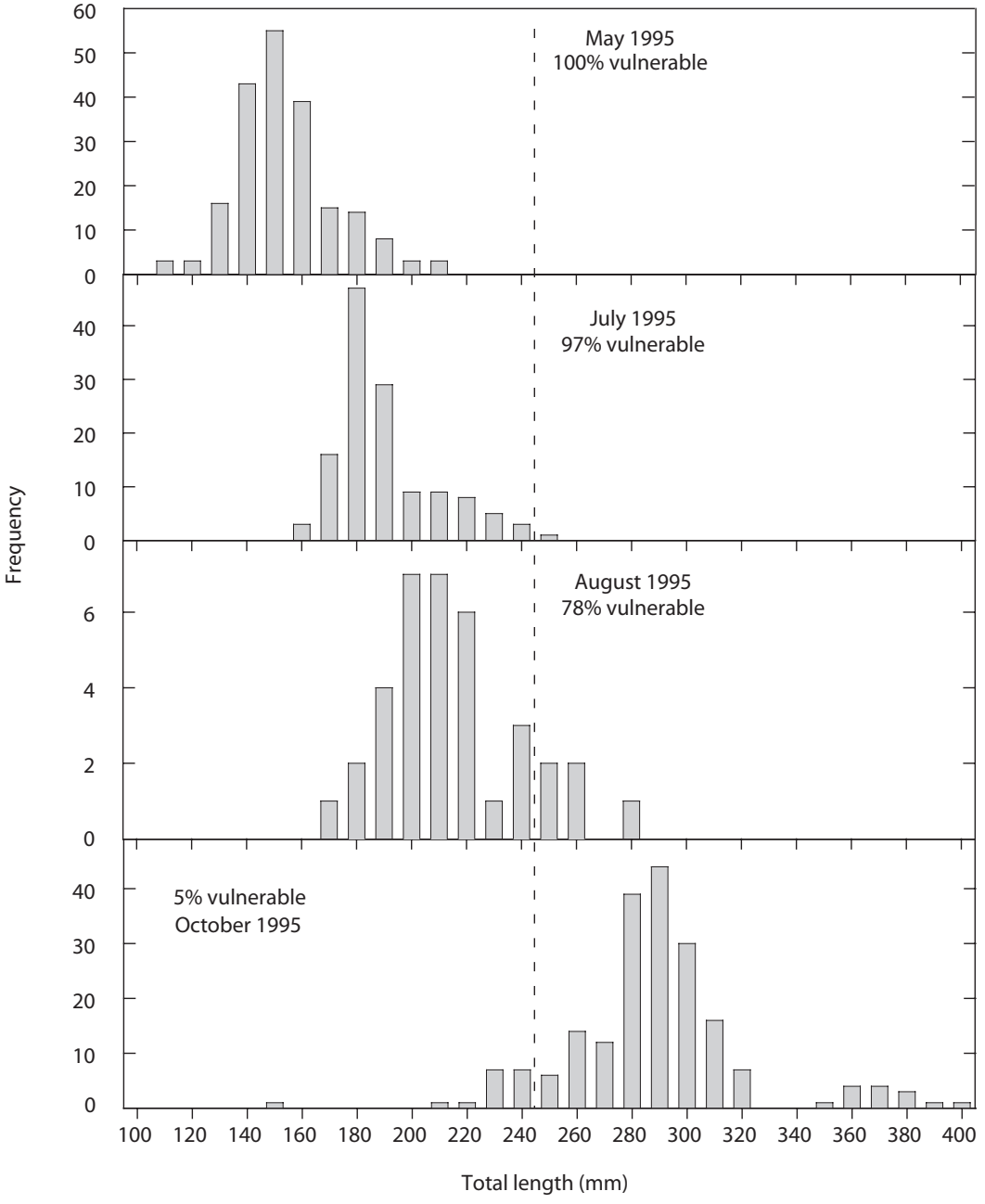


Figure 16.11 Temporal change in length frequency for kokanee stocked during May in Flathead Lake. The dashed vertical line is an arbitrary size vulnerability threshold that represented a 50% reduction in the number of lake trout still large enough to eat kokanee that size and a sixfold reduction in numerical losses due to predation. The size vulnerability threshold was based on a regression ($TL_{prey} = -100 + 0.65 \cdot TL_{lake\ trout}$, where TL is total length; $r^2 = 0.925$; $P < 0.0001$) of predator size versus maximum (95th percentile) of prey size in stomachs (see Chapter 11). All kokanee were vulnerable in May, but only 5% of the survivors remained below the vulnerability threshold by October.

Box 16.10 Case Study of Biomanipulation in Lake Mendota

Lake Mendota is a 4,000-ha eutrophic lake in south-central Wisconsin. The lake hosts a diverse assemblage of cool- and warmwater fishes. Expanding agricultural activities and urban development of the watershed increased external nutrient loading to the lake, contributing to nuisance algal blooms. Because nutrient inputs were mainly from nonpoint sources, and therefore difficult to control, limnologists and fishery scientists tried an innovative experiment to evaluate the potential for manipulating predation (biomanipulation) as a water quality management tool in a large, urban lake (Kitchell 1992). This "top down" experiment involved stocking nearly three million fingerling walleye and northern pike during 1987–1989 and protecting them with restrictive harvest regulations in an attempt to shift the balance among predators and their prey (mainly planktivorous fishes such as yellow perch and cisco). A major thrust of the study was predicting and quantifying the predatory impact by the stocked piscivores.

To forecast and then evaluate predator consumption rates, bioenergetics models (Chapter 12) were used to estimate the biomass of prey consumed by predator populations under various assumptions about predator stocking and mortality rates (Johnson et al. 1992a, 1992b). Applying the models to study predator–prey interactions required detailed field estimates made with a variety of analytical approaches covered in this book.

Analytical Methods

As can be seen below, estimating population-level effects of predators requires considerable field data and an involved set of numerical analyses (see figure below). However, these kinds of ecological questions are central to effective fisheries management, with implications that extend to the ecosystem scale (Kitchell et al. 1994).

1. Size-specific mark–recapture estimates (Chapter 8) of predator abundance were obtained using a suite of sampling gears aimed at minimizing sampling bias.
2. The abundance of size-classes of predators was converted to abundance by age using age–length keys (DeVries and Frie 1996) derived from scale sample analysis (Chapter 5). Scale samples also provided mean length at age of each predator, and these lengths were converted to weights at age using species- and sex-specific length–weight regressions (Chapter 10).
3. Fishing mortality rate (F) was computed from creel survey estimates of harvest by age and the mark–recapture abundance estimates (Johnson et al. 1992a; Chapter 6). Total mortality rate (Z) could not be estimated from a catch-curve analysis because of highly variable recruitment; thus the natural mortality rate (M) could not be estimated by difference ($M = Z - F$). Instead, natural mortality rates from the literature were used.
4. Diet (Chapter 11) was determined from a sampling program, stratified by predator size and season (Chapter 3). Energy density of predator and prey was also required for bioenergetics modeling (Chapter 12).
5. Spatial distribution and thermal history was determined from seasonally and depth- stratified gill-net surveys and radio telemetry (Chapter 14).

Each of the demographic, diet, and distribution inputs to a population-level analysis of consumption demand are associated with some degree of uncertainty. A formal treatment of this issue could involve a Monte Carlo simulation of key sources of error in input data. The parameters of the bioenergetics model itself are not known precisely either, but this uncertainty has been addressed (e.g., Bartell et al. 1986). In many cases uncertainty in field data probably exceeds that of the model.

Fisheries scientists should consider the degree of precision required for the questions at hand and allocate field sampling effort accordingly. For instance, fish abundance is likely to be much more

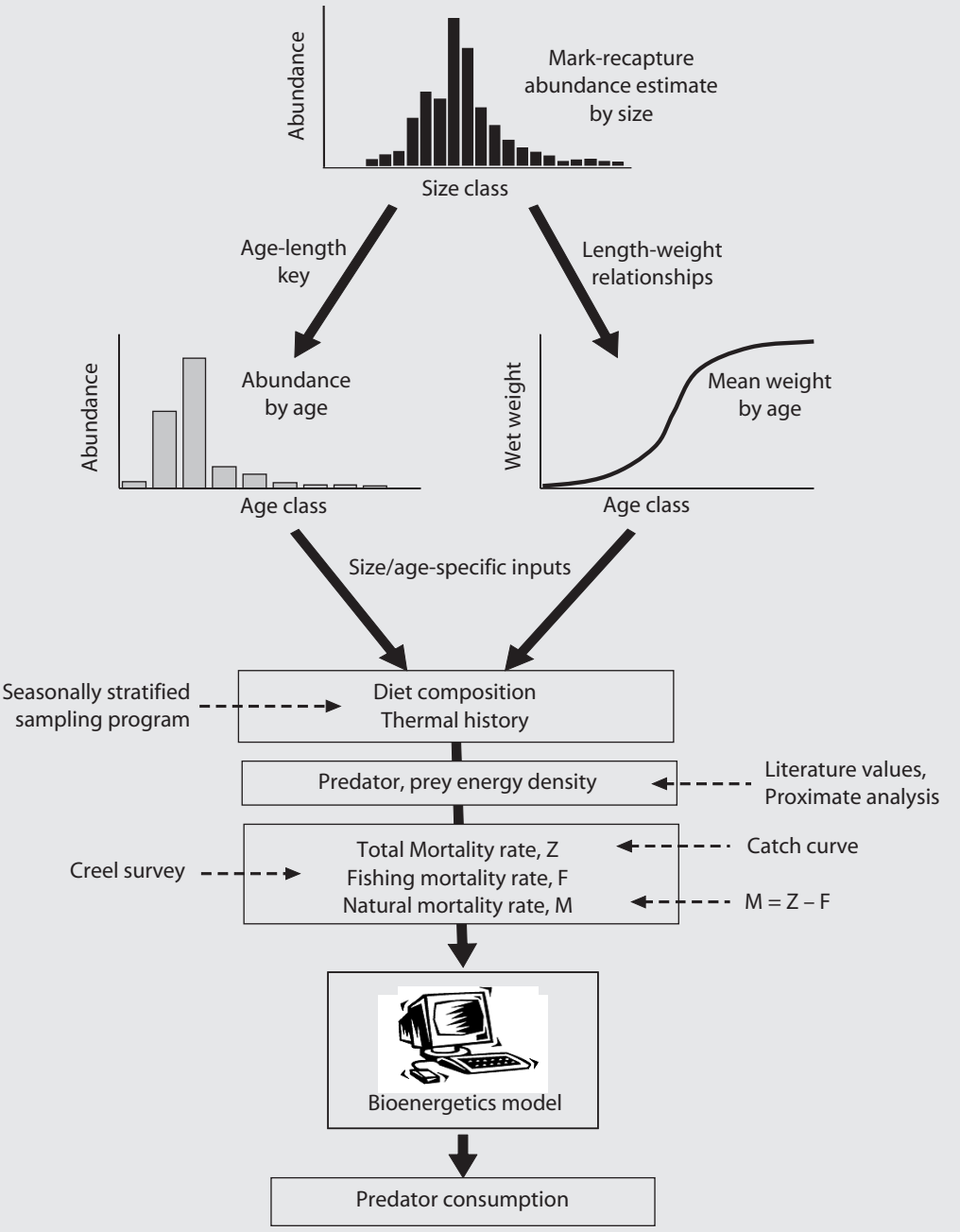


Figure Schematic of analytical methods used to generate data input to bioenergetics model, which, in turn, were used to estimate the biomass of prey consumed by predator populations under various assumptions about predator stocking and mortality rates.

(Box continues)

Box 16.10 *(continued)*

difficult to estimate precisely than is growth rate, especially in large systems. The level of effort devoted to estimating predator diet composition should also be suited to the question of interest. In some studies it may be sufficient to determine the fraction of fish versus invertebrates in the diet, which could be known with much less sampling effort than if species level diet composition were required.

Major Findings

The diets of piscivores varied seasonally and with the size of the predator. Age-0 sunfishes were the dominant prey for both piscivores in the fall (Johnson et al. 1992b). Yellow perch were important to the diet of northern pike and walleye greater than 304 mm in total length (TL), and the relative contribution of yellow perch to their diets depended on year-class strength of yellow perch. Ciscoes were not found in predator diets after the large summer kill in 1987 (Rudstam et al. 1993). However, given the extremely large size structure of the cisco population in the late 1980s, few ciscoes would have been within the gape limits of the piscivores, and distribution studies showed low spatial overlap among the pelagic cisco and the more littoral and benthic walleye and northern pike. Although stocked walleyes were relatively rare in walleye and northern pike guts, projections of total biomass of walleye consumed suggested that piscivory (Polis et al. 1989) was a significant source of mortality for walleye fingerlings. Simulations with an individual-based model of growth of age-0 walleye (Madenjian et al. 1991) predicted that if walleye fingerlings were stocked slightly earlier in the season or 20% larger (60 versus 50 mm TL) at the traditional time, then the proportion of fish in the walleye's diet, and their first year growth, would be enhanced greatly, with an expected increase in first-year survival of the cohort.

Biomass of both piscivore species increased rapidly from less than 1 kg/ha to 4–6 kg/ha throughout the 12-year evaluation period (Lathrop et al. 2002). Annual consumption by walleye and northern pike increased from about 5 kg/ha in 1987 to 10–29 kg/ha (mean 17 kg/ha) during 1989–1998. Premanipulation projections of piscivore biomass and consumption were higher; the shortfall was largely due to predation effects on both small and large piscivores. Low survival rates of stocked walleyes were attributed partly to the piscivory described above, with previously stocked piscivores preying on subsequent cohorts (Johnson et al. 1996). Dramatic increases in predation on larger piscivores by humans and the concomitant increases in fishing mortality during the experiment (Lathrop et al. 2002; Johnson and Carpenter 1994) also limited the effectiveness of the stocking program.

Total planktivore standing stock dropped from 140 kg/ha in 1986 to 50 kg/ha in 1987, primarily due to a large die-off of cisco. Planktivore biomass remained very low (20–40 kg/ha) during 1988–1998. Thus, piscivore consumption represented about half of the planktivore standing stock during the experiment, suggesting that piscivory may have suppressed planktivore populations. However, a large year-class of yellow perch was produced in 1997, indicating that under favorable conditions (abundant zooplankton food resources and weather conditions conducive to spawning) piscivores were not able to control planktivore recruitment.

Dramatically lower planktivory by fishes during the 1990s resulted in a trophic cascade (Carpenter et al. 1985) with higher densities of large cladoceran grazers (mainly *Daphnia pulicaria*), lower algal densities, and improved water clarity (Lathrop et al. 2002). Regardless of the efficacy of the biomanipulation effort to improve water quality, the Lake Mendota experiment stands as a clear example of the importance of a quantitative understanding of fish predator–prey interactions at multiple trophic levels.

Box 16.11 Case Study of Predation Losses Imposed by Lake Trout on Stocked Kokanee in Flathead Lake

Flathead Lake, Montana, historically supported one of the largest kokanee fisheries in North America. The kokanee population crashed in the mid-1980s, coincident with the establishment of high densities of the opossum shrimp, which had invaded from a lake higher in the watershed where they had previously been introduced (Beattie and Clancey 1991). Also coincident with the opossum shrimp increase was a marked increase in lake trout catch and reciprocal declines in native bull trout and westslope cutthroat trout. Federal, state, and tribal managers attempted to re-establish the kokanee population by stocking up to 1 million yearling kokanee each spring, but they needed to know whether any of the proposed stocking strategies would result in reasonable adult returns to the fishery and spawning traps. Of primary concern was whether predation by lake trout would prevent sufficient kokanee recruitment to satisfy a viable fishery and egg-taking operation.

The objective of this study (Beauchamp 1996) was to estimate the predation losses imposed by lake trout on 800,000 yearling kokanee stocked in June. The diet, distribution, size structure, and growth of lake trout and kokanee were obtained by sampling randomly selected locations in five regions of the lake and four depth intervals per location by means of overnight sets with sinking experimental variable-mesh gill nets. Sampling was conducted monthly during May–August, then once per season during fall, winter, and early spring. This provided diet and distribution patterns for lake trout before and after the kokanee release. The proportional weight contribution of each prey type in the diet of lake trout was estimated by season and size-class. The size structure of the lake trout population was corrected for size-selective bias for the array of mesh sizes used (Rudstam et al. 1984; Hansen et al. 1997). Length at age for lake trout was determined by measuring annual growth increments on otoliths, and lengths were converted to weights using a length–weight regression from this population. Abundance of lake trout was estimated in a separate study (Deleray et al. 1999) by use of several methods, including mark–recapture, hydroacoustics, and depletion estimators. Bioenergetics models were used to estimate monthly and seasonal consumption rates for individual lake trout from each age-class by fitting annual size and growth, using the monthly and seasonal change in proportional diet composition for each size-class (Boxes 16.3 and 16.8), and computing the thermal experience from seasonal vertical distribution patterns and temperature profiles (as in Box 16.9). Individual consumption was multiplied by the abundance of lake trout from each age-class (from abundance estimates and size structure data) to expand to seasonal population level predation estimates on kokanee and other key prey (Figure 16.8).

Bioenergetic model simulations suggested that lake trout predation imposed serious losses on the kokanee population in Flathead Lake, accounting for 87% of the total number stocked within the first year of their release. The heaviest predation in 1994 occurred during the first month after stocking 800,000 kokanee (120 mm in fork length [FL]) in June (351,000 kokanee eaten). Kokanee losses during this acute predation period exceeded total predation losses accrued during July–September (263,000 eaten). Lake trout in the 626–750 mm and 501–625 mm (TLs) size-classes were responsible for more than 64% of the estimated predation, and 376–500 mm lake trout consumed another 21% (Figure 16.8). Kokanee disappeared from the diets of progressively larger predators over time, suggesting that the kokanee could rapidly outgrow the smaller, more abundant predators (Figures 16.8, 16.9). The potential change in predation losses was computed as a function of increasing prey size, either through growth or by stocking kokanee at a larger size. The change in predation losse was based on the size-structured abundance of predatory lake trout N_{pred}, the size-specific bioenergetic consumption demand of lake trout, and the predator–prey size relationship,

(Box continues)

Box 16.11 *(continued)*

which indicated that lake trout could consume salmonid-shaped prey up to 50% their own body length (Figure 16.9). In Flathead Lake, the estimated predator population of nearly 900,000 lake trout of 200 mm or greater (TL) was capable of eating 100-mm (FL) kokanee, but the number of potential predators declined sharply as prey size increased (Figure 16.10, upper panel). Under the observed seasonal and size-specific diet composition and consumption rate patterns, lake trout could consume an estimated 3,500 kg or 450,000 100-mm kokanee per month (Figure 16.10, lower panel). As prey size increases, the number of kokanee that could potentially be lost to lake trout predation declines dramatically for two reasons: first, because of the sharp decline in N_{pred} (Figure 16.10, upper panel); second, because prey body mass increases rapidly with increasing length, thus fewer prey are required to satisfy predator demand (Figure 16.10, lower panel). In this case, when the mean size at release was increased from 120 mm (FL) in 1994 to 150 mm in 1995, the initial predation losses should have declined from 351,000 kokanee eaten per month to about 120,000 eaten per month for a nearly threefold reduction in predation rate (Figure 16.10, lower panel). Predation losses would then decline in subsequent months as kokanee grew. If fisheries scientists wished to limit the initial loss rate to only 20,000 kokanee per month, they would have to release kokanee at a length of nearly 240 mm (FL) (the arbitrary vulnerability size threshold in Figures 16.10, 16.11). The monthly change in length frequencies for stocked kokanee in 1995 indicated that all the kokanee stocked in May were below this vulnerability threshold; despite an apparent growth rate of 15 mm/month, 78% were still below the threshold in August, but only 5% were below the threshold in October (Figure 16.11).

Different predation scenarios were modeled to examine the effects of different dietary responses by lake trout, different assumptions about the abundance and size structure of lake trout, and different stocking rates for kokanee. For example, a worst-case scenario could be constructed by assuming the predators fed exclusively on kokanee and achieved their physiological maximum consumption rate (i.e., *p*-value = 1.0 in the Wisconsin bioenergetics model; see Chapter 12) to determine an upper limit to predation losses. When modeling the effects of either an acute predation response (the diet of lake trout was composed of 100% kokanee) or a higher chronic predation response (i.e., the observed initial proportion of kokanee in the diet for each size-class of lake trout was sustained throughout the year), kokanee survival over the first year in the lake declined from 13.2% in the nominal run to 4.6% in the chronic predation scenario, whereas no kokanee survived past midsummer in the acute predation scenario. Lake trout abundance might have been underestimated in model simulations because size and abundance were based on a hydroacoustic survey in August 1995. Because standard hydroacoustic methods cannot detect fish 1 m or less from the bottom, some fraction (e.g., 10–50%) of the predator population might not have been detected. When simulations increased the lake trout population by 10%, survival for a release of 800,000 kokanee dropped from 13.2% to 4.2%; no survival was predicted if the lake trout population was 50% larger than the acoustic-based estimate. Model simulations suggested that the kokanee mitigation program could not meet its harvest or egg-taking goals under the current stocking regime of releasing 800,000–1,000,000 yearling kokanee in late spring.

Predation losses alone accounted for nearly all of the kokanee stocked (87%), but other sources of mortality (other predators or disease) could also reduce adult returns. The prohibitive cost and insufficient hatchery capacity prevented the production of more or larger kokanee to reduce predator demand enough to achieve acceptable and sustainable egg-taking and harvest goals. Based on this analysis, the kokanee mitigation program was terminated because of the unsustainably high predation losses.

consumption per standard unit of a size-structured predator population. For instance, we could create a standard population of 1,000 predators, varying from the youngest to the oldest age-classes that ate the prey species of interest. These 1,000 predators could be allocated into size-classes in proportion to the size structure observed in the population. Consumption by individuals in each age-class would be multiplied by the corresponding number of predators allocated to each age-class from the pool of 1,000 predators to estimate total predation by the size-structured population of 1,000 predators and the relative magnitude of predation exerted by each age- or size-class (Beauchamp et al. 1995; Beauchamp and Van Tassell 2001). Predation losses could then be reported in terms of numbers or biomass of prey consumed per 1,000 predators per year (or over other time scales). Although predator abundance estimates might be lacking, fisheries scientists often have some sense of their abundance, at least within an order of magnitude. Given this information, fisheries scientists can decide whether predation rates are severe enough to warrant further attention. If so, they will either have sufficient information to proceed with management actions or the rationale for justifying further examination into the abundance or dynamics of the predator population.

16.6.2.7 Predation versus Prey Supply

Population level predation rates can be compared with the abundance, biomass, or production of prey populations to determine whether predation represents a significant source of mortality for prey (Kitchell and Crowder 1986; Stewart and Ibarra 1991), the prey represent a sustainable source of food for the predator (Ney 1990; Cyterski et al. 2003), or potential bottlenecks in prey supply might develop during particular periods or locations (Johannsson et al 1994; Rand et al. 1995; Beauchamp et al. 2004). If growth or reproduction significantly alters the abundance or biomass of prey during the period of interest, then it may be more relevant to compare predation to prey production rather than to prey biomass (Figure 16.12). Comparing predation losses to the biomass of available and vulnerable prey represents a more severe estimate of predation mortality and thus provides a more conservative basis for managing the impacts of predation on sensitive prey species than if prey production were included in the analysis (Figure 16.12).

16.6.3 Prey Encounter Rate Models

These models combine the search volume of a predator with the densities of prey that overlap in time and space to estimate the encounter rate for the fraction of prey that are actually available (Box 16.2) during foraging periods (Gerritsen and Strickler 1977). Encounter rates can be calculated separately for different temporal–spatial cells to account for variability in factors that influence prey detection limits or localized differences in prey density. Encounter rate models are conceptually attractive because they link localized environmental conditions and prey densities to foraging success at temporal and spatial scales that are relevant to predators and consistent with primary sensory mechanisms involved in prey detection (e.g., visual, tactile, chemical, pressure, electrical, and sound).

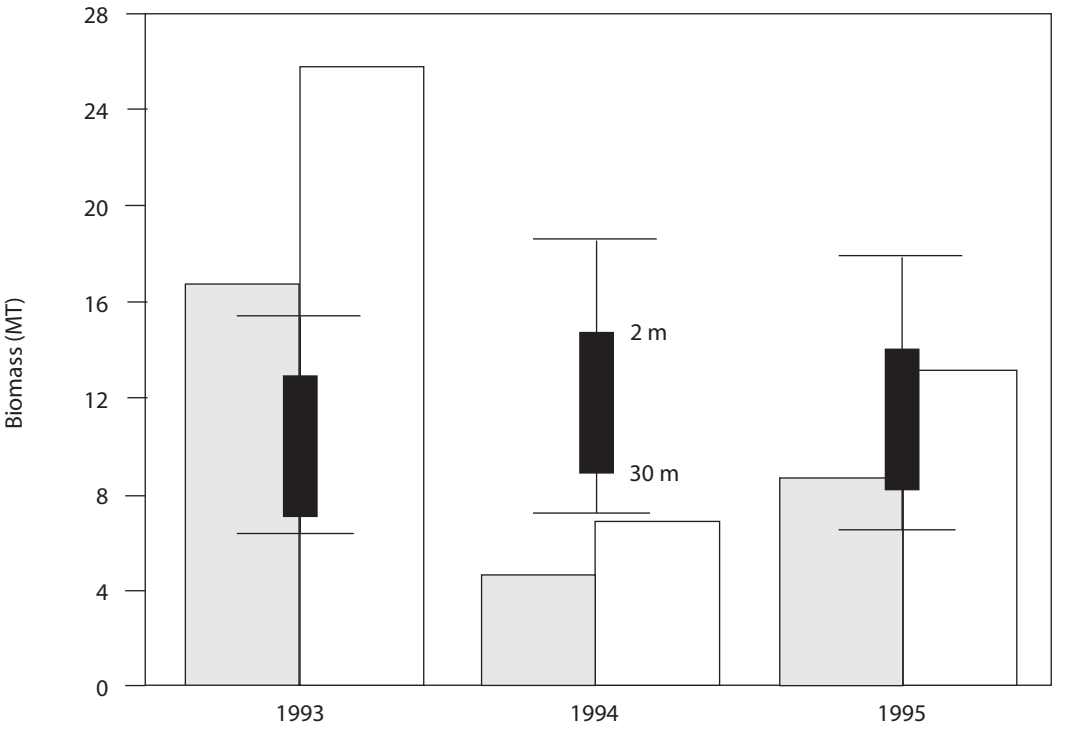


Figure 16.12 Standing stock (metric tons [MT], gray bars) and production (white bars) of rainbow smelt and walleye consumption (vertical lines) of rainbow smelt in Horsetooth Reservoir, Colorado. Uncertainty in walleye consumption is shown as heavy vertical lines representing the range of consumption under two hypothesized thermal regimes (temperatures measured at 2 m and 30 m) and thin vertical lines representing the 95% confidence interval around walleye abundance (adapted from Johnson and Goettl 1999).

Encounter rates can be equated to consumption in some cases; however, a probability of consumption would generally be applied to encounters. These probabilities could either be fixed proportions or functions of other factors that mediate capture success following detection. Encounter rates are essential elements of optimal foraging models and, though generally ignored, should also be incorporated into estimates of prey selectivity and functional responses (Koski and Johnson 2002) in environments where overlap between predators and prey vary through time and space. This approach is particularly valuable when applied to systems in which prey densities and detection capabilities vary through time or space. Encounter rate and bioenergetics models have been combined to estimate the profitability of various feeding positions in streams by drift-feeding fishes (Hughes and Dill 1990; Hill and Grossman 1993), growth potential of piscivores in lakes or estuaries (Brandt et al. 1992; Goyke and Brandt 1993), growth or survival in individual-based models (Breck 1993; DeAngelis et al. 1993), and variation in prey

encounter rates by pelagic planktivores (Mason and Patrick 1993) or piscivores with time and depth in lakes and reservoirs (Beauchamp et al. 1999; Mazur and Beauchamp 2006; Box 16.12).

16.6.4 Community and Ecosystem Level Models

Because many fish species and humans are such potent predators, predation (or harvest) is a central process in the structure and function of aquatic ecosystems. Most fishes, and indeed all harvested species, are both predators and prey. Human interference with natural predator–prey systems may have sweeping and often unexpected effects on food webs and ecosystems. Overharvest of apical predators can allow less desirable competing species to expand (e.g., Fogarty and Murawski 1998) or in some cases may lead to a shift in community structure (Scheffer et al. 2001). Release of prey species from predation as top predators are overexploited can promote depensatory recruitment in the predator as expanding prey species compete with or prey upon juvenile predators (Walters and Kitchell 2001; Post et al. 2002). On the other hand, overprotection of predators in recreational fisheries may contribute to unsustainable management strategies when predation rates exceed prey populations' replacement rates (Johnson and Martinez 2000). And finally, overharvest of prey species may have unintended consequences for higher trophic levels (Cury et al. 2000).

Clearly a more inclusive and holistic view is needed to understand and manage fisheries better, but with holism can come greater complexity. Analytical and statistical tools developed for single species or simple predator–prey systems may not be adequate when expanding the fisheries scientist's purview to the ecosystem scale. A modeling package developed by University of British Columbia's Fishery Centre, Ecopath with Ecosim (Christensen 2001), is an exciting and widely used (Christensen and Pauly 1993) framework to cope with complexity of fished ecosystems. Basic analytical features of the package are shown in Box 16.13.

16.7 SUMMARY

Predator–prey interactions can be studied at the level of individual predators and prey up through ecosystem level effects. Investigations are often most effective when integrating field sampling, natural or laboratory experiments, and modeling in an interactive or complementary fashion. The question and temporal or spatial scale of interest will determine the most appropriate mix of these methods. Body size relationships, variability of processes at different temporal and spatial scales, and effects of habitat and environmental conditions consistently emerge as important factors affecting predator–prey interactions and should be considered as a conceptual framework for addressing any question of interest. Interpolation or extrapolation errors can be minimized by analyzing interactions at appropriate scales and by stratifying along important dimensions of variability for the process of interest.

Box 16.12 Visual Encounter Rate Model

Pelagic piscivores (e.g., salmons, trouts, walleye, and striped bass) feed visually in lakes and reservoirs where visual foraging conditions can be very dynamic. Visual search volumes change as photic conditions vary by depth and time of day. Prey densities also change dramatically by depth and time as prey fishes undergo diel vertical migrations. Under these conditions, visual encounter rates need to be calculated separately for different time and depth intervals.

Beauchamp et al. (1999) modeled prey encounter rates, $ER_{z,t}$ (prey/h), for each depth z and diel period t as the product of depth and temporally explicit search volumes, $SV_{z,t}$, and the vertical density distribution of prey fishes, $PD_{z,t}$, which was obtained for each 5-m depth interval from hydroacoustic surveys during each diel period t (daylight, dusk, and night).

$$ER_{z,t} = SV_{z,t} \cdot PD_{z,t}.$$

Within each depth and time cell, search volume was modeled as a cylinder with length equal to the average swimming speed, SS_t (m/·s), of predators (from telemetry and laboratory studies in the literature) during each diel period multiplied by the duration of each period:

$$SV_{z,t} = SS_t \cdot \pi \cdot RD_{z,t}^2.$$

The circular cross section of the cylinder has a radius equal to the reaction distance, $RD_{z,t}$ (m), to prey. Reaction distance changes as a function of light intensity, and light changes with depth and diel period; RD is further reduced by turbidity, which can change seasonally and among locations. For an ambient light intensity, $I_{z,t}$, of 17.8 lx or less,

$$RD = 0.120 \cdot I_{z,t}^{0.4747} \cdot NTU^{-0.624}.$$

For an $I_{z,t}$ greater than 17.8 lx,

$$RD = RD_{max} = 0.478 \cdot NTU^{-0.624}.$$

"Clear water" has a minimum turbidity of 0.3 nephelometric turbidity unit (NTU). Ambient light intensity declines exponentially with depth z and time t:

$$I_{z,t} = I_{0,t} \cdot e^{z \cdot -k},$$

where $I_{0,t}$ is surface light intensity (lux) at time t, and k is the light extinction coefficient (m^{-1}). Surface light intensity can be measured directly or approximated using a computer program by Janiczcek and deYoung (1987).

The visual encounter rate model described above was applied to the diel distribution patterns of prey in Alturas Lake, Idaho (Figure 16.1), using data inputs for the model summarized in the Table below. We assumed that turbidity (0.34 NTU) remained constant throughout the water column. The light extinction coefficient $k = -0.1535$ was applied to the average surface light intensity during daylight, mid-dusk, and night periods to estimate light intensity within each depth × time cell. Swimming speeds (m/h) for piscivores were computed for each diel period from telemetry and laboratory results reported in the literature (Henderson and Northcote 1985; Beauchamp et al. 1999; Baldwin et al. 2002).

Table Inputs for the visual encounter rate model applied to Alturas Lake, Idaho. Given are surface light intensity, I_0; light extinction coefficient, k; nephelometric turbidity unit, NTU; hours in each light period; and the swimming speed, SS, of piscivorous trouts.

Parameter	Daylight	Dusk	Night
I_0(lx)	35,842.29	11.326	0.014
k (m^{-1})	–0.1535	–0.1535	–0.1535
NTU	0.34	0.34	0.34
h/period	14	3	7
SS (m/h)	1,062	846	144

Based on the conditions summarized in the table above, light levels were computed for the midpoint of each 5-m depth interval. Reaction distances were computed for the light level at each depth interval with turbidity of 0.34 NTU. Search volumes (m^3/h) were applied to depth-specific densities of the smaller two size-classes of kokanee (3–18 cm TL; Figure 16.1) obtained by hydroacoustic and midwater trawl surveys (Beauchamp et al. 1997). For each diel period, prey encounter rates were estimated by multiplying search volume and the associated prey density for each depth cell. The results of these calculations are presented below in separate tables for each diel period.

Table Estimates of prey encounter rates for each diel period by 5-m depth intervals.

Depth interval (m)	Mid-interval depth (m)	Light (lx)	Reaction distance, RD (m)	Search volume, SV (m^3/h)	Prey density, PD (prey/1,000m^3)	Encounter rate, ER (prey/h)
			Daylight			
0–5	2.5	24,419.071	0.94	2,929.8	0.000	0.000
5–10	7.5	11,334.329	0.94	2,929.8	0.537	1.573
10–15	12.5	5,260.929	0.94	2,929.8	0.179	0.524
15–20	17.5	2,441.907	0.94	2,929.8	0.716	2.097
20–25	22.5	1133.433	0.94	2929.8	0.358	1.048
25–30	27.5	526.093	0.94	2929.8	0.358	1.048
30–35	32.5	244.191	0.94	2929.8	0.358	1.048
35–40	37.5	113.343	0.94	2929.8	0.358	1.048
40–45	42.5	52.609	0.94	2929.8	1.224	3.586
45–50	47.5	24.419	0.94	2929.8	1.224	3.586
50–60	55.0	7.722	0.62	1285.8	0.000	0.000
			Dusk			
0–5	2.5	7.716	0.62	1023.6	0.000	0.000
5–10	7.5	3.582	0.43	493.9	0.918	0.454
10–15	12.5	1.662	0.30	238.3	0.566	0.135
15–20	17.5	0.772	0.21	115.0	1.997	0.230
20–25	22.5	0.358	0.14	55.5	2.468	0.137
25–30	27.5	0.166	0.10	26.8	1.146	0.031
30–35	32.5	0.077	0.07	12.9	0.544	0.007
35–40	37.5	0.036	0.05	6.2	0.544	0.003

(Box continues)

Box 16.12 *(continued)*

Table (*continued*)

Depth interval (m)	Mid-interval depth (m)	Light (lx)	Reaction distance, RD (m)	Search volume, SV (m^3/h)	Prey density, PD (prey/1,000m^3)	Encounter rate, ER (prey/h)
			Dusk (*continued*)			
40–45	42.5	0.017	0.03	3.0	0.224	0.001
45–50	47.5	0.008	0.02	1.5	0.224	0.000
50–60	55.0	0.002	0.01	0.5	0.000	0.000
			Night			
0–5	2.5	0.010	0.03	0.3	0.000	0.000
5–10	7.5	0.005	0.02	0.1	0.544	0.000
10–15	12.5	0.002	0.01	0.1	0.492	0.000
15–20	17.5	0.001	0.01	0.0	2.542	0.000
20–25	22.5	0.000	0.01	0.0	5.843	0.000
25–30	27.5	0.000	0.00	0.0	4.339	0.000
30–35	32.5	0.000	0.00	0.0	2.069	0.000
35–40	37.5	0.000	0.00	0.0	2.069	0.000
40–45	42.5	0.000	0.00	0.0	0.892	0.000
45–50	47.5	0.000	0.00	0.0	0.892	0.000
50–60	55.0	0.000	0.00	0.0	0.000	0.000

Interpretation

High water transparency and associated low light extinction maintained maximum reaction distances (0.94 m) and search volumes (2,929.8 m^3/h) down to 50 m during daylight. Corresponding kokanee densities in the upper water column were low during daylight compared with dusk and night periods, and the highest daylight density occurred below 40 m (Figure 16.1, table above). Kokanee avoided the upper water column during daylight where predator densities were at least four times higher (particularly densities of rainbow trout and northern pikeminnow), and the visual search volumes of the piscivores were maximized. During daylight, the prey encounter rates for pelagic piscivores varied from 0.5 to 2.1 kokanee/h at depths above 40 m but increased to 3.6 kokanee/h below 40 m. Prey encounter rates during dusk and night periods were considerably lower (0.00–0.45 kokanee/h) than during daylight despite markedly higher prey fish densities in the water column. At dusk, the highest encounter rates occurred at 5-10 m, whereas the highest prey density was at 15–25 m. The model predicted no prey encounters at night at any depth.

These analyses demonstrated that kokanee reduced predation risk by undergoing diel vertical migrations. The fraction of prey fish actually available to piscivores was considerably lower than the abundance measured by standard assessment methods, and prey availability changed with time and depth. An important insight from this analysis was that prey encounters were relatively rare events. The visual encounter rate model enabled a quantitative evaluation of predation risk by prey fish and foraging opportunities by the piscivores. This approach can compare the potential impact of piscivores in waters of different transparency, productivity, and predator–prey assemblages (Beauchamp et al. 1999; Mazur and Beauchamp 2006). This could be a useful tool for evaluating the feasibility of introducing or enhancing predator or prey species in candidate waters. Further refinements of this approach could incorporate species-specific differences in reaction distance, swimming speed, and capture success after encounter, potentially as functions of light, turbidity, temperature, or other factors (Sweka and Hartman 2001; DeRobertis et al. 2003; Mazur and Beauchamp 2003, 2006).

Box 16.13 Framework of Ecopath with Ecosim Model for Fished Ecosystems

Ecopath Master Equation I

$$B_i \cdot (P/B)_i \cdot EE_i = Y_i + \Sigma B_j \cdot (Q/B)_j \cdot DC_{ij},$$

where for each functional group *i*,
B_i = biomass of *i*;
B_j = biomass of *j* consumers of *i*;
$(P/B)_i$ = production to biomass ratio;
EE_i = fraction of production consumed or harvested (ecotrophic efficiency);
Y_i = biomass harvested (or otherwise lost from system);
$(Q/B)_j$ = food consumed per unit biomass of *j*; and
DCij = contribution of *i* to diet of *j*.

Ecosim

$$\frac{dB_i}{dt} = g_i \cdot \sum Q_{ij} - \sum Q_{ji} + I_i - (M_i + F_i + E_i) \cdot B_i,$$

where,
dB/dt = rate of change in biomass;
B_i = biomass of *i*;
g = growth efficiency;
F = fishing mortality rate;
M = natural mortality rate (excluding predation);
E = emigration rate;
I = immigration rate; and
Qij(*Qji*) = consumption rate of type *i* (or *j*) biomass by type *j* (or *i*) organisms.

Ecopath Master Equation II

$$Q = P + U + R,$$

where
Q = consumption;
P = production;
U = unassimilated food; and
R = respiration.

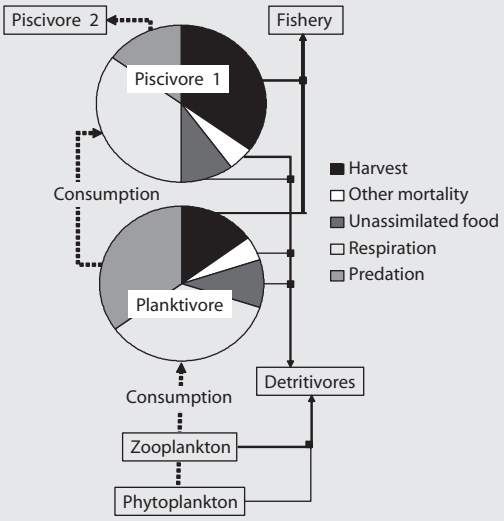


Figure Model of a trophic system. Pies represent biomasses of two fish species, one a piscivore and the other a planktivore, and arrows represent the flows of biomass in the system.

Ecopath (Polovina 1984; Christensen and Pauly 1992; Christensen 2001) organizes biomasses and flows into a static (i.e., baseline or historical average) picture of the ecosystem based on principles of mass-balance and thermodynamics as constraints. The model is solved as a system of simultaneous linear equations. A simple model of a trophic system is represented in the figure above; the pies represent biomasses of two fish species, one a piscivore and the other a planktivore, and arrows represent the flows of biomass in the system. The configuration of the food web and parameterization of flows can be facilitated by stable isotope analysis (e.g., Saito et al. 2001; see also Chapter 11). Both fish are harvested and both are prey for other species. Other components of the food web are simplified for clarity. By converting to differential equations and making the model dynamic, Ecosim (Walters et al. 1997; Pauly et al. 2000) allows the user to examine the trophic implications of a variety of fisheries policy options. Recent improvements to the model (Walters et al. 2000) incorporate compensatory responses in fish populations arising from changes in prey supply that may be direct or indirect effects of harvest. Ecosim can be a useful tool for fisheries scientists concerned about multi-species implications of a fishery management policy and for fisheries scientists who want to design adaptive management experiments at the ecosystem scale.

16.8 REFERENCES

Abbey, D. H., and W. C. MacKay. 1991. Predicting growth of age-0 yellow perch from measures of whole lake productivity. Freshwater Biology 26:519–525.

Anderson, R. O., and A. S. Weithman. 1978. The concept of balance for coolwater fish populations. Pages 371–381 *in* R. L. Kendall, editor. Selected coolwater fishes of North America. American Fisheries Society, Special Publication 11, Bethesda, Maryland.

Appenzeller, A. R., and W. C. Leggett. 1995. An evaluation of light-mediated vertical migration of fish based on hydroacoustic analysis of the diel vertical migration by rainbow smelt *(Osmerus mordax)*. Canadian Journal of Fisheries and Aquatic Sciences 52:504–511.

Atema, J. 1980. Chemical sense, chemical signals and feeding behavior in fishes. Pages 57–101 *in* J. E. Bardach, J. J. Magnuson, R. C. May, and J. M. Reinhart, editors. Fish behavior and its use in the capture and culture of fishes. ICLARM International Center for Living Aquatic Resource Management) Conference Proceedings 5, Manila, Philippines.

Baldwin, C. M., D. A. Beauchamp, and C. P. Gubala. 2002. Seasonal and diel distribution and movement of cutthroat trout from ultrasonic telemetry. Transactions of the American Fisheries Society 131:143–158.

Baldwin, C. M., D. A. Beauchamp, and J. J. Van Tassell. 2000. Bioenergetic assessment of temporal food supply and consumption demand by salmonids in Strawberry Reservoir food web. Transactions of the American Fisheries Society 129:429–450.

Bartell, S. M., J. E. Breck, R. H. Gardner, and A. L. Brenkert. 1986. Individual parameter perturbation and error analysis of fish bioenergetics models. Canadian Journal of Fisheries and Aquatic Sciences 43:160–168.

Beattie, W. D., and P. T. Clancey. 1991. Effects of *Mysis relicta* on the zooplankton community and kokanee population of Flathead Lake, Montana. Pages 39–48 *in* T. P. Nesler and E. P. Bergerson, editors. Mysids in fisheries: hard lessons from headlong introductions. American Fisheries Society, Symposium 9, Bethesda, Maryland.

Beauchamp, D. A. 1990. Seasonal and diel food habits of rainbow trout stocked as juveniles in Lake Washington. Transactions of the American Fisheries Society 119:475–482.

Beauchamp, D. A. 1994. Spatial and temporal dynamics of piscivory: implications for food web stability and the transparency of Lake Washington. Lake and Reservoir Management 9(1):151–154.

Beauchamp, D. A. 1996. Estimating predation losses under different lake trout population sizes and kokanee stocking scenarios in Flathead Lake. Final Report to Montana Department of Fish, Wildlife, and Parks, Kalispell.

Beauchamp, D. A., C. M. Baldwin, J. L. Vogel, and C. P. Gubala. 1999. Estimating diel, depth-specific foraging with a visual encounter rate model for pelagic piscivores. Canadian Journal of Fisheries and Aquatic Sciences 56(Supplement 1):128–139.

Beauchamp, D. A., and eight coauthors. 2004. Temporal–spatial dynamics of early feeding demand and food supply of sockeye salmon fry in Lake Washington. Transactions of the American Fisheries Society 133:1014–1032.

Beauchamp, D. A., M. G. LaRiviere, and G. L. Thomas. 1995. An evaluation of competition and predation as limits to juvenile kokanee and sockeye salmon production in Lake Ozette, Washington. North American Journal of Fisheries Management 15:193–207.

Beauchamp, D. A., and seven coauthors. 1997. Hydroacoustic assessment of abundance and diel distribution of sockeye salmon and kokanee in the Sawtooth Valley Lakes, Idaho. North American Journal of Fisheries Management 17:253–267.

Beauchamp, D. A., D. J. Stewart, and G. L. Thomas. 1989. Corroboration of a bioenergetics model for sockeye salmon. Transactions of the American Fisheries Society 118:597–607.

Beauchamp, D. A., and J. J. Van Tassell. 2001. Modeling trophic interactions of bull trout in Lake Billy Chinook, Oregon. Transactions of the American Fisheries Society 130: 204–216.

Beauchamp, D. A., S. A. Vecht, and G. L. Thomas. 1992. Temporal, spatial, and size-related foraging of wild cutthroat trout in Lake Washington. Northwest Science 66:149–159.

Beckman, B. R., D. A. Larsen, B. Lee-Pawlak, S. Moriyama, and W. W. Dickoff. 1998. Insulin-like growth factor-I and environmental modulation of growth during smoltification of spring Chinook salmon *(Oncorhynchus tschawytscha)*. General and Comparative Endocrinology 109:325–333.

Begon, M., J. L. Harper, and C. R. Townsend. 1996. Ecology: individuals, populations and communities, 3rd edition. Blackwell Scientific Publications, Oxford, UK.

Bell, W. J. 1990. Searching behaviour: the behavioural ecology of finding resources. Chapman and Hall. New York.

Boisclair, D. 1992. An evaluation of the stereocinematographic method to estimate fish swimming speed. Canadian Journal of Fisheries and Aquatic Sciences 49:523–531.

Boisclair, D., and W. C. Leggett. 1989a. Among population variability of fish growth: I. Influence of quantity of food consumed. Canadian Journal of Fisheries and Aquatic Sciences 46:457–467.

Boisclair, D., and W. C. Leggett. 1989b. Among population variability of fish growth: II. Influence of fish growth. Canadian Journal of Fisheries and Aquatic Sciences 46:468–482.

Boisclair, D., and W. C. Leggett. 1989c. Among population variability of fish growth: III. Influence of fish community. Canadian Journal of Fisheries and Aquatic Sciences 46:1539–1550.

Boisclair, D., and W. C. Leggett. 1989d. Importance of activity in bioenergetics models applied to actively foraging fish. Canadian Journal of Fisheries and Aquatic Sciences 46:1859–1867.

Boisclair, D., and W. C. Leggett. 1991. If computers could swim or fish could be programmed: a response to comments by Hewett et al. (1991). Canadian Journal of Fisheries and Aquatic Sciences 48:1337–1344.

Bowen, S. H. 1996. Quantitative description of the diet. Pages 513–522 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques. 2nd edition. American Fisheries Society. Bethesda, Maryland.

Brandt, S. B. 1996. Acoustic assessment of fish abundance and distribution. Pages 385–432 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques. 2nd edition. American Fisheries Society, Bethesda, Maryland.

Brandt, S. B., and K. J. Hartman. 1993. Innovative approaches with bioenergetics models—future applications to fish ecology and management. Transactions of the American Fisheries Society 122:731–735.

Brandt, S. B., D. M. Mason, and E. V. Patrick. 1992. Spatially-explicit models of fish growth rate. Fisheries (Bethesda) 17(2):23–35.

Brandt, S. B., and six coauthors. 1991. Acoustic measures of abundance and size of pelagic planktivores in Lake Michigan. Canadian Journal of Fisheries and Aquatic Sciences 48:894–908.

Breck, J. E. 1993. Foraging theory and piscivorous fish: are forage fish just big zooplankton? Transactions of the American Fisheries Society 122:902–911.

Breck, J. E., and M. J. Gitter. 1983. Effect of fish size on the reactive distance of bluegill *(Lepomis macrochirus)* sunfish. Canadian Journal of Fisheries and Aquatic Sciences 40:162–167.

Bremigan, M. T., and R. A. Stein. 1994. Gape-dependent larval foraging and zooplankton size: implications for fish recruitment across systems. Canadian Journal of Fisheries and Aquatic Sciences 51:913–922.

Brett, M. T., and C. R. Goldman. 1996. A meta-analysis of the freshwater trophic cascade. Proceedings of the National Academy of Sciences of the USA 93(15):7723–7726.

Brett, J. R., and T. D. D. Groves. 1979. Physiological energetics. Pages 279–352 *in* W. S. Hoar, D. J. Randall, and J. R. Brett, editors. Fish physiology, volume 8. Academic Press, New York.

Brodeur, R. D., R. C. Francis, and W. G. Pearcy. 1992. Food consumption of juvenile coho *(Oncorhynchus kisutch)* and chinook *(O. tshawytscha)* on the continental shelf off Washington and Oregon. Canadian Journal of Fisheries and Aquatic Sciences 49:1670–1685.

Brooks, J. L., and S. I. Dodson. 1965. Predation, body size, and composition of plankton. Science 150:28–35.

Bryan, S. D., C. A. Soupir, W. G. Duffy, and C. E. Freiburger. 1996. Caloric densities of three predatory fishes and their prey in Lake Oahe, South Dakota. Journal of Freshwater Ecology 11:153–161.

Buijse, A. D., and R. P. Houthuijzen. 1992. Piscivory, growth, and size-selective mortality of age 0 pikeperch *(Stizostedion lucioperca)*. Canadian Journal of Fisheries and Aquatic Sciences 49:894–902.

Burczynski, J. J., and R. L. Johnson. 1986. Application of dual-beam acoustic survey techniques to limnetic populations of juvenile sockeye salmon *(Oncorhynchus nerka)*. Canadian Journal of Fisheries and Aquatic Sciences 43:1776–1788.

Burczynski, J. J., P. H. Michaletz, and G. M. Marrone. 1987. Hydroacoustic assessment of the abundance and distribution of rainbow smelt in Lake Oahe. North American Journal of Fisheries Management 7:106–116.

Burke, B. J., and J. A. Rice. 2002. A linked foraging and bioenergetics model for southern flounder. Transactions of the American Fisheries Society 131:120–131.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York.

Carpenter, S. R. 1988. Complex interactions in lake communities. Springer-Verlag. New York.

Carpenter, S. R., and J. F. Kitchell. 1988. Consumer control of lake productivity. BioScience 38:764–769.

Carpenter, S. R., and J. F. Kitchell. 1993. The trophic cascade in lakes. Cambridge University Press, New York.

Carpenter, S. R., J. F. Kitchell, and J. R. Hodgson. 1985. Cascading trophic interactions and lake productivity. BioScience 35:634–639.

Cartwright, M. A., D. A. Beauchamp, and M. D. Bryant. 1998. Quantifying cutthroat trout (*Oncorhynchus clarki*) predation on sockeye salmon (*Oncorhynchus nerka*) fry using a

bioenergetics approach. Canadian Journal of Fisheries and Aquatic Sciences 55:1285–1295.

Charnov, E. L. 1976. Optimal foraging: attack strategy of a mantid. American Naturalist 110:141–151.

Christensen, B. 1996. Predator foraging capabilities and prey anti-predator behaviours: pre- versus postcapture constraints on size-dependent predator–prey interactions. Oikos 76:368–380.

Christensen, V. 2001. Download Ecopath with Ecosim Software. Welcome to Ecopath and Ecosim. Available at http://www.ecopath.org/ (August 2002).

Christensen, V., and D. Pauly. 1992. ECOPATH II—a software for balancing steady-state ecosystem models and calculating network characteristics. Ecological Modelling 61: 169–185.

Christensen, V., and D. Pauly, editors. 1993. Trophic models of aquatic ecosystems. ICLARM (International Center for Living Aquatic Resource Management) Conference Proceedings 26, Manila, Philippines.

Clark, C. W., and D. A. Levy. 1988. Diel vertical migrations by juvenile sockeye salmon and the antipredation window. American Naturalist 131:271–290.

Cochran, P. A., and I. R. Adelman. 1982. Seasonal aspects of daily ration and diet of largemouth bass, *Micropterus salmoides,* with an evaluation of gastric evacuation rates Environmental Biology of Fishes 7:265–275.

Collins, N., and S. G. Hinch. 1993. Diel and seasonal variation in foraging activities of pumpkinseeds in an Ontario pond. Transactions of the American Fisheries Society 122:357–365.

Collins, N., S. G. Hinch, and K. A. Baia. 1991. Non-intrusive time-lapse video monitoring of shallow water aquatic environments. Canadian Technical Report of Fisheries and Aquatic Sciences 1821.

Confer, J. I., and P. I. Blades. 1975. Omnivorous zooplankton and planktivorous fish. Limnology and Oceanography 20:571–579.

Conover, W. J., and R. L. Iman. 1981. Rank-transformations as a bridge between parametric and nonparametric statistics. American Statistician 35:124–129.

Cooper, W. E., and L. B. Crowder. 1979. Patterns of predation in simple and complex communities. Pages 257–268 *in* R. H. Stroud and H. Clepper, editors (1979).

Coutant, C. C. 1985. Striped bass, temperature, and dissolved oxygen: a speculative hypothesis for environmental risk. Transactions of the American Fisheries Society 114: 31–61.

Coutant, C. C., R. B. McLean, and D. L. DeAngelis. 1979. Influences of physical and chemical alterations on predator–prey interactions. Pages 57–68 *in* R. H. Stroud and H. Clepper, editors (1979).

Crowder, L. B., J. J. Magnuson, and S. B. Brandt. 1981. Complementarity in the use of food and thermal habitat by Lake Michigan fishes. Canadian Journal of Fisheries and Aquatic Sciences 38:662–668.

Cummins, K. W., and J. C. Wuycheck. 1971. Caloric equivalents for investigations in ecological energetics. Mitteilungen Internationale Vereinigung fur Theoretische und Angewandte Limnologic 18.

Cury, P., and six coauthors. 2000. Small pelagics in upwelling systems: patterns of interaction and structural changes in "wasp-waist" ecosystems. ICES Journal of Marine Science 57:603–618.

Cyterski, M., J. Ney, and M. Duval. 2003. Estimation of surplus biomass of clupeids in Smith Mountain Lake, Virginia. Transactions of the American Fisheries Society 132: 361–370.

DeAngelis, D. L., B. J. Shuter, M. S. Ridgway, and M. Scheffer. 1993. Modeling growth and survival of in an age-0 fish cohort. Transactions of the American Fisheries Society 122:927–941.

Deleray, M., L. Knotek, S. Rumsey, and T. M. Weaver. 1999. Flathead Lake and River system fisheries status report. Montana Department of Fish, Wildlife, and Parks. Dingell-Johnson Report Number F-78-R-1–5, Kalispell.

DeRobertis, A., C. H. Ryer, A. Veloza, and R. D. Brodeur. 2003. Differential effects of turbidity on prey consumption of piscivorous and planktivorous fish. Canadian Journal of Fisheries and Aquatic Sciences 60:1517–1526.

Dettmers, J. M., D. R. DeVries, and R. A. Stein. 1996. Reservoir biomanipulation: quantifying piscivore effects and prey productivity across multiple trophic levels. Transactions of the American Fisheries Society 125:491–504.

DeVries, D. R., and R. V. Frie. 1996. Determination of age and growth. Pages 483–512 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Diana, J. S. 1979. The feeding pattern and daily ration of a top carnivore, the northern pike *(Esox lucius)*. Canadian Journal of Zoology 57:2121–2127.

Dill, L. M., and A. H. G. Fraser. 1984. Risk of predation and the feeding behavior of juvenile coho salmon *(Oncorhynchus kisutch)*. Behavioral Ecology and Sociobiology 16: 65–71.

Eby, L. A., L. G. Rudstam, and J. F. Kitchell. 1995. Predator responses to prey population dynamics: an empirical analysis based on lake trout growth rates. Canadian Journal of Fisheries and Aquatic Sciences 52:1564–1571.

Eck, G. W., and E. H. Brown, Jr. 1985. Lake Michigan's capacity to support lake trout *(Salvelinus namaycush)* and other salmonines: an estimate based on the status of prey populations in the 1970s. Canadian Journal of Fisheries and Aquatic Sciences 42: 449–454.

Eggers, D. M. 1977. Factors in interpreting data obtained by diel sampling of fish stomachs. Journal of the Fisheries Research Board of Canada 34:290–294.

Eggers, D. M. 1978. Limnetic feeding behavior of juvenile sockeye salmon in Lake Washington and predator avoidance. Limnology and Oceanography 23:1114–1125.

Einfalt, L. M., and D. H. Wahl. 1997. Prey selection by juvenile walleye as influenced by prey morphology and behavior. Canadian Journal of Fisheries and Aquatic Sciences 54:2618–2626.

Eklov, P., and S. Diehl. 1994. Piscivore efficiency and refuging prey: the importance of predator search mode. Oecologia 98:344–353.

Eklov, P., and S. F. Hamrin. 1989. Predatory efficiency and prey selection: interactions between pike *Esox lucius*, perch *Perca fluviatilis*, and rudd *Scardinus erythrophthalmus*. Oikos 56:149–156.

Elliott, J. M., and L. Persson. 1978. The estimation of daily rates of food consumption for fish. Journal of Animal Ecology 47:977–991.

Fogarty, M. J., and S. A. Murawski. 1998. Large-scale disturbance and the structure of a marine system: fishery impacts on Georges Bank. Ecological Applications 8(Supplement):S6–S22.

Forseth, T., B. Jonsson, R. Naeumann, and O. Ugedal. 1992. Radioisotope method for estimating food consumption by brown trout *(Salmo trutta)*. Canadian Journal of Fisheries and Aquatic Sciences 49:1328–1335.

Fraser, D. F., and J. F. Gilliam. 1992. Nonlethal impacts of predator invasion: facultative suppression of growth and reproduction. Ecology 73:959–970.

Gabelhouse, D. W., Jr. 1984. A length-categorization system to assess fish stocks. North American Journal of Fisheries Management 4:273–283.

Gerritsen, J., and J. R. Strickler. 1977. Encounter probabilities and community structure in zooplankton: a mathematical model. Journal of the Canadian Fisheries Research Board of Canada 34:73–82.

Gislason, H., and M. M. Sinclair. 2000. Ecosystem effects of fishing—introduction. ICES Journal of Marine Science 57:466–467.

Gotelli, N. J. 1998. A primer of ecology, 2nd edition. Sinauer Associates, Sunderland, Massachusetts.

Goyke, A. P., and S. B. Brandt. 1993. Spatial models of salmonine growth rates in Lake Ontario. Transactions of the American Fisheries Society 122:870–883.

Gregory, R. S. 1994. The influence of ontogeny, perceived risk of predation and visual ability on the foraging behavior of juvenile chinook salmon. Pages 271–284 *in* Stouder et al. (1994).

Hambright, K. D. 1991. Experimental analysis of prey selection by largemouth bass: role of predator width and prey body depth. Transactions of the American Fisheries Society 120:500–508.

Hansen, M. J., D. Boisclair, S. B. Brandt, and S. W. Hewett. 1993. Applications of bioenergetics models to fish ecology and management—where do we go from here? Transactions of the American Fisheries Society 122:1019–1030.

Hansen, M. J., C. P. Madenjian, J. H. Selgeby, and T. E. Hesler. 1997. Gillnet selectivity of lake trout *(Salvelinus namaycush)* in Lake Superior. Canadian Journal of Fisheries and Aquatic Sciences 54:2483–2490.

Hanson. P. C., T. B. Johnson, D. E. Schindler, and J. F. Kitchell. 1997. Fish bioenergetics 3.0. University of Wisconsin, Sea Grant Institute, WISCU-T-97-001, Madison.

Hartman, K. J., and S. B. Brandt. 1995. Estimating energy density of fish. Transactions of the American Fisheries Society 124:347–355.

Hartman, K. J., and F. J. Margraf. 1992. Effects of prey and predator abundances on prey consumption and growth of walleyes in western Lake Erie. Transactions of the American Fisheries Society 121:245–260.

He, X., and J. F. Kitchell. 1990. Direct and indirect effects of predation on a fish community: a whole-lake experiment. Transactions of the American Fisheries Society 119:825–835.

He, E., and W. A. Wurtsbaugh. 1993. An empirical model of gastric evacuation rates for fish and an analysis of digestion in piscivorous brown trout. Transactions of the American Fisheries Society 122:717–730.

Henderson, M. A., and T. G. Northcote 1985. Visual prey detection and foraging in sympatric cutthroat trout (*Salmo clarki*) and Dolly Varden (*Salvelinus malma*). Canadian Journal of Fisheries and Aquatic Sciences 42:785–790.

Hewett, S. W., and B. L. Johnson. 1987. A generalized bioenergetics model of fish growth for microcomputers. University of Wisconsin, Sea Grant Institute, WIS-SG-87-245, Madison.

Hewett, S. W., and B. L. Johnson. 1992. Fish bioenergetics model 2: an upgrade of a generalized bioenergetics model of fish growth for microcomputers. University of Wisconsin, Sea Grant Institute, WIS-SG-92-250, Madison.

Hewett, S. W., C. E. Kraft, and B. L. Johnson. 1991. Consumption, growth, and allometry: a comment on Boisclair and Leggett (1989a, 1989b, 1989c, 1989d). Canadian Journal of Fisheries and Aquatic Sciences 48:1334–1337.

Hill, J., and G. H. Grossman. 1993. An energetic model of habitat use for rainbow trout and rosyside dace. Ecology 74:685–698.

Holling, C. S. 1959. The components of predation as revealed by a study of small mammal predation of the European pine sawfly. Canadian Entomologist 91: 293–320.

Holling, C. S. 1966. The functional response of invertebrate predators to prey density. Memoirs of the Entomological Society of Canada 48:3–86.

Hora, S. C., and W. J. Conover. 1984. The F-statistic in the two-way layout with rank-score transformed data. Journal of the American Statistical Association 79:668–677.

Howick, G. L., and W. J. O'Brien. 1983. Piscivorous feeding behavior of largemouth bass: an experimental analysis. Transactions of the American Fisheries Society 112:508–516.

Hughes, N. F., and L. M. Dill. 1990. Position choice by drift-feeding salmonids: model and tests for Arctic grayling in subarctic mountain streams, interior Alaska. Canadian Journal of Fisheries and Aquatic Sciences 47:2039–2048.

Huston, M. A. 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: synthesis of comments. Ecology 80:1088–1089.

Ibrahim, A. A., and F. A. Huntingford. 1989. Laboratory and field studies of the effect of predation risk on foraging in three-spine sticklebacks *(Gasterosteus aculeatus)*. Behaviour 109:46–57.

Jakobsen, P. J., G. H. Johnsen, and P. Larsson. 1988. Effects of predation risk and parasitism on the feeding ecology, habitat use, and abundance of lacustrine threespine stickleback *(Gasterosteus aculeatus)*. Canadian Journal of Fisheries and Aquatic Sciences 45: 426–431.

Janiczcek, P. M. J., and A. J. deYoung. 1987. Computer programs for sun and moon illuminance with contingent table and diagram. U.S. Naval Observatory Circulation 171:1–32.

Janssen, J. 1997. Comparison of response distance to prey via the lateral line in the ruffe and yellow perch. Journal of Fish Biology 51:921–930.

Jenkins, R. M. 1979. Predator–prey relations in reservoirs. Pages 123–134 *in* R. H. Stroud and H. Clepper, editors (1979).

Jenkins, R. M., and D. I. Morias. 1978. Predator–prey relations in predator-stocking-evaluation reservoirs. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 30(1976):141–157.

Johannsson, O. E., L. G. Rudstam, and D. C. Lasenby. 1994. *Mysis relicta*: assessment of metalimnetic feeding and implications for competition with fish in Lakes Ontario and Michigan. Canadian Journal of Fisheries and Aquatic Sciences 51:2591–2602.

Johnson, B. M., and S. R. Carpenter. 1994. Functional and numerical responses: a framework for fish-angler interactions? Ecological Applications 4:808–821.

Johnson, B. M., and J. P. Goettl, Jr. 1999. Food web changes over fourteen years following introduction of rainbow smelt into a Colorado reservoir. North American Journal of Fisheries Management 19:629–642.

Johnson, B. M., and P. J. Martinez. 2000. Trophic economics of lake trout management in reservoirs of differing productivity. North American Journal of Fisheries Management 20:115–131.

Johnson, B. M., P. J. Martinez, and J. D. Stockwell. 2002. Tracking trophic interactions in coldwater reservoirs using naturally occurring stable isotopes. Transactions of the American Fisheries Society 131:1–13.

Johnson, B. M., and six coauthors. 1992b. Piscivores and their prey. Pages 319–352 *in* J. F. Kitchell, editor. Food web management: a case study of Lake Mendota, Wisconsin. Springer-Verlag, New York.

Johnson, B. M., R. S. Stewart, S. J. Gilbert, C. Luecke, and J. F. Kitchell. 1992a. Forecasting the effects of gamefish stocking and harvest regulations on the planktivore community in a eutrophic lake. North American Journal of Fisheries Management 12:797–807.

Johnson, B. M., M. T. Vogelsang, and R. S. Stewart. 1996. Enhancing a walleye population by stocking: effectiveness and constraints on recruitment. Annales Zoologici Fennici 33:577–588.

Jones, M. L., J. F. Koonce, and R. O'Gorman. 1993. Sustainability of hatchery-dependent salmonine fisheries in Lake Ontario: the conflict between predator demand and prey supply. Transactions of the American Fisheries Society 122:1002–1018.

Juanes, F. 1994. What determines prey size selectivity in piscivorous fishes? Pages 79–100 *in* Stouder et al. (1994).

Juliano, S. A. 1993. Nonlinear curve fitting: predation and functional response curves. Pages 159–182 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Kerfoot, W. C., and A. Sih. 1987. Predation: direct and indirect impacts on aquatic communities. University Press of New England, Hanover, New Hampshire.

Kitchell, J. F. 1979. In search of an optimal environment. Pages 31–34 *in* R. H. Stroud and H. Clepper, editors (1979).

Kitchell, J. F., editor. 1992. Food web management: a case study of Lake Mendota, Wisconsin. Springer-Verlag, New York.

Kitchell, J. F., and L. B. Crowder. 1986. Predator–prey interactions in Lake Michigan: model predictions and recent dynamics. Environmental Biology of Fishes 16:205–211.

Kitchell, J. F., L. A. Eby, X. He, D. E. Schindler, and R. A. Wright. 1994. Predator–prey dynamics in an ecosystem context. Journal of Fish Biology 45(Supplement A):209–226.

Kitchell, J. F., D. J. Stewart, and D. Weininger. 1977. Applications of a bioenergetics model to yellow perch *(Percaflavescens)* and walleye *(Stizostedion vitreum vitreum)*. Journal of the Fisheries Research Board of Canada 34:1992–1935.

Kolar, C. S., and D. H. Wahl. 1998. Daphnid morphology deters fish predators. Oecologia 116:556–564.

Koski, M., and B. M. Johnson. 2002. Functional response of kokanee *(Oncorhynchus nerka)* feeding on *Daphnia* under a range of light intensities. Canadian Journal of Fisheries and Aquatic Sciences 59(4):707–716.

LaBar, G. W. 1993. Use of bioenergetics models to predict the effect of increased lake trout predation on rainbow smelt following sea lamprey control. Transactions of the American Fisheries Society 122: 942–950.

Lane, E. D., M. C. S. Kingsley, and D. E. Thornton. 1979. Daily feeding and food conversion efficiency of the diamond turbot: an analysis based on field data. Transactions of the American Fisheries Society 108:530–535.

Lathrop, R. C., and nine coauthors. 2002. Stocking piscivores to improve fishing and water clarity: a synthesis of the Lake Mendota biomanipulation project. Freshwater Biology 47:2410–2424.

Lima, S. L., and L. M. Dill. 1990. Behavioral decisions made under the risk of predation: a review and prospectus. Canadian Journal of Zoology 68:619–640.

Link, J. S., and L. P. Garrison. 2002. Changes in piscivory associated with fishing induced changes to the finfish community on Georges Bank. Fisheries Research 55:71—86.

Luecke, C., and D. Brandt. 1993. Estimating the energy density of daphnid prey for use with rainbow trout bioenergetics models. Transactions of the American Fisheries Society 122:386–389.

Luecke, C., T. C. Edwards, Jr., M. W. Wengert, Jr., S. Brayton, and R. Schneidervin. 1994. Simulated changes in lake trout yield, trophies, and forage consumption under various slot limits. North American Journal of Fisheries Management 14:14–21.

Luecke, C., J. A. Rice, L. B. Crowder, S. E. Yeo, and F. P. Binkowski. 1990a. Recruitment mechanisms of bloater in Lake Michigan: and analysis of the predatory gauntlet. Canadian Journal of Fisheries and Aquatic Sciences 47:524–532.

Luecke, C., M. J. Vanni, J. J. Magnuson, and J. F. Kitchell. 1990b. Seasonal regulation of *Daphnia* populations by planktivorous fish—implications for the spring clear water phase. Limnology and Oceanography 35:1091–1100.

Luecke, C., and W.A. Wurtsbaugh. 1993. Effects of moonlight and daylight on hydroacoustic estimates of pelagic fish abundance. Transactions of the American Fisheries Society 122:112–120.

Lundvall, D., R. Svanback, L. Persson, and P. Bystrom. 1999. Size-dependent predation in piscivores: interactions between predator foraging and prey avoidance abilities. Canadian Journal of Fisheries and Aquatic Sciences 56:1285–1292.

Luo, J. G., K. J. Hartman, S. B. Brandt, C. F. Cerco, and T. H. Rippetoe. 2001. A spatially-explicit approach for estimating carrying capacity: an application for the Atlantic menhaden *(Brevoortia tyrannus)* in Chesapeake Bay. Estuaries 24(4):545–556.

Madenjian, C. P., B. M. Johnson, and S. R. Carpenter. 1991. Stocking strategies for fingerling walleyes: an individual based model approach. Ecological Applications 1:280–288.

Madenjian, C. P., and D. V. O'Connor. 1999. Laboratory evaluation of a lake trout bioenergetics model. Transactions of the American Fisheries Society 128:802–814.

Mason, D. D., and E. V. Patrick. 1993. A model for the time–space dependence of feeding for pelagic fish populations. Transactions of the American Fisheries Society 122:884–901.

Mayer, C. M., and D. H. Wahl. 1997. The relationship between prey selectivity and growth and survival in a larval fish. Canadian Journal of Fisheries and Aquatic Sciences 54:1504–1512.

Mazur, M. M., and D. A. Beauchamp. 2003. A comparison of visual prey detection among species of piscivorous salmonids: effects of light and low turbidities. Environmental Biology of Fishes 67:397–405.

Mazur, M. M., and D. A. Beauchamp. 2006. Linking piscivory to spatial–temporal distributions of pelagic prey fish with a visual foraging model. Journal of Fish Biology.

Mesa, M. G. 1994. Effects of multiple acute stressors on the predator avoidance ability and physiology of juvenile Chinook salmon. Transactions of the American Fisheries Society 123:786–793.

Mesa, M. G., T. P. Poe, D. M. Gadomski, and J. H. Petersen. 1994. Are all prey created equal? A review and synthesis of differential predation on prey in substandard condition. Journal of Fish Biology 45(Supplement A):81–96.

Milinski, M. 1993. Predation risk and feeding behaviour. Pages 285–305 *in* T. J. Pitcher, editor. Behaviour of teleost fishes. Chapman and Hall, London.

Miller, T. J., L. B. Crowder, J. A. Rice, and E. A. Marschall. 1988. Larval size and recruitment mechanisms in fishes: toward a conceptual framework. Canadian Journal of Fisheries and Aquatic Sciences 45:1657–1670.

Mittlebach, G. 1981. Foraging efficiency and body size: a study of optimal diet and habitat use by bluegills. Ecology 62:1370–1386.

Mittlebach, G. G. 1984. Predation and resource partitioning in two sunfishes (Centrarchidae). Ecology 65:499–513.

Mittlebach, G. G. 1986. Predator-mediated habitat use: some consequences for species interactions. Environmental Biology of Fishes 1:150–169.

Mittlebach, G. G. 1988. Competition among refuging sunfishes and effects of fish density on littoral zone invertebrates. Ecology 69:614–623.

Mittlebach, G. G., and C. W. Osenberg. 1994. Using foraging theory to study trophic interactions. Pages 45–59 *in* Stouder et al. (1994).

Mittlebach, G. G., and L. Persson. 1998. Ontogeny of piscivory and its ecological consequences. Canadian Journal of Fisheries and Aquatic Sciences 55:1454–1465.

Moyle, P. B., and T. Light. 1996a. Biological invasions of fresh water: empirical rules and assembly theory. Biological Conservation 78:149–162.

Moyle, P. B., and T. Light. 1996b. Fish invasions in California: do abiotic factors determine success? Ecology 77:1666–1670.

Nesler, T. P., and E. P. Bergersen, editors. 1991. Mysids in fisheries: hard lessons from headlong introductions. American Fisheries Society, Symposium 9, Bethesda, Maryland.

Ney, J. J. 1990. Trophic economics in fisheries: assessment of demand–supply relationships between predators and prey. Reviews in Aquatic Sciences 2:55–81.

Ney, J. J. 1993. Bioenergetics modeling today—growing pains on the cutting edge. Transactions of the American Fisheries Society 122:736–748.

Norberg, R. A. 1977. An ecological theory on the foraging time and energetics and choice of the optimal food searching method. Journal of Animal Ecology 46:511–529.

Northcote, T. G. 1988. Fish in the structure and function of freshwater ecosystems: a "top-down" view. Canadian Journal of Fisheries and Aquatic Sciences 45:361–379.

O'Brien, W. J., H. I. Browman, and B. I. Evans. 1990. Search strategies in foraging animals. American Scientist 78:152–160.

O'Brien, W. J., N. A. Slade, and G. L. Vinyard. 1976. Apparent size as the determinant of prey selection by bluegill sunfish *(Lepomis macrochirus)*. Ecology 57:1304–1310.

Paine, R. T. 1980. Food webs: linkage, interaction strength and community infrastructure. Journal of Animal Ecology 49:667–685.

Parkinson, E., B. E. Reiman, and L. Rudstam. 1994. Comparison of acoustic and trawl methods for estimating density and age composition of kokanee. Transactions of the American Fisheries Society 123:841–854.

Pauly, D., V. Christensen, and C. Walters. 2000. Ecopath, ecosim, and ecospace as tools for evaluating ecosystem impacts of fisheries. ICES Journal of Marine Science 57:697–706.

Persson, L. 1991. Behavioral response to predators reverses the outcome of competition between prey species. Behavioral Ecology and Sociobiology 28:101–105.

Persson, L., P. Bystrom, and E. Wahlstrom. 2000. Cannibalism and competition in Eurasian perch: population dynamics of an ontogenetic omnivore. Ecology 81:1058–1071.

Persson, L., S. Diehl, L. Johansson, G. Andersson, and S. F. Hamrin. 1991. Shifts in fish communities along the productivity gradient of temperate lakes—patterns and the importance of size-structured interactions. Journal of Fish Biology 38:281–293.

Persson, L., and P. Eklov. 1994. Prey refuges affecting interactions between piscivorous perch and juvenile perch and roach. Ecology 76:70–81.

Peterman, R. M., and M. Gatto. 1978. Estimation of functional responses of predators on juvenile salmon. Journal of the Fisheries Research Board of Canada 35:797–808.

Peterson, B. J., and B. Fry. 1987. Stable isotopes in ecosystem studies. Annual Review of Ecology and Systematics 18:293–320.

Ploskey, G. R., and R. M. Jenkins. 1982. Biomass model for reservoir fish and fish–food interactions, with implications for management. North American Journal of Fisheries Management 2:105–121.

Poe, T. P., H. C. Hansel, S. Vigg, D. E. Palmer, and L. A. Prendergast. 1991. Feeding of predaceous fishes on out-migrating juvenile salmonids in John Day Reservoir, Columbia River. Transactions of the American Fisheries Society 120:405–420.

Polis, G. A., C. A. Myers, and R. D. Holt. 1989. The ecology and evolution of intraguild predation: potential competitors that eat each other. Annual Review of Ecology and Systematics 20:297–330.

Polovina, J. J. 1984. Model of a coral reef ecosystem I. The ECOPATH model and its application to French Frigate Shoals. Coral Reefs 3:1–11.

Pope, J. G. 1991. The ICES multispecies assessment working group: evolution, insights, and future problems. ICES (International Council for the Exploration of the Sea) Marine Science Symposium 193:22–33.

Post, J. R., and eight coauthors. 2002. Canada's recreational fisheries: the invisible collapse? Fisheries (Bethesda) 27(1):6–17.

Pyke, G. H. 1984. Optimal foraging theory: a critical review. Annual Review of Ecology and Systematics 15:523–575.

Rand, P. S., B. F. Lantry, R. O'Gorman, R. W. Owens, and D. J. Stewart. 1994. Energy density and size of pelagic prey fishes in Lake Ontario, 1978–1990—implications for salmonid energetics. Transactions of the American Fisheries Society 123:519–534.

Rand, P. S., D. J. Stewart, B. F. Lantry, and L. G. Rudstam. 1995. Effect of lake-wide planktivory by the pelagic prey fish community in Lakes Michigan and Ontario. Canadian Journal of Fisheries and Aquatic Sciences 52:1546–1563.

Rice, J. A., and P. A. Cochran. 1984. Independent evaluation of a bioenergetics model for largemouth bass. Ecology 65:732–739.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Ricker, W. E. 1992. Back-calculation of fish lengths based on proportionality between scale and length increments. Canadian Journal of Aquatic and Fisheries Sciences 49: 1018–1026.

Rieman, B. E., R. C. Beamesderfer, S. Vigg, and T. P. Poe. 1991. Estimated loss of juvenile salmonids to predation by northern squawfish, walleyes, and smallmouth bass in John Day Reservoir, Columbia River. Transactions of the American Fisheries Society 120:448–458.

Rudstam L. G., R. C. Lathrop, and S. R. Carpenter. 1993. The rise and fall of a dominant planktivore: direct and indirect effects on zooplankton. Ecology 74:303–319.

Rudstam, L. G., J. J. Magnuson, and W. M. Tonn. 1984. Size selectivity of passive fishing gear: a correction for encounter probability applied to gill nets. Canadian Journal of Fisheries and Aquatic Sciences 41:1252–1255.

Ruggerone, G. T., and D. E. Rogers. 1992. Predation on sockeye salmon fry by juvenile coho salmon in Chignik Lakes, Alaska: implications for salmon management. North American Journal of Fisheries Management 12:87–102.

Saito, L., B. M. Johnson, J. Bartholow, and R. B. Hanna. 2001. Assessing ecosystem effects of reservoir operations using food web–energy transfer and water quality models. Ecosystems 4:105–125.

Sandheinrich, M. B., and G. J. Atchison. 1990. Sublethal toxicant effects on fish foraging behavior: empirical versus mechanistic approaches. Environmental Toxicology and Chemistry 9:109–119.

Santucci, V. J., Jr. and D. H. Wahl. 1993. Factors influencing survival and growth of stocked walleye *(Stizostedion vitreum)* in a centrarchid-dominated impoundment. Canadian Journal of Fisheries and Aquatic Sciences 50:1548–1558.

Santucci, V. J., Jr., D. H. Wahl, and T. Storck. 1994. Growth, mortality, harvest, and cost effectiveness of stocked channel catfish in a small impoundment. North American Journal of Fisheries Management 14:781–789.

SAS Institute. 1990. SAS/STAT user's guide, volume 1 and 2, version 6, 4th edition. SAS Institute, Cary, North Carolina.

Savino, J., E. A. Marschall, and R. A. Stein. 1992. Bluegill growth, as modified by plant density: an explanation of underlying mechanisms. Oecologia 89:153–160.

Savino, J., and R. A. Stein. 1982. Predator–prey interaction between largemouth bass and bluegills as influenced by simulated, submersed vegetation. Transactions of the American Fisheries Society 111:255–266.

Schael, D. M., J. A. Rice, and D. J. Degan. 1995. Spatial and temporal distribution of threadfin shad in a southeastern reservoir. Transactions of the American Fisheries Society 124:804–812.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folkes, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature (London) 413:591–596.

Scheiner, S. M. 1993. MANOVA: multiple response variables and multispecies interactions. Pages 94–112 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York.

Schindler, D. E., J. F. Kitchell, and R. Ogutu-Ohwayo. 1998. Ecological consequences of alternative gill net fisheries for Nile perch in Lake Victoria. Conservation Biology 12: 56–64.

Schindler, D. E., and M. D. Scheuerell. 2002. Habitat coupling in lake ecosystems. Oikos 98: 177–189.

Sih, A., and R. D. Moore. 1990. Interacting effects of predator and prey behavior in determining diets. Pages 771–796 *in* R. N. Hughes, editor. Behavioural mechanisms of food selection. NATO Advanced Studies Institute Series G Ecological Sciences Volume 20, Springer-Verlag, Berlin.

Smagula, C. M., and I. R. Adelman. 1982. Day-to-day variation in food consumption by largemouth bass. Transactions of the American Fisheries Society 111:543–548.

Sogard, S. M. 1994. Use of suboptimal habitats by fishes: consequences to growth and survival. Pages 103–131 *in* Stouder et al. (1994).

SPSS. 1999. SPSS base 10.0 for Windows user's guide. SPSS, Chicago.

Stables, T. B., and G. L. Thomas. 1992. Acoustic measurement of trout distributions in Spada Lake, Washington, using stationary transducers. Journal of Fish Biology 40:191–203.

Stahl, T. P., and R. A. Stein. 1994. Influence of larval gizzard shad *(Dorosoma cepedianum)* density on piscivory and growth of young-of-year *(Stizostedion vitreum* × *S. canadense)*. Canadian Journal of Fisheries and aquatic Sciences 51:1993–2002.

Stein, R. A. 1979. Behavioral response of prey to fish predators. Pages 343–354 *in* R. H. Stroud and H. Clepper, editors (1979).

Stein, R. A., R. F. Carline, and R. S. Hayward. 1981. Largemouth bass predation on stocked tiger muskellunge. Transactions of the American Fisheries Society 110:604–612.

Stephens, D. W., and J. R. Krebs. 1986. Foraging theory. Princeton University Press, Princeton, New Jersey.

Stewart, D. J., and M. Ibarra. 1991. Predation and production by salmonid fishes in Lake Michigan, 1978–88. Canadian Journal of Fisheries and Aquatic Sciences 48:909–922.

Stewart, D. J., J. F. Kitchell, and L. B. Crowder. 1981. Forage fishes and their salmonid predators in Lake Michigan. Transactions of the American Fisheries Society 110:751–763.

Stockwell, J. D., K. L. Bonfantine, and B. M. Johnson. 1999. Kokanee foraging: a *Daphnia* in the stomach is worth two in the lake. Transactions of the American Fisheries Society 128:169–174.

Stockwell, J. D., and B. M. Johnson. 1997. Refinement and calibration of a bioenergetics-based foraging model for kokanee *(Oncorhynchus nerka)*. Canadian Journal of Fisheries and Aquatic Sciences 54:2659–2676.

Stockwell, J. D., and B. M. Johnson. 1999. Field evaluation of a bioenergetics-based foraging model for kokanee *(Oncorhynchus nerka)*. Canadian Journal of Fisheries and Aquatic Sciences 56(Supplement 1):140–151.

Stouder, D. J., K. L. Fresh, and R. J. Feller, editors. 1994. Theory and application in fish feeding ecology. Belle W. Baruch Library in Marine Science 18, University of South Carolina Press, Columbia.

Stroud, R. H., and H. Clepper, editors. 1979. Predator–prey systems in fisheries management. Sport Fishing Institute, Washington, D.C.

Summerfelt, R. C., and G. E. Hall. 1987. Age and growth. Iowa State University Press, Ames.

Sweka, J. A., and K. J. Hartman. 2001. Effects of turbidity on prey consumption and growth in brook trout and implications for bioenergetics modeling. Canadian Journal of Fisheries and Aquatic Sciences 58:386–393.

Swenson, A. W., and L. L. Smith, Jr. 1973. Gastric digestion, food consumption, feeding periodicity and growth efficiency in walleye *(Stizostedion vitreum vitreum)*. Journal of the Fisheries Research Board of Canada 30:1327–1336.

Swingle, H. S. 1950. Relationships and dynamics of balanced and unbalanced fish populations. Alabama Polytechnic Institute Agricultural Experiment Station, Bulletin 274, Auburn.

Thompson, G. L. 1991. A unified approach to rank tests for multivariate and repeated measures designs. Journal of the American Statistical Association 86:410–421.

Tonn, W. M., C. A. Paszkowski, and I. J. Holopainen. 1992. Piscivory and recruitment: mechanisms structuring prey populations in small lakes. Ecology 73:951–958.

Trudel, M., A. Tremblay, R. Schetagne, and J. B. Rasmussen. 2000. Estimating food consumption rates of fish using a mercury mass balance model. Canadian Journal of Fisheries and Aquatic Sciences 57:414–428.

Van Den Avyle, M. J. 1993. Dynamics of exploited fish populations. Pages 105–136 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America. American Fisheries Society, Bethesda, Maryland.

Vander Zanden, M. J., B. J. Shuter, N. P. Lester, and J. B. Rasmussen. 2000. Within- and among-population variation in the trophic position of a pelagic predator, lake trout *(Salvelinus namaycush).* Canadian Journal of Fisheries and Aquatic Sciences 57: 725–731.

Vogel, J. L., and D. A. Beauchamp. 1999. Effects of light, prey size, and turbidity on reaction distances of lake trout *(Salvelinus namaycush)* to salmonid prey. Canadian Journal of Fisheries and Aquatic Sciences 56:1293–1297.

Wahl, D. H., and R. A. Stein. 1988. Selective predation by three esocids: the role of prey behavior and morphology. Transactions of the American Fisheries Society 117: 142–151.

Wahl, D. H., and R. A. Stein. 1989. Comparative vulnerability of three esocids to largemouth bass *(Micropterus salmoides)* predation. Canadian Journal of Fisheries and Aquatic Sciences 46:2095–2103.

Wahl, D. H., and R. A. Stein. 1991. Food consumption and growth of three esocids: field tests of a bioenergetics model. Transactions of the American Fisheries Society 120: 230–246.

Wahl, D. H., and R. A. Stein. 1993. Comparative population characteristics of muskellunge *(Esox masquinongy),* northern pike *(E. lucius),* and their hybrid *(E. masquinongy* × *E. lucius).* Canadian Journal of Fisheries and Aquatic Sciences 50:1961–1968.

Wahl, D. H., R. A. Stein, and D. R. DeVries. 1995. An ecological framework for evaluating the success and effects of stocked fishes. Pages 176–189 *in* H. L. Schramm, Jr., and R. G. Piper, editors. Uses and Effects of Cultured Fishes in Aquatic Ecosystems. American Fisheries Society, Symposium 15, Bethesda, Maryland.

Walters, C. J., V. Christensen, and D. Pauly. 1997. Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments. Reviews in Fish Biology and Fisheries 7:139–172.

Walters, C. J., and F. Juanes. 1993. Recruitment limitation as a consequence of natural selection for use of restricted feeding habitats and predation risk taking by juvenile fishes. Canadian Journal of Fisheries and Aquatic Sciences 50:2058–2070.

Walters, C., and J. F. Kitchell. 2001. Cultivation/depensation effects on juvenile survival and recruitment: implications for the theory of fishing. Canadian Journal of Fisheries and Aquatic Sciences 58:39–50.

Walters, C., D. Pauly, V. Christensen, and J. F. Kitchell. 2000. Representing density dependent consequences of life history strategies in aquatic ecosystems: EcoSim II. Ecosystems 3:70–83.

Webb, P. W. 1986. Locomotion and predator-prey relationships. Pages 24–41 *in* M. E. Feder, editor. Predator-prey relationships. University of Chicago Press, Chicago.

Werner, E. E. 1974. The fish size, prey size, handling time relation in several sunfishes and some implications. Journal of the Fisheries Research Board of Canada 31:1531–1536.

Werner, E. E., and J. F. Gilliam. 1984. The ontogenetic niche and species interactions in size-structured populations. Annual Review of Ecology and Systematics 15:393–425.

Werner, E. E., J. F. Gilliam, D. J. Hall, and G. G. Mittlebach. 1983. An experimental test of the effects of predation risk on habitat use in fish. Ecology 64:1540–1548.

Werner, E. E., and D. J. Hall. 1974. Optimal foraging and the selection of prey by the bluegill sunfish *(Lepomis macrochirus)*. Ecology 55:1042–1052.

Werner, E. E., and D. J. Hall. 1988. Ontogenetic habitat shifts in bluegill: foraging rate–predation risk trade-off. Ecology 69:1352–1366.

Werner, E. E., and G. G. Mittelbach. 1981. Optimal foraging: field tests of diet choice and habitat switching. American Zoologist 21:813–829.

Wiley, M. J., R. W. Gordon, S. W. Waite, and T. Powless. 1984. The relationship between aquatic macrophytes and sportfish production in Illinois ponds: a simple model. North American Journal of Fisheries Management 4:111–119.

Winemiller, K. O., and D. B. Jepsen. 1998. Effects of seasonality and fish movement on tropical river food webs. Journal of Fish Biology 53(Supplement A):267–296.

Wright, D. I., and W. J. O'Brien. 1984. The development and field test of a tactical model of the planktivorous feeding of white crappie (*Pomoxis annularis*). Ecological Monographs 54:65–98.

Wurtsbaugh, W. A., and H. W. Li. 1985. Diel migrations of a zooplanktivorous fish (*Menidia beryllina*) in relation to the distribution of its prey in a large eutrophic lake. Limnology and Oceanography 30:565–576.

Yule, D. L. 2000. Comparison of horizontal acoustic and purse-seine estimates of salmonid densities and sizes in eleven Wyoming waters. North American Journal of Fisheries Management 20:759–775.

Zaret, T. M. 1980. Predation and freshwater communities. Yale University Press. New Haven.

Zaret, T. M., and R. T. Paine. 1973. Species introduction in a tropical lake. Science 182: 449–455.

17 Habitat Evaluation

Tammy J. Newcomb, Donald J. Orth, and Dean F. Stauffer

17.1 INTRODUCTION

The demand for clean freshwater resources continues to grow with increasing needs for instream (hydropower) and offstream (agricultural, municipal, and industrial) uses of water. These activities influence water quality and habitat availability in river and lake habitats as well as the hydrologic processes that support them. Financial, cultural, and biological stakes can be high in water use and allocation decisions, so it is important to be able to quantify fish habitat made available by the presence of water as well as the quality of the habitat provided by that water. Even in water-rich places where allocation is not an issue, many human uses on the landscape can result in reductions in water quality or degradation of physical habitat.

Fish habitat is composed of physical and biological components required to support fish growth, survival, and reproduction. Habitat components can include specific attributes of a location occupied by a fish or the suite of areas required to complete life histories and sustain a population. More specific terms for fish habitat include essential fish habitat, critical habitat, and preferred habitat. Essential fish habitat is defined by the 1996 Magnuson–Stevens Fishery Conservation and Management Act as "those waters and substrate necessary to fish for spawning, breeding, feeding or growth to maturity" (16 U.S.C. §§ 1801 to 1882). Critical habitat is a term associated with the Endangered Species Act of 1973 (16 U.S.C. §§ 1531 to 1543) and represents the area required to conserve an endangered species. In a broader context, critical habitat areas provide habitat for sensitive life stages such as spawning or early life history (Pitlo 1989). Finally, preferred habitat is defined as those areas that organisms select with greater frequency than those areas occur in the environment (Johnson 1980).

Defining and evaluating fish habitat requires determination of ranges for specific habitat parameters that delineate suitable habitat conditions for a fish species or assemblage of fishes. Fish habitat is assessed to measure baseline conditions, monitor habitat availability, prioritize habitat for protection or enhancement, and establish instream flows for conservation of aquatic biota. Further-

more, fish habitat assessment provides the basis for manipulating aquatic habitats when the desired outcomes include increasing fish abundance or ensuring their sustainability (Orth and White 1993). Results from habitat assessment can lead to actions such as improving water quality, providing a limiting habitat feature such as spawning gravel or cover, or improving conditions for the production of forage.

Procedures used for analyzing fish habitat data should be determined a priori based on the goals and objectives of the habitat assessment and, thus, reasons for data collection. Ideally, careful project scoping and design dictate which habitat parameters are measured. For example, an assessment of baseline conditions is usually desired prior to implementing habitat management prescriptions, and thus before and after analyses are critical for measuring the effect of aquatic habitat improvement (Kondolf and Micheli 1995). Therefore, habitat variables measured should be those that would exhibit a response either physically or biologically to the habitat manipulation. Monitoring can be used to track changes over time, as in the case of monitoring for habitat degradation or documenting a long-term response of the biota to restoration activities.

Whereas tools to assess water quality, water quantity, and biological integrity are readily available and established, measures of habitat quality and approaches to assessment and evaluation are still developing (Maddock 1999). Some of the most basic relationships such as habitat availability and fish standing stock are still not well defined for most species or fish communities. Furthermore, only recently have fisheries scientists begun to understand the importance of temporal stream channel dynamics and flow variability on habitat and fish populations (Poff and Ward 1989; Palmer and Poff 1997).

Often the first step in a habitat evaluation project is to classify stream channels, habitat units within the stream, or lake types (Bisson et al. 1981; Hawkins et al. 1993; Rosgen 1994; Montgomery and Buffington 1997). Classification is used to group streams, lakes, or reaches that may be responding to similar physical processes and is a good starting point for formulating conceptual models regarding fish habitat or population response to habitat. However, taxonomic approaches usually do not identify limiting habitat factors, and a higher resolution is required for recommending management actions. The objective of this chapter is to present approaches used to analyze and interpret data collected for fish habitat assessment in both lotic and lentic systems.

17.2 MEASUREMENT OF HABITAT PARAMETERS

Time, money, and personnel resources limit the amount and type of data that are collected. Therefore, the number of habitat parameters and quantity of data collected should be carefully specified as needed to meet study objectives. Greater statistical power is achieved when collecting a larger number of samples with a few well-defined parameters versus collecting few samples and a larger number of parameters that may or may not provide meaningful information.

17.2.1 Qualitative versus Quantitative Data

Habitat data can be measured through qualitative or quantitative approaches. Qualitative data are often recorded using categorical procedures and can be descriptive, ordinal (categorical), and nominal (yes–no or presence–absence) data (Zar 1996; Table 17.1). Common variables measured qualitatively include substrate composition, substrate embeddedness, cover types, streambank condition, or habitat units (e.g., riffle, pool, cove, and island complex), and each parameter varies in measurement subjectivity. For example, when categorizing stream bottom embeddedness as high, medium, or low, the differences between high and low may be obvious, but medium to low and medium to high are subject to interpretation. Inherently, qualitative data are less sensitive than are quantitative data in capturing changes or trends in habitat over time. Categorical data can be assigned ranks, which allows for greater statistical analyses (Table 17.2).

Continuous data, which may be measured on ratio and interval scales (Zar 1996), are measured along a continuum and include variables such as velocity, depth, dissolved oxygen (DO), and substratum particle size (Table 17.1). Ratio data have a true zero and meaningful relationships between two values. For example, 4 cubic feet per second (cfs) is twice as much water flowing as 2 cfs. Interval data do not have a true zero nor the meaningful relationship between values;

Table 17.1 Common parameters measured in lotic and lentic habitats. Parameters are noted (superscripts) as to whether they are most commonly measured as ratio (r) or ordinal (o) types of data, with some listed as more than one type depending on study requirements.

Lotic habitats	Lentic habitats
Riparian data	Depth[r]
Canopy[ro]	Physical structure[o]
Shading[ro]	Vegetation[ro]
Bank slope[r]	Turbidity[r]
Erosion[o]	Substrate[o]
	Dissolved oxygen[r]
Channel data	Temperature[a]
Discharge[r]	Bank slope[r]
Stream surface gradient[r]	Shoreline length[r]
Depth[r]	Secchi disk depth[r]
Velocity[r]	Nutrients (e.g., P and N)[r]
Substrate[ro]	Total dissolved solids[r]
Bank-full width[r]	
Streambed elevation[r]	
Bed load movement[r]	
Temperature[a]	
Fish cover[o]	
Large woody debris[o]	

[a] Temperature as measured in °C or °F is interval scale data without a true zero. When temperature is converted to Kelvin, the data has a true zero and can be considered as ratio data. This distinction is important for assumptions inherent in many statistical computations.

Table 17.2 Example of qualitative categorization (rating) of stream substrate embeddedness and streambank condition estimated by semi-quantitative description (based on Pfankuch 1975; Platts et al. 1983).

Embeddedness	
Categorical rating	Proportion of gravel, rubble, and boulder particles covered by fine sediment
1	>75%
2	50–75%
3	25–50%
4	5–25%
5	<5%

Streambank Condition		
Score	Estimated bank slope	Rating
2	<30%	Excellent
4	30–40%	Good
6	40–60%	Fair
8	>60%	Poor

examples include compass points, time, and temperature data. Temperature data (°C or °F) can be considered ratio data only when converted to Kelvin (1°C = 274.15 K; 1°F = 255.93 K). Although the data are more costly to acquire, quantitative or semi-quantitative approaches are more cost effective in the quality of the information that is collected (Milner et al. 1985). Furthermore, quantitative data are more useful in developing strong quantitative relationships between fish abundance or community composition and habitat.

17.2.2 Spatial and Temporal Considerations

Spatial scale considerations are important for answering questions regarding cause and effect, and multiple scales of observation may be required to gain a full understanding of fish–habitat relationships. Habitat scales include microhabitat (e.g., depth, cover, substrate, and velocity at a specific fish location), mesohabitat or channel unit (e.g., coves, island complexes, littoral zone, pool, riffle, run, or glide), and macrohabitat (variables that range longitudinally over a larger area in streams or stratify vertically in reservoirs and lakes, such as DO, pH, and temperature). Measures of microhabitat parameters may be used to evaluate preference or probability of fish distribution within a reach (Rabeni and Sowa 1996), whereas qualities of mesohabitat throughout a stream system can dictate population dynamics such as spawning, mortality, and growth. Additionally, although the interest for analyzing habitat may be at the site level, such as for restoration projects, longitudinal stream level effects may be important when considering relationships among

many sites and fish population characteristics (Dunham and Vinyard 1997). Watershed scale considerations are discussed in Chapter 18, and thus this chapter focuses on micro and mesohabitat analyses.

Temporal scale is important when evaluating changes in habitat use by fishes during important life history phases and for assessing physical processes that dictate channel condition and development of habitat areas. For example, assessment of spawning habitat and shifts in habitat use due to ontogenetic changes would require additional habitat parameters in early life history or a more frequent assessment of use due to rapid changes. Because these early life history shifts are vulnerable stages, they can be an important focus for assessing habitat availability and potential changes in habitat over a limited temporal scale. From a geomorphic perspective, temporal scales play an important role in modifying or creating habitat features by the frequency of hydrological events such as floods and effective stream discharge for forming channel shape. Finally, lentic systems change vertically in habitat availability through seasons because of changes in depth, temperature, and DO. Thus when conducting a habitat investigation and analysis, it is important to address the temporal and spatial scale issues in the study design and in the interpretation of the results

17.2.3 Unbiased Sampling Approaches

Unbiased approaches to sampling and measuring habitat attributes and repeatability in the protocol and measures are necessary when comparing results from year to year or stream to stream. Many habitat parameters are often estimated rather than measured. Further, even though some parameters are measured, biases may be inherent in the instrumentation or procedural approach. Habitat features that can be measured with good precision and repeatability include streambank measures (e.g., vegetative stability and undercut bank), stream width and depth, riparian measures (e.g., streamside cover and habitat type), and substrate embeddedness (Platts 1981). Habitat measures with low precision and low repeatability include bank-full width, bank–bank width, and proportions of sediment type (Platts 1981). Training observers prior to conducting assessments can increase precision and repeatability, preventing serious errors in interpretations that result from biased data (Hannaford et al. 1997). Those interpreting habitat analyses should consider the possible sources of bias or error based on the protocols for measuring parameters. Correction factors based on the associated error determined between actual measures and estimated values for a subsample of observations can be applied to reduce investigator bias (Dolloff et al. 1993) (Box 17.1).

17.3 LOTIC VERSUS LENTIC HABITAT ANALYSES

Contemporary paradigms in fish ecology assume that fundamental differences exist between the underlying mechanisms that limit fish communities and influence population dynamics in lakes and streams. Lentic environments tend to be

Box 17.1 Corrections to Visual Estimations of Habitat Data

A measure of habitat characteristics is desired for the River Raisin. To increase efficiency and reduce time, the manager desires to estimate area associated with instream mesohabitat channel types, which were determined to be riffle, shallow pool, and deep pool. The sampling team is instructed to determine area for each mesohabitat unit by estimating each unit's average depth and width. To validate the team's estimates, 20% of each mesohabitat channel unit type (or a minimum of 10) shall be measured for average channel width and depth to calculate mesohabitat area.

Once the data were collected, the measured (m_i) and estimated (x_i) data for each mesohabitat type were plotted to check for data errors and to confirm that there was a high correlation between the measured and observed areas. Only the riffle habitat is shown below.

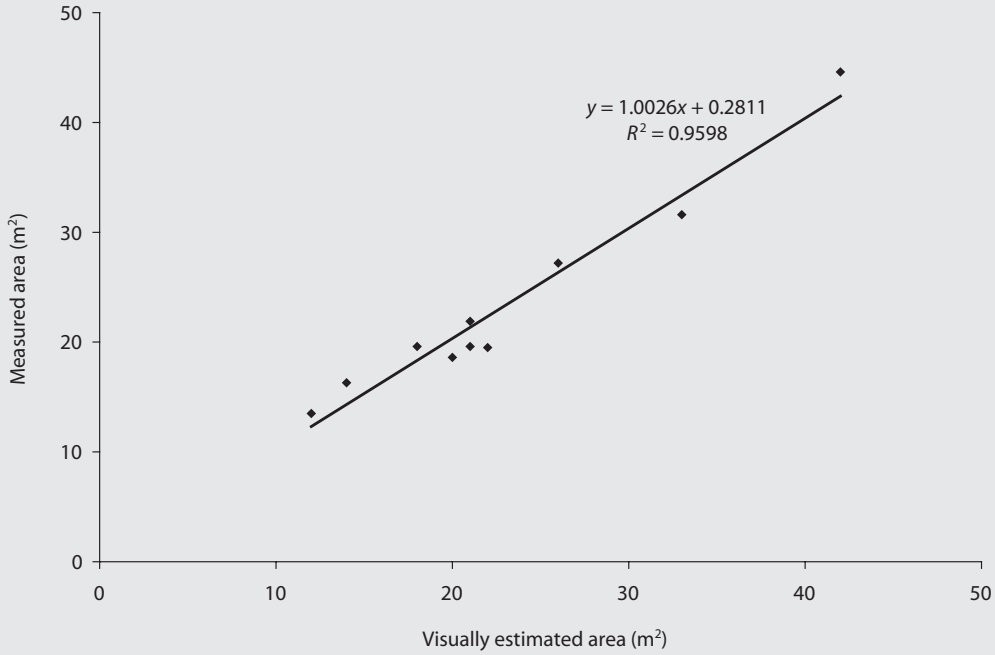


Figure Correlation of area data between the measured and observed riffle mesohabitats.

In this example, the correlation is high and the y-intercept is very near 0. The next step is then to determine the calibration ratio ($\hat{C}$) for riffles. The $\hat{C}$ is calculated by summing m_i for all units and dividing by the sum of x_i for those same units. In this example, $\hat{C}$ = 232.4/229 = 1.015.

Table Measured (m_i) and estimated (x_i) data for riffle mesohabitat. The calibrated area is given by $\hat{C}x_i$, where $\hat{C} = 1.015$. The sum of $(m_i - \hat{C}x_i)$ is used to calculate the variance of the estimated area.

Riffle number and sum	x_i	m_i	$\hat{C}x_i$	$(m_i - \hat{C}x_i)^2$
1	22	19.5	22.3	8.0
2	33		33.5	
3	21	19.6	21.3	3.0
4	18		18.3	
5	14		14.2	
6	20	18.6	20.3	2.9
7	21		21.3	
8	23		23.3	
9	42	44.6	42.6	3.8
10	24		24.4	
11	26	27.2	26.4	0.7
12	14		14.2	
13	13		13.2	
14	12	13.5	12.2	1.7
15	22		22.3	
16	33	31.6	33.5	3.6
17	21		21.3	
18	18	19.6	18.3	1.8
19	21	21.9	21.3	0.3
20	23		23.3	
21	26		26.4	
22	14	16.3	14.2	4.4
Sum	481	232.4		30.2

The $\hat{C}$ is multiplied by the estimated area to determine the calibrated area ($\hat{C}x_i$). The total area of riffles in River Raisin can then be determined by summing x_i and multiplying by $\hat{C}$. Thus the amount of riffle area ($\hat{A}_{\text{riffles}}$) in River Raisin is estimated to be 488.2 m^2. A variance can be calculated on this estimate by

$$\hat{V}(\hat{A}_{\text{riffles}}) = \frac{N(N-n)}{n(n-1)} \sum_{i=1}^{n} (m_i - \hat{C}x_i)^2,$$

where N is the sample size for x_i and n is the sample size for m_i.

$$\hat{V}(\hat{A}_{\text{riffles}}) = \frac{22(22-10)}{10(10-1)} 30.2 = 88.6.$$

And finally, a 95% confidence interval on this value can be calculated by

$$\hat{A}_{\text{riffles}} \pm t_{(0.05\ n-1)} \sqrt{\hat{V}(\hat{A}_{\text{riffles}})},$$

which equals 488.2 m^2 ± (2.262)(9.4) for a confidence interval of 466.9 – 509.5 m^2.

relatively stable, and thus community composition and abundance results from lake productivity, competition, and predation. Conversely, in lotic environments, stream flow dictates the availability of habitat area. Therefore, a combination of habitat persistence and variability is believed to be the dominant force shaping stream fish communities. Habitat quality and availability in streams varies longitudinally whereas in lakes, water column stratification of temperature and DO define available habitat. Because of these differences, stream fisheries scientists and lake fisheries scientists often diverge in their approaches to habitat assessment and analysis.

With the exception of water quality parameters (e.g., nitrogen, phosphorous, and turbidity), evidence for relationships between physical lake habitat and fish communities is sparse. Because the lentic fish community structure tends to be influenced by predator–prey interactions, bioenergetics approaches are traditionally more useful for evaluating stock abundance and limitations to growth. However, as development continues along shorelines, physical habitat in lentic environments is becoming more prominent in management approaches. For example, spawning habitat along shorelines can be severely reduced by the construction of barrier walls that may act to increase the energy of wave action along nearshore nursery or incubation areas. Spawning reefs also may be a definitive habitat feature that dictates recruitment of juveniles to a population.

17.4 WATER QUALITY AS A HABITAT PARAMETER

Dissolved oxygen (DO) and water temperature are two commonly measured macrohabitat water quality variables. In lakes, DO and temperature stratify vertically. In reservoirs, as summer progresses, available habitat for cold- or coolwater species will horizontally compress toward the dam in deeper water, resulting in a habitat "squeeze" (Coutant 1985). Temperature and DO measurements can be collected from throughout a reservoir or lake to develop an isopleth and make conclusions about available habitat during potentially stressful periods such as hot summers or ice cover in winter. Data analysis can be conducted as a time series (documenting increases or decreases in habitat area over time). In rivers, DO and temperature vary longitudinally; however, in some cases where groundwater input is sufficient and mixing is negligible, pools can stratify, providing pockets of thermal refuge (Nielsen et al. 1994).

17.4.1 Analysis of Temperature and Dissolved Oxygen Data

Recent advances in technology have made it easy and relatively inexpensive to monitor temperature and DO on a continuous basis. Empirical data collected at point locations provide station-specific habitat conditions, but the amount of data produced by a large number of data loggers in a study area can be overwhelming. For example, 7 months of water temperature collected at 10-min intervals in a single location results in more than 31,000 observations. While the objectives for the sampling should dictate the time step (interval) of measures, generally, hourly

measurements provide a satisfactory trade-off between effort to analyze the data, desired download intervals, and the amount of information acquired. Objectives for collecting temperature data should be clearly defined, and those objectives will also dictate the placement of the loggers. Programs such as SAS (2004) or macros in Microsoft Excel can be used to minimize the time required to reduce data to an appropriate format for further analysis.

Data collected at point locations can be analyzed to determine biologically relevant statistics such as daily mean, daily flux, or daily maximum or minimum temperatures. The coefficient of variation can be calculated to obtain a measure of variability for a given time-step (e.g., one day), but the data must first be transformed to ratio data by converting it to Kelvin. To investigate the biological relevance of thermal information, Pearson's correlation coefficient, simple linear regression, and multiple regression can be used to relate habitat conditions with measures of fish growth, abundance, or productivity. Temperature can also be related to an index of interest such as an environmental severity index (Seelbach 1993), degree-day accumulation (Bovee et al. 1994), or duration of time above a threshold (Lohr et al. 1996). Data on maximum temperatures can be compared with biologically relevant values such as upper incipient lethal levels or critical thermal maxima for a species. In most applications, temperature data are used to provide descriptive statistics of the thermal habitat that fish occupy. For example, Wehrly et al. (2003) used thermal characteristics throughout Lower Michigan to examine fish abundance and fish community patterns across the landscape and determine a baseline characterization for comparison with future changes that may occur with the thermal habitat.

17.4.2 Modeling Approaches to Quantifying Habitat Defined by Water Quality Parameters

Empirical measurements of temperature and DO as outlined in the previous section provide measures of conditions at specific locations, but use of those data to model conditions throughout a study area can provide a quantitative assessment of DO, water quality, or thermal habitat properties. By combining the quality of macrohabitat with species-specific requirements, the quantity of macrohabitat can also be determined. In addition, the effects of population-altering phenomena such as the presence of thermal barriers that limit fish movement and the suitability of temperature ranges and fluctuations for fish survival, egg incubation, growth, feeding, and spawning can be estimated. Furthermore, modeling stream temperatures under current or future altered conditions can be used to assist in defining instream flows, managing a coldwater fishery, regulating dam releases, or assessing the anticipated thermal effects prior to logging, urbanization, changes in hydropower operations, or future conditions related to global warming (Stefan et al. 2001).

Methods used to evaluate thermal patterns in streams include two basic approaches that differ in levels of complexity and predictability. The two general categories of stream temperature models are empirical, or statistical, models and stochastic models (Bartholow 1989). Empirical models are developed using

measured observations, such as water temperature, air temperature, and discharge, and applying regression or harmonic analyses to develop a predictive model. Empirical models are very useful for filling in missing data or estimating historical temperatures when no water temperature data existed. Although straightforward, empirical models generally do not consider the physical relationships associated with heat flux or heat transport. Thus, empirical models are limited in their abilities to predict changes in thermal patterns that result from changes in the physical stream properties or surrounding landscape, such as channel width or riparian conditions.

Stochastic models incorporate physical process relationships in an energy budget to predict instream temperatures. By incorporating an energy budget, stochastic models predict water temperature on the basis of gains and losses in thermal energy from processes such as radiation, convection, conduction, and evaporation. Relevant physical parameters (e.g., stream gradient, discharge, humidity, and shading) are incorporated into stochastic models to address the energy flux processes. Because of their predictive abilities in response to a change in the ecosystem, these models are very powerful for fisheries scientists having to make difficult decisions regarding instream flow habitat management. The tradeoff, however, is that stochastic models are usually much more difficult to develop and require more data than do statistical models (Bartholow 1989). Many of the data requirements are now easier to meet, however, in the form of online databases, through geographical information system (GIS) analysis of aerial photos, and through the use of advanced technology, such as Doppler or hydroacoustic equipment, to collect the information.

The stream reach model (SSTEMP) and the stream network model (SNTEMP) (Theurer et al. 1984; Bartholow 1989), the enhanced stream water quality model (QUAL2E; USEPA 1995), and the Tennessee Valley Authority (TVA) river modeling system (ADYN and RQUAL; Hauser and Walters 1995) are examples of readily available stochastic stream temperature models (Table 17.3). The SNTEMP and SSTEMP models are similar in their algorithms and prediction of water temperatures based on an energy budget. The SSTEMP model is used to evaluate short stream reaches over a limited number of time periods. This straightforward, interactive model presents a simplified modeling approach and is useful in sensitivity analyses for specific parameters. The SSTEMP model can be very useful in predicting temperatures over a limited reach for definition of thermal habitat characteristics.

The SNTEMP model requires the development of a conceptual stream network model and establishment of multiple input data files. The network model is very powerful when considering several channel reaches with tributary inputs over a lengthy time period. The network model software also provides postsimulation statistical evaluations that are useful for evaluating model sensitivity and predictive ability. Common applications of the SNTEMP model include assessment of the effects of altered thermal regimes as a result of changes in hydrology or hydropower management (Theurer et al. 1984; Bartholow 1991; Zedonis 1997; Krause et al. 2004).

Table 17.3 Comparison of stream temperature and water quality models that are readily available to analyze stream temperature and analyze the potential impacts of management scenarios or alterations in hydrology or the riparian area. Models are identified as follows: SSTEMP = stream reach model (Theurer et al. 1984; Bartholow 1989); SNTEMP = stream network model (Theurer et al. 1984; Bartholow 1989); QUAL2E = enhanced stream water quality model (USEPA 1995); and RQUAL (RQUAL + ADYN) = Tennessee Valley Authority (TVA) river modeling system (Hauser and Walters 1995).

	Models			
Model capabilities	SSTEMP	SNTEMP	QUAL2E	RQUAL
Time step for prediction	Daily	Daily	Daily	Hourly
Reach versus basin network	Reach	Basin	Basin	Basin
Predict multiple water quality parameters	No	No	Yes	Yes
Analyze alternative shade scenarios	Yes	Yes	No	No
Predicts maximum water temperature	Yes	Yes	No	Yes

The QUAL2E model does not incorporate the influence of shade on water temperature, whereas SNTEMP and RQUAL do. Thus, QUAL2E may be better applied for modeling large and wide rivers, which have less shade influence than do smaller streams (USEPA 1995). The SNTEMP, SSTEMP, and QUAL2E models are steady-state models and assume that flow is constant over a 24-h period and input parameters are daily average values. However, because of these assumptions, the models cannot consider daily variation, which can either help or hinder the model from providing accurate daily mean predictions depending on meteorological conditions. If daily variations are dramatic enough to be considered important to fish survival, a model capable of predicting temperature multiple times per day should be used. For most management situations daily averages are suitable. The QUAL2E model can also function in a quasi-dynamic mode, which still assumes steady flow but accounts for the influence of diel climate fluctuation through input of meteorological parameters at 3-h time steps (USEPA 1995).

The TVA river modeling system consists of the ADYN hydrologic flow model and the RQUAL water quality model. The ADYN + RQUAL model is one of a few available dynamic models and is capable of modeling flows fluctuating within a 24-h period. The model requires frequent input parameters (e.g., hourly), and thus the data requirements are greatly increased over a steady-state model (Hauser and Walters 1995). The model is useful for analyzing trade-offs with hydropower operations that fluctuate flows within a 24-h period (Krause et al. 2005).

The QUAL2E and RQUAL models can also predict water quality parameters including DO, nitrogen (organic, ammonia, nitrite, and nitrate) and phosphorous (organic and dissolved) concentrations, algae as chlorophyll a, an arbitrary, nonconservative carbonaceous biochemical demand (CBOD), up to three conservative minerals, and coliform bacteria (Brown and Barnwell 1987; Hauser and Walters 1995; USEPA 1995).

Models SNTEMP and QUAL2E are well documented and easily obtained (available at http://smig.usgs.gov/cgi-bin/SMIC/browse_models), training is available,

and their use is prevalent in the literature (Theurer et al. 1984; Lifton et al. 1985; Brown and Barnwell 1987; Wilson et al. 1987; Bartholow 1989, 1991; Waddle 1989; Sullivan et al. 1990; USEPA 1995; Zedonis 1997).

17.5 HABITAT–FISH POPULATION INDICES

Identifying a predictable relationship between fish populations and habitat is highly desired for addressing the potential for a fishery and expectations of production, identifying degraded habitat conditions, or assessing implications for human development or, conversely, habitat restoration efforts. When the desire is to develop a foundation for making management decisions over a large spatial scale, an empirical model can be developed using a variety of approaches and statistical analyses. However, in the regulatory process for setting stream flows or reservoir levels or allocating water withdrawal, the analytical approaches must be quantitative in nature. These approaches must also have the capability to evaluate potential trade-offs in alternate management scenarios. Thus standardized approaches such as the instream flow incremental methodology or the habitat evaluation procedures (HEP) are often used.

Empirical relationships of fish populations and habitat characteristics can be derived from biological and physical habitat data collected over the geographic range of interest. The most common form of habitat models are those for which the investigator measures multiple habitat variables, sometimes over multiple scales (e.g., basin variables, channel variables, and microhabitat variables), for a large number of streams or lakes and measures fish abundance simultaneously. Simple, multiple, and logistic regression are often used as well as other multivariate approaches such as principal components analysis, linear discriminant analysis, or correspondence analysis (Fraley and Graham 1981; Parsons et al. 1981; Anderson and Nehring 1985; Watson and Hillman 1997; Claramunt and Wahl 2000). Using these approaches, the investigator attempts to develop a predictive relationship between abiotic variables and fish abundance or attributes of a fish population such as growth. Fish–habitat index models are used to predict potential fish standing crop for inventory or planning, to increase the understanding between functional processes and standing crop, to diagnose stream health, and to predict responses to proposed management actions (Binns and Eiserman 1979; Milner et al. 1985; McClendon and Rabeni 1987; Fausch et al. 1988; Modde et al. 1991).

When developing fish abundance indices, the target organisms and species must be defined as well as the area of interest for the model application. Using knowledge of the biological requirements of the organisms, the investigator identifies key habitat parameters that are relevant mechanistically to fish abundance (e.g., spawning gravel, sedimentation levels, and cover). The developed models are usually limited in that they are region specific, but they can be used to determine important physical habitat variables and identify limiting habitat features.

If a fish–habitat index is desired for a large region, a hierarchical approach can be used to classify subwatersheds on the basis of land types or regions and to group together streams or lakes that would be expected to be similar on the basis

of climate, soil, and topography. At the stream mesohabitat level, instream habitats or larger channel units such as pool, riffle, run, and glide (e.g., Bisson et al. 1981; Hawkins et al. 1993) are measured against elements of fish growth or standing stock. Channel units may be divided further into specific types; for example, pools can be further categorized as plunge pool, scour pool, deep pool, or shallow pool depending on their perceived function for acting as fish habitat. Habitat variables such as water quality measurements (e.g., pH, alkalinity, phosphate, and nitrate), presence of cover, substrate, turbidity, and bank condition can be measured within the specific habitat types for characterizing those features. A two-factor analysis of covariance can be used to test for differences in fish abundance or biomass with land type or ecoregion and channel unit (Modde et al. 1991). Perhaps the most common approach, however, is to use a linear model (either multivariate or univariate) to relate aspects of fish abundance to stream habitat. The linear aspect between the predictor and response variables of these models in some cases may be a detriment to understanding fully the interactions between habitat quality and fish populations, in which case a regression tree analysis may be used (Stoneman and Jones 1996).

Logistic regression is a flexible statistical method that is often used in habitat evaluation. It is a useful technique to identify threshold relationships and to select key habitat variables that relate strongly to presence and absence of a target organism. The key advantage of this technique is that the dependent variable has only two classes (presence = 1 or absence = 0). Presence and absence are easier to measure than is fish abundance or other metrics of population status.

The generalized form is

$$P = \frac{e^{a + b_1x_1 + b_2x_2 + \ldots b_nx_n}}{(1 - e^{a + b_1x_1 + b_2x_2 + \ldots b_nx_n})}, \tag{17.1}$$

where P is the probability of presence; a, b_1, b_2, . . . , b_n are parameters; and x_1, x_2, . . . , x_n are independent variables. In order for this procedure to be applied successfully, the investigator must systematically sample the target organisms over a wide range of habitat conditions including habitats where the organism is absent. Ideally, both used and unused habitats are sampled equally to provide unbiased estimates of parameters. The logistic regression may include multiple variables, interaction terms, and polynomial terms (Box 17.2).

Key metrics used to evaluate a fitted logistic regression model include (1) proportion of observations correctly classified, (2) sensitivity, or the proportion of presences correctly classified, (3) specificity, or the proportion of absences correctly classified, (4) false positives, or the proportion of presences incorrectly classified, and (5) false negatives, or the proportion of absences incorrectly classified. The best model for habitat evaluation procedures should have a high sensitivity and a low false positive rate.

Common issues associated with these types of models include a lack of statistical rigor in assessing colinearity or analysis of residuals, small sample size ($n > 20$

Box 17.2 A Habitat Evaluation Model Based on Logistic Regression

Smith (1999) used a team of snorkelers to identify locations where male darters showed both behavioral and morphological indicators of spawning. Fifteen habitat variables were measured at all spawning locations ($n = 87$) and a systematic sample of available microhabitat ($n = 146$). All fifteen variables showed different distributions based on two-sample Kolmogorov–Smirnov tests ($P < 0.05$). This test, while demonstrating differences, does not identify which variable is most useful to identify essential habitats. This can be done with PROC LOGIST in SAS (2004).

Results of initial model performance with three variables were evaluated with the chi-square test statistic for each model coefficient estimated by maximum likelihood technique. Two of the three habitat variables for prediction of egg-clusterer spawning habitat in the upper Roanoke River basin were significant.

Table Logistic regression of three habitat variables as predictors of spawning habitat for darters.

Variable	Mean	Parameter estimate	SE	Probability of greater chi-square
Intercept		–3.457	1.092	0.0015
Diameter of spawning rock (mm)	52	0.044	0.007	0.0001
Percent embeddedness	11	–0.465	0.080	0.0001
Percent silt	23	0.011	0.193	0.9539

	Proportion Correctly Classified
Sensitivity	0.93
Specificity	0.96
False positive	0.09
False negative	0.03

recommended), too few sample locations in comparison with the number of variables measured, and error in measuring habitat variables (Fausch et al. 1988). The most serious drawback to the empirical fish–habitat index approach is the need for recognition that for the predictive model developed, the underlying assumption is that habitat predominately dictates population size or species presence in a stream or lake. This assumption should be seriously considered when choosing sampling units to include in the development of the model, and it is important to validate and test the model by applying it to a stream or lake unit not used in the development of the model. Also, at times, population size may not adequately reflect habitat quality, thus measures of survival, growth, or reproduction may be more suitable (VanHorne 1983).

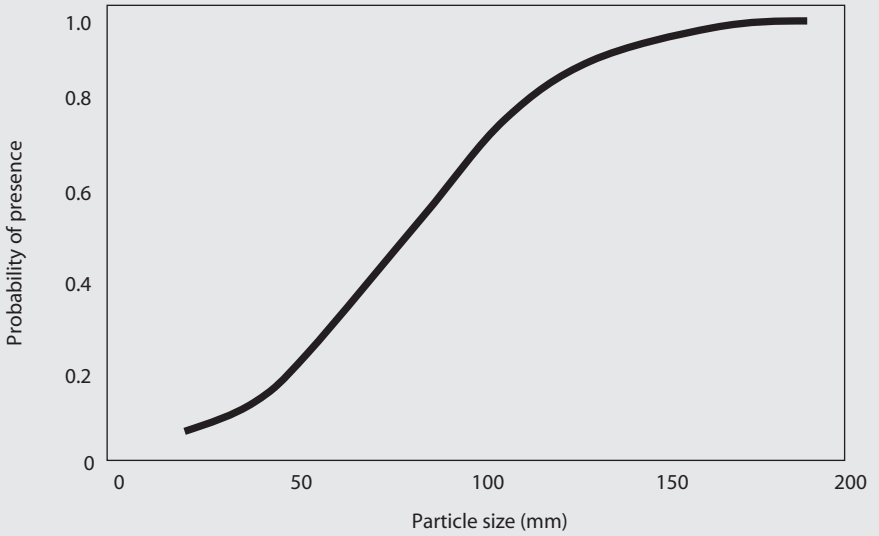


Figure The fitted relation between probability of spawner presence and particle size (mm) of the spawning rock for egg-clustering darters in the upper Roanoke River basin (Smith 1999).

The results (Table and Figure) support the hypothesis that large cobbles are important components of spawning habitat for these darters. Furthermore, the logistic model provides quantification of the gradual threshold effect by use of a continuous variable, diameter of the spawning rock. The values of high sensitivity and low false positives suggest that this model should perform well in similar situations. In this test, the empirical model development and transferability testing approach with logistic regression provides a much more reliable habitat evaluation tool than would any qualitative descriptor.

17.6 ASSESSMENT OF FISH HABITAT USE AND HABITAT QUALITY

Understanding how fishes use their habitat is a key element to meaningfully describing and quantifying habitat used by a single species, life stage, fish community, or guild. Habitat analyses are also used to determine features of critical habitat where population bottlenecks may occur, such as when spawning habitat may be limiting the production of age-0 fish. Results of fish habitat use studies are used to make qualitative statements about the suitability of a particular habitat, and these evaluations can be applied to quantify and characterize available habitat in a stream, lake, or even estuary (Rubec et al. 1998). Identifying fish use of habitat is key for many approaches to habitat assessment and for making

assumptions about habitat gains and losses at the population level. The analyses used to make conclusions about fish habitat use vary with the intended purpose of the information.

17.6.1 Habitat Suitability Criteria for a Single Species

The most basic habitat use question asks what type of habitat a species uses or apparently prefers. Habitat suitability criteria (HSC) are developed to characterize fish use of habitat and can be applied to a study area, either in an informal grid–transect approach in a stream or in a more formal habitat evaluation procedure (HEP) or physical habitat simulation analysis, to determine quality and quantity of habitat. In some cases, HSC are presented in a "blue-book" format that represents a synthesis of pooled data from a variety of different sources (Terrell and Carpenter 1997). One drawback to using pooled data is that the HSC may be so broadly defined from many different systems that when applied to a particular stream or lake, they are insensitive to describing high-quality habitat in that stream or lake. Furthermore, it is likely invalid to assume that all criteria are readily transferable from one water body to another without some form of validation or testing of the HSC (section 17.6.4). In some cases, observational data used to develop HSC reflect what was available for use by fish within the study area rather than reflecting preferred or optimal habitat. Thus, HSC developed from observations and measures in a degraded stream channel may reflect what fish use under those conditions rather than what they would use in a system that was not degraded. A comparison of the geomorphic properties between the study stream and a reference stream with high-quality habitat could be done to address this concern.

There are three types of criteria than can be developed and analyzed (Bovee 1986). Type I criteria are based on expert opinion or professional judgment. One structured and scientific approach that uses professional judgment is called the Delphi technique (Zuboy 1981). In this approach, study participants remain anonymous to one another, and all participants are asked to produce HSC on the basis of their professional experience and observations. Mail surveys achieve anonymity, and once initial HSC are collected, the information is developed into a group-defined set of HSC, and these are then sent back to all participants for concurrence. Several iterations may need to occur before all participants agree that the resulting criteria are a reasonable representation of habitat used and required by a species. At least eight panelists are recommended as a minimum number of participants (Hodgetts 1977). The Delphi approach can provide very good HSC for an interim period or when HSC information is needed quickly, but a subsequent field validation study (section 17.6.4) of the HSC should occur in the field to verify the resulting criteria (Crance 1987).

Type II criteria are developed by observations of individual fish and the local characteristics of the habitat in the area occupied by the fish and hence are called utilization criteria. An unobtrusive method is used to observe fish (e.g., snorkeling, observation tower and polarized binoculars, or throwable anode with small shocking area), and at each fish location, microhabitat parameters are measured

(depth, substrate, velocity, and cover are characteristics typically measured). The observations are then presented as a distribution of fish use of a single habitat variable. Theoretically, the habitat conditions where the most fish are observed should represent the optimum habitat. Assumptions of this method require that a range of all habitat types in the study stream are sampled and that the fish population is near carrying capacity.

When a fish species uses a particular microhabitat feature more than it is available in the study stream it can indicate a preference for that habitat type (Johnson 1980; Manly et al. 2002), and thus type III criteria are those that are adjusted for availability of habitat types and are called preference criteria. Preference criteria can be created by adjusting habitat use information with measures of habitat availability by use of an electivity preference function. Electivity functions have been found to result in criteria that are biased toward the least available type of habitat and thus could result in an underestimate of suitable habitat. Habitat suitability criteria that are adjusted for habitat availability can be very different from those based on use (Baldridge and Amos 1981). Bovee et al. (1998) suggest using a stratified (by habitat type) equal effort approach to fish observations, thereby accounting for habitat availability within the sampling approach. For example, if four habitat types are found in a stream (e.g., deep pool, shallow pool, run, and riffle) all are sampled with an equal amount of effort (e.g., same area is sampled for each habitat type) to gather fish observations with an equal chance of observations in all habitat types.

There are several approaches to developing HSC from observation data, and selection of the analytical approach requires careful thought and application. Because HSC are the biological underpinning in conclusions reached from habitat analyses, shifts in the values of what is determined to be high-quality habitat can result in significant effects in the determination of the habitat area that is modeled, thus resulting in wide differences between recommended lake or reservoir levels or instream flows (Cheslak and Garcia 1988). The first step in analyzing data that has been collected for the development of HSC is through histogram analysis to achieve an understanding of the underlying data distribution. Types of responses that can be observed by plotting frequency of use data are monotonic (increasing or decreasing over the range of observations), unimodal (one curve), bimodal (two curves), or polymodal (more than two curves). In the creation of frequency histograms, the width of the bins, or intervals, of the habitat variable should be small enough to present the information meaningfully, but the width is obviously limited by the level of precision in the measurements (e.g., velocity measured only to the nearest 1 cm/s or 0.1 cm/s). One of the biggest problems with using the frequency analysis approach is the subjectivity in the method, beginning with the decision on the size of the bin or interval widths for a frequency histogram. One objective approach to determining bin widths is to use the Sturges (1926) equation:

$$C = \frac{R}{(1 + 3.222 \cdot \log_{10} N)}, \tag{17.2}$$

where, C is the optimal interval size, R is the range of the observed habitat variable (e.g., $\text{velocity}_{max} - \text{velocity}_{min}$), and N is the number of observations (Cheslak and Garcia 1988). Once an initial histogram is developed, it can be smoothed by adding together bins (incrementally, two or three at a time), summing their frequency, and plotting the bar over the mean of the two variable measures (Slauson 1988). Once an endpoint is reached with an adequately smoothed histogram, the frequencies on the y-axis can be normalized to 1.0, and lines can be drawn from the midpoints of the histogram bars (i.e., averaged frequencies) to develop a curve with associated habitat suitability scores along that continuous curve. Cheslak and Garcia (1988) found that the use of a running 3-point mean was an effective approach after appropriate bin sizes were determined. Smoothness in this case is obtained to the detriment of accuracy as it may inadequately reflect the source data. Furthermore, there is no approach to determining a measure of error by means of the residuals with this approach (Slauson 1988) (Box 17.3).

A second approach to developing HSC retains the original frequency data in the form of a scatterplot, and a curve is fit using nonlinear regression analysis. Quadratic, cubic, or higher-order polynomials can be fit to the data. This approach provides statistical measures of reliability such as an R^2 value to determine goodness of fit. The intercept can be forced through 0 when, in the case of depth, a 0 value for water depth indicates that fish will not be present. A forward selection model can be used to help determine the appropriate order of the polynomial model to fit the data. Furthermore, an F-test or t-test can be used to determine if the polynomial model significantly fits the data, and polynomials are added until it is appropriate to fail to reject the null hypothesis (Zar 1996). If there is a large discrepancy in the numbers of observations for each interval (e.g., 100 versus 6), it is appropriate to use log-transformed data for analyses, and 1 can be added to all observations to avoid taking the log of 0 (Zar 1996).

Alternatively, nonparametric tolerance limits (NPTLs) offer a nearly distribution-free approach to defining habitat use. Note that in this case, the term tolerance does not imply biological tolerance, but rather it is a statistical term similar to a confidence interval. However, whereas a confidence interval defines a range within which an unknown population parameter lies, a tolerance interval is the range in which a certain proportion of the population lies (Slauson 1988). Nonparametric tolerance limits are based on population theory and assume that for any given number of observations, there is, with certainty, a proportion (P) of the population that will lie within given percentiles (50%, 75%, and 95%) when the data are ordered from lowest to highest. Suitability is assigned on the basis of

$$\text{SI} = 2(1 - P) \,, \tag{17.3}$$

where P values of 0.50, 0.75, and 0.95 correspond to suitability index (SI) values of 1.0, 0.5, and 0.1 (Bovee 1986). Somerville (1958) provides the tolerance limit values for large sample sizes (Table 17.4) and Slauson (1988) provides values for small sample sizes (Table 17.5). Both can be used to determine the number of observations that are outside the 50, 75, or 95% range of the population for a

Box 17.3 Development of Habitat Suitability Criteria by Use of Frequency Analysis

A stratified, random, equal-effort approach was used to sample all habitat types and observe subadult smallmouth bass (150–200 mm, total length) in a river. Depth (m) and mean column velocities (m/s) were recorded for each observation. Over 300 observations were measured ($N = 340$).

Table Abridged depth and velocity data for 340 observations of habitat in a river.

Depth (m)	Mean column velocity (m/s)
0.09	0.00
0.58	0.07
2.13	0.03
.	.
.	.
.	.
2.20	1.00

First, an initial frequency histogram is developed using the Sturges equation:

$$C = \frac{R}{(1 + 3.222 \cdot \log_{10} N)}.$$

Then the observations are ordered from least to greatest and the value of the range (R) is calculated by subtracting the least from the greatest.

$$C_{\text{depth}} = \frac{(2.20 - 0.09)}{[1 + 3.222 \cdot \log_{10}(340)]}, \text{ and } C_{\text{velocity}} = \frac{(1.00 - 0)}{[1 + 3.222 \cdot \log_{10}(340)]}.$$

Calculated bin widths are 0.23 m for depth and 0.11 m/s for velocity. The following histograms result.

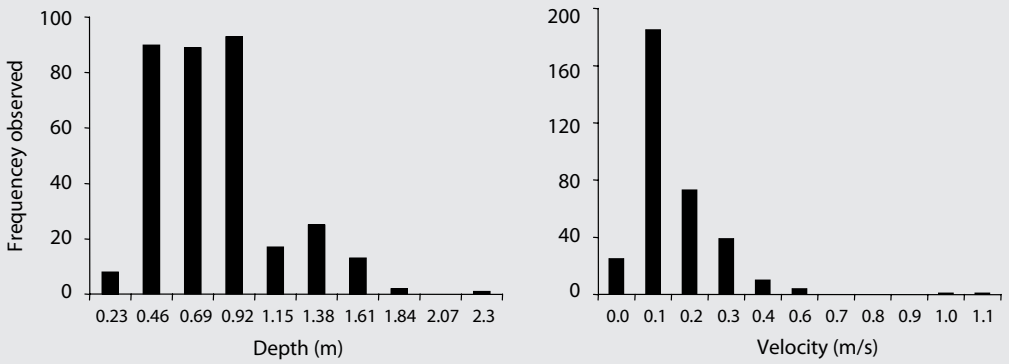


Figure Frequency histograms with bin widths of 0.23 m for depth and 0.11 m/s for velocity.

(Box continues)

Box 17.3 *(continued)*

The next step (if desired at this point) is to smooth the histogram distribution by using a 3-point mean. The mean of three observations is calculated and a new distribution is determined. The investigator must decide how to handle the tails of the distribution. No more than three passes are recommended to smooth the histogram to reduce the error associated with this process (Cheslak and Garcia 1988).

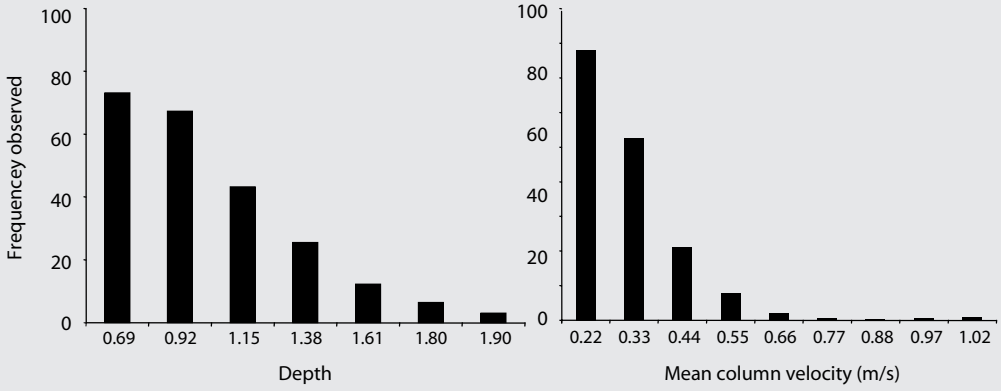


Figure Smoothed histogram obtained by using a 3-point mean. The mean of three observations is calculated and a new distribution is determined.

given confidence interval. Unlike curve fitting, this approach can be used with small data sets and can be replicated by different HSC developers given the same set of data. An assumption is, however, that the habitat observations defined by the central 50, 75, or 95% are biologically relevant (Bovee et al. 1994; Newcomb et al. 1995). Additionally, the investigator must know the distribution of the data to determine if a one-sided approach of the NPTLs is applicable. For example, when organisms require low to 0 stream velocities, the distribution of observations will be heavily skewed toward 0, thus indicating that 0 is a desirable condition, and therefore the limits should be applied to the faster velocities. Criteria can also be categorized within optimal ($\geq$0.5 or $\geq$0.75), suitable ($\geq$0.75 or $\geq$0.1), or unsuitable ($\leq$0.1) ranges and used as nominal data for further analyses (Box 17.4).

All HSC are only as good as the data used to develop them, and several issues have arisen regarding the scale of data collection for fish habitat use relative to the scale at which habitat is modeled (Williams 1996). Use of binary criteria (such as those developed with NPTLs) may address scale issues and associated error in habitat measurements by more broadly defining optimal or useable habitat rather than by a narrow optimum point on an HSC curve. The quality of the resulting criteria depends on several factors including the number of observations (>150

Once the histograms are smoothed, they can be normalized to 1.0 by dividing the y-value for each observation by the maximum y-value in the chart. A curve can be drawn by creating an *x–y* plot with the histogram data. At this time, it is also important to understand the original frequency distribution to determine whether or not the curve could have an optimum value at 0.

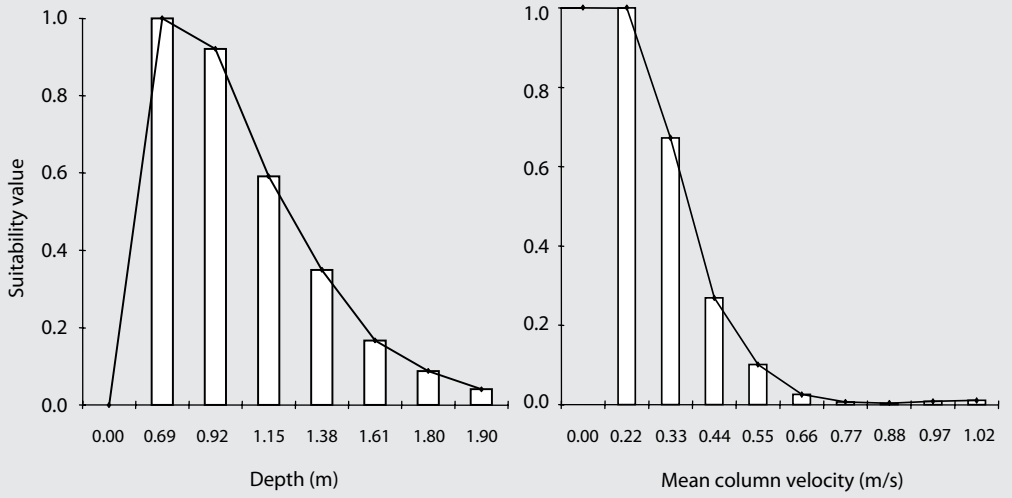


Figure Normalized histogram obtained by dividing the *y*-value for each observation by the maximum *y*-value in the data.

for developing a smooth curve), state of the stream habitat (degraded or not) from which the fish is able to choose, and the sampling design (stratified random sampling or adjusted for availability). It is also important to note that curve development should be used only for continuous and not categorical data.

Categorical data are used also to establish a suitability index value. For instance, ratings of excellent, good, average, and below average could also correspond to rating values of 4, 3, 2, and 1. To this, a habitat suitability index (HSI) value of 1.0 (excellent, 4), 0.75 (good, 3), 0.50 (average, 2), and 0.25 (below average, 1) could be applied (USFWS 1981). Categorical assignments are often used to describe sediment (e.g., modified Wentworth scale) or instream cover.

17.6.2 Guild- or Community-Based Approaches

The use of single-species information to optimize instream habitat may result in habitat limitations for biologically or recreationally important species. Additionally, the assessment of HSC for multiple species and life stages can lead to a dizzying array of curves to reconcile. An alternative approach to habitat assessment is to use a group of fishes with similar habitat requirements. Root (1967) defines a guild as a group of species that utilize similar resources in similar ways. Habitat

Table 17.4 Nonparametric tolerance limit values for large sample sizes (table from Somerville 1958). Given is the proportion (P) of the population (that is, number of observations) that will lie within the tails of the distribution for given percentiles (0.50–0.99) and different confidence levels (γ). Permission for reprint provided by the Institute of Mathematical Statistics.

	P																									
	$\gamma = 0.50$					$\gamma = 0.75$					$\gamma = 0.90$					$\gamma = 0.95$					$\gamma = 0.99$				
n	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99
50	25	12	5	2	0	22	10	3	1		20	9	2	1		19	8	2			16	6	1		
55	28	14	5	3	0	25	12	4	2		23	10	3	1		21	9	2			19	7	1		
60	30	15	6	3	0	27	13	4	2		25	11	3	1		24	10	2	1		21	8	1		
65	33	16	6	3	0	30	14	5	2		27	12	4	1		26	11	3	1		23	9	2		
70	35	17	7	3	1	32	15	5	2		30	13	4	1		28	12	3	1		25	10	2		
75	38	19	7	4	1	35	16	6	2		32	14	4	1		30	13	3	1		27	10	2		
80	40	20	8	4	1	37	17	6	3		34	15	5	2		33	14	4	1		30	11	2		
85	43	21	8	4	1	39	19	7	3		37	16	5	2		35	15	4	1		32	12	3		
90	45	22	9	4	1	42	20	7	3		39	17	5	2		37	16	5	1		34	13	3	1	
95	48	24	9	5	1	44	21	7	3		41	18	6	2		39	17	5	2		36	14	3	1	
100	50	25	10	5	1	47	22	8	3		44	20	6	2		42	18	5	2		38	15	4	1	
110	55	27	11	5	1	51	24	9	4		48	22	7	3		46	20	6	2		43	17	4	1	
120	60	30	12	6	1	56	27	10	4		53	24	8	3		51	22	7	2		47	19	5	1	
130	65	32	13	6	1	61	29	11	5		58	26	9	3		56	25	8	3		52	21	6	2	
140	70	35	14	7	1	66	31	12	5	1	62	28	10	4		60	27	8	3		56	23	6	2	
150	75	37	15	7	1	71	34	12	6	1	67	31	10	4		65	29	9	3		61	26	7	2	
170	85	42	17	8	2	81	39	14	7	1	77	35	12	5		74	33	11	4		70	30	9	3	
200	100	50	20	10	2	95	46	17	8	1	91	42	15	6		88	40	13	5		84	36	11	4	
300	150	75	30	15	3	144	70	26	12	2	139	65	23	10	1	136	63	22	9	1	130	58	19	7	
400	200	100	40	20	4	193	94	36	17	3	187	89	32	15	2	184	86	30	13	1	177	80	27	11	
500	250	125	50	25	5	242	118	45	22	3	236	113	41	19	2	232	109	39	17	2	224	103	35	14	1
600	300	150	60	30	6	292	143	55	26	4	284	136	51	23	3	280	133	48	21	2	272	126	44	18	1
700	350	175	70	35	7	341	167	65	31	5	333	160	60	28	4	328	156	57	26	3	319	149	52	22	2
800	400	200	80	40	8	390	192	74	36	6	382	184	69	32	5	377	180	66	30	4	367	172	61	26	2
900	450	225	90	45	9	440	216	84	41	7	431	208	79	37	5	425	204	75	35	4	415	195	70	30	3
1000	500	250	100	50	10	489	241	94	45	8	480	233	88	41	6	474	228	85	39	5	463	219	79	35	3

Table 17.5 Nonparametric tolerance limit values for 50 habitat use observations or less. Given is the proportion (*P*) of the population (that is, number of observations) that will lie within the tails of the distribution for given percentiles (0.50–0.95) and different confidence levels. Tolerance values are derived from Murphy (1948) as presented in Slauson (1988).

	Confidence level 0.90				Confidence level 0.95			
n	0.50	0.75	0.90	0.95	0.50	0.75	0.90	0.95
15	6	2			5	2		
20	8	3	1		7	2		
25	10	4	1		9	3		
30	12	5	1		11	4	1	
35	14	6	2		14	5	1	
40	16	7	2	1	15	6	2	
45	18	8	3	1	17	7	2	

use data can be measured as outlined above for several species, and guilds can be defined using cluster analysis or other multivariate approaches. Examples of groups include riffle, run, pool, and stream margin guilds (Leonard and Orth 1988). Habitat suitability criteria can be developed from measurements on species within each of the identified guilds. When using guild- or community-based approaches to determining instream flow needs, a representation of different guild types should be evaluated as the use of only pool or only riffle species can result in flow recommendations that are erroneously high or low depending on the guild used (Leonard and Orth 1988).

In more productive, often warmwater, habitats, there are many fish species, and habitat evaluation must be either extremely general (e.g., rapid bioassessment procedures habitat quality index; Barbour et al. 1999) or focused on one or a few target organisms. Either approach is flawed. If all fish species are grouped into some guild typology, numerous species can be simultaneously evaluated with guild-specific criteria (Box 17.5).

17.6.3 Sample Sizes Required for Developing Habitat Suitability Criteria

The number of observations required to develop an accurate representation of fish use of their habitat is an important consideration. Collecting habitat use information can involve a significant amount of financial and personnel resources. Bovee (1986) generally recommended sample sizes of around 150 to gain a good perspective of the distribution of the fish use of habitat. However, in reality, the greater the variability in habitat use, the larger the sample size required to capture this information, and conversely, fish species with very specific habitat preferences may require fewer observations. One valuable approach that may be useful for project scoping is based on a standard normal distribution (Ott 1993). Sample size can be determined according to the following:

$$n = (z_{\alpha/2})^2 \frac{\sigma^2}{E^2}, \tag{17.4}$$

Box 17.4 Use of Nonparametric Tolerance Limits to Develop Habitat Suitability Criteria

Data are first ordered from lowest to highest for the parameter of interest. Here we will use depth and velocity data from Box 17.3. It is useful to develop a histogram to determine the distribution of data and determine if a one-tailed or two-tailed application of the tolerance limits is required.

Once the data are ordered, refer to Tables 17.4 or 17.5 to determine the interval ranges. Using a confidence level of 0.95, we find that we need to interpolate between the values of 300 and 400 to find the correct values. Simple linear regression can be used to interpolate between the two points in the chart.

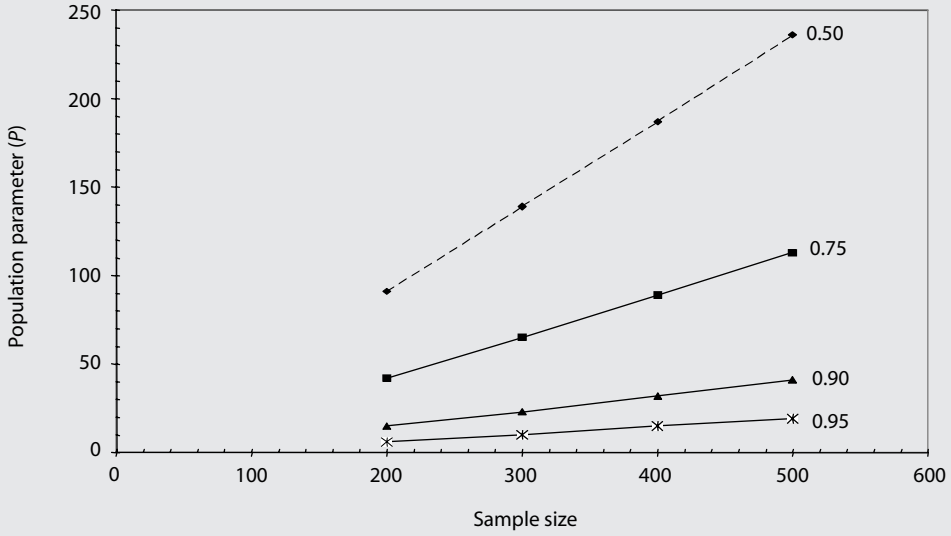


Figure Linear interpolation of nonparametric tolerance limits for intermediate sample sizes (see Table 17.4).

From this, we find that our population proportion *(P)* values for a sample size of 340 and a confidence level of 0.95 to be 45 observations for the central 0.90, 75 for the central 0.75, and 160 for the central 0.50 of the population.

From the original histograms plotted according to the Sturges equation (see Box 17.3), we see that both distributions are two tailed, although the velocity distribution is skewed toward 0. Histogram analysis shows that using a running mean, the velocity distribution could be interpreted as being a one-tailed distribution. For the purposes of illustration, we will assume that depth is a two-tailed distribution and velocity is a one-tailed distribution. We then take the ordered data and determine the values that are within the desired 0.50, 0.75, and 0.90 ranges by subtracting the designated number of observations from the tails.

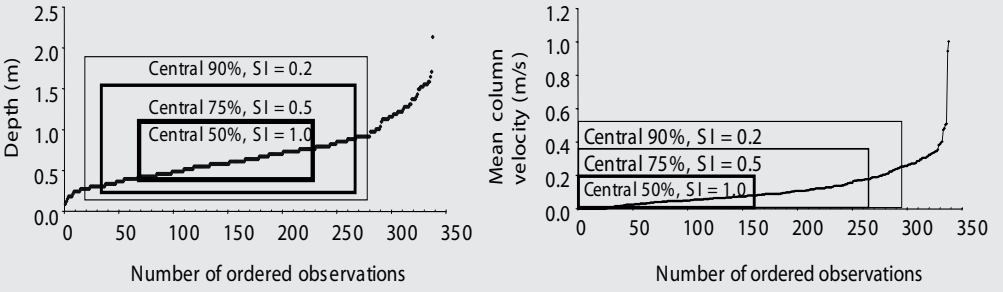


Figure Application of nonparametric tolerance limits to depth and velocity data. Depth is assumed to be a two-tailed distribution and velocity a one-tailed distribution. Using the ordered data, determine the values that are within the desired 0.50, 0.75, and 0.90 ranges by subtracting the designated number of observations from the tails.

From this, we determine the suitability index, SI = 2(P – 1). Our habitat suitability criteria for depth and mean column velocity for subadult smallmouth bass based on nonparametric tolerance limits are given in the table below.

Table Habitat suitability criteria for depth and mean column velocity for subadult smallmouth bass. The suitability index, SI, is given by 2(P – 1).

Depth (m)		Mean Column Velocity (m/s)	
SI	Range	SI	Range
1.0	0.43–0.85	1.0	0.0–0.09
0.5	0.34–0.42 and 0.86–1.16	0.5	0.10–0.17
0.2	0.27-0.33 and 1.17–1.28	0.2	0.18–0.24

For graphical display, tolerance limits are best presented with a histogram or a modified stem plot. A curve indicates interpolation between two points, and this is not appropriate for these tolerance limits.

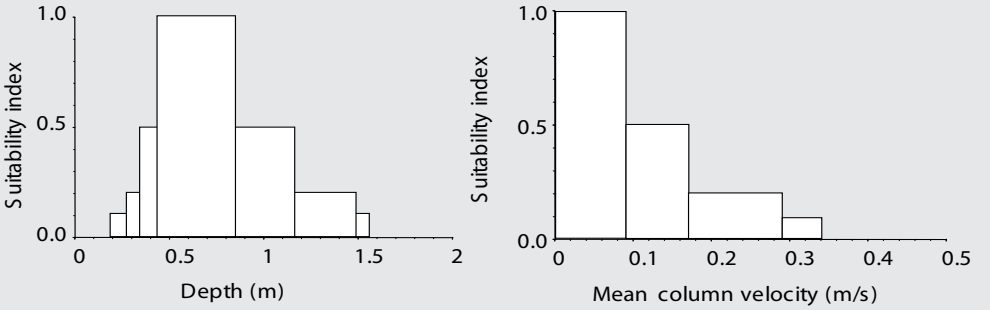


Figure Graphical display of habitat suitability criteria for smallmouth bass.

Box 17.5 Use of Guilds in Developing Habitat Suitability Criteria

In this example, four habitat use guilds (Vadas and Orth 2000) were used to describe the habitat use for 37 species of fish in a warmwater stream. Discriminant analysis of habitat use data reported by Persinger (2003) illustrates the significant differences in habitat variables among the four habitat use guilds. The majority of habitat use measurements at fish locations were correctly assigned to guilds based on linear discriminant functions. The table below shows the number of observations and percent classified into each habitat use guild (Persinger 2003). Only the pool–run guild (a transition group) had the majority of observations misclassified.

Table Number of observations (habitat use measurements at fish locations) and percent classified (in parentheses) assigned to each predetermined habitat use guild (Persinger 2003). Pool–cover is a combination habitat type that is a mixture of open pool and pool habitat with a significant amount of cover.

Fish observations by assigned habitat guild type	Habitat type				
	Riffle	Fast generalist	Pool–run	Pool–cover	Total
Riffle	217 (64.4)	90 (26.7)	11 (3.3)	19 (5.6)	337 (100)
Fast generalist	120 (32.6)	157 (42.7)	22 (6.0)	69 (18.8)	368 (100)
Pool–run	59 (25.1)	82 (34.9)	21 (8.9)	73 (31.1)	235 (100)
Pool–cover	47 (12.8)	83 (22.6)	21 (5.7)	216 (58.9)	367 (100)

The derived habitat–stream discharge relation for the riffle guild of fishes was very similar to the relations for two different species within this guild (see Figure below).

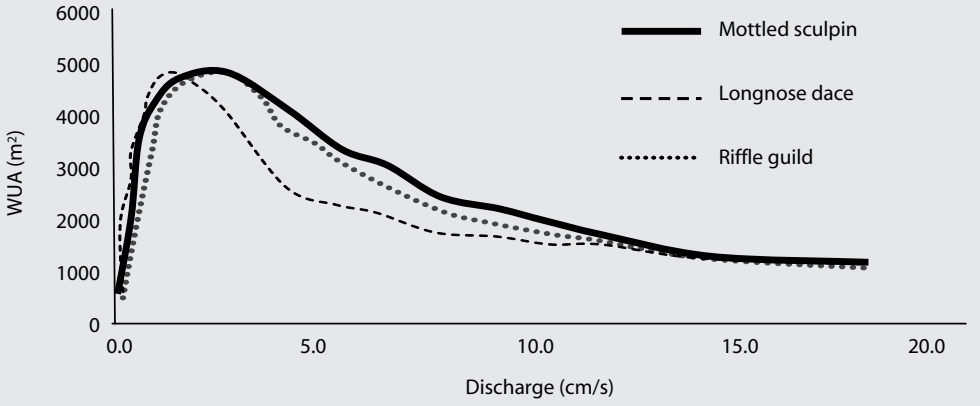


Figure Physical habitat modeling results for mottled sculpin, longnose dace, and the riffle guild at a study site on the North Fork Shenandoah River based on only the depth and velocity criteria (from Vadas and Orth 2000). Shown is the weighted usable area (WUA) versus discharge.

where n = recommended sample size, z = the critical value derived from the standard normal distribution, σ^2 = assumed population variance, and E = the desired precision in terms of a plus-or-minus bound on the true mean. A pilot sample is required to gather information on a selected number of individuals to generate the standard deviation. Alternatively, the information could be obtained from a previous study of similar parameters. However, once again this effort is aimed at deriving information about the mean rather than about the range of habitat usage.

17.6.4 Verification and Transferability of Habitat Suitability Criteria

Collecting site-specific criteria can be expensive and time consuming, so often a measure of transferability is desired to evaluate the potential for HSC developed in one stream to describe fish habitat use in other streams. The testing for transferability is a question of goodness of fit of the distribution of the data. The developed criteria should adequately describe not only the "best" habitat for the species or community but also the appropriate range of habitat values. Thus we are interested in both population metrics for the purposes of quantifying habitat. Data collected from an aquatic system that has the functional hydrologic and instream habitat characteristics of a high-quality system can be used as a standard for comparison. Subjective approaches to assessing the transferability of HSC from one location include visual inspection of one set of frequency diagrams with another. Other statistical approaches include tests of point estimates and frequency distributions (analysis of variance or Kruskal–Wallis and Kolmogorov–Smirnov or chi-square). The abbreviated convergence approach (Bovee 1986) is a visual method that involves collecting a subsample of fish habitat use information and overlaying the frequency histograms on the developed criteria. If the tails and center of the distribution correspond, one could conclude that the HSC are appropriate for use in the new stream. On the other hand, if they don't overlap, it doesn't necessarily mean that the HSC are invalid for use, but negative results could be a product of a small sample size. Statistically, a Kolmogorov–Smirnov test can be used to test for differences in habitat variables with continuous distributions such as depth and velocity (Box 17.6) and a chi-square or *G*-test can be used to test for differences in categorical variables (Sokal and Rohlf 1995).

One transferability test that is relatively straightforward uses a one-sided chi-square to test for random or nonrandom selection of habitat locations by fish (Thomas and Bovee 1993). In this case, sample locations in the stream or lake are assigned a categorical rating (e.g., optimum or usable or suitable or unsuitable) based on single parameters or an index (section 17.6.5). Sample locations can be determined randomly or with a grid sampling design. Locations are noted as occupied or not occupied by fish and recorded accordingly. Those locations that are rated higher should be occupied at a higher frequency than those locations that are not; if the null hypothesis is rejected the criteria are determined acceptable for describing habitat quality in a particular stream. To avoid unacceptable error rates with this approach, sample sizes should be at least 55 occupied locations and 200 unoccupied locations. Thus this approach has been criticized because of its

Box 17.6 Comparison of Two Distributions by Means of Kolmogorov–Smirnov Analysis

Velocity observations were collected from river 1 for development of habitat suitability criteria. Later, it was desired to use the criteria from river 1 to describe habitat use in river 2. A small sample of observations was collected from river 2 and compared with river 1 by means of a Kolmogorov–Smirnov analysis. The hypothesis is that the two samples are distributed identically.

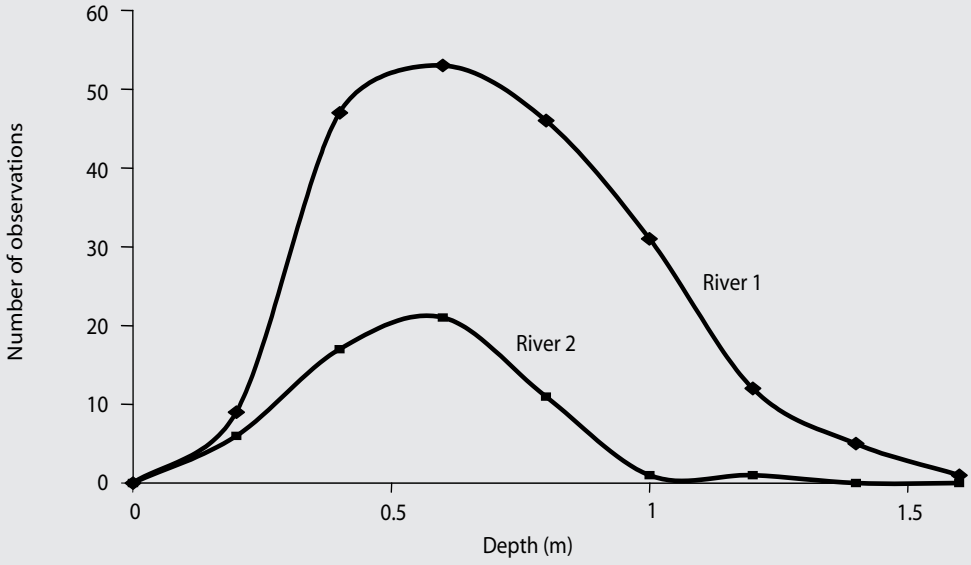


Figure Frequency distribution of number of observations at each depth for two rivers.

dependence on the large number of "no fish" observations required. However, in any approach to measurement of habitat availability, a large number of observations is required to describe the habitat adequately, and thus while the method quantifies the number of observations, the amount isn't unusual in determining general habitat availability (Box 17.7).

17.6.5 Habitat Suitability Index Models

Habitat suitability index (HSI) models were developed to represent the total habitat quality of a site for a species of interest compared with the optimum conditions for that species; HSI values vary from 0 to 1 (USFWS 1981). An HSI model is a composite of all the HSC that are important life requisites for a species. For example, juvenile channel catfish may have food, cover, and water quality life requisites. Once life stages and life requisites are determined, specific, measurable attributes of the habitat for each life requisite are identified. Next, a suitability index

Program

The SAS program for a Kolmogorov–Smirnov test follows.

```
PROC IMPORT OUT= WORK.TROUT
DATAFILE= "D:\K-Sdata.txt"
DBMS=TAB REPLACE;
GETNAMES=YES;
DATAROW=2;
RUN;
PROC NPAR1WAY DATA=WORK.TROUT;
CLASS RIVER;
VAR DEPTH;
RUN;
QUIT;
```

Results

The results of the Kolmogorov–Smirnov (KS) two-sample test (asymptotic; KSa) were as follows.

KS	0.135017	*D*-statistic	0.306473
KSa	2.247135	*P* > KSa	<0.0001

The Kolmogorov–Smirnov (KS) statistic represents the maximum deviation from the empirical distribution function. The KSa is the asymptotic KS statistic, which corrects for sample size and equals the square root of the sample size multiplied by the KS statistic. The KSa value is used when the product of the sample sizes is less than 10,000. The two-sample Kolmogorov–Smirnov test statistic is given by *D*, and the *P*-value is given for KSa. In this example, the two distributions were found to be significantly different, and thus river 1 data may not adequately describe the habitat use in river 2.

(SI) curve (also known as HSC) or relationship is developed for each habitat attribute representing the suitability of the habitat on a 0–1 scale for the species at different values of the habitat variable (section 17.6.1). Therefore, if water quality is a life requisite for channel catfish juveniles, then total dissolved solids might be a suitable variable to measure to represent water quality suitability. The number of total variables in an HSI model varies but should adequately reflect how the animal responds to its habitat.

To apply an HSI model, the investigator must sample all the habitat attributes of the model within the river reach or lake area to be evaluated. The suitability of the site based on each individual habitat variable is then determined from the SI relationship. These values then must be combined to reflect the overall HSI of the site. Usually, the life requisite suitability values (LRSI) are calculated, and the HSI value is determined from the LRSIs. Suitability index values determined from the HSC may be combined in several ways to generate the LRSIs. These approaches include the following.

Box 17.7 Transferability Testing of Habitat Suitability Criteria by Means of a One-Sided Chi-Square Analysis

Habitat suitability criteria were developed for smallmouth bass in Cedar Creek using nonparametric tolerance limits (Thomas and Bovee 1993). A flow alteration is proposed for Town Creek, a warmwater stream of reasonable comparison with Cedar Creek in size, hydrology, and species community. There is interest in using Cedar Creek criteria for quantifying habitat in Town Creek, but first a test of transferability must be done. Optimum habitat is defined by that with a suitability value of 1, usable habitat is that rated as 0.5 or greater, suitable habitat is rated by any values greater than or equal to 0.1, and unsuitable habitat is that with values that fall outside the 0.1 rating (as in the last figure in Box 17.4).

A 300-m section of stream was sampled using a homogenous cell grid of 1 m^2, resulting in 1,500 observations. The stream reach was snorkeled, and smallmouth bass were located in 180 locations. Depth and velocity were recorded in all cells and at the fish locations. Depth and velocity were multiplied together to formulate a composite cell value. Only cells with both optimum depth and velocity were categorized as optimum cells. If either value was unacceptable, the cell was rated as unacceptable. Any other combinations were rated as useable.

Verifying transferability of the criteria requires hypotheses tests that first identify the use of suitable (optimum + useable) versus unsuitable and then optimum versus useable. Thus, this test looks at the full range of habitat use as well as the central location as identified by the optimum habitat classification.

The following data resulted.

Table Smallmouth bass presence (occupied) versus absence (unoccupied) in suitable and unsuitable habitat.

	Suitable	Unsuitable	Total
Occupied cells	92 (cell a)	88 (cell b)	180
Unoccupied cells	733 (cell c)	587 (cell d)	1,320
Total	825	675	1,500

1. Arithmetic mean ($\Sigma SI_i/n$). This approach is the best when variables combine to provide a single resource, but all are not required to have suitable habitat. The only way that habitat suitability can be 0 is if all the variables in the equation are 0.
2. Geometric mean ($[SI_1 \cdot SI_2 \cdot SI_n]^{1/n}$). This approach is best used when a 0 suitability for any variable will result in an overall value of 0.0. However, for the same set of habitat values, the geometric mean will be less than or equal to the arithmetic mean.
3. Minimum function. For $i = 1$ to n, the smallest value is used (e.g., the lowest SI value of variable 1 to 3, SIV_1, SIV_2, or SIV_3). This approach is applied when the concept of limiting factors is believed to apply (USFWS 1980).

To test the null hypothesis, H_o, that suitable cells are occupied in equal proportion to unsuitable cells, versus the alternative hypothesis, H_a, that suitable cells are occupied in greater proportion, the following computation is made:

$$T = [N^{0.5}(ad - bc)]/[(a + b)(c + d)(a + c)(b + d)]^{0.5}.$$

$$T = [1{,}500^{0.5}(92 \cdot 587 - 88 \cdot 733)]/[(180)(1{,}320)(825)(675)]^{0.5} = -1.1.$$

The significance level of T is determined from the standard normal distribution table. In this example, there is a 27% probability of a greater value of T, and the null hypothesis fails to be rejected.

If the suitable versus unsuitable test was significant, then a second test of optimal habitat versus useable habitat is required. To test the H_o that optimal cells are occupied in equal proportion as useable cells versus the H_a, that optimal cells are occupied in greater proportion, the following table is constructed and a T-value can be computed as above.

Table Smallmouth bass presence (occupied) versus absence (unoccupied) in optimal versus useable habitat.

	Optimal	Useable	Total
Occupied cells	30 (a)	62 (b)	92
Unoccupied cells	173 (c)	560 (d)	733
Total	203	622	825

4. Additive function (ΣSIV_i). This approach is the best when two or more variables are supplemental (e.g., different food resources summed together to define total food). If the final value is greater than 1, it should be scaled back to 1.

17.7 QUANTIFICATION OF FISH HABITAT

The approaches outlined in section 17.6 illustrate how to define and qualify fish habitat, but a further step is required to quantify habitat areas. Combinations of habitat suitability criteria and indices are used in different modeling applications to predict the amount of area that can be defined as quality habitat in a lake, reservoir, or river.

17.7.1 Transect- and Grid-Based Approaches

Often when studying fish populations, investigators desire to compare sites, reaches, streams, or lakes that they have sampled for aquatic biota to compare for relative differences in aquatic habitat. A simple, straightforward approach is to use a transect-based method that defines sampling locations in a systematic fashion. For example, in sampling a small pond, six transects could be located, and measurements of DO, temperature, vegetation, and pond depth could be recorded every 2 m laterally across the stream and then at specified depths at each lateral location. Transect spacing is important and should be considered in light of the habitat available (see Chapter 3 for systematic sampling). If the habitat is very diverse, then a greater number of transects and point location measurements should be recorded. In general, the measurements taken at one location should characterize the area bounded by the cell, which is bounded by locations halfway between transect locations and each lateral measurement. Analyses of the data can include parametric analyses for continuous data for comparison between locations and chi-square analysis for categorical data. When significant differences are found in a chi-square analysis, it is often desirable to use a decomposition of the chi-square to determine the location of the differences between the categories. Transect approaches provide the basis for data collection in the following habitat evaluation procedure and in habitat-based hydrologic modeling for streams.

17.7.2 Habitat Evaluation Procedure

The habitat evaluation procedure (HEP) was developed to document the quality and quantity of habitat for a species (USFWS 1980). The HEP approach is intended for use in evaluating baseline habitat conditions with HSI models and in assessing the impacts of various actions that may occur, such as development of a dam or habitat management activities. The HSI values for a selected species are determined empirically and then modeled for various scenarios based on expected changes in habitat conditions. The analysis unit used in HEP is the habitat unit (HU), which is determined as the product of the HSI of the site multiplied by the area (HSI × area = HU). Thus, 1 HU represents one spatial unit (e.g., acre or hectare) of the optimal habitat for the species. Habitat units are determined for the baseline conditions and for future conditions under various scenarios and provide information concerning expected impacts on habitat potential. The HEP process also can be used to evaluate potential mitigation and compensation scenarios to alleviate habitat impacts (Box 17.8).

17.7.3 One- and Two-Dimensional Hydrologic Models

Water allocation and hydropower management are two resource uses for which it is advantageous to evaluate incrementally the trade-offs of numerous flow regimes. Water for irrigation, municipal use, and generation of electricity has tremendous socioeconomic value, and although justification and legal precedence for instream

Box 17.8 Application of a Habitat Suitability Index (HSI) Model and the Habitat Evaluation Procedure (HEP)

To illustrate the application of HSI models in a HEP context, we will use a portion of a bluegill model (Stuber et al. 1982). This model is somewhat complex, with five life requisites modeled for the riverine version of the model, and four life requisites represented by the lacustrine model. However, it is not unreasonable to apply a life requisite submodel when that particular life requisite is considered to be the most limiting aspect of habitat quality for a species. Accordingly, here we will use the life requisite of cover in the lacustrine model to represent overall bluegill habitat quality. This implies that cover is most limiting to the species in this particular instance.

Lacustrine Bluegill Cover Life Requisite

The three variables that constitute the cover life requisite for lacustrine bluegill habitat are percent cover of logs and other objects (V_1), percent cover of aquatic vegetation (V_2), and percent littoral area (V_3). For clarity, variables have been renumbered from the original model.

Suitability index (SI) curves are developed for each variable based on the available data.

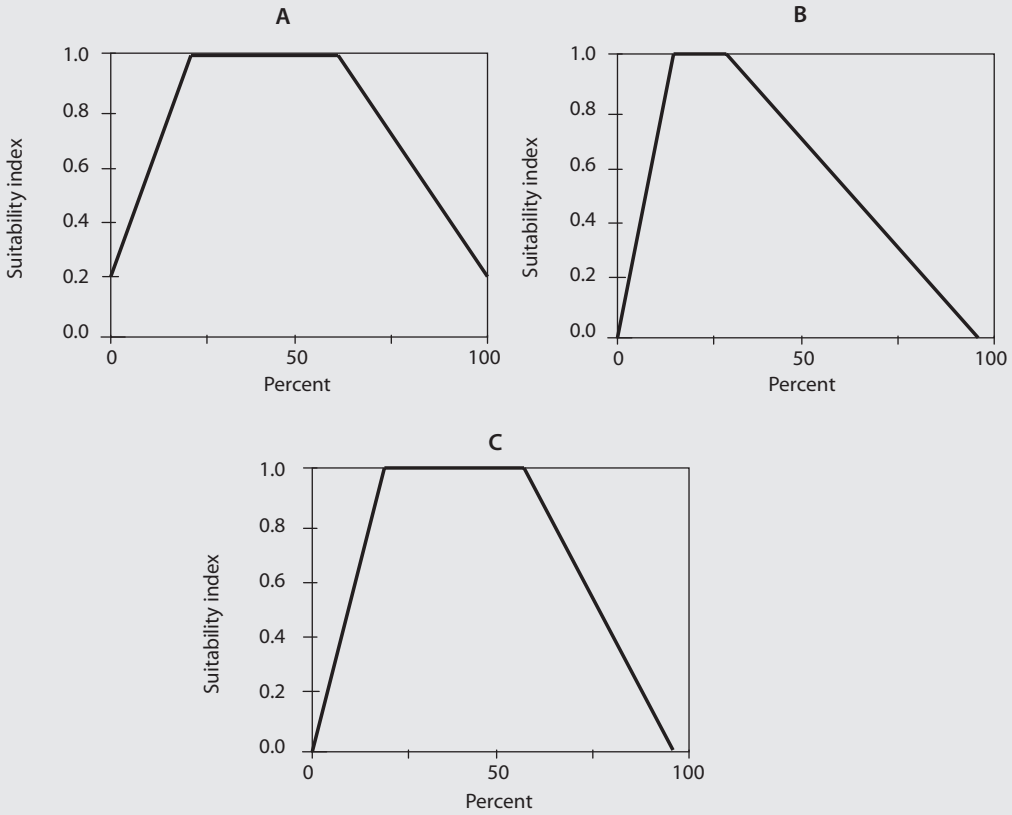


Figure Suitability indices for cover life requisite for lacustrine bluegill habitat: **(A)** percent cover (e.g., logs, brush, and debris) within pools or littoral areas during summer (V_1); **(B)** percent cover (aquatic vegetation—submersed, dense stands, and finely divided leaves) (V_2); and **(C)** percent littoral area during summer stratification (V_3).

(Box continues)

Box 17.8 *(continued)*

The life requisite suitability index (LRSI) is found as a weighted geometric mean of the suitability of these three variables:

$$LRSI_{cover} = (SI_1 \cdot SI_2 \cdot SI_3^2)^{0.25}.$$

Note that the developers of the model assumed that the percent littoral area is more important than the other two variables in determining overall cover quality, and this V_3 was weighted accordingly. This $LRSI_{cover}$ value can be taken to be the HSI value if cover is assumed to be limiting.

Application of the Model to Data

Now, assume that we were interested in the value of bluegill habitat in a 50-ha impoundment. The first step would be to sample and estimate values of the three variables for the site appropriately. Once this was done, it is necessary to determine the SI associated with each habitat value from the above curves. For example, assume we collected the following data.

Table Data necessary to determine LRSI for bluegill in a 50-ha impoundment.

Variable	Field value (%)	Associated SI value
V_1 Percent cover (e.g., logs, brush, and debris) within pools or littoral areas during summer	10	0.40
V_2 Percent cover (aquatic vegetation—submersed, dense stands, and finely divided leaves)	50	0.71
V_3 Percent littoral area during summer stratification	30	1.0

Our LRSI based on cover is then calculated as

$$(0.4 \cdot 0.71 \cdot 1^2)^{0.25} = 0.73.$$

uses of water is increasing, detailed models assist in evaluating costs and benefits associated with different flow regimes.

Hydrologic models use engineering principles of flow to model lotic habitat at different flows. Within the one-dimensional (1-D) instream flow incremental methodology (IFIM), the physical habitat simulation system (PHABSIM) offers three different hydrologic models (HEC-2, IFG-4, and WSP) to predict depths and velocities in defined habitat cells within a study area (Bovee 1986). The hydrologic model is paired with biological information (HSC; section 17.4) to produce weighted usable area (WUA) or, if binary criteria are used, useable area (UA). For comparisons between sites of unequal area and for describing habitat availability, percent useable area (PUA) can be defined by 100 · (WUA/area) (LeClerc et al. 1995). Other hydrologic models also use criteria to quantify habitat under altered flow scenarios (Heggenes et al. 1996).

Thus, we can say that the LRSI for the impoundment for bluegill is 0.73 on a 0.0 to 1.0 scale. Within the HEP, the common unit of currency and comparison is the habitat unit (HU), which is found as

$$HU = HSI \cdot area.$$

Thus, 1 HU represents one unit (e.g., acre or hectare) of area of optimal (HSI = 1.0) habitat. In our bluegill example, the number of bluegill habitat units for the impoundment would be: $50 \cdot 0.73 = 36.5$ HU. This represents a baseline measurement that can be used for comparison among areas or to compare with projected future conditions.

Use of HEP to Assess Effects of Habitat Change

We now know the baseline conditions for the impoundment under consideration (36.5 HU). We can assess potential effects of habitat changes by projecting expected habitat conditions in the future, determining the habitat suitability based on those conditions, and then comparing habitat units between baseline conditions and the expected conditions.

For example, assume that the impoundment was expected to undergo a drawdown in water level. Say that this change resulted in a reduction in area from 50 ha to 40 ha. We might predict that as a result V_1 and V_2 will not change much, so their SI values remain at 0.4 and 0.71, respectively. However, there may be a change in the percent littoral area from 30 to 15%. The associated SI value is 0.75. The calculated LRSI then becomes

$$(0.4 \cdot 0.71 \cdot 0.75^2)^{0.25} = 0.63.$$

The HU associated with the new conditions is $40 \cdot 0.63 = 25.2$ HU. Thus, the net impact of the drawdown can be estimated as a loss of $36.5 - 25.2 = 11.3$ HU.

Recent applications of engineering principles to habitat models include two-dimensional (2-D) models. The approach uses finite elements to define instream habitat. The 2-D models are advantageous in that the spatial resolution of the model can be adapted to the scale of fish habitats, islands can be readily incorporated into the modeling assessment, and there is additional availability of flow resistance correction functions (LeClerc et al. 1995; Ghanem et al. 1996). The approach also allows for analysis of other features such as distance to cover, resting habitats, or feeding habitats. Similar to the stochastic temperature modeling, this approach requires a large amount of data collection and data reduction, which are easier to accomplish as the technology and software advance.

Based on the output from either a 1-D or 2-D hydrologic model, a discharge–WUA relationship can be created. Often this is where the process of analysis ends, and judgments are made about the flow that provides the greatest WUA for the

target species. However, stopping at this point diminishes the value of the incremental models and their output. Further analyses can be conducted to determine actual habitat availability based on historical hydrographs and the habitat models to create a time series of habitat available for additional analyses (Milhous et al. 1990). Additionally, a subsampling routine (i.e., bootstrap) has been used to establish confidence intervals on the WUA output (Williams 1996). For measures of duration or magnitude of available habitat or periods of limited habitat, habitat duration curves (Sale et al. 1981; Milhous et al. 1990) or a continuous-under-threshold habitat-duration curve (Capra et al. 1995) can be developed to analyze temporal features of habitat availability.

17.7.3.1 *Habitat Duration Curves and Exceedance Thresholds*

Habitat duration curves are similar to flow frequency duration curves. Using the WUA function, a historical flow record can be converted to daily values of WUA, depicting available habitat over time. From this record, the daily habitat values are ranked by sorting from largest to smallest. Then, the following formula is used on each value:

$$f(h) = 100 \cdot \frac{r(h)}{n-1}, \tag{17.5}$$

where $f(h)$ represents the frequency at which that ranked habitat value is met or exceeded, $r(h)$ is the rank of habitat value h, and n is the number of events in the period of record. Sale et al. (1981) proposed a nomograph approach to setting flow standards under the duration curve analysis. The y-axis contains the WUA values, the lower x-axis contains the discharge, and a WUA curve is plotted. The upper x-axis is labeled with the probability of exceedance, and the WUA exceedance curve is plotted. A threshold for exceedance of habitat can be established by the stakeholders. The value on the WUA flow curve that corresponds to the exceedance value can then be used to establish the minimum desired flow conditions. This approach can be used with either a single species or guild WUA relationship, but multiple nomographs may need to be considered for setting a final flow.

17.7.3.2 *Continuous-under-Threshold Habitat-Duration Curve*

The approach for developing the continuous-under-threshold habitat-duration curve, or CUT curve, is outlined in Capra et al. (1995). Briefly, a threshold value for WUA is determined (this step establishes the magnitude of the habitat limitation). For some period of time (e.g., an annual increment may be most appropriate for evaluating limitations on a species' life history), a time series is evaluated for periods when the WUA is less than the threshold (e.g., 1, 2, or 4 d to establish the duration). Then, the threshold periods are sorted in descending order from the longest to the shortest durations. These values are plotted on the y-axis while the cumulative percentage (Cp_i) of the number of threshold days (d_i) on the total number of days (D) is plotted on the x-axis ($Cp_i = d_i/D + Cp_{i-1}$, where $d_o = 0$). For

comparison, several curves can be drawn on the same figure. This approach still needs to be verified with species other than trout, and some approach for determining thresholds should be evaluated. However, it is a promising approach that moves away from the static approaches to determining flow requirements in streams. But still, the approach must be considered as a bottleneck analysis rather than a unilateral standing stock prediction.

17.7.4 Approaches Based Only on Stream Flow Data

Office approaches are often used when little information is available on stream habitat in a channel or when a quick answer is required. Drawbacks to these methods, however, are that they are not well justified biologically outside the areas or species for which they were developed, nor do they allow for the evaluation of trade-offs when flows must be specified.

17.7.4.1 The Tennant Method

One straightforward office approach valued for its simplicity is the Montana method or Tennant method (Tennant 1975; Table 17.6). In his work, Tennant outlined proportions of a stream's mean annual flow that provided habitat from poor to outstanding ratings; these ratings, however, were based entirely on subjective interpretation of photographs. Although this is a convenient desktop approach to determining flows, assumptions about the conditions provided by each of the proportional categories should be investigated in the field. At best, the Montana method provides a guideline for further investigations.

17.7.4.2 New England Base Flow Method

This approach uses the median monthly flow for the low-flow month of the river, which is typically August in North America. The approach assumes that median historical flows during the low-flow month will sustain indigenous aquatic organisms throughout the year. The drawbacks to this approach are field testing in

Table 17.6 Instream flows for fish, wildlife, recreation, and related environmental resources (adapted from Tennant 1975).

	Recommended base flows	
Description of flow	October–March	April–September
Flushing or maximum	200% of the average flow	
Optimum range	60–100% of the average flow	
Outstanding	40%	60%
Excellent	30%	50%
Good	20%	40%
Fair or degrading	10%	30%
Poor or minimum	10%	10%
Severe degradation	10% of average flow to 0 flow	

different regions is lacking, the nonincremental approach leaves no room for negotiating, and long-term flow data are required.

17.7.4.3 Index of Hydraulic Alteration

The index of hydraulic alteration allows for calculation of the expected variability in flow regimes based on past hydrological patterns (Richter et al. 1997). This approach assumes that natural flow variability is inherent to the health of river ecosystems and channel morphology. The use of this method allows for a recognition of changes in flow patterns as well as providing a target for flow restoration. Sixty-seven index values are identified in this approach.

17.7.4.4 Approximation of Optimum Habitat Based on Weighted Useable Area

Weighted useable area (WUA) is the product of habitat modeling based on a hydraulic habitat model meshed with species models of habitat use (section 17.7.2). The output value is a function of discharge and is used to negotiate for instream flows that approach a species or life stage optimum. Both the hydrologic model and the species models can be complex and expensive to develop. Recent work conducted on salmonid species in the Pacific Northwest has found that mean annual discharge (MAD) is a good approximate of the optimum WUA values (Hatfield and Bruce 2000). Further, the authors provide regression equations for species and life stages, but they caution against using them outside the region. This approach could be used for project scoping or initial evaluation of flow alternatives.

17.8 FUTURE DIRECTIONS FOR HABITAT DATA ANALYSIS

Technology is providing the capability to collect large amounts of information and analyze physical processes in ways that we haven't been able to in the past. The use of hydroacoustics to sample stream and lake depths, bottom type, and vegetation types provides GIS-linked spatial coverage at a large scale with greater precision ($<m^2$ in some cases) than that with traditional physical measurements (e.g., Fisher and Rahel 2004). Additionally, Doppler technology allows the measurement of velocity profiles in flowing water. This creates the ability not only to measure discharge in a quick fashion but also to provide greater detail in describing microhabitat environments. Continuous measurement water temperature and water quality devices provide managers with the ability to analyze real-time fish macrohabitat for incorporation into long-term monitoring efforts as well as modeling. All of these improvements in technology are heading toward new pathways for analysis of ecosystem processes in both lake and stream environments.

Future development and standardization of approaches to habitat evaluation will result from technological advances in the gathering of field measurements and the standardization of large data sets. These advancements will hopefully lead to a better understanding of the linkages between habitat and sustainable fish communities.

17.9 REFERENCES

Anderson, R. M., and R. B. Nehring. 1985. Impacts of stream discharge on trout rearing habitat and trout recruitment in the South Platte River, Colorado. Pages 59–64 38 *in* F. W. Olson, R. G. White, and R. H. Hamre, editors. Symposium on small hydropower and fisheries. American Fisheries Society, Western Division and Bioengineering Section, Bethesda, Maryland.

Armantrout, N. B., editor. 1981. Acquisition and utilization of aquatic habitat inventory information. American Fisheries Society, Western Division, Bethesda, Maryland.

Baldridge, J. E., and D. Amos. 1981. A technique for determining fish habitat suitability criteria: a comparison between habitat utilization and availability. Pages 251–258 *in* Armantrout (1981).

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Bartholow, J. M. 1989. Stream temperature investigations: field and analytic methods. Instream Flow Paper Number 13. U.S. Fish and Wildlife Service Biological Report 89(17).

Bartholow, J. M. 1991. A modeling assessment of the thermal regime for an urban sport fishery. Environmental Management 15:833–845.

Binns, N. A., and F. M. Eiserman. 1979. Quantification of fluvial trout habitat in Wyoming. Transactions of the American Fisheries Society 108:215–228.

Bisson, P. A., J. L. Nielsen, R. A. Palmason, and L. E. Grove. 1981. A system of naming habitat types in small streams, with examples of habitat utilization by salmonids during low streamflow. Pages 62–73 *in* Armantrout (1981).

Bovee, K. D. 1986. Development and evaluation of habitat suitability criteria for use in the instream flow incremental methodology. Instream Flow Information Paper Number 21. U.S. Fish and Wildlife Service Biological Report 86(7).

Bovee, K. D., B. L. Lamb, J. M. Bartholow, C. B. Stalnaker, J. Taylor, and J. Henriksen. 1998. Stream habitat analysis using the instream flow incremental methodology. U.S. Geological Survey, Biological Resources Division, Information and Technical Report USGS/BRD-1998-0004, Fort Collins, Colorado.

Bovee, K. D., T. J. Newcomb, and T. G. Coon. 1994. Relations between habitat variability and population dynamics of bass in the Huron River, Michigan. U.S. Department of the Interior, National Biological Survey, Biological Report 21, Washington, D.C.

Brown, L. C., and T. O. Barnwell Jr. 1987. The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual. U.S. Environmental Protection Agency, Environmental Research Laboratory EPA/600/3-87/007, Athens, Georgia.

Capra, H., P. Breil, and Y. Souchon. 1995. A new tool to interpret magnitude and duration of fish habitat variations. Regulated Rivers: Research and Management 10:281–289.

Cheslak, E. F., and J. C. Garcia. 1988. An evaluation of the effects of various smoothing and curve-fitting techniques on the accuracy of suitability functions. Pages 259–283 *in* K. D. Bovee and J. R. Zuboy, editors. Proceedings of a workshop on the development and evaluation of habitat suitability criteria. U.S. Fish and Wildlife Service Biological Report 88(11).

Claramunt, R. M., and D. H. Wahl. 2000. The effects of abiotic and biotic factors in determining larval fish growth rates: a comparison across species and reservoirs. Transactions of the American Fisheries Society 129:835–851.

Coutant, C. C. 1985. Striped bass, temperature, and dissolved oxygen: a speculative hypothesis for environmental risk. Transactions of the American Fisheries Society 114:31–61.

Crance, J. H. 1987. Habitat suitability index curves for paddlefish, developed by the Delphi technique. North American Journal of Fisheries Management 7:123–130.

Dolloff, C. A., D. G. Hankin, and G. H. Reeves. 1993. Basinwide estimation of habitat and fish populations in streams. U.S. Forest Service, Southeastern Forest Experiment Station, General Technical Report SE-83, Asheville, North Carolina.

Dunham, J. B., and G. L. Vinyard. 1997. Incorporating stream level variability into analyses of site level fish habitat relationships: some cautionary examples. Transactions of the American Fisheries Society 126:323–329.

Fausch, K. D., C. L. Hawkes, and M. G. Parsons. 1988. Models that predict standing crop of stream fish from habitat variables: 1950–1985. U.S. Forest Service, Pacific Northwest Research Station, General Technical Report PNW-GTR-213, Portland, Oregon.

Fisher, W. L., and F. J. Rahel. 2004. Geographic information systems in fisheries. American Fisheries Society, Bethesda, Maryland.

Fraley, J. J., and P. J. Graham. 1981. Physical habitat, geologic bedrock types and trout densities in the tributaries of the Flathead River drainage, Montana. Pages 178–185 *in* Armantrout (1981).

Ghanem, A., P. Steffler, F. Hicks, and C. Katopodis. 1996. Two-dimensional hydraulic simulation of physical habitat conditions in flowing streams. Regulated Rivers: Research and Management 12:185–200.

Hannaford, M. J., M. T. Barbour, and V. H. Resh. 1997. Training reduces observer variability in visual-based assessments of stream habitat. Journal of the North American Benthological Society 16:853–860.

Hatfield, T., and J. Bruce. 2000. Predicting salmonid habitat–flow relationships for streams from western North America. North American Journal of Fisheries Management 20:1005–1015.

Hauser, G. E., and M. C. Walters. 1995. TVA river modeling system, version 1.0. Tennessee Valley Authority Report WR28-1-590-164.

Hawkins, C. P., and 10 coauthors. 1993. A hierarchical approach to classifying stream habitat features. Fisheries (Bethesda) 18:3–11.

Heggenes, J., S. J. Saltveit, and O. Lingaas. 1996. Predicting fish habitat use to changes in water flow: modelling critical minimum flows for Atlantic salmon *Salmo salar*, and brown trout, *S. trutta*. Regulated Rivers: Research and Management 12:331–344.

Hodgetts, R. M. 1977. Applying the Delphi technique to management gaming. Simulation 29:209–212.

Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 6:65–71.

Kondolf, G. M., and E. R. Micheli. 1995. Evaluating stream restoration projects. Environmental Management 19:1–15.

Krause, C. W., B. Lockard, T. J. Newcomb, D. Kibler, V. Lohani, and D. J. Orth. 2004.

Predicting influences of urban development on thermal habitat in a warm water stream. Journal of the American Water Resources Association 40:1645–1658.

Krause, C. W., T. J. Newcomb, and D. J. Orth. 2005. Thermal habitat assessment of alternative flow scenarios in a tailwater fishery. River Research and Applications 21:581–593.

LeClerc, M., A. Boudreault, J. A. Bechara, and G. Corfa. 1995. Two-dimensional hydrodynamic modeling: a neglected tool in the instream flow incremental methodology. Transactions of the American Fisheries Society 124:645–662.

Leonard, P. M., and D. J. Orth. 1988. Use of habitat guilds of fishes to determine instream flow requirements. North American Journal of Fisheries Management 8:399–409.

Lifton, W. S., K. A. Voos, and D. Gilbert. 1985. The simulation of the Pit 3, 4, and 5 hydroelectric project using the USFWS instream temperature model. Pages 1805–1814 *in* Waterpower 1985, volume 3. Proceedings of an international conference on hydropower, Las Vegas, Nevada, September 25–27, 1985. American Society of Civil Engineers, Reston, Virginia.

Lohr, S. C., P. A. Byorth, C. M. Kaya, and W. P. Dwyer. 1996. High temperature tolerances of fluvial arctic grayling and comparisons with summer river temperatures of the Big Hole River, Montana. Transactions of the American Fisheries Society 125:933–939.

Maddock, I. 1999. The importance of physical habitat assessment for evaluating river health. Freshwater Biology 41:373–391.

Manly, B. F. J., L .L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies, 2nd edition. Kluwer Academic Publishers, Dordrecht, The Netherlands.

McClendon, D. D., and C. F. Rabeni. 1987. Physical and biological variables useful for predicting population characteristics of smallmouth bass and rock bass in an Ozark stream. North American Journal of Fisheries Management 7:46–56.

Milhous, R. T., J. M. Bartholow, M. A. Updike, and A. R. Moos. 1990. Reference manual for generation and analysis of habitat time series, version II. Instream Flow Information Paper Number 27. U.S. Fish and Wildlife Service Biological Report 90(16).

Milner, N. J., R. J. Hemsworth, and B. E. Jones. 1985. Habitat evaluation as a fisheries management tool. Journal of Fish Biology 27(Supplement A):85–108.

Modde, T., R. C. Ford, and M. G. Parsons. 1991. Use of a habitat-based stream classification system for categorizing trout biomass. North American Journal of Fisheries Management 11:305–311.

Montgomery, D. R., and J. M. Buffington. 1997. Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin 109(5):596–611.

Murphy, R. B. 1948. Nonparametric tolerance limits. Annals Mathematics and Statistics 19:581–589.

Nielsen, J. L., T. E. Lisle, and V. Ozaki. 1994. Thermally stratified pools and their use by steelhead in northern California streams. Transactions of the American Fisheries Society 123:613–626.

Newcomb, T. J., S. A. Perry, and W. B. Perry. 1995. Comparison of habitat suitability criteria for smallmouth bass (*Micropterus dolomieu*) from three West Virginia rivers. Rivers 5:170–183.

Orth, D. J., and R. J. White. 1993. Stream habitat management. Pages 205–230 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management of North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Ott, R. L. 1993. An introduction to statistical methods and data analysis. Duxbury Press, Belmont, California.

Palmer, M. A., and N. L. Poff. 1997. Heterogeneity in streams: the influence of environmental heterogeneity on patterns and processes in streams. Journal of the North American Benthological Society 16:169–173.

Parsons, M. G., J. R. Maxwell, and D. Heller. 1981. A predictive fish habitat index model using geomorphic parameters. Pages 85–91 *in* Armantrout (1981).

Persinger, J. 2003. Development of habitat suitability criteria for individual species and habitat guilds in the Shenandoah River basin. Master's thesis. Virginia Polytechnic Institute and State University, Blacksburg.

Pfankuch, D. J. 1975. Stream reach inventory and channel stability evaluation. U.S. Forest Service R1-75-002, Government Printing Office 696-260/200, Washington, D.C.

Pitlo, J., Jr. 1989. Walleye spawning habitat in Pool 13 of the upper Mississippi River. North American Journal of Fisheries Management 9:303–308.

Platts, W. S. 1981. Stream inventory garbage in—reliable analysis out: only in fairy tales. Pages 75–84 *in* Armantrout (1981).

Platts, W. S., W. F. Megahan, and G. W. Minshall. 1983. Methods for evaluating stream, riparian, and biotic conditions. U.S. Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report INT-138, Ogden, Utah.

Poff, N. L., and J. V. Ward. 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Sciences 46:1805–1818.

Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. Canadian Journal of Fish and Aquatic Sciences 53 (Supplement 1):252–259.

Richter, B. D., J. V. Baumgartner, R. Wigington, and D. P. Braun. 1997. How much water does a river need? Freshwater Biology 37:231–249.

Root, R. B. 1967. The niche exploitation pattern of the blue-gay gnatcatcher. Ecology 37:317–350.

Rosgen, D. 1994. A classification of natural rivers. Catena 22:169–199.

Rubec, P. J., M. S. Coyne, R. H. McMichael Jr., and M. E. Monaco. 1998. Spatial methods being developed in Florida to determine essential fish habitat. Fisheries (Bethesda) 23:21–25.

Sale, M. J., S. F. Railsback, and E. E. Herricks. 1981. Frequency analysis of aquatic habitat: a procedure for determining instream flow needs. Pages 340–346 *in* Armantrout (1981).

SAS Institute. 2004. SAS/Stat(R) 9.1 user's guide. SAS Institute, Cary, North Carolina.

Seelbach, P. W. 1993. Population biology of steelhead in a stable-flow, low gradient tributary of Lake Michigan. Transactions of the American Fisheries Society 122:179–198.

Slauson, W. L. 1988. Constructing suitability curves from data. Pages 225–255 *in* K. D. Bovee and J. R. Zuboy, editors. Proceedings of a workshop on the development and evaluation of habitat suitability criteria. U.S. Fish and Wildlife Service Biological Report 88(11).

Smith, R. 1999. Differential stability of spawning microhabitats of warmwater stream fishes. Master's thesis. Virginia Polytechnic Institute and State University, Blacksburg.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. Freeman, New York.

Somerville, P. N. 1958. Tables for obtaining non-parametric tolerance limits. Annals of Mathematical Statistics 29:599–601.

Stefan, H. G., X. Fang, and J. G. Eaton. 2001. Simulated fish habitat changes in North American lakes in response to projected climate warming. Transactions of the American Fisheries Society 130:459–477.

Stoneman, C. L., and M. L. Jones. 1996. A simple method to classify stream thermal stability with single observations of daily maximum water and air temperatures. North American Journal of Fisheries Management 16:728–737.

Stuber, R. J., G. Gebhart, and O. E. Maughan. 1982. Habitat suitability index models. U. S. Fish and Wildlife Service, Office of Biological Services, Western Energy and Land Use Team, Washington, D.C.

Sturges, H. A. 1926. The choice of a class interval. Journal of the American Statistical Association 21:65–66.

Sullivan, K., J. Tooley, K. Doughty, J. E. Caldwell, and P. Knudsen. 1990. Evaluation of prediction models and characterization of stream temperature regimes in Washington. Timber/Fish/Wildlife Report TFW-WQ3–90-006, Washington Department of Natural Resources, Olympia.

Tennant, D. L. 1975. Instream flow regiments for fish, wildlife, recreation, and related environmental resources. U.S. Fish and Wildlife Service, Billings, Montana.

Terrell, J. W., and J. Carpenter. 1997. Selected habitat suitability index model evaluations. U.S. Geological Survey, Biological Resources Division, USGS/BRD/ITR-1997-0005, Washington, D.C.

Theurer, F. D., K. A. Voos, and W. J. Miller. 1984. Instream water temperature model. Instream Flow Paper Number 16, U.S. Fish and Wildlife Service FWS/OBS-84/15.

Thomas, J. A., and K. D. Bovee. 1993. Application and testing of a procedure to evaluate transferability of habitat suitability criteria. Regulated Rivers: Research and Management 8:285–294.

USEPA (U.S. Environmental Protection Agency). 1995. QUAL2E windows interface user's guide. USEPA document EPA/823/B95/003, Washington, D.C..

USFWS (U.S. Fish and Wildlife Service). 1980. Habitat evaluation procedures (HEP). Division of Ecological Services, ESM 102, Washington, D.C.

USFWS (U.S. Fish and Wildife Service). 1981. Standards for the development of habitat suitability index models. Division of Ecological Services, ESM 103, Washington, D.C.

Vadas, R. L., Jr., and D. J. Orth. 2000. Habitat use of fish communities in a Virginia stream system. Environmental Biology of Fishes 59:253–269.

VanHorne, B. 1983. Density as a misleading indicator of habitat quality. Journal Wildlife Management 47:893–901.

Waddle, T. J. 1989. Water temperature data analysis and simulation for the Salmon River, Osewgo County, New York, Summer, 1986. U.S. Fish and Wildlife Service, National Ecology Research Center, Fort Collins, Colorado.

Watson, G., and T. W. Hillman. 1997. Factors affecting the distribution and abundance of bull trout: an investigation at hierarchical scales. North American Journal of Fisheries Management 17:237–252.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2003. Classifying regional variation in thermal regime based on stream fish community patterns. Transactions of the American Fisheries Society 132:18–38.

Williams, J. G. 1996. Lost in space: minimum confidence intervals for idealized PHABSIM studies. Transactions of the American Fisheries Society 125:458–465.

Wilson, W. J., M. D. Kelly, and P. R. Meyer. 1987. Instream temperature modeling and fish impact assessment for a proposed large scale Alaska hydroelectric project. Pages 183–206 *in* J. F. Craig and J. B. Kemper, editors. Regulated streams: advances in ecology. Plenum, New York.

Zar, J. H. 1996. Biostatistical analysis. Prentice-Hall, Upper Saddle River, New Jersey.

Zedonis, P. 1997. A water temperature model of the Trinity River. U.S. Department of the Interior, Coastal California Fish and Wildlife Office, Arcata.

Zuboy, J. R. 1981. A new tool for fishery managers: the Delphi technique. North American Journal of Fisheries Management 1:55–59.

18 Watershed Level Approaches

Frank J. Rahel and Donald A. Jackson

18.1 INTRODUCTION

Assessing fish habitat requirements is a major focus of fisheries management and aquatic conservation efforts. Historically, we have tended to think of fish habitat in terms of local conditions such as water depth, current velocity, and cover. However, fish habitat can be viewed at a variety of spatial scales varying from microhabitat conditions to stream channel units or lake zones up to watershed level characteristics (Fausch et al. 2002). Furthermore, these scales form a hierarchy such that local habitat conditions often are the result of processes that operate at much larger spatial scales. For example, regional geology and glacial history can influence the productivity and morphology of lakes (Riera et al. 2000). In streams, basin shape and geology interact with riparian vegetation to determine the types of habitats present (Frissell et al. 1986; Modde et al. 1991). In describing how stream features are controlled by characteristics of the drainage basin, Hynes (1975) noted, "in every respect, the valley rules the stream." This chapter describes approaches used to relate both large-scale habitat features and habitat patchiness to fish abundance patterns in lakes and streams.

18.1.1 The Watershed

A watershed is an area drained by surface and groundwater flow. A drainage basin is a watershed that collects and discharges its surface streamflow through one outlet or mouth. The term catchment refers to a subdrainage or the land area draining toward a specified point of interest within the drainage basin. Watersheds of large rivers are commonly called basins, such as the Missouri River basin or the Ohio River basin. Watersheds exist within a hierarchical framework, such that catchments exist within watersheds that, in turn, are part of basins.

Many state and federal management agencies within the USA use a hierarchical system of hydrologic units developed by the U.S. Geological Survey (Seaber et al. 1987). These units are identified by unique hydrologic unit boundary (HUB) codes that are based on dividing the USA and the Caribbean into 21 major regions, 222 subregions, and 2,149 cataloging units. At the regional level, the HUBs

define zoogeographic provinces that share a similar geologic and evolutionary history and thus similar species pools. However, HUBs are not true topographic watersheds in that many are downstream segments of larger watersheds or collections of several adjacent small watersheds. They are useful to management agencies because it is impossible to divide the country into a finite number of true watersheds at any hierarchical level and because HUB units at any given level are similar in size, which is not true of topographic watersheds. However, because HUBs are not topographically based, they do not always integrate natural and anthropogenic influences acting upstream. As noted by Griffith et al. (1999), studies relating biotic conditions to large-scale habitat features should also consider ecoregion effects. Ecoregions are areas of relative homogeneity of ecological systems caused by similarity in soils, vegetation, climate, geology, physiography, and biogeographic history (Abell et al. 2000). Ecoregion boundaries do not necessarily coincide with watershed boundaries and thus can contribute to within watershed variability in ecological communities.

18.1.2 Importance of Watershed Scale Analyses in Fisheries Management

Many of the most pressing environmental problems involve habitat alterations at large spatial scales. These include regional phenomena such as climate change or acid precipitation as well as diffuse impacts across watersheds due to livestock grazing, timber harvest, road building, and water development. Dealing with these issues requires a large-scale perspective, and thus resource management agencies are increasingly embracing an ecosystem approach to habitat management (Schramm and Hubert 1999). The goal of ecosystem management is to preserve ecosystem integrity while maintaining sustainable benefits for human populations. Achieving this goal requires an understanding of physical and biological interactions occurring over spatial scales from microhabitat to the watershed level. Watersheds provide a natural framework for implementing ecosystem management because habitat conditions at any point on a stream or lake reflect the integration of characteristics in the watershed above that point (Montgomery et al. 1995; Allan 2004). Thus watersheds, rather than political or administrative units, become the fundamental unit for studying large-scale geomorphological and ecological processes.

Another reason for the interest in watershed analysis is an increased awareness that large-scale geomorphological or geological features often determine local habitat characteristics in aquatic systems. For example, basin geology has a large influence on water chemistry and thus on characteristics such as pH that determine the types of fish species that can live in a lake (Rahel and Magnuson 1983). Even within a watershed, a lake's position relative to hydrologic flow patterns can determine whether it receives most of its water via precipitation or groundwater. Groundwater-dominated lakes tend to have higher cation concentrations than do precipitation-dominated lakes and therefore have higher productivity and are more likely to contain calcium-limited species such as crayfish and snails (Riera et al. 2000). Streams also reflect the geomorphology and geologic conditions of their

watershed. In the North Fork Humboldt River drainage of Nevada, for example, trout were absent from drainages dominated by volcanic or detrital geologic deposits but were abundant in nearby drainages dominated by sedimentary deposits (Nelson et al. 1992). This pattern reflected the fact that volcanic and detrital rocks produced fine particles upon weathering that resulted in highly embedded stream substrates. Trout require clean gravel for spawning and therefore could not reproduce in such streams. Sedimentary rocks weathered to produce large particles that resulted in low embeddedness of stream substrates and thus provided a suitable habitat for trout reproduction. The distribution of brown trout in the Black Hills National Forest also was strongly influenced by watershed characteristics, especially the land type association. Brown trout were most abundant in limestone canyon streams, apparently because this land type produced the greatest amount of pool habitat favored by this species (Modde et al. 1991).

Another consideration in the move toward watershed analysis is the realization that large-scale anthropogenic perturbations often are a dominant influence on fish abundance and community composition at the site level (Allan 2004; Marchetti et al. 2004). For example, Allan et al. (1997) calculated the extent of agricultural land use upstream of sampling sites at four spatial scales varying from local (a 30-m buffer zone extending 150 m upstream of the sample site) to regional (entire area of agricultural land upstream of the sample site). The health of fish assemblages (measured by an index of biotic integrity) was most strongly correlated with the amount of agricultural land at the largest spatial scale, suggesting that regional rather than local land use was the primary factor influencing fish assemblages. Other large-scale perturbations that can have negative effects on fish populations are livestock grazing and timber harvest. Both activities can expedite runoff and negatively affect stream base flows by compacting soil layers and decreasing filtration (Isaak and Hubert 2001a). Finally, considering habitat at the watershed scale is important because many fish species require different habitats for spawning, rearing, summer feeding, and winter refuge. Often, these habitats are widely dispersed across the watershed and, therefore, migration among habitats is necessary to meet all life history needs (Schlosser and Angermeier 1995; Schrank and Rahel 2004). Dams, road crossings, and irrigation canals that prevent fish from accessing critical habitats such as spawning grounds can disrupt migrations. Such fragmentation of drainage networks also can prevent fish from recolonizing areas where populations have been extirpated by harsh conditions (Scheurer et al. 2003).

18.1.3 Watershed Features Related to Fish Abundance

A large number of watershed features have been related to the distribution and abundance of fish populations. These can be grouped into six general categories that reflect geological–hydrological processes, topographic attributes, climate factors, vegetative or land use categories, disturbance features, and habitat patch characteristics (Table 18.1). Overviews of important watershed features and how they are measured can be found in Maxwell et al. (1995), Johnson and Gage (1997), Wesche and Isaak (1999), and Allan (2004) as well as in the studies discussed in

Table 18.1 Examples of watershed characteristics related to the distribution and abundance of fish populations.

Category	Examples
Geology–hydrology	
Surficial or bedrock geology	Nelson et al. 1992; Wiley et al. 1997
Groundwater discharge patterns	Baxter et al. 1999; Riera et al. 2000
Topographic attributes	
Basin area	Gresswell et al. 1997; Porter et al. 2000
Basin relief	Lanka et al. 1987
Mean basin elevation	Gresswell et al. 1997
Drainage density	Lanka et al. 1987
Watershed area or lake volume	Prepas et al. 2001
Climate factors	
Thermal zones	Rahel et al. 1996; Rahel and Nibbelink 1999; Torgersen et al. 1999
Precipitation zones or water yield	Gresswell et al. 1997
Vegetative or land use categories	
Proportion of watershed in various vegetation categories	Isaak and Hubert 2001a; Wall et al. 2004
Proportion of watershed in various land use categories	Wang et al. 1997; Marchetti et al. 2004
Disturbances	
Road density	Moyle and Randall 1998; Baxter et al. 1999; Dunham and Rieman 1999
Livestock grazing intensity	Isaak and Hubert 2001a, b
Degree of urbanization (e.g., population density or percent impervious land cover)	Jessup 1998; Wang et al. 2003; Scheuerell and Schindler 2004
Extent of water development (e.g., number of dams, diversions, or reservoirs)	Moyle and Randall 1998; Schrank et al. 2001; Marchetti et al. 2004
Habitat patchiness and juxtaposition	
Size of suitable habitat patches	Rieman and McIntyre 1995; Eros and Grossman 2005
Connectedness to other habitat patches	Kruse et al.1997; Dunham and Rieman 1999; Riera et al. 2000; Olden et al. 2001
Juxtaposition relative to other aquatic habitats (e.g., downstream link number)	Osborne and Wiley 1992; Cumming 2004

the remainder of this chapter. The development of geographic information system (GIS) technology has resulted in numerous digital databases that greatly facilitate the measurement of watershed features (Fisher and Rahel 2004).

18.1.4 Issues of Spatial Scale in Watershed Level Analyses

The influence of watershed characteristics on aquatic biota can be evaluated at multiple spatial scales. For streams, the largest spatial scales typically involve the entire watershed or the subbasin that lies upstream of the area of interest and the smallest scales involve localized conditions of the streambank (Allan 2004). Intermediate scales can be defined based on buffer zones that extend both laterally

into the riparian zone and longitudinally upstream for various distances. Identifying an appropriate spatial scale for analysis can be problematic because there can be conflicting results regarding the relative importance of whole watershed versus riparian land cover in determining stream properties (Allan 2004). Whole watershed influences are likely to be greatest for properties such as flow variability or nutrient concentrations, which are largely determined by processes operating across the entire landscape. By contrast, riparian influences are likely to be greatest for properties such as woody debris inputs or stream temperatures, which are mainly determined by processes operating close to the stream (Barton et al. 1985). For variables in the latter category, buffer widths of 100–200 m along each bank are most commonly used, but Frimpong et al. (2005) cautioned against arbitrary designation of a buffer zone. They provided a procedure for identifying the optimal buffer in both the lateral and longitudinal dimension to maximize the explanatory power of stream biota–land cover association models.

For lakes, the largest scale typically involves geologic or land use features measured for the entire watershed, and the smallest scale is typically the immediate shoreline surrounding the lake (Jennings et al. 2003). In contrast to streams, studies of lake–terrestrial linkages have generally not examined intermediate scales that involve measuring land cover or land use features at increasing distances inland from the shoreline. However, the issue of buffer zones can be important when considering the impacts of logging on water quality in lakes (Steedman 2000).

18.2 QUANTIFICATION OF SPECIES OCCURRENCE ACROSS THE WATERSHED

18.2.1 Detection of Species Presence versus Absence at the Watershed Scale

Sometimes it is important to be able to detect the presence of a species in a watershed without having to quantify its abundance. Furthermore, one might want to have a high probability of detecting the species even though it is not possible to sample all the habitats within the watershed. Examples include surveys to determine if a species is extinct (Reed 1996) or cases in which habitat models predicting species presence versus absence are being developed (Dunham and Rieman 1999). Rieman and McIntyre (1995) provided an example involving the occurrence of bull trout in naturally fragmented habitat patches of varied size. Habitat patches were defined as contiguous stream areas believed suitable for spawning and rearing of bull trout based on having suitable summer water temperatures. Because of cost considerations, only 450 m of stream (partitioned among several sites) could be sampled in each patch. The authors were interested in determining the probability that their sampling protocol would detect bull trout in a patch. They assumed a minimum detection probability of 0.25 for each site based on the efficiency of electrofishing or snorkeling in detecting fish. The authors also assumed a Poisson sampling distribution and a minimum population of 15 fish per 1,000 m. The expected minimum number of fish detected by their sampling was calculated as $\mu = 1.69 = ([15 \text{ fish}/1{,}000 \text{ m}] \times 450 \text{ m sampled} \times 0.25)$. The probability of detecting no fish was $P(0) = e^{-\mu} = 0.18$; thus the probability of detecting

one or more fish was $P(1 \text{ or more}) = 1 - P(0) = 0.82$. The authors believe the probability of 0.82 for detecting the presence of bull trout in a patch was likely on the low side because they purposely chose conservative estimates for the minimum detection probability and minimum fish population likely to be present. Thus, the authors believed their sampling protocol was sufficient for determining the presence of bull trout in habitat patches of varied size.

Being able to determine the probability of detecting a species is especially relevant when surveying for endangered species. Failure to find the species could result in habitat conversion that would doom any remaining individuals. Reed (1996) discussed statistical approaches for increasing the confidence that a species is absent from a site. MacKenzie et al. (2002) presented a method to estimate the proportion of sites occupied by a species when the probability of detection is less than one, a common situation when sampling fishes. If one is interested in determining the proportion of sites occupied by a species but can sample only a portion of the sites, then a probability-based sampling protocol is recommended for choosing which sites to sample (Olsen et al. 1999). Ellison and Agrawal (2005) summarized a suite of papers that provide additional guidelines regarding statistical issues in detecting rare species.

18.2.2 Quantification of Fish Abundance at the Watershed Scale

Sometimes it is important to estimate the total fish population present in a watershed rather than just presence or absence of species. For example, one might wish to assess the impact of basinwide land use practices on fish populations or to estimate the population size for conservation purposes (Hankin and Reeves 1988; Kruse et al. 2001). Three approaches have been suggested to quantify fish habitat conditions or fish population abundance at the watershed scale: a comprehensive census; the representative reach extrapolation technique; and the basinwide visual estimation technique (Dolloff et al. 1997). As the name implies, a comprehensive census involves visiting and measuring all habitats and counting every fish in the watershed. Although highly accurate, this approach is impractical for all but very small watersheds. The other two techniques are based on a two-phase sampling design whereby a stream is subsampled in the first phase and habitat or fish abundance estimates are extrapolated to the entire stream in the second phase. Methods of estimating fish populations at a whole-lake scale are discussed elsewhere in this text (Chapters 7 and 8).

The representative reach extrapolation technique involves measuring habitat and fish abundance in reaches (typically 30–300 m long) of a stream and then extrapolating the estimates to the entire watershed. The key feature is that reaches selected as representative must be "typical" for the stream. However, this selection relies heavily on the professional experience and intuition of fisheries scientists and is difficult to implement because one or a few reaches may not capture the true range of habitat conditions or fish abundance present in the drainage (Dolloff et al. 1993).

Hankin and Reeves (1988) proposed an alternative two-phase sampling approach termed the basinwide visual estimation technique. They noted that the variance associated with extrapolating fish abundance or habitat estimates to

unsampled areas (phase two) is much greater than the variance from estimating abundances at sampling locations (phase one). The solution to this problem was to sample more of the stream in the first phase but in a manner that was more cost-effective than a comprehensive inventory. They proposed that rather than estimating fish or habitat conditions for an arbitrary distance, sections sampled in the first phase should be equivalent to natural habitat units (e.g., pools or riffles) and should be independent samples drawn from within strata constructed based on stream geomorphology and location in the watershed. To make the method more cost-effective than a comprehensive survey, habitat is assessed by walking the entire stream and visually estimating the area of each habitat unit along with other features of interest such as pool depth or woody debris occurrence. At systematic intervals, such as every tenth habitat unit, quantitative habitat measurements are made of the same habitat features. This allows one to develop calibration ratios to correct for observer biases and to allow estimation of sampling variances. Fish population estimates are made by snorkeling a portion of the habitat units. These observations also are calibrated against a more accurate method, such as multiple-pass depletion by electrofishing at a predetermined fraction of the units sampled by snorkeling. In the second step, estimates are determined for each stratum using the calibration ratios and then combined to provide a basinwide estimate of habitat abundance or fish population size. An example of using the basinwide visual inventory technique to estimate fish abundance is given in Box 18.1.

Dolloff et al. (1997) found that the basinwide visual inventory technique was superior to the representative reach extrapolation technique for assessing habitat conditions in three watersheds in the Appalachian Mountains, USA. This was because the basinwide approach produced a complete census of the areas of various habitat types, whereas the values generated by the representative reach technique often were over- or underestimates of these parameters. This lack of accuracy was due to representative reaches seldom containing habitat types in the same proportion as streams across the watershed. Hankin and Reeves (1988) reported that for the same cost, the basinwide visual inventory technique produced fish population estimates with a lower variance than did the representative reach extrapolation technique. The average time to estimate the fish population in a habitat unit was only 0.5 staff-hours using visual methods compared with 10 staff-hours using depletion removal electrofishing. This meant that fish population estimates could be done in many more habitat units by means of the visual approach than by electrofishing. Consequently, the variance of the estimate of fish abundance for the entire stream was based on a larger proportion of habitat units than if habitat units had to be sampled by electrofishing. Toepfer et al. (2000) extended the Hankin and Reeves two-phase approach to situations where visual assessment of habitat or fish abundance is difficult (e.g., large streams with limited public access) or where there are longitudinal gradients in fish abundance within habitat types. They also used a GIS to aid in basinwide calculations of fish abundance. A two-phase approach was used by Young and Guenther-Gloss (2004) to determine population size of greenback cutthroat trout in streams and relate population characteristics to recovery criteria for this federally threatened species.

Box 18.1 Basinwide Estimation of a Fish Population

The premise of the basinwide visual estimation technique is that there is a consistent relationship between relatively accurate estimates of fish abundance as determined by labor-intensive depletion electrofishing and less accurate estimates as determined by less-intensive visual methods such as snorkeling. If this is true, then we can calculate a calibration ratio and correct for bias associated with visual observations. This would allow a large number of habitat units to be surveyed by the less-intensive visual method and then have these observations converted into population abundance estimates. This process is illustrated by the following hypothetical example from Dolloff et al. (1993). The object is to estimate how many coho salmon are present in a stream that contains 1,000 habitat units, 500 of which are pools. The example focuses on estimating the number of coho salmon in pools. Similar analyses would be done for other habitat types and summed to give a total population estimate. Counts of coho salmon were made in 20% of the pools ($n = 100$) by snorkeling every fifth pool in an upstream direction. The average coho salmon abundance in the 100 pools was 34.43 fish. Afterward, 10% of the snorkeled pools were sampled by multiple-pass depletion electrofishing, and a population estimate was calculated for each pool (see table). A plot of electrofishing estimates versus diver counts indicates a strong correlation, but diver counts appear to overestimate the actual abundance of coho salmon in a pool, perhaps because some fish are counted twice (see figure).

Table Coho salmon counts determined by snorkeling during the first phase (x_i) and population estimates determined by electrofishing during the second phase (y_i).

	Coho salmon count	
Pool number	x_i	y_i
10	20	26
20	2	2
30	14	13
40	32	22
50	27	21
60	27	25
70	5	3
80	59	69
90	17	12
100	20	17
Total number of fish	223	210

18.3 RELATIONSHIPS BETWEEN FISH OCCURRENCE AND WATERSHED FEATURES

There are many techniques available for relating the occurrence or abundance of species to large-scale habitat features. Which technique is most appropriate depends upon the nature of the response (dependent) and predictor (independent) variables. Sometimes we have a single response variable that is categorical, that is, units can be assigned to only one of a finite number of mutually exclusive

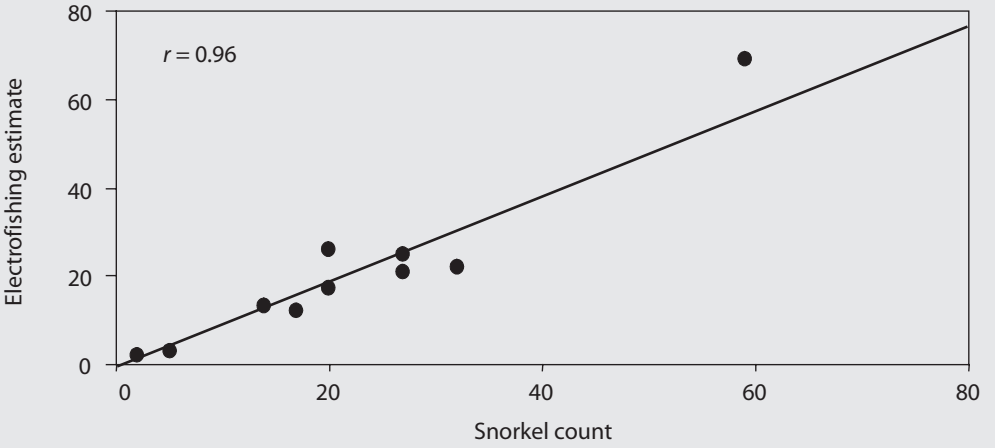


Figure Electrofishing estimates versus snorkel counts for a hypothetical population of coho salmon in pools.

Adjusting the snorkeling numbers to account for this overestimate will provide a better estimate of the number of coho salmon in pools in the drainage. The first step is to calculate the calibration ratio *(R)* by dividing the total number of fish estimated by multiple-pass removal electrofishing by the total number of fish counted during snorkeling:

$$R = \Sigma y_i / \Sigma x_i = 210/223 = 0.94.$$

The estimated mean number of fish per habitat unit is the product of the calibration ratio from the second-phase sample multiplied by the mean snorkel count from the first phase with the subscripts *d* and *r* indicating the double sampling and ratio estimation:

$$y_{d,r} = (R)(\bar{x}) = (0.94)(34.43) = 32.36.$$

The total number of habitat units (N) multiplied by the estimated mean number of fish per habitat unit provides an estimate of the total number of fish (Y) in that habitat type:

$$Y = (N)(y_{d,r}) = (500)(32.36) = 16{,}180 \text{ coho salmon in pools within the drainage.}$$

Methods for calculating confidence limits for the population estimate are given in Dolloff et al. (1993) and Hankin and Reeves (1988).

categories. Examples include species occurrence coded as present (1) or absent (0); relative abundance categories (e.g., low, moderate, or high abundance); or assemblage types (e.g., minnow versus sunfish assemblages). Techniques that can be used to compare habitat features among categories or develop a model that will predict category membership based on habitat features include logistic regression, classification and regression trees, and artificial neural networks.

In other cases, the response variable may be continuous, such as fish abundance in kilograms per hectare or number of fish species present at a site. Techniques that

can be used to relate a single continuous response variable to habitat features include multiple regression, path analysis, classification and regression trees, and artificial neural networks. Finally, there are situations in which we may be interested in relating multiple response variables to multiple habitat features. For example, we may have data on fish species abundances and habitat characteristics for a set of sites. We may wish to identify gradients of fish assemblage change across the sites and relate these to habitat gradients. Ordination methods and canonical correspondence analysis are appropriate techniques for such analyses. In the following sections, these techniques are discussed with a focus on how they can be used to identify large-scale habitat features important to fishes.

It is important to examine the underlying relationships among variables. In many cases, simple bivariate plots will provide insight into the association between the dependent variables, for example, species occurrence or abundance and the independent (predictor) variables such as habitat conditions. Such plots may also indicate whether variable transformations are needed to meet the underlying assumptions of the statistical methods being used. Many of the methods discussed in this chapter require that relationships between variables be linear, variance be homogeneous along the range of the relationship, residuals follow a normal distribution, and observations be independent of one another. Meeting such assumptions is essential when hypotheses are being tested statistically. Failure to meet the assumptions means that probability values and interpretations about whether or not the association differs from random may not be accurate. It is not necessary to meet these assumptions when the purpose is purely exploratory (i.e., no formal testing is employed), but most of these methods provide more reliable estimates when the assumptions are met. Therefore, we strongly recommend that researchers carefully evaluate their data using graphical displays and diagnostic tests to ensure that assumptions (e.g., homogeneity of variance or normality) of the statistical methods have been met.

18.3.1 Logistic Regression

Logistic regression is used to model a categorical dependent variable as a function of one or more predictor variables. A common situation involves relating the presence or absence of a species to basin scale habitat features. In this situation, the dependent variable can have two values: 0 if the species is absent and 1 if the species is present. Logistic regression is used to estimate the probability of a species being present as a function of the predictor (habitat) variables. This typically results in a sigmoidal (S-shaped) curve for which the probability of species occurrence gradually approaches 0 and 1 at the ends of the predictor variable range (Box 18.2). Although discriminant analysis also can assign membership in categories based on predictor variables, logistic regression is preferred when the predictor variables are not multivariate normal. In addition, logistic regression provides a probabilistic prediction of occurrence that often reflects a species' association with environmental gradients (Rieman and McIntyre 1995). Hosmer and Lemeshow (2000) give details on logistic regression and the diagnostic statistics used to evaluate model fit.

Box 18.2 Logistic Regression

Logistic regression predicts the probability of a species being present on a 0 to 1 scale using the model

$$\text{Probability of occurrence} = e^{\mu}/(1 + e^{\mu}),$$

where e is the base of natural logarithm. The linear model, μ, is given by

$$\mu = b_0 + b_1(x_1) + b_2(x_2) + \ldots + b_m(x_m),$$

where b_0 is the regression constant, b_m are the regression coefficients, and x_m are the independent or predictor variables.

As an example, Rieman and McIntyre (1995) used logistic regression to model the probability of bull trout occurrence as a function of habitat patch size in the Boise River basin, Idaho. Habitat patches were defined as high-elevation watersheds (above 1,600 m elevation) that had thermally suitable conditions for bull trout in summer. Such watersheds varied in area and were isolated from other suitable watersheds by intervening valleys where streams became too warm for bull trout. The probability of bull trout being present was predicted from the natural logarithm of patch area (x) by the equation

$$\text{Probability (occurrence)} = \frac{e^{-13.293 + 1.689x}}{1 + e^{-13.293 + 1.689x}}.$$

As shown in the figure below, the logistic regression model (solid line) and the empirical data (open bars) indicated that the probability of observing bull trout exceeded 0.80 for patches over 8,000 ha, but was less than 0.10 for patches less than 1,000 ha.

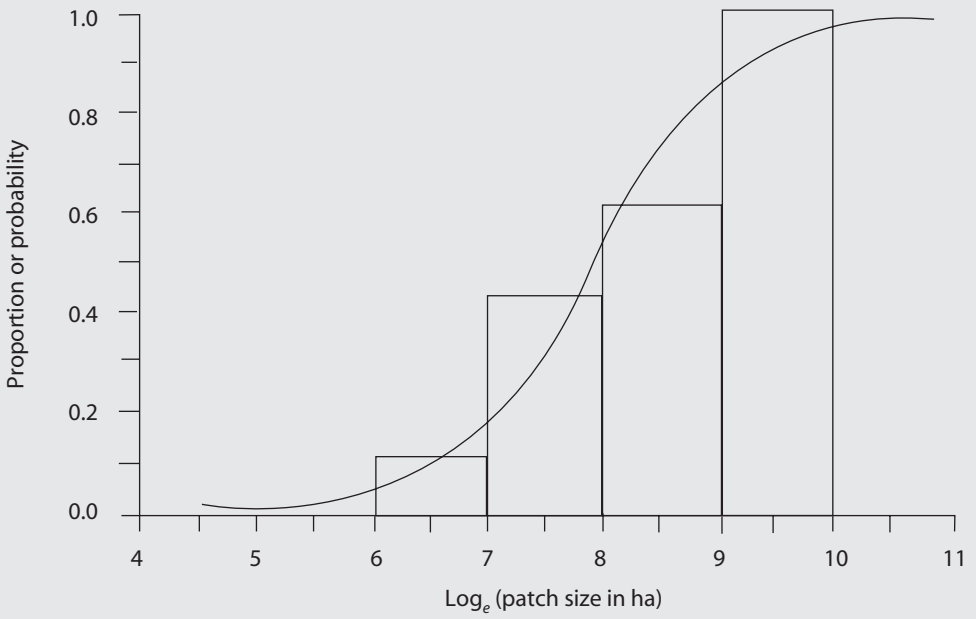


Figure Probability of observing bull trout versus patch size (log$_e$hectares).

Logistic regression has become one of the most widely used methods of relating fish occurrence to habitat features measured at a variety of spatial scales. Rieman and McIntyre (1995) found that local-scale variables such as reach width or gradient were not useful in predicting the occurrence of bull trout, but patch size (measured as the area of contiguous stream above 1,600 m elevation) was a good predictor (see Box 18.2). Schrank et al. (2001) found that both local features (e.g., pool length) and landscape features (e.g., number of impoundments per hectare in the watershed) were useful for predicting the probability that the endangered Topeka shiner had been extirpated from sites in Kansas. Other studies that incorporated large-scale habitat features in logistic regression models to predict fish presence versus absence include Kruse et al. (1997), Dunham and Rieman (1999), Porter et al. (2000), and Wall et al. (2004).

Logistic regression can be adapted to cases in which the dependent variable exists as more than two categories. As an example, Harig and Fausch (2002) used habitat features to predict the likelihood that a translocated population of cutthroat trout would fall into one of three categories: the population likely would be extirpated, the population would survive at low abundance, or the population would thrive and attain high abundance.

18.3.2 Classification and Regression Trees

Classification and regression trees (CART) are nonparametric approaches to describing variation of a single response variable in terms of one or more explanatory variables. The response variable can be categorical (classification trees) or continuous (regression trees), and the explanatory variables can be categorical or continuous. For example, we might wish to explain the presence versus absence of a particular fish species (a categorical variable) among a set of lakes in relation to habitat features such as lake size and isolation, in which case a classification tree would be appropriate. Conversely, we might wish to explain patterns in fish species richness (a continuous variable) among the lakes in relation to the same habitat features, in which case a regression tree would be appropriate. In either case, the tree is started by splitting the data into two mutually exclusive groups based on one of the explanatory variables. The goal is to make each group as homogeneous as possible for the response variable. When the response variable is categorical, the homogeneity of the groups is related to how well sites are assigned to their correct category. For example, a group for which 95% of the sites belong to the same category, such as all have the species present, is more homogeneous than a group for which only 75% of the sites have the species. When the response variable is continuous, the homogeneity of groups can be measured by the sums of squares about the group mean. The splitting procedure is applied to each of the two groups separately using another of the explanatory variables to again produce groups that are as homogeneous as possible.

The data set continues to be divided through recursive binary partitioning until there is no more than a single site in a group or until there is no variation among

the sites within a group. The tree is then pruned back to a size that captures most of the important relationships between the response variable and the explanatory variables. Trees are represented graphically, with the root node, which represents the undivided data, at the top, and the branches and leaves beneath. Each leaf represents one of the final groups. An example of using classification and regression trees to determine the influence of habitat variables on fish species richness and fish assemblage types in small lakes is shown in Box 18.3.

Classification and regression trees have advantages over parametric statistical techniques such as multiple regression, analysis of variance, and linear discriminant analysis because the method is applicable to unbalanced data structures as well as situations in which the relationships between variables are strongly nonlinear or involve high-order interactions. Unlike linear or logistic regression, regression trees automatically identify interactions and display them in an easily interpreted visual format. In addition, tree-based models are insensitive to monotonic transformations of the predictor variables because they rely solely on the rank ordering of variables.

Classification and regression trees have only recently been applied to fisheries research (Magnuson et al. 1998; Rathert et al. 1999; Kolar and Lodge 2002). Rieman et al. (1997) used classification trees to examine the status of bull trout in relation to 28 landscape features across 4,462 watersheds in the Columbia River basin, Oregon. They found bull trout were more likely to occur in colder, higher-elevation, low- to mid-order watersheds with low road densities. In an

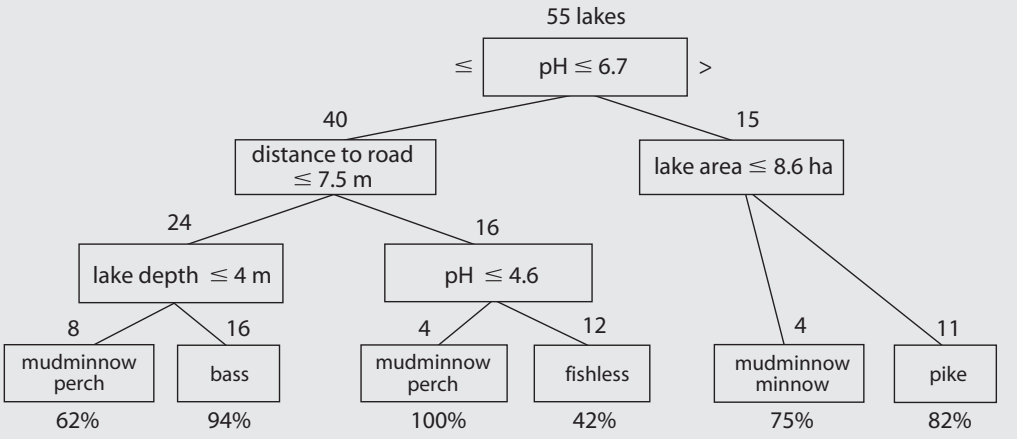
Box 18.3 Classification and Regression Trees

Magnuson et al. (1998) used classification and regression trees to examine relations between ten habitat variables and fish assemblages in small Wisconsin lakes. Figure **(A)** shows a classification tree relating the type of fish assemblage to habitat features. There were five assemblage categories based on which taxa were dominant. Figure **(B)** shows a regression tree relating fish species richness to habitat features. Each sorting node contains a variable and its sorting criterion. Lakes with values equal to or less than the sorting criterion go to the left; those with greater values than the sorting criterion go to the right. The number of lakes at each node is given above the node. In Figure **(A),** each terminal node has an assigned assemblage type; the same assemblage type may be reached by multiple routes. The percentages of lakes correctly classified in each terminal node are given below the node. Overall, 74% of the lakes were correctly classified to assemblage type. Note that lakes with a northern pike assemblage tend to have high pH and be large. Different habitat variables are important in determining different fish assemblage types. For example, note that distance to a road plays a role in determining if a lake is likely to have largemouth bass but is not important in determining if the lake will have northern pike. In Figure **(B),** the mean richness for the lakes of each terminal node is given within the node. The psuedo-R^2 value that measures the amount of variation explained by the model was 0.90. Lakes with the lowest richness (average of 1.3 species) have conductivity of 25 µS or less, are small, have a low stream gradient connecting them to the next lake downstream, and are far from roads.

(Box continues)

Box 18.3 *(continued)*

A. Fish Assemblage Type



B. Species Richness

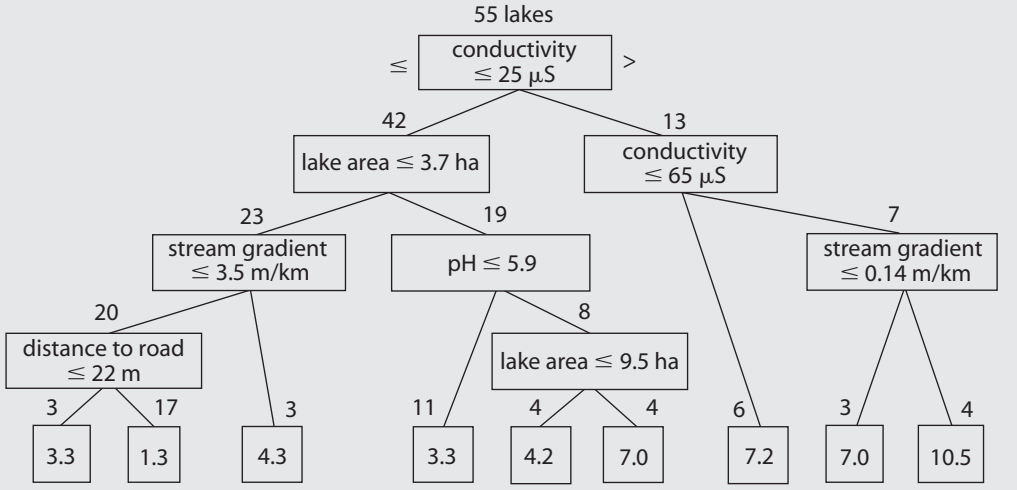


Figure **(A)** A classification tree relating the type of fish assemblage to habitat features. Each terminal node has an assigned assemblage type named after the characteristic fish species (bass for largemouth bass, minnow for several species of minnows, mudminnow for central mudminnow, perch for yellow perch, and pike for northern pike). The percentages of lakes correctly classified in each terminal node are given below the node. **(B)** A regression tree relating fish species richness to habitat features. Each sorting node contains a variable and its sorting criterion. The mean richness for the lakes of each terminal node is given within the node.

analysis based on regression trees, Greenfield et al. (2001) determined that mercury concentrations in yellow perch in Wisconsin lakes were related more closely to within-lake chemistry conditions than to watershed features such as the amount of wetlands in the drainage basin. Further information on CARTs, including pruning methods and software packages, is available in Breiman et al. (1984) and De'ath and Fabricius (2000).

18.3.3 Artificial Neural Networks

Artificial neural networks (ANNs) are a class of methods that mimic the ability of the human brain to recognize patterns. Researchers developed models that represented the linkages between brain neurons such that errors encountered in a model would provide feedback, alter the strength of the association between neurons, and lead to a refinement of the model that minimized the errors. The idea is that by using a training data set, the model could "learn" through modifying the weighting of the different variables and their interactions to enhance the predictive capabilities of the final model. There are different forms of ANNs and we present the one-hidden-layer, feed-forward model that is one of the approaches best suited for fisheries applications. (For details of the computations, we refer readers to Lek and Guégan [1999] or Olden and Jackson [2001]). Figure 18.1 shows a representation of the one-hidden-layer, feed-forward model consisting of an input layer (a set of p neurons each representing one of the predictor variables used in the analysis), an output layer (response variable), and a hidden layer (a series of intermediate connections between the predictors and response variables). There are connections between neurons from different layers (i.e., input, hidden, and output) but not within layers. These connections can vary in strength and in their influence (i.e., positive or negative) in the same way that signals or activity between brain neurons can vary in their strength. Similarly, the strength of the connections between layers is analogous to regression coefficients in that they quantify the association between the predictor variables and the response variable(s). As in biological nervous systems, the artificial neural network receives signals from other neurons or outside through synaptic connections. The neuron processes this information and sends an output signal to other neurons in the network.

To interpret ANNs, the combination of magnitude and type of signal (±) going from the input variable to the hidden layer and from the hidden layer to the output variable must be considered. The general rule is that multiplication of the input to hidden by the hidden to output neural pathway determines the overall effect that each input variable has on the response variable. Consider in Figure 18.1 (top panel) that the first input neuron (i.e., X_1) represents a habitat variable (water depth) that has a strong positive state (i.e., the magnitude and direction of the association) with the first hidden neuron but a weak negative state with the second hidden neuron. In turn, both hidden neurons have a strong positive state with the output neuron (e.g., occurrence of bull trout). Now we must consider the complete path from the input to the output neurons. Water depth has a strong positive connection weight with the first hidden neuron, which in turn has a

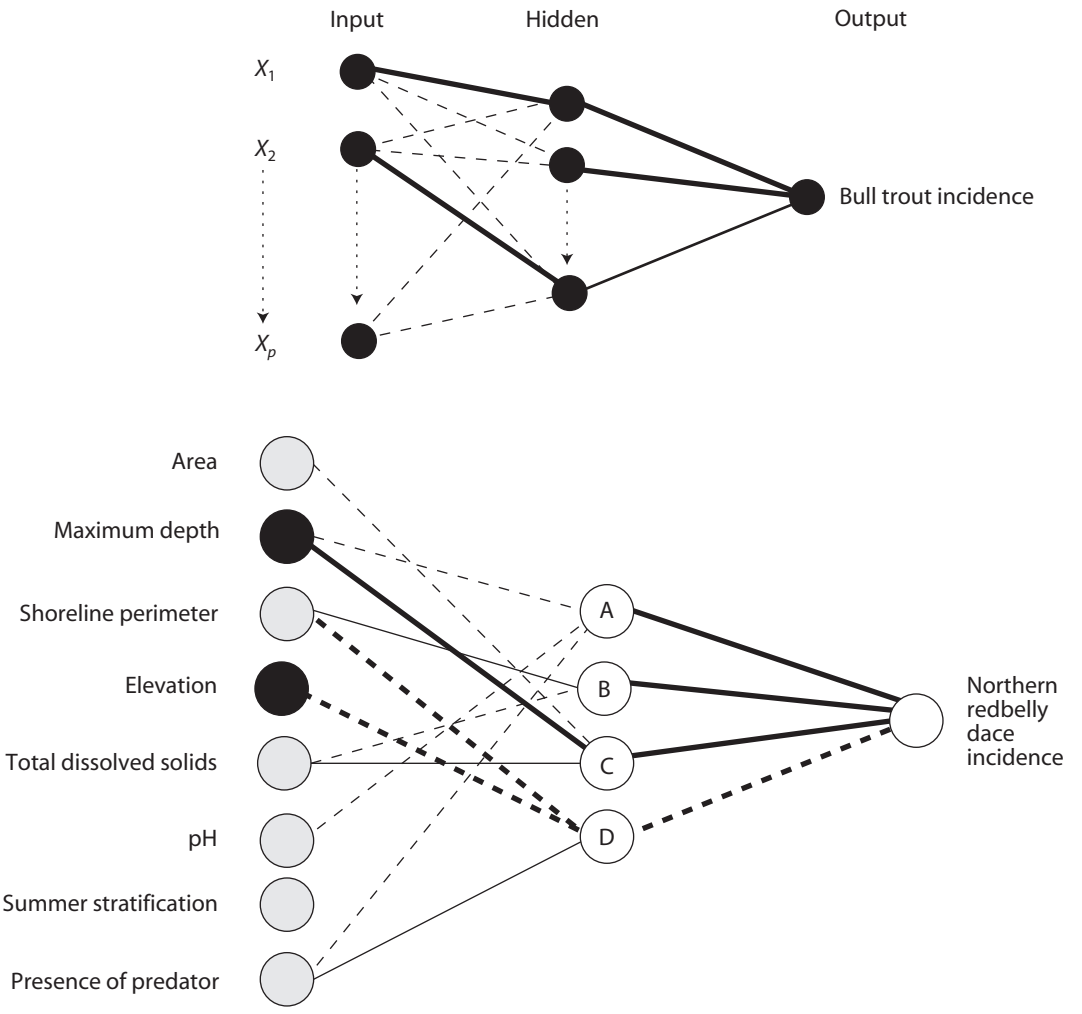


Figure 18.1 Examples of a one-hidden-layer, feed-forward neural network design. Solid lines indicate positive connections, dashed lines indicate negative connections, and line thickness indicates strength of the connections. Top panel shows how the incidence of bull trout could be related to habitat variables (X_1 to X_p). Bottom panel shows a neural network predicting the incidence of northern redbelly dace among 128 lakes in Ontario, Canada. Nonsignificant connections were identified using a randomization test and deleted to simplify the network (Olden and Jackson 2001). Gray shading of the input neurons indicates their overall contribution is negatively associated with northern redbelly dace incidence whereas black shading represents a positive association.

strong positive connection weight with bull trout occurrence. For this case, both are strong and positive leading to a large, positive outcome. This part of the network indicates that as water depth increases, the probability of bull trout occurrence increases. However, this considers only part of the way in which water depth contributes to the probability of bull trout occurrence. There is a

second path through a weak negative association with the second hidden neuron, which has a large positive signal with bull trout incidence. This combination of a small negative value with a large positive value leads to an outcome of a small negative contribution of water depth to bull trout occurrence. In assessing how a particular predictor variable relates to the response variable, we sum the outcomes through the various hidden neuron pathways (see Olden and Jackson 2001, 2002a for examples). From this simplified example, the sum would be a strong positive result, leading us to conclude that bull trout incidence is positively associated with water depth.

As the influence of predictor variables may be expressed through various neurons in the hidden layer, several predictors often will be associated with each hidden neuron. This leads to several possible outcomes. First, all predictors may have similar connection weights (either all positive or negative), thereby strengthening their individual respective effects. More commonly, we find a mix of positive and negative connection weights that represent negative interactions between these predictors. Olden and Jackson (2001) found such positive and negative interactions in the probability of occurrence of northern redbelly dace from 128 locations within a watershed in Ontario, Canada. The probability of occurrence was negatively associated with the occurrence of predatory fish species, but this effect was decreased due to the interaction with site elevation and shoreline perimeter (Figure 18.1, bottom panel). In calculating the overall sum of the pathways by which the predator variable was linked to the response variable, they found the sum was negative, and this is indicated by the gray shading of the input neuron (in contrast variables having positive associations with northern redbelly dace occurrence are shaded in black). Therefore, the overall predation effect was negative, but it was moderated by other environmental conditions, especially shoreline perimeter. A likely explanation was that as shoreline perimeter increases for a given lake area, the shoreline becomes more convoluted, increasing the potential for the presence of protected embayments and patchy nearshore habitats that provide refuge from predators (Olden and Jackson 2001).

Network diagrams often prove difficult to interpret as every neuron at one layer is connected to every neuron in adjacent layers. Complicated figures result when more than a few variables are included, and interpretation becomes challenging. Olden and Jackson (2001, 2002a) showed how to determine whether the connections contained nonrandom information by using a randomization test. This procedure allows nonsignificant variables to be identified (e.g., summer stratification) and nonsignificant connections to be removed from the analysis, thereby simplifying the network diagrams.

Artificial neural networks provide much greater flexibility in the form of the model constructed than do traditional approaches such as logistic regression analysis. These methods are robust to many forms of data varying from binary to continuous variables and to many forms of error distributions. Many fisheries data sets do not meet the assumptions for traditional regression methods, such as linear relationships between variables, homogeneity of variance, or normal distribution

of errors. In such situations, ANNs offer an attractive alternative approach. In general, well-resolved solutions offer greater predictive power than do traditional methods when such assumptions are not met and generally provide comparable or better results even when the assumptions are met.

Previously there was limited use of ANNs in ecology because of the limited availability of suitable software. However, the procedure now is available in many commercial and freeware software packages. The method has great promise given its flexibility in modeling relationships of very different forms in addition to simple presence–absence models. Brosse et al. (1999) used ANNs to assess fish abundance and spatial occupancy and found that various predator and prey assemblages tended to be separated in space. Laë et al. (1999) used ANNs to predict fish yield in a set of African lakes from habitat variables such as lake depth, catchment area, lake surface area, and conductivity. Because of their flexibility and power to determine the relative importance of predictor variables, ANNs have great potential in fisheries research (Olden and Jackson 2001, 2002a). Formal comparisons of ANN, CART, logistic regression, and linear discriminant analysis for predicting fish species composition have shown the ANN method to be superior overall, in particular when trying to model rare species (Olden and Jackson 2002b).

18.3.4 Multiple Linear Regression

Multiple linear regression is commonly used to quantify the relationship between a dependent variable such as the abundance of a species and a set of predictor variables such as watershed characteristics. As with simple linear regression, the goal is to estimate parameters for a linear model that best relates abundance to habitat features given a particular set of assumptions. One assumption is that the expected value of the residuals (observed minus predicted values) is zero (i.e., $E[e] = 0$), which requires that the relationship between the dependent variable and each predictor variable is linear. This can be examined with a bivariate scatterplot, and if curvature is evident, one can either transform the variables or allow for nonlinear components in the model. Another assumption is that residuals are normally distributed, which can be assessed using a histogram. Additional assumptions are that the residuals are distributed with equal variance (homoscedasticity) and are independent. These assumptions can be checked by plotting the residuals against each predictor variable and the fitted values from the model. If no assumptions are violated, the residuals should be randomly distributed and have a mean of zero. Formal statistical tests have been developed for each assumption, and these are included in most statistical packages. In many cases, data transformations may help in meeting these assumptions (Sokal and Rohlf 1995).

A final assumption is that the predictor variables are not correlated. Correlation among the predictor variables is called multicollinearity and can lead to unreliable results from the multiple-regression model. The problem is that two or

more variables share a considerable amount of variation with the response variable. Estimating the regression coefficients for these variables and their significance becomes problematic, and the resulting model may exclude important variables. As the coefficients may be incorrect, the true predictive capability of the model may be impaired. One clue that multicollinearity may be an issue presents itself when variable coefficients change substantially, particularly in their signs, when used in combination relative to when the variables are used independently. Other signs of multicollinearity include having a highly significant overall model *F*-test even though the *t*-tests for most of the regression coefficients (betas) are nonsignificant and having the opposite signs on regression coefficients from what was expected. Multicollinearity also can be evaluated using the variance inflation factor (VIF), which reflects the correlation between predictors in their association with the response variable. A VIF value of 1 indicates that all variation related to a predictor is unique information, whereas increasing VIF values indicate increasing degrees of multicollinearity. Values greater than 5 (or sometimes 10) are suggested as levels at which one may want to consider removing predictors from the model. If one wants to retain all variables in the analysis, an alternative approach is to use a principal component analysis to extract a reduced number of uncorrelated habitat variables and then use these component scores as predictor variables in a regression analysis (e.g., Braaten and Guy 1999; Whittier et al. 2002). MacNally (2000) provides an insightful discussion about the use of multiple-regression approaches in ecology and conservation biology.

As an example of the application of multiple regression for identifying fish–habitat associations, consider how the abundance of a species is related to five environmental variables measured at 20 sites across a watershed (species A in Box 18.4). For purposes of this example, we have a limited number of observations (sites), which will reduce the statistical power of our model and increase the error associated with estimates for each variable. An initial multiple-regression model included the variables basin, current velocity, and urban development. The variable termed basin is different from the others in that it is a binary variable. Using this dummy variable in a multiple regression allows us to compare whether the general regression model is similar or different between the two subbasins. In the final multiple-regression model, there was a difference ($P = <0.001$) in how the abundance of species A varied between the two subbasins. Sites located in the basin identified as 1 had 2.9 more individuals of species A relative to sites in the basin identified as 0, given identical conditions for depth. Therefore, inclusion of such dummy variables can allow one to detect effects of habitat features that exist as categories rather than as continuous variables.

There are several strategies available for selecting predictor variables. A backward stepwise multiple-regression model begins with all possible predictor variables included in the analysis. Based on some predefined criterion such as alpha level or *F*-value, each variable is considered and removed from the model if it does not meet this criterion. The variables that remain are significant predictors of the species' abundance. An alternative approach is a forward stepwise multiple regression, in

Box 18.4 Multiple Regression Analysis of Species Abundance and Environmental Conditions

Below is a hypothetical data set showing the abundance of seven fish species and the values for five environmental variables across 20 stream sites.

Table Hypothetical data set for seven fish species across 20 stream sites. Sampling of the watershed was focused on two subbasins (coded as 0 or 1). At each site, the average current velocity (cm/s), stream width (nearest m), and water depth (cm) were measured. The variable urban represents the percent of the landscape in urban development upstream of the site.

	Species abundance							Environmental variables				
Site	A	B	C	D	E	F	G	Basin	Velocity	Width	Depth	Urban
1	10	9	8	7	6	5	4	0	0	8	10	10
2	9	8	7	6	5	4	3	0	1	6	20	3
3	8	7	6	5	4	3	2	0	2	5	30	8
4	7	6	5	4	3	2	1	0	3	2	40	4
5	6	5	4	3	2	1	0	0	4	1	50	5
6	5	4	3	2	1	0	0	0	0	8	60	5
7	4	3	2	1	0	0	0	0	1	6	70	6
8	3	2	1	0	0	0	0	0	2	5	80	8
9	2	1	0	0	0	0	0	0	3	2	90	9
10	1	0	0	0	0	0	0	0	4	1	100	10
11	13	3	9	8	6	3	3	1	0	4	10	8
12	13	3	3	3	3	7	5	1	1	3	20	2
13	11	9	5	5	8	9	8	1	2	2	30	4
14	10	9	3	9	5	7	9	1	3	1	40	4
15	8	6	9	8	2	5	5	1	4	0	50	5
16	7	2	5	7	4	6	2	1	0	4	60	6
17	7	3	2	5	7	6	9	1	1	3	70	7
18	6	5	7	6	5	3	5	1	2	2	80	2
19	5	7	5	4	3	5	8	1	3	1	90	8
20	4	3	3	3	8	9	4	1	4	0	100	10

Multiple-regression analysis could be used to determine which environmental variables are related to the abundance of a given species. Below are the results for an initial multiple linear regression model that relates the abundance of species A to the five environmental variables.

Table Summary of the multiple linear regression model for the abundance of species A based on five variables. The variation inflation factor (VIF) is a measure of collinearity between predictor variables.

Regression Model					
Source	*df*	Sum of squares	Mean square	*F*-value	*P* > *F*
Model	5	214.19	42.84	217.16	<0.001
Error	14	2.76	0.20		
Total	19	216.95			
Root mean square error		0.444			
R^2		0.987			
Adjusted R^2		0.983			

Parameter Estimates					
Variable	Parameter estimate	SE	*t*-value	$P > t$	VIF
Intercept	11.520	1.160	9.93	<0.001	0.00
Basin	2.776	0.426	6.52	<0.001	4.60
Velocity	–0.065	0.235	–0.28	0.785	11.20
Width	–0.041	0.158	–0.26	0.798	14.60
Depth	–0.101	0.004	–23.68	<0.001	1.52
Urban	–0.021	0.044	–0.47	0.647	1.26

Because the *P*-value for the overall model is less than 0.001, we would conclude that the abundance of species A is related to these stream environmental variables. The model also explains a high amount of variation (adjusted $R^2 = 0.983$). However, the variation inflation factor (VIF) values for velocity and width are large, suggesting a high degree of collinearity between some variables. Thus, the assumption that predictor variables are uncorrelated is likely violated. We can examine a correlation matrix of the predictor variables and see that there is a strong correlation ($r = –0.82$) between stream width and current velocity. Thus, these two variables have a similar pattern of variation with the dependent variable (i.e., the abundance of species A). Because width has the highest variance inflation factor (VIF) value, we omit it from the model and redo the analysis. The new results are as follows.

Table Summary of the multiple linear regression model for the abundance of species A based on four variables.

Regression Model					
Source	*df*	Sum of squares	Mean square	*F*-value	$P > F$
Model	4	214.17	53.54	289.41	<0.001
Error	15	2.78	0.19		
Total	19	216.95			
Root mean square error		0.430			
R^2		0.987			
Adjusted R^2		0.984			

Parameter Estimates					
Variable	Parameter estimate	SE	*t*-value	$P > t$	VIF
Intercept	11.231	0.318	35.29	<0.001	0.00
Basin	2.874	0.199	14.45	<0.001	1.07
Velocity	–0.008	0.078	–0.10	0.920	1.33
Depth	–0.101	0.004	–24.68	<0.001	1.53
Urban	–0.022	0.042	–0.52	0.611	1.25

(Box continues)

Box 8.4 *(continued)*

All VIF values are now close to 1, indicating that collinearity among the remaining four environmental variables is no longer a problem. However, two of the variables (velocity and urban) have coefficients that are not significantly different from 0 ($P > 0.05$). Thus, we could eliminate these variables to produce a final model as follows.

Table Summary of the multiple linear regression model for the abundance of species A based on two variables.

Regression Model					
Source	*df*	Sum of squares	Mean square	*F*-value	$P > F$
Model	2	214.12	107.06	644.09	<0.001
Error	17	2.83	0.17		
Total	19	216.95			
Root mean square error		0.408			
R^2		0.987			
Adjusted R^2		0.985			

Parameter Estimates					
Variable	Parameter estimate	SE	*t*-value	$P > t$	VIF
Intercept	11.117	0.217	51.23	<0.001	0.00
Basin	2.900	0.182	15.91	<0.001	1.00
Depth	–0.102	0.003	–32.17	<0.001	1.00

The final model explains almost 99% of the variation in the abundance of species A across the 20 sites. The root mean square error (0.408) provides a measure of the average error associated with the estimated values. This is considered a better measure of model performance than R^2 when comparing across different data sets or models because R^2 is greatly influenced by the number of variables, atypical observations, and the range of variation present within variables.

For site 5, the value for basin is 0 and for depth is 50; our estimate for the abundance of species A is 11.117 + 2.900(0) – 0.102(50) = 6.02. The observed value is 6, indicating the model closely estimates the abundance of species A at this site. Examination of the residuals could be used to assess potential bias or errors in the model and the relative degree of departure of observed from predicted values for any given site.

which all variables are excluded initially. The variable explaining the greatest amount of variation in species abundance is included if it meets an a priori criterion, again based on alpha level or *F*-value. After a variable is entered, all variables are reconsidered as to whether they explain significant amounts of variation in addition to what is explained by the preceding variables. Variables will be entered in sequence until those remaining no longer meet the entry criterion. In choosing different values for entry and deletion, the modeler must decide to accept a

greater error in either including unimportant variables or excluding important variables (type I and II error rates). Often the implications of this decision are not recognized nor can the decision be justified. MacNally (2000) provides a detailed discussion about evaluating models.

Another common approach, called best subsets regression, involves evaluating all possible multiple-regression models and choosing the best model. There are different forms of output from different statistical packages for this approach. Some evaluate and provide the best model containing a single predictor, the best two-predictor model, and so on, until the full set of variables is included. This provides the researcher with several competing models that vary in their complexity. The researcher then must choose a model from the set based on its statistical significance, its ability to predict an independent set of observations (i.e., cross validation), or other measures based on information criteria.

Researchers often do not recognize that they may be evaluating a large number of regression models when running stepwise or best subset approaches (Olden and Jackson 2000). Although researchers have become more aware of the implications on their overall error rates when comparing many correlations or *t*-tests, and use either different analytical approaches or a posteriori corrections (e.g., Bonferroni correction), few people consider such problems with multiple-regression approaches. In fact, the number of possible regression models is $2^p - 1$, where p equals the number of independent predictor variables included. Thus, if one includes eight predictors into one of these approaches, 255 regression models would be evaluated with only the best one being presented. Therefore, one's confidence in the associated significance of the results may be overstated if there are several models that differ little in their explanatory power. When comparing regression models or attempting to pick a "best" model, it is recommended that various quality-of-fit measures be used. Examples are the adjusted R^2, which takes into account the number of variables in the model, or the Akaike information criterion (AIC), which compares the information gained by including additional predictors in the model (Anderson and Burnham 2001; Burnham and Anderson 2002). The AIC approach is generally preferred over the adjusted R^2 and essentially involves a penalty being imposed for adding each additional predictor to a model. In the AIC, there is a trade-off related to the additional information gained by adding predictors to a model relative to the penalty imposed by including them. The AIC balances these two effects to aid in selecting the optimal model that includes the fewest predictors. In many cases, the AIC approach will identify a set of competing models that have a similar level of statistical support.

Another common problem is how to evaluate the value and utility of regression models. Many conclude that models that are statistically significant, have high R^2 values, or have a strong relationship between the predicted and observed values for each observation are good models. However there is a degree of circularity here in that one uses the same data set to generate the model and then to evaluate the model (i.e., lack of independence). The model may not perform as well when evaluated with an independent set of observations; that is, the model is biased. Using an independent set of observations or splitting the data into a training data

set and a testing data set provide more reliable measures of the model's utility. In many cases, investigators do not have a sufficient number of observations to allow these approaches so other cross-validation techniques, such as jackknifing, may be appropriate (Manly 1997). In applying the jackknife, an observation is removed, the model is developed using the remaining $(n - 1)$ observations, and then the model is used to predict the value for the omitted observation. Each observation is deleted in turn and a summary made of the overall results. This provides a more reliable estimate of the model and reduces the bias (e.g., Olden and Jackson 2000).

Multiple regression is one of the most commonly used approaches for relating fish abundance to watershed features. Lanka et al. (1987) used this approach to relate trout abundance to habitat features in small Rocky Mountain streams and found that models based on drainage basin morphology predicted trout standing stock as well as models based on site level characteristics. Diamond and Serveiss (2001) used forward stepwise multiple regression to account for variation in the health of fish and mussel assemblages in the Clinch River and Powell River basins, Virginia, and found that coal mining and land use conditions were significant predictors of assemblage status. Multiple-regression analysis indicated few relations between stream habitat conditions and the abundance of American eel in streams in Maryland (Wiley et al. 2004). The lack of relationships supported earlier reports that the American eel is a habitat generalist. Interestingly, one of the few factors correlated with American eel abundance was distance to the nearest dam, suggesting that American eels were accumulating downstream of these structures because of impeded migration. Frimpong et al. (2005) used multiple regression to optimize the riparian buffer parameter in models that related stream fish community attributes to land cover features in a watershed.

18.3.5 Path Analysis

Path analysis is related to multiple-regression analysis in that both procedures attempt to explain variation in a dependent variable as a function of several predictor variables. However, unlike multiple-regression analysis, path analysis allows one to examine chains of causality whereby the influence of a predictor variable on the dependent variable is mediated by the effect of an intermediate variable (Sokal and Rohlf 1995; Legendre and Legendre 1998). For example, Isaak and Hubert (2001b) found that watershed slope influenced maximum stream water temperatures in montane landscapes through the intermediate variable of riparian tree abundance. The causal mechanism was that steeper watersheds created more mesic conditions for trees that, in turn, shaded streams from the warming effects of sunlight. In a path analysis, researchers must first develop an a priori hypothesis about causal linear relationships among a set of variables. An assumption is that both direct and indirect effects can influence the dependent variable. The relative strengths of these effects are quantified using least-squares regressions for response variables in a manner that is structured by the hypothesis (see Sokal and Rohlf 1995 for details). Alternatively, the effects can be estimated using relationships among the various correlation coefficients. An assumption is that

not all of the variance in the response variables can be explained and there will be other variables not considered that explain additional portions of the variance in the dependent variable.

As an example of path analysis (Box 18.5), consider that stream depth has a direct negative effect on the abundance of a fish species. Similarly, current velocity also has a direct negative effect on the abundance of the species. However, stream depth may also have an indirect effect through a corresponding influence on current velocity. In the example data set, stream depth and current velocity are positively correlated; hence as streams get deeper, they also have faster current velocities. Using a path analysis, we can partition the amount of variation in the abundance of the species that can be explained into a direct component due to the effect of current velocity and an indirect component of stream depth that is mediated through current velocity. As shown in Box 18.5, stream depth has an overall negative effect on the abundance of species A (–0.898), with the direct effect (-0.816) being much larger than the indirect effect (–0.082).

Path analysis allows one to test specific hypotheses about causal relationships among variables and to estimate the amount of variation explained by direct and indirect pathways. It permits one to display these relationships in a much more detailed manner than is provided by multiple regressions. However, due to the complexity of the calculations and numerous direct and indirect components, it is obvious that such models could not include more than a few variables.

Isaak and Hubert (2001a) used path analysis to explore the direct and indirect effects of large-scale habitat features such as watershed size, basin elevation, and watershed slope on reach scale components of fish habitat, such as stream width, in Rocky Mountain streams. Cumming (2004) demonstrated that low-head dams reduced fish species richness in Wisconsin streams; however, a path analysis indicated that the effect of dams on species richness was small in comparison to the influence of water quantity and summer water temperatures. Cumming noted that, from a management perspective, these results imply that reduction of water volume and increased water temperature are greater threats to fish assemblages than is the decrease in stream connectivity caused by low-head dams. Other examples of the use of path analysis in fisheries studies include Sheldon and Meffe (1995) and Eros and Grossman (2005). Path analysis can be implemented using packages such as SPSS or with specialized packages such as AMOS (Miles 2000).

18.3.6 Ordination Techniques

As noted in the discussion of multiple regression (section 18.3.4), having many correlated habitat variables complicates an analysis by potentially inflating the type I error rate and making it difficult to identify causal relationships. However, another class of techniques is ideal for summarizing information from large numbers of correlated variables. Ordination methods are designed to summarize patterns of variation among variables and order the observations such that one can determine their relative similarity based on large amounts of information. For example, we may want to determine how a suite of correlated habitat features

Box 18.5 Path Analysis

The general idea is to identify a set of causal mechanisms relating the abundance of species A (y_1) to current velocity (y_2) and stream depth (y_3). Not all the variation in the dependent variable (species A abundance) can be explained by the other variables, so there is a residual amount of variation (μ_1) associated with the response variable. Similarly, not all of the variation in current velocity is explained by stream depth, so there is an unexplained component associated with current velocity (μ_2).

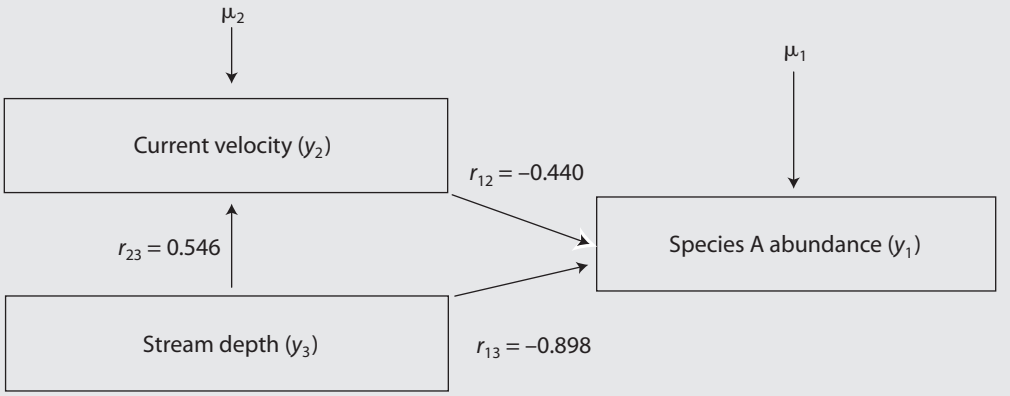


Figure This path analysis diagram is based on data from Box 18.4 and examines the abundance of species A in relation to current velocity and stream depth. Residual variation is represented by μ.

The effect of stream depth on the abundance of species A is due to both a direct effect (DE) and an indirect effect (IE) that is mediated through the current velocity. For this particular path analysis configuration, we need to calculate the correlations between all variables plus the path regression coefficients associated with the variable stream depth (i.e., p_{31} and p_{32}). The path regression coefficients can be estimated using standardized regression coefficients from a multiple regression ($\beta_{31.2}$) or directly from the associated correlation coefficients as follows.

The correlation coefficients are

$r_{12} = -0.440$, $r_{13} = -0.898$, and $r_{23} = 0.546$.

The path coefficients are

$p_{21} = r_{12} = -0.440$;
$p_{31} = \beta_{31.2} = (r_{13} - r_{23}r_{12})/(1 - r^2_{12}) = \{-0.898 - (0.546 \times -0.440)/[1 - (-0.440)^2]\} = -0.816$; and
$p_{32} = \beta_{32.1} = (r_{23} - r_{13}r_{12})/(1 - r^2_{12}) = \{0.546 - (-0.898 \times -0.440)/[1 - (-0.440)^2]\} = 0.187$.

The effect of stream depth on species A abundance is the sum of the DE and IE and equals

$r_{13} = \text{DE} + \text{IE} = p_{31} + p_{32}p_{21} = -0.816 + (0.187 \times -0.440) = -0.816 - 0.082 = -0.898$.

Therefore, we can conclude that the abundance of species A is strongly influenced by stream depth but that its influence is mainly due to a direct effect with relatively little of the effect of stream depth being explained through the mediating effect of current velocity.

may be related to the abundance of one or more species. There are many forms of ordination analysis and interested readers can find detailed discussions in Digby and Kempton (1987) or Legendre and Legendre (1998).

18.3.6.1 Principal Component Analysis

The simplest ordination method is principal component analysis (PCA), and it is similar to a multiple regression. In PCA, we calculate a linear model that summarizes the greatest amount of variation in the data. As a simple example, consider the variables current velocity, stream width, stream depth, and urban development from our previous example of factors influencing fish abundance at sites across a watershed (Box 18.6). We can use PCA to summarize the overall relationship among the sites, that is, to identify which sites are similar and which are different in regards to the habitat variables. The PCA calculates the first principal component (also referred to as a principal axis or eigenvector) by fitting a linear relationship to all of the variables such that this line summarizes the greatest amount of variance. In a multiple-regression analysis, we have a dependent variable, and it is assumed to contain greater amounts of error than do the predictor variables. However, in a PCA, we do not have a dependent variable and all variables are considered in a similar way. Therefore, we fit the line differently. In traditional regression analysis, we fit a line to minimize the sum of the squared deviation of the dependent variable for each observation (i.e., each residual). This means the deviations are measured in a vertical direction, whereas in PCA, lacking a dependent variable, we fit the line to minimize the sum of the squared deviations perpendicular to the fitted line (a form of model II regression analysis; see Sokal and Rohlf 1995). This results in a line that summarizes the general pattern or trend of the variables. Projecting each point onto this line allows us to order the sites based on this general pattern.

In a regression analysis, we obtain coefficients that relate the original variables to the regression line. Multiplying the observed habitat values for a site by their respective regression coefficients gives the site's position on the regression line (see Box 18.4 for multiple linear regression). In PCA, we obtain analogous coefficients, called eigenvector coefficients, that tell us how each of the original variables relates to the new line or first principal component (Box 18.6). The magnitude and sign of these coefficients indicate their relative importance and whether they tend to be more important at sites positioned at one end of the principal component or at the other end. Additional principal components are generated by summarizing remaining variation in the original data set that is independent of the variation explained by the first principal component. For principal component one (PC1) in our example, we see that current velocity, stream width, and stream depth have eigenvector coefficients of similar magnitude, but stream width is opposite in sign relative to the other two. This signifies that current velocity and stream depth have a negative association with stream width on the first principal component axis. Sites positioned toward the positive end have higher current velocity and greater stream depth but smaller widths (this finding is consistent with the ideas discussed in the path analysis section). Sites positioned toward the

Box 18.6 Principal Component Analysis of Habitat Data

Here an example of principal components analysis based on the data in Box 18.4 is presented.

1. Plot the variables against one another to determine if they have a linear relationship; if not, use appropriate data transformation, for example, logarithmic, to produce linear bivariate relationships.
2. Variables should be standardized if they were measured in different units. This gives each variable equal weighting in the analysis. This step may be done directly or will be implicit when the correlation matrix is calculated.
3. Based on the strength of the relationships summarized in the correlation matrix, a new axis is created that represents a linear combination of the standardized variables. The underlying relationship among the observations remains unchanged, but the observations are ordered based on the principal trend (i.e., principal component 1) in the data.
4. Unexplained variation in the data, that is deviations of points away from principal component 1, is summarized by fitting axes in succession such that they are at right angles to each of the preceding axes. This ensures that the patterns of variation summarized by different axes are uncorrelated.
5. There will be as many principal components as original variables. However, often most of the variation among sites can be summarized in the first few principal component axes, thus reducing the dimensionality of the data.

Table Summary of steps in principal component (PC) analysis based on data from Box 18.4.

Correlation Matrix

	Velocity	Width	Depth
Width	–0.825		
Depth	0.546	–0.412	
Urban	0.125	0.043	0.363

Eigenvalues of the Correlation Matrix

PC	Eigenvalue	Proportion of variance	Cumulative proportion of variance
1	2.251	0.563	0.563
2	1.136	0.284	0.847
3	0.459	0.115	0.962
4	0.153	0.038	1.00

Eigenvectors

Variable	PC1	PC2	PC3
Velocity	0.616	–0.187	–0.222
Width	–0.568	0.381	0.297
Depth	0.511	0.336	0.784
Urban	0.191	0.841	–0.497

Principal Component Scores

Observation	PC1	PC2
1	–2.439	1.726
2	–1.890	–0.854
.		
.		
.		
19	1.702	0.530
20	2.678	1.006

The score for each observation *i* on a principal component is calculated by multiplying the eigenvector coefficient for a variable by the standardized value from the original data and summing across the variables (the standardized value is obtained by subtracting the observed value from the mean and dividing by the SD). For example, for site 1 on PC1,

$$Y_i = 0.616(\text{velocity}) + -0.568(\text{width}) + 0.511(\text{depth}) + 0.191(\text{urban}) = -2.439.$$

For site 1 on PC2,

$$Y_i = -0.187(\text{velocity}) + 0.381(\text{width}) + 0.336(\text{depth}) + 0.841(\text{urban}) = 1.726.$$

These values or coordinates can be plotted to display the relationship among sample sites. The scores for the first two axes are plotted below. The proximity of the points is a measure of their similarity to one another. Thus sites 10 and 20 are similar in their overall habitat conditions but are different from sites 1 and 2.

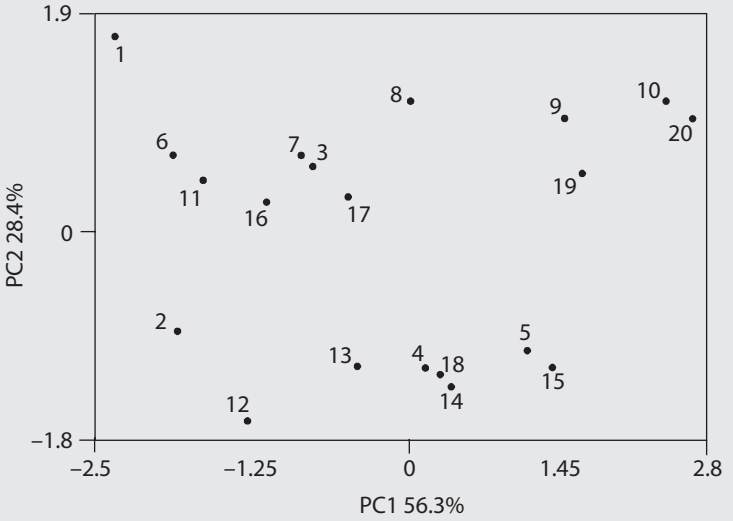


Figure Scores for first two principal components of habitat variables affecting abundance of seven fish species among 20 sites. The amount of the total variation summarized by each principal component is given on respective axes.

negative end of the first axis are wider but have lower current velocity and less depth. The second axis (PC2) is dominated by the contrast in degree of urban development. Sites positioned toward the negative end of PC2 have low values for urban development. It is important to emphasize that we have not altered the relationship among the observations but simply have expressed the relationships on new axes that are linear combinations of the original habitat variables. Sites that are positioned close together on the plot tend to share similar environmental conditions. In our example (Box 18.6), sites 10 and 20 have higher current velocities and are deeper but narrower relative to sites 6 and 11. Sites 14 and 18 have intermediate values for those variables but low values for urban development. This ordering is the basic goal of ordination methods and allows us to integrate large numbers of correlated variables to determine the relative similarity among sampling locations.

A PCA produces measurements called eigenvalues, which quantify the amount of variation summarized by each principal component and which are similar to the r^2 in a correlation analysis. In our example, the first principal component had an eigenvalue equal to 2.251 and explained 56.3% of the total variation. Variation unexplained by this first component is summarized on subsequent principal component axes (i.e., 28.4, 11.5 and 3.8%, respectively). All variation in the original data will be summarized when all principal components are considered. In most cases, we use PCA to reduce large data sets into a few dimensions that we interpret as habitat gradients. Suggestions for identifying the number of principal component axes that summarize ecologically meaningful patterns are found in Jackson (1993). Peres-Neto et al. (2003) evaluated ways to determine which variables are contributing meaningfully to individual principal component axes.

As habitat features (e.g., current velocity and urban development) are often measured in different units, we generally standardize the data before running a PCA by centering (i.e., subtracting the mean of each variable and dividing by the standard deviation). This gives all variables an equal weighting or importance in the analysis and is implicit when we use a correlation matrix in the analysis. In addition to summarizing the similarity of sites based on their environmental characteristics, PCA provides a set of new variables that can be related to species abundance data via graphical or regression approaches. As the new axes are orthogonal, there are no problems with multicollinearity, and the new axes can be used as predictor variables in a multiple regression. For example, Braaten and Guy (1999) used PCA to summarize seven habitat variables in two principal component axes representing gradients of temperature (PC1) and turbidity–discharge–depth (PC2) in tributaries of the Missouri River. A multiple-regression analysis then predicted fish abundance based on a tributary's position along these two habitat gradients. Sheldon and Meffe (1995) reduced sixteen habitat variables to four principal components prior to doing path analysis on relationships between fish assemblage attributes and habitat gradients. Rahel (1984) used PCA to summarize environmental conditions from 43 Wisconsin bog lakes and to contrast the differences in their fish communities relative to their environmental conditions. Whittier et al. (2002) used PCA to combine data on land use, road density, and human

population density in a watershed into a new variable termed the watershed disturbance index. They found that the number of invasive fish species in streams increased with the degree of watershed disturbance.

Principal components analysis works well when there are strong linear relationships between variables. Where these relationships are nonlinear or random, the ability to summarize the variation is impaired, and the resulting PCA will be less informative or potentially misleading about the relationships among variables. Environmental variables often show such linear relationships, or can be transformed to provide linear relationships, but species abundance data seldom show such features. Therefore, PCA often is not an appropriate approach for summarizing species or site patterns in fish communities. Input variables should not be derived from one another (e.g., ratio variables, percentages, or proportions) and included in a PCA. As the PCA is designed to summarize linear relationships among variables, including variables that are derived from one another will inflate the amount of variation summarized (Jackson 1997), leading one to overestimate the strength of the patterns in the data. In addition, multivariate approaches are generally preferable to trying to combine many variables into a single measure (e.g., diversity indices or indices of biotic integrity) as the multivariate measures retain more statistical and biological information and have better statistical properties.

18.3.6.2 Correspondence Analysis

When the abundances of species are related to each other in a nonlinear manner, linear approaches, such as the correlation used in PCA, will not adequately capture the relationship. In such cases, it would be appropriate to use approaches based on measures of species association that do not assume linear relationships. One such approach is correspondence analysis (CA, also called reciprocal averaging), which is based on a chi-square distance measure rather than a correlation measure of association among species. This ordination method generally works better for data sets having numerous zero abundance values or having nonlinear relationships in species abundance (Legendre and Legendre 1998). In CA, we do not need to standardize the data, as most statistical packages implement the calculations on the original data values.

We can plot the CA results from both the analyses of species and the sampling sites together on a graph called a biplot (Figure 18.2). In our example, the first axis distinguishes sites dominated by species A and B from sites dominated by species E, F and G. By positioning vectors from the origin to each species, we can assess the relationships between species in the ordination space. Angles close to 0° (e.g., between species F and G) indicate strong positive correlations between the species occurrences whereas angles approaching 180° (e.g., species B and E) indicate strong negative correlations. The length of the vectors is proportional to the importance of each variable on the ordination plot. The position of species relative to the sampling sites on the plot is a measure of their association. Species will be positioned close to sites where they tend to be abundant and positioned away from sites where they are absent or in low abundance. We can correlate site

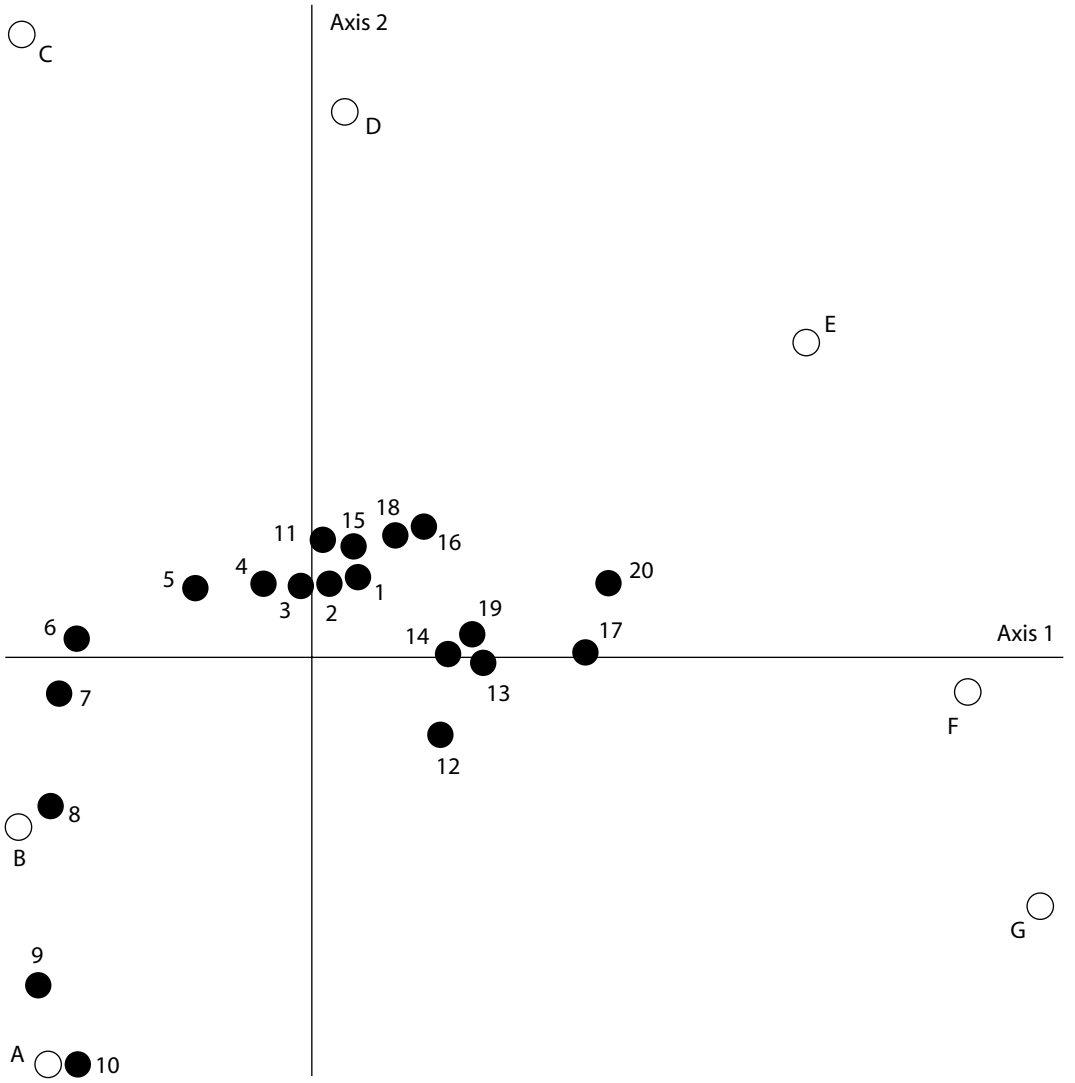


Figure 18.2 Biplot graph of the correspondence analysis results for the species abundance data in Box 18.4. Sampling locations are numbered with sites 1–10 being from one subbasin and sites 11–20 from the second subbasin. Letters indicate the position of each species on the biplot.

scores on axis 1 with the environmental variables to determine if the similarity among the sites is associated with the environmental conditions. This is called indirect gradient analysis because we are assessing the relationship between assemblage composition and environmental conditions across sites through an indirect approach. We initially analyze our community data to determine the patterns in species occurrences among sites and then subsequently determine if there is an association with the environmental conditions. This contrasts with a direct

gradient analysis in which we determine environmental gradients among the sites and then ask how the species composition relates to these gradients (see the discussion of canonical correspondence analysis in section 18.3.7). In our example, the site scores on axis 1 have correlations of r equal to –0.03, –0.34, –0.20, and –0.18 with current velocity, stream width, stream depth, and urban development, respectively. Thus, wide and deep sites are located on the left end of the first axis (e.g., sites 6 and 7). Site scores on axis 2 have correlations of r equal to –0.30, 0.13, –0.48, and –0.37 for velocity, stream width, stream depth, and urban development, respectively. Thus, sites with deep water and extensive urban development in their watershed are located on the bottom portion of the second axis (e.g., sites 9 and 10).

Jackson and Harvey (1989) used a CA to examine similarities among fish assemblages in 286 lakes in Ontario and to convert fish presence–absence relationships into continuous variables suitable for use with other statistical analyses. Jackson and Harvey (1993) used CA to show there was a high degree of concordance between the similarity of Ontario lakes based on their fish species composition and similarity based on their benthic invertebrate communities. Marsh-Matthews and Matthews (2000) used a variant of CA to determine that large-scale geographic factors were more important than local habitat conditions in explaining differences in fish assemblages among stream sites.

18.3.7 Canonical Correspondence Analysis

An alternative approach to indirect gradient analysis is canonical correspondence analysis (CCA; Legendre and Legendre 1998), a form of direct gradient analysis that involves linking species composition to environmental conditions across sample sites. A CCA combines features of multiple regression and ordination analysis in that it develops a predictive model linking the pattern of species abundance with the environmental conditions. As in correspondence analysis, the axes represent gradients of community change. The goal is to constrain the ordination of the species abundances by the environmental conditions. This means that the pattern in the species and the sites will be linked directly to the habitat features included in the analysis and will provide the best summary of relationships between the fish community and environmental conditions given the underlying model. In doing a CCA, we can make individual plots showing the patterns involving fish species, sampling sites, or habitat features; biplots related to pairs of these variables; or triplots that relate habitat features to the patterns in the fish species and the sampling sites (Figure 18.3). In a biplot or triplot, the strength of the correlation of an environmental variable is reflected in the length of the line, and its association with a particular axis is reflected in the acuteness of the angle to the other variables or an axis. Vectors with a small angle between them are positively correlated, angles approaching 180° indicate a strong negative correlation, and angles approximating 90° indicate that the variables are uncorrelated. In Figure 18.3, we see that stream width is negatively correlated with current velocity, stream depth, and the amount of urban development. Species A is most abundant at sites

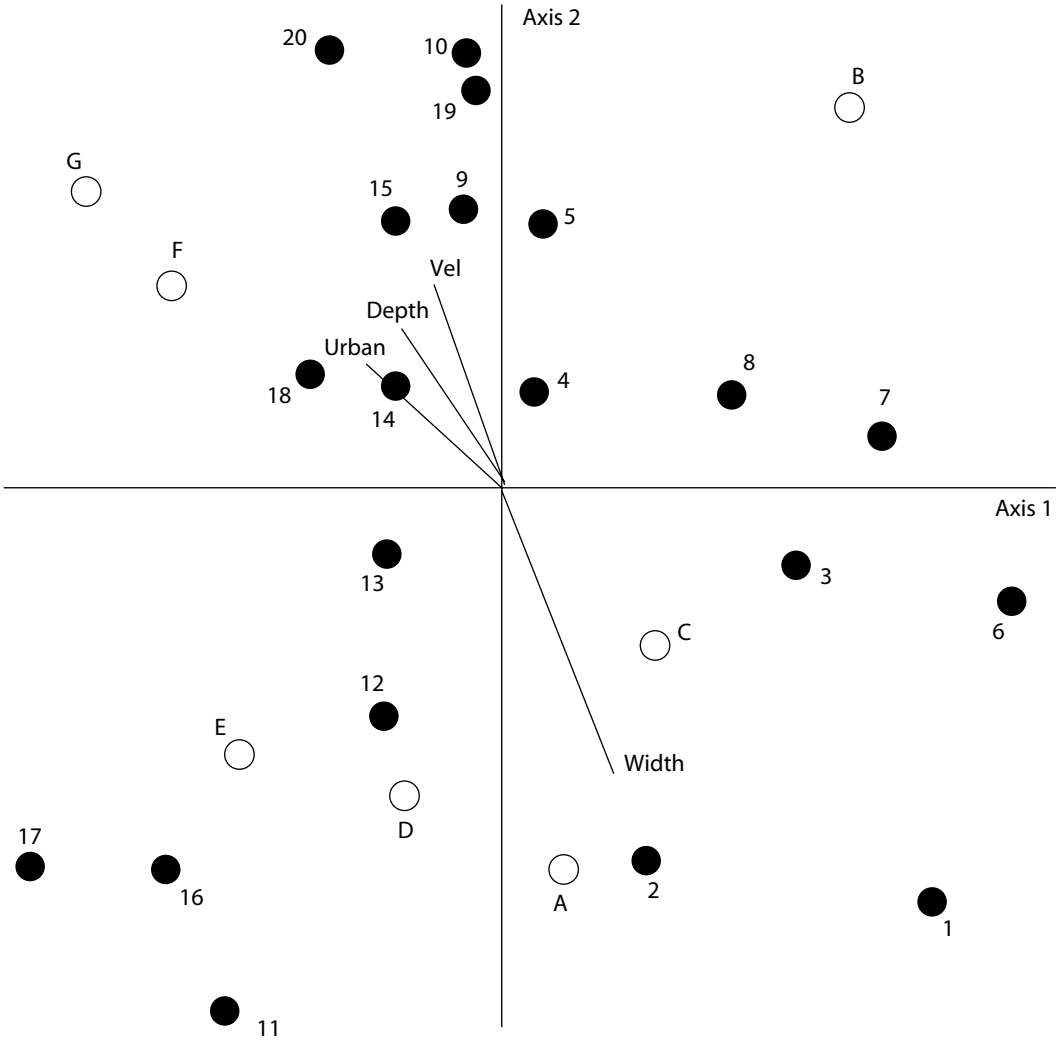


Figure 18.3 Triplot graph of the canonical correspondence analysis results based on species abundance and habitat data in Box 18.4. Sampling locations are numbered with sites 1–10 being from one subbasin and sites 11–20 from the second subbasin. Letters indicate the position of each species (A–G) on the triplot. The four environmental variables (stream width, stream depth, velocity, and amount of urban development) are identified with their respective vectors indicating the associated correlation between variables and association with the species and sites. Vectors with a small angle between them are positively correlated, angles approaching 180° indicate a strong negative correlation, and angles approximating 90° indicate that the variables are uncorrelated.

that are wide and have slow current velocity, shallow depth, and low values for urban development. Species G is most abundant at narrow sites that are deep, have high current velocity, and high values for urban development. The abundance of species E, on the other hand, tends to be uncorrelated with the environmental conditions measured in this example.

Comparing the results from CA and CCA can be revealing. If species abundances are strongly related to the habitat features, then both techniques will show similar patterns in species associations and explain similar amounts of variation in fish assemblage structure across sites. If substantially less variation is summarized with the CCA compared with the CA, this indicates that the environmental variables do not explain much of the variation in fish assemblage composition across sites.

Canonical correspondence analysis is the most commonly used method of community analysis currently. Williams et al. (2002) used CCA to demonstrate that timber-harvesting practices and drainage basin differences in water chemistry and riparian vegetation were related to trophic and taxonomic characteristics of stream fish assemblages in the Ouachita Mountains, Arkansas. Quist et al. (2004a) found that stream geomorphology and thermal conditions were important factors structuring stream fish assemblages in Rocky Mountain streams. Marchetti et al. (2004) used CCA as a means of determining whether native versus nonnative fish assemblages differed in their relationship to environmental conditions.

18.4 STATISTICAL ANALYSES OF SPATIAL RELATIONS

18.4.1 Spatial Autocorrelation and the Importance of Spatial Location in Watershed Analyses

In fisheries studies, we often wish to relate the characteristics of fish assemblages to environmental conditions across a range of sites. The standard approach to analyzing such data is to regress or correlate the dependent variables against the corresponding environmental variables from the sites. We calculate the slope or correlation coefficient and determine whether it meets some level of statistical significance based on the associated degrees of freedom. For many studies, this type of analysis may be appropriate. However, in other cases, the patterns we observe and the strength of the association between variables may be strongly influenced by the location of the sampling points relative to one another (Williams et al. 2002; Grenouillet et al. 2004). Sample sites that are close together may be similar because of environmental gradients not considered in the analysis. Such sites do not function as true replicates, and their similarity due to spatial proximity can confound our ability to detect the influence of other environmental factors on species assemblage characteristics. Situations for which the spatial location of sites is of concern lead us to the topic of spatial autocorrelation.

We begin with a simple example to illustrate the issues and possible solutions for dealing with spatial autocorrelation. Consider a situation in which fish species richness and water alkalinity, a chemical parameter often associated with biological productivity, were measured at 10 locations. A plot of the locations in an x and y geographical coordinate framework along with the alkalinity and species richness at each location are shown in Box 18.7 (panel A). We show the same points positioned within a lake (panel B) or along a stream drainage (panel C) to illustrate how such samples might be located in two familiar situations. We need to be concerned about the spatial relationships of these samples because spatial correlations

Box 18.7 Example of Spatial Autocorrelation between Fish Species Richness and Alkalinity

Table Ten site locations in an *x* and *y* geographical coordinate framework along with the alkalinity (measured as mg/L $CaCO_3$) and fish species richness at each location.

	Geographic coordinates		Measured variables	
Sampling site	*x*	*y*	Alkalinity	Species richness
1	1	9	1.0	2
2	2	7	1.5	3
3	1	8	3.0	3
4	1.5	6	2.0	3
5	4	3	8.0	7
6	5	2	9.0	11
7	4.5	2.5	9.0	9
8	8	7	3.0	4
9	8.5	8	3.5	6
10	9.5	9	2.0	5

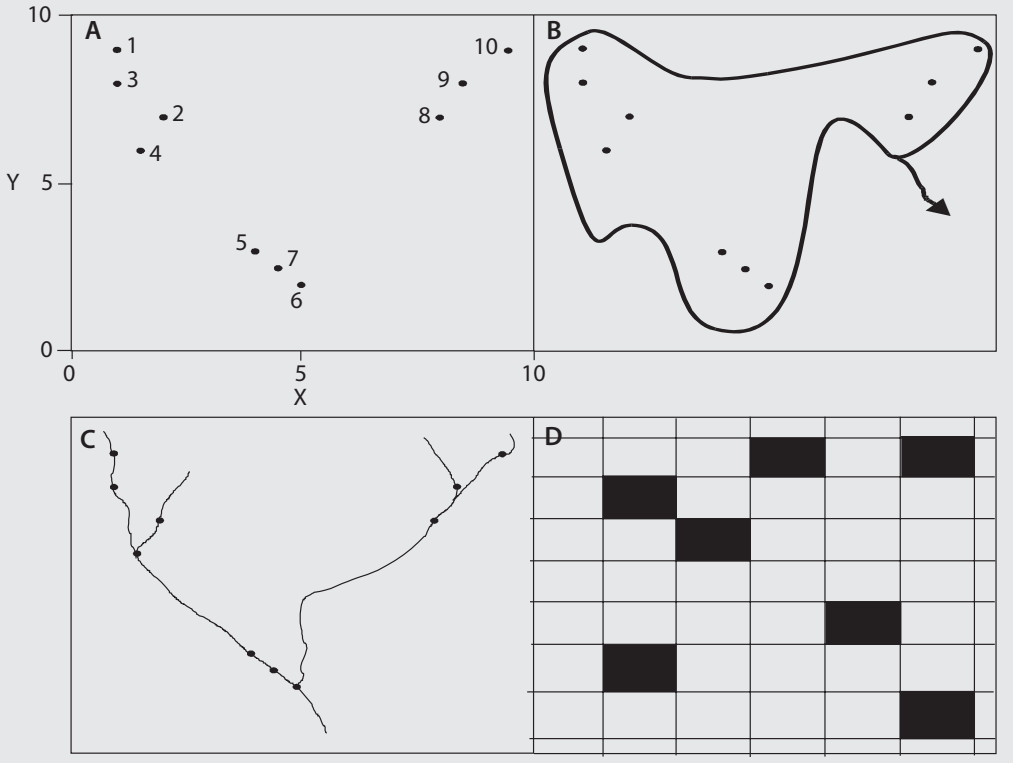


Figure **(A)** Plot of the ten site locations in an *x* and *y* geographical coordinate framework; the same points positioned within a lake **(B)** or along a stream drainage **(C)**; and the checkerboard-type pattern indicative of a negative autocorrelation at small distances for species presence (darkened rectangle) versus absence (open rectangle) data **(D)**.

can violate some of the fundamental assumptions in standard statistical tests. A basic assumption is that our observations are independent of one another. The stream example provides a situation where an effect occurring at one site is likely to have a direct impact on a site located downstream. Therefore, the observations are not independent. Similarly, sites within one of the embayments in a lake could be influenced by factors that are not important elsewhere in the lake (e.g., sediment characteristics that are determined by wave actions caused by fetch effects). In these two examples, sites in close proximity would exhibit positive spatial correlation because their attributes would tend to be more similar than for pairs of sites distributed at random.

Spatial correlation is not always positive. If we partitioned a stream into rectangular quadrants and mapped the fine-scale spatial distribution of trout within a stream, we might find that behavioral interactions tend to keep trout spaced apart, likely due to their competition for resources. Trout tend to remain in specific areas, such as downstream of rocks, and defend the surrounding waters from other trout. Therefore, at a fine spatial scale we may find trout occupying a site and excluding fish from the surrounding sites, thereby causing those sites to be empty. This would result in a pattern of negative association in occurrence for adjacent locations or, alternatively, a negative spatial autocorrelation related to small distances. A negative autocorrelation at a small distance for presence–absence data tends to result in a checkerboard-type pattern (see panel D of Box 18.7 or the results in Figure 18.4 for the negative correlation at the intermediate distance around 12 km).

When the attributes of sites exhibit positive or negative spatial correlations, then the sites are not independent observations for the purpose of statistical tests. As the degrees of freedom in our statistical tests are based on the independence of the observations, a lack of independence leads us to overestimate the degrees of freedom (or underestimate in the case of negative spatial correlation). The redundancy in the data due to spatially proximate sites results in the effective sample size being smaller than the actual number of observations. Consequently, we will tend to reject the null hypotheses in our tests more frequently than we really should; that is, our type I error rates are incorrect. This means that relationships between habitat variables and fish assemblage characteristics deemed statistically significant may not be significant at the stated alpha levels. Given that standard approaches may not be appropriate, we need to consider the following alternative methods of analysis.

18.4.2 Use of Distance Matrices to Assess the Geographic Proximity or Ecological Similarity among Sites

In some cases, we may want to make pairwise comparisons between sites for a set of variables. For example, how far apart geographically are the sites, how different are sites in terms of their species richness, or how different are the various populations in their genetic relatedness? This leads us to creating distance matrices (or alternatively similarity matrices, such as a correlation matrix) of our sites

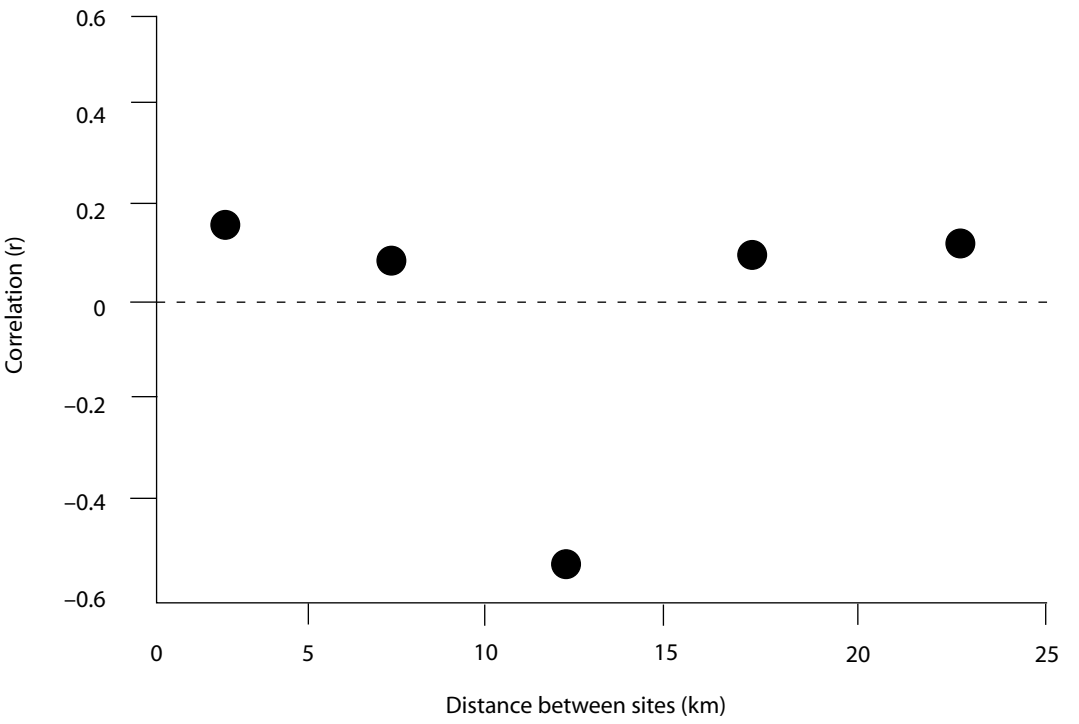


Figure 18.4 Correlogram of species richness. There are weak positive correlations between the species richness values at sites located close together or far apart but a strong negative correlation in species richness values for sites separated by intermediate distances.

that serve as the basis for many multivariate techniques (e.g., principal coordinate analysis or cluster analysis). Many types of distance or similarity measures have been developed that are appropriate for working with different types of data (e.g., presence–absence, ordinal, or continuous) or that emphasize different attributes of the data. Measures commonly used with presence–absence data include the Jaccard, simple matching, and phi coefficients (see Chapter 15). For continuous data, Euclidean distance (Chapter 15) may be suitable for small geographic distances where distances can be calculate directly from Universal Transverse Mercator (UTM; see Chapter 14) coordinates or from latitude and longitude. However, corrections must be employed for larger distances due to errors in measuring straight-line distances over a curved space. Other measures such as Bray–Curtis (Chapter 15) or chord distance may be more appropriate for species data for which some form of standardized relative abundance comparison may be desired. Legendre and Legendre (1998) describe the numerous measures available.

Having chosen a particular distance measure, we begin with n sampling points, potentially lacking independence among their attributes, and create a distance

matrix having $n(n-1)/2$ elements or values (see Box 18.8). These values are not independent as a large distance between sites 1 and 10 and a small distance between sites 1 and 2 means that sites 2 and 10 cannot be located close together. Although we might be interested in assessing how these distance matrices relate to one another, clearly we cannot use the total number of elements to estimate our degrees of freedom. We are left with two main solutions, namely we summarize and interpret the results graphically without using inferential statistics, (i.e., no hypotheses are tested formally) or we must use statistical approaches that are not compromised by the lack of independence among points or spatial autocorrelation.

18.4.3 Comparison of Distance Matrices by Means of Graphical Approaches

Simple graphical approaches can be used to indicate if spatial autocorrelation exists within a dataset and at what spatial scale(s) autocorrelation is important. One measure of whether a variable has a spatial pattern is obtained by correlating the distance matrix of between-site differences in the variable of interest with a distance matrix of inter-site geographical distances. If several variables are measured this way, a ranking of the resulting correlations provides a relative measure of which variables show the greatest spatial autocorrelation. However, as pointed out earlier, we cannot test whether these correlations differ from random by means of standard approaches (but we can test using the Mantel test discussed in section 18.4.4).

Another approach is to divide the geographic distance matrix into distance classes, for example, the sites that fall between 0 and 5 km apart, between 5 and 10 km, and so on. Then we can ask how the variable of interest (e.g., species abundance) is related between sites within each distance class. If sites located close together tend to have similar abundance values, we will find a strong positive correlation in abundance at small distances. If sites far apart have very different abundance values, then a strong negative autocorrelation would exist for large distances. Generally, this type of analysis is presented graphically as the average correlation in the variable of interest versus the distance classes. This type of graph, called a correlogram, shows how the pattern of autocorrelation changes across a range of geographic distances (Figure 18.4). The pattern we observe in the correlogram may help support or reject specific hypotheses about our ecological system. In our example (Figure 18.4), we see weak positive correlations in species richness at small and large distance classes but a much stronger negative correlation at intermediate distances. This form of a relationship might arise when fish assemblages are sampled in two tributaries and their downstream confluence in a drainage basin. The distance between sites is measured as the separation along the drainage network. Samples taken close to one another along the drainage network tend to have similar habitats and fish communities. Therefore, they have a positive association. The greatest distance will occur in comparisons of a headwater site from one branch with a headwater site in the other branch. These

Box 18.8 Estimation of Distances between Samples and Calculation of Distance Matrices

A standard way of estimating the distance between the observations *i* and *j* is the Euclidean distance (D_{ij}), or minimum straight-line distance between them.

$$D_{ij} = \sqrt{\sum_{k=1}^{P} (X_{ik} - X_{jk})^2},$$

where *X* is the observed values for the *p* variables measuring the geographic location for each site. In the case of sites 1 and 2 from Box 18.7, the Euclidean distance is

$$D_{ij} = \sqrt{\sum_{k=1}^{2} (X_{ik} - X_{jk})^2} = \sqrt{(1-2)^2 + (9-7)^2} = \sqrt{5} = 2.24.$$

If we calculate all pairwise distance values, we can fill out the following table.

Table Pairwise Euclidean distance values for geographic coordinate data in Box 18.7. Note that the distance between a location and itself is 0. The table is symmetric; therefore, we only need to present one half of the matrix in most cases.

	Site								
Site	2	3	4	5	6	7	8	9	10
1	2.24	1	3.04	6.71	8.06	7.38	7.28	7.56	8.5
2	0	1.41	1.12	4.47	5.83	5.15	6	6.58	7.76
3		0	2.06	5.83	7.21	6.52	7.07	7.5	11.85
4			0	3.91	5.32	4.61	6.58	7.28	8.54
5				0	1.41	7.07	5.66	6.72	8.14
6					0	7.07	5.83	6.94	8.32
7						0	5.70	6.80	8.20
8							0	1.12	2.5
9								0	1.41

sites likely have similar habitat and fish communities, thereby leading to a positive correlation. However, the intermediate distances typically compare sites from the headwaters with sites farther downstream near the area of confluence. Sites from these two areas will tend to have the most different habitats and likely the most divergent fish communities. As a result, comparisons of sites at this intermediate spatial scale tend to have negative associations. Typically the correlation values obtained are not assessed statistically due to their lack of independence, but suitable randomization tests can be used that correctly evaluate the null hypothesis for each of these correlations (see Legendre and Legendre 1998 for details).

We can calculate a similar table of values comparing the difference in species richness between sites.

Table Pairwise Euclidean distance values for species richness data in Box 18.7.

	Site								
Site	2	3	4	5	6	7	8	9	10
1	1	1	1	5	9	7	2	4	1
2	0	0	0	4	8	6	1	3	0
3		0	0	4	8	6	1	3	0
4			0	4	8	6	1	3	0
5				0	4	2	3	1	4
6					0	2	7	3	6
7						0	5	3	6
8							0	2	1
9								0	1

Are the patterns in species richness related to the geographic distance among sites? For example, if two sites are close to one another (corresponding to a small value in the spatial distance matrix), do they tend to have a similar species richness (corresponding to a small distance, or difference, in the species richness matrix)? An overall measure of whether the two matrices show similar patterns is provided by the sum of the cross products of the two matrices (Mantel 1967):

$$Z = \sum_{\substack{j=1,\\ k=j+1}}^{n} X_{j,k}\, Y_{j,k}\,.$$

In comparing the distance matrices for species richness versus the geographic location for each site, we obtain a Mantel Z value of 916.4 and a standardized correlation coefficient, r, of 0.32, both of which have an associated probability of 0.026. This indicates that there is a strong spatial pattern in the species richness values with sites close together tending to have a similar number of species.

Somers and Jackson (1993) used a correlogram to study the spatial relationships of mercury concentrations in 30 populations of lake trout from Ontario. They used randomization methods to assess the significance of the individual distance classes and found significant positive spatial autocorrelation in mercury concentration at large distances, a trend toward positive autocorrelation at small distances, and nonsignificant negative relationships at intermediate distances. Their study involved ten lakes from each of three geographical regions. There was considerable similarity in the patterns of mercury concentration among lakes within each region, thereby producing a strong correlation in mercury concentrations

for small distances. The two most distant groups of lakes also had similar concentrations of mercury in fish, leading to strong positive correlations at large distances. However, the geographically intermediate set of lakes had quite different patterns in the mercury concentrations relative to either of the other sets. This led to the negative correlations at intermediate distances. Nash et al. (1999) used correlograms, also with tests of significance, to compare biomass of fish in rivers versus lakes over various distances. The incorporation of spatial autocorrelation into their analyses led to some major differences between their conclusions and those from an earlier study by Randall et al. (1995). In particular, Nash et al. found no statistically significant difference in fish biomass between lakes and rivers, which was in contrast to the results reported by the earlier study that did not account for spatial autocorrelation among sample sites.

In many cases, the variable of interest may not have a continuous distribution or even be quantitative. For example, consider the spatial pattern shown in panel D of Box 18.7, where we have the presence (P, darkened rectangle) or an absence (A, open rectangle) of an attribute in a spatial grid. We want to determine if there is a spatial component to values in the cells (e.g., the presence or absence of a trout species or the type of land use within a watershed) and if the spatial pattern varies across spatial scales. Two adjacent cells will be one of the following cases: both present (PP); both absent (AA); or a combination (PA or AP). We need to quantify the relative frequency of these three cases and their likelihood given the number of individual cases of P and A. In its simplest form, we estimate the probability of these results occurring if the underlying pattern were random. This analysis is termed the joint count approach. We can calculate the number of each type of outcomes or connections between adjacent cells. These can be compared against expected values to determine whether the number of PP, AA, or PA combinations differs from a purely random scenario (see Griffith 1987 for details).

Often we are interested in whether these patterns change depending on the particular distance used in the comparison. For example, if we consider a checkerboard, the pattern is negatively associated at a distance of one unit, positively associated at distances of two units, negatively at three units, and so on. These results can be plotted to show the proportion of joint counts in particular categories (e.g., PP) over various distance intervals in a manner similar to the correlogram. There are two main metrics, Moran's I and Geary's C, used with these types of patterns to quantify the association between cells over various distances. Details are provided by Griffith (1987) or Legendre and Legendre (1998). If we are working with continuous response variables, Moran's I is similar to the correlation coefficient and bounded between –1 and +1. Increasing values of Moran's I indicate positive autocorrelation between sites whereas Geary's C would tend toward 0 under such conditions. The lower bound for Geary's C is 0 and indicates a strong positive association. As negative spatial autocorrelation increases, Geary's C increases but has no upper bound. It is normally most informative to plot Moran's I or Geary's C in a correlogram manner as this allows interpretation of how the patterns of autocorrelation vary across spatial scales. The designation of distances classes can be made so that equal distance intervals are considered or so that

equal numbers of observations occur within each interval. The choice of distance classes can influence the outcome in a correlogram, so it is important to examine the results for varied distance classes.

18.4.4 Comparison of Distance Matrices by Means of the Mantel Test

We have stressed the difficulty in knowing the appropriate degrees of freedom to use in tests for which there is a lack of independence among the observations (Fortin and Gurevitch 2001; Ver Hoef and Cressie 2001). Fortunately, alternatives have been developed that permit us to evaluate the strength of the association between matrices and whether it differs from random. Dutilleul (1993) developed an analytical approach that can be used with correlations between variables (e.g., species richness and alkalinity) that corrects for the spatial autocorrelation in the test of the null hypothesis. Mantel (1967) developed an approach, now called the Mantel test, to test for spatial or temporal autocorrelation. In our example from Box 18.8, we can test for spatial autocorrelation by determining if sites close to each other in the spatial distance matrix also have a corresponding small distance, or difference, in the distance matrix based on species richness. An overall measure of whether the two matrices show similar patterns is provided by the sum of the cross products of the two matrices (Mantel 1967):

$$Z = \sum_{\substack{j=1,\\ k=j+1}}^{n} X_{j,k} Y_{j,k} \,. \tag{18.1}$$

This measure is an unstandardized correlation between the two distance matrices **X** and **Y.** The larger Z is, the greater the match between a given pair of matrices. However, Z increases as the size of the matrices increase and varies depending on the distance measures used. As well, we know that the distances between points are not independent. Mantel (1967) proposed an analytical solution to test the null hypothesis based on some additional assumptions that are not always met. Therefore, this test is now evaluated using randomization methods that allow us to calculate what values of Z are likely when the null hypothesis of no spatial autocorrelation is true. This randomization approach commonly involves thousands of iterations (e.g., Jackson and Somers 1989), and then the observed value of Z is compared to the distribution of randomized Z values. The proportion of randomized values equal to or greater than the observed value represents the resulting probability value of the test. The Mantel test allows us to go beyond the simple correlation summary of various matrices or the simple graphical approach of the correlogram. We can use the test to infer whether there is a nonrandom linear relationship between distance matrices.

As an example of using the Mantel test, Jackson and Harvey (1989) tested whether spatial or environmental factors were more closely associated with attributes of lake fish assemblages in Ontario. Rodriguez and Lewis (1997) found no significant spatial relationship between the similarity of fish species assemblages and spatial location in the Orinoco River basin. Cattaneo et al. (2003) used

a series of Mantel tests to relate synchrony in population fluctuations of brown trout to stream connectivity. More detailed considerations of the Mantel test can be found in Legendre and Legendre (1998), Fortin and Gurevitch (2001), and Peres-Neto and Jackson (2001). The Mantel test is included in commercial packages such as NT-SYS (Rohlf 1993) or free packages available on the World Wide Web (e.g., http://www.bio.umontreal.ca/legendre/indexEnglish.html).

The Mantel test allows us to determine whether two distance matrices match more closely than expected due to chance. However, in many cases we may be interested in testing the association of two variables (e.g., species richness and alkalinity) rather than their association with their spatial configuration. This can be done directly using the correlation procedure developed by Dutilleul (1993) or through a modification of the Mantel test. In this latter approach, we use a partial Mantel test to control or remove the variation in species richness and alkalinity that is due to spatial position. We can regress the distance matrix based on species richness against the distance matrix based on sampling locations. Then the residuals from this regression are calculated. Next, the distance matrix based on alkalinity is regressed against the same distance matrix based on sampling locations and the residuals are determined. The two distance matrices of residuals then are analyzed through the standard Mantel test. If the Mantel test indicates a significant probability of correlation (e.g., $P < 0.05$), we conclude the two variables (e.g., species richness and alkalinity) are significantly associated even after removing the influence of spatial autocorrelation. Hinch et al. (1994) used the partial Mantel test to test the relationship between lake fish assemblages and several matrices of environmental variables while controlling for spatial relationships among the environmental variables. The results differed for various between-matrix comparisons depending on whether the spatial autocorrelation was included or removed prior to the analyses. The partial Mantel approach was used to factor out the influence of spatial relationships before doing a principal coordinates ordination on the fish assemblages. Hinch et al. (1994) found major differences in their interpretation of fish assemblage patterns due to the strong spatial autocorrelation of abundances for brown bullhead and pumpkinseed whereas white sucker abundance was not spatially autocorrelated.

Peres-Neto and Jackson (2001) demonstrated that a test based on a randomized Procrustean rotation provides improved statistical power over the Mantel test. The test is used to compare sets of variables such as morphological, genetic, environmental, or spatial information collected at a series of locations. As with the Mantel test, the Procrustean rotation test determines whether the pattern between two data sets is consistent with the null hypothesis of random association. Peres-Neto (2004) used Procrustes analysis to compare fish morphology among stream fish assemblages with respect to spatial and environmental conditions.

18.4.5 Consideration of Spatial Autocorrelation in Regression Analysis

Because information on fish–habitat relations is often spatial in nature, there is growing interest in using spatially explicit regression approaches to identify habitat

features that influence the occurrence and abundance of fish. The typical approach is to do a traditional (nonspatial) regression relating the property of interest (often the abundance of individual species or a group of species) to habitat characteristics. The residuals from this regression are then tested for spatial autocorrelation using a statistic such as Moran's *I* and a spatial weights matrix based on the distance between sample locations. The distance between sample locations can be determined as the geographic straight-line distance for lakes or the ocean. However, studies involving stream sites also should consider the distance along the drainage network because the similarity in fish assemblages among sites may depend more on the distance an organism or disturbance has to travel along a stream corridor than the overland distance between the sites.

Tiffan et al. (2002) tested whether spatial correlation among habitat features influenced logistic regression models that predicted the probability of Chinook salmon presence among sites in the Columbia River. In addition to habitat conditions at a site, they included the distance of each site from the downstream end of the study reach as a predictor variable. This variable was not statistically significant in their model, indicating that spatial autocorrelation was not a problem in their data set.

If spatial autocorrelation is detected, one has four options. First, a plot of residuals across the study area may suggest some geographic gradients in abiotic factors not considered in the original analysis. For example, a latitudinal gradient among the residuals may indicate that a climate variable should be added to the regression model to account for the fact that the growing season declines with increased latitude in the northern hemisphere. Or there may be underlying elevational or geologic gradients across the study area that could be incorporated into the analysis. If no new variables can be identified that remove the autocorrelation, then the second option is to adjust the regression model to account for the fact that the error terms (i.e., the residuals) are not independent. This is done by adjusting the covariance matrix used in the general least-squares method to account for the fact that off-diagonal elements (assumed to be zero in traditional regression analysis when error terms are not correlated) will have nonzero values that increase as the spatial separation between pairs of sampling locations decreases. Further information is presented by Odland (1988).

Adjusting the regression model typically leads to increased standard errors for the regression model parameter estimates and may cause some of the original variables to be dropped from the regression model. The parameter estimates themselves usually are not changed greatly. In general, failure to account for spatial autocorrelation will result in an increased chance of type I statistical errors (i.e., finding regression model parameters to differ significantly from zero when, in fact, they do not).

Isaak and Hubert (2001a) contrasted nonspatial and spatial regression analyses in predicting stream habitat attributes from watershed characteristics. They used watershed characteristics such as watershed size, basin elevation, basin slope, and vegetation characteristics to predict three reach scale characteristics: base flow stream width, stream alkalinity, and stream slope. Thirteen of fifteen least-squares

regressions had spatially correlated residuals. When the regressions were spatially adjusted, the magnitude of standardized regression coefficients tended to decline slightly indicating that the impact of these predictor variables would be overestimated if spatial autocorrelation were ignored.

A third option for dealing with spatial autocorrelation in regression analysis is to include a categorical variable in the regression analysis along with the site level characteristics that are used as predictor variables. This allows one to assign clumped sites to a similar category or group and then to test for the relative importance of category effects (indicative of spatial autocorrelation) versus site level effects in explaining the variation in the dependent variable. Dunham and Vinyard (1997) provide an example of this approach for relating variation in fish abundance to both stream level and site level effects. Their data consisted of multiple sample sites for each of several streams. At each site both fish abundance and local habitat features were measured. When the analysis was done considering only site level variability (i.e., without a categorical variable identifying individual streams), wetted stream width was a strong predictor of cutthroat trout abundance. However, when a categorical variable for stream identity was added to the regression, it explained the majority of the variation in trout abundance and site level effects such as wetted stream width became nonsignificant. The reason for the loss of significance of site level effects was that variables such as wetted stream width were highly confounded among streams. Some streams contained only narrow sample sites and other streams contained only wide sample sites, and this made it impossible to separate out site level influences on fish abundance from larger-scale influences operating at the stream level.

A fourth option can be used for data that consist of groups of spatially clumped sample sites that show autocorrelation within groups but not across groups. In such a case, a solution is simply to average all sites within a group into a single composite site for both the dependent and independent variables. Nash et al. (1999) used this approach for a comparison of fish abundance patterns in lakes versus streams. The data came from widely distributed locations around the world but were highly clumped in that a group of lakes or streams had been sampled at each location. When Nash et al. (1999) formed a composite site for the lakes or rivers at each location, they found that differences in fish assemblage attributes between lakes and rivers were less pronounced than when all sites at a location were considered as independent samples.

18.4.6 Conclusions Regarding Spatial Analyses in Fisheries Applications

Spatial patterns sometimes confound relationships between fish assemblage characteristics and environmental features, but such problems may not be recognized unless researchers examine the data for spatial autocorrelation. Researchers should consider carefully the design of their study in order to minimize spatial autocorrelation. For example, a spatially blocked or nested design may reduce the impact of spatial autocorrelation by ensuring that sampling sites are spread across the study area. This reduces the probability of having many sites in clumped

patterns as aggregated sampling tends to emphasize positive or negative spatial autocorrelation at small scales and tends to complicate the assignment of suitable degrees of freedom. In some studies, biologists may be interested primarily in the spatial aspects of the data as these may be important in relating fish abundance to environmental quality in river systems or may strongly determine spawning migration patterns. The spatial connectivity of these systems may be the information of interest. Alternatively, for some studies, spatial aspects may be confounding information about fish–habitat relations. We also caution that researchers should not necessarily be alarmed if they find strong autocorrelation. In some cases, researchers may be interested in longitudinal patterns of species richness within a river system. Attributes of downstream sites are not likely to be independent of conditions upstream. Therefore, we may not want to test the significance of our null hypotheses formally, but it is appropriate to provide summary statistics (e.g., correlation or regression coefficients) for variables of interest as these help us understand how the variables are related. We may be able to provide an excellent predictive model of species richness within a river (without the need to assign a probability to this model) but may not be able to generalize beyond the bounds of the system. When we want to control for spatial autocorrelation and apply inferential statistics, we now have a set of tools suitable for addressing many of our questions. Spatial autocorrelation is neither inherently good nor bad, but we must consider the impact that it may have on our application of inferential statistics and the conclusions made.

18.5 ADDITIONAL FISHERIES ISSUES AT THE WATERSHED SCALE

18.5.1 Quantification of the Ecological Condition of Watersheds

There is a long history of trying to assess the well-being of aquatic ecosystems based on the status of fish assemblages (Simon 1999). A popular approach has been to summarize information on species abundances and guild composition in an index of biotic integrity (IBI; see Chapter 15). Typically IBIs assess aquatic conditions at the site scale, but there is growing interest in evaluating biological integrity at the watershed scale (Jensen and Bourgeron 2001; Saunders et al. 2002). One approach is simply to average site IBI scores across a basin (Steedman 1988). Another approach is to develop indices of watershed condition that explicitly consider large-scale characteristics. Moyle and Randall (1998) developed such an index they termed the watershed index of biotic integrity (W-IBI). To use the W-IBI, each watershed was scored along six metrics thought to reflect overall ecosystem health (Table 18.2). For each metric, the condition of the watershed was scored as 1 (poor), 3 (intermediate), or 5 (good), and the sum of the scores became the W-IBI for that watershed. The higher the W-IBI score, the better the biological condition of the watershed. A negative correlation between W-IBI score and watershed characteristics such as the abundance of dams, reservoirs, and roads close to streams indicated that the index could be used to rank watersheds based on the degree of human-related perturbation. Being able to quantify the biological status

Table 18.2 Watershed level metrics used to develop an index of biotic integrity for watersheds (W-IBI) in California (Moyle and Randall 1998). The ratings used to assess watershed condition based on the scoring criteria are poor (1), moderate (2), and good (3).

	Scoring criteria		
Metric	1	2	3
Native ranid frogs	No or few recent records of frogs	Frogs present in multiple locations but not abundant	Frogs abundant and widely distributed
Native fishes	Native fishes absent or rare or watershed was historically fishless but now contains fish	Native fishes present in only part of native range but still easy to find	Native fishes throughout watershed at historical abundances
Native fish assemblages	Assemblages consist mainly of nonnative species, native species missing, or both	Remnants of native assemblages, one or more nonnative species, or both	Native assemblages largely intact
Anadromous fishes	Anadromous fishes absent or rare	Anadromous fishes present mainly below dams, abundances reduced from historic levels, or both	Anadromous fishes present throughout original range
Trout	Range expanded through introductions, mix of native and nonnative species, or native populations depleted	Range somewhat expanded or native populations reduced and nonnative species present	Trout mainly native species within native range
Stream fish abundance	Overall fish abundance substantially reduced from historic levels or fish abundant in historically fishless waters	Fish abundances reduced somewhat from historical levels	Abundances similar to historical levels

of watersheds is important for prioritizing watersheds for conservation or rehabilitation purposes (Moyle and Randall 1998; Saunders et al. 2002).

Another example of an approach to summarize habitat conditions at a watershed scale is the watershed disturbance index developed by Whittier et al. (2002). The authors used PCA (section 18.3.6) to combine data on watershed land use, road density, and human population density into a single metric of disturbance and found that the number of invasive species in streams was positively correlated with the degree of watershed disturbance. Van Kirk and Benjamin (2001) developed indices of salmonid population status and hydrologic integrity for 41 watersheds in the Greater Yellowstone Ecosystem. The status of salmonid populations was correlated with hydrologic integrity when both nonnative and native species were considered together. However, the status of native salmonid populations considered separately was not strongly correlated with hydrologic integrity. This indicated that other factors, especially introduction of nonnative salmonid species, was more important than were physical factors in determining the status of native salmonids across watersheds in the Greater Yellowstone Ecosystem.

Although much effort has been applied to developing sampling protocols for determining the number of fish species at the site scale, relatively little work has been done in determining fish species richness at the watershed scale. Smith and Jones (2005) provided guidelines for determining riverine fish species composition at the watershed scale and emphasized the value of plotting species accumulation curves to determine adequate inventory completion. They also indicated the value of combining random site selection with targeted site sampling to ensure capture of species associated with rare habitats.

18.5.2 Identification of Sufficient Habitat for Conservation of Stream Fishes

Identifying the amount of habitat necessary to sustain a population is a basic problem in conservation biology, particularly when trying to establish new populations. Population genetic theory suggests that an effective population size of at least 500 breeding adults is necessary to prevent extinction from stochastic population processes or inbreeding. An effective population size of 500 could be achieved if 250 males and 250 females each contributed equally to the genetic make-up of the next generation. For stream fishes, the length of stream necessary to sustain a given number of breeding adults can be estimated based on demographic data and population density. For example, Hilderbrand and Kershner (2000) estimated that a population of 2,500 cutthroat trout (>75 mm total length) would be needed to attain an effective population size of 500 breeding adults. The abundance of cutthroat trout in their Rocky Mountain streams varied from 0.1 fish to 0.3 fish per meter of stream length. Thus, the length of stream needed to support 2,500 cutthroat trout would be 25 km at low fish abundance and 8 km at high fish abundance. Using this approach, Kruse et al. (2001) found that only 7 of 23 headwater populations of cutthroat trout in the Absaroka Mountains of Wyoming contained sufficient habitat to maintain an effective population size of 500 or more breeding individuals. They concluded that isolating headwater populations to prevent

invasion by nonnative trout would not be an effective long-term conservation strategy because small size would make most populations vulnerable to extinction.

Harig and Fausch (2002) examined the minimum habitat requirements for establishing translocated populations of cutthroat trout. Rather than estimate stream length needed to maintain a certain number of fish, they used an empirical approach to relate the persistence of previously translocated populations to various habitat factors. Using logistic regression, they found that watershed area was a useful predictor of the probability of success of a translocated population. In particular, they estimated that the chance of a translocation being successful was greater than 50% when watershed area was greater than 14.7 km^2. The explanation was that watersheds of that size would contain sufficient stream habitat to support a large population of cutthroat trout and have enough lower elevation streams where thermal conditions would be optimal to ensure recruitment of young fish.

18.5.3 Approaches from Landscape Ecology: Influence of Patch Size, Isolation, and Landscape Position on Fish Distribution and Abundance

Landscape ecology encompasses both the study of ecological phenomena at large spatial scales as well as the study of habitat patchiness. Recently, there has been interest in understanding how landscape concepts such as patch size, isolation, or juxtaposition influence the distribution and abundance of fishes (Schlosser 1995; Fausch et al. 2002). Habitat patches are areas of suitable habitat surrounded by unsuitable habitat. Patches can consist of subcatchments in stream networks (Rieman and McIntyre 1995); embayments in reservoirs (Phillips et al. 1997); pools in streams (Lonzarich et al. 2000; Eros and Grossman 2005); stream reaches with suitable temperatures (Torgersen et al. 1999); or macrophyte beds in lakes (Chick and McIvor 1994).

Rieman and McIntrye (1995) studied large-scale distribution patterns of bull trout in Idaho and defined a habitat patch as a stream catchment above 1,600 m elevation that contained thermally suitable habitat for this species. Using logistic regression, they found patch size was a good predictor of the occurrence of bull trout (Box 18.2). In a similar analysis, Dunham et al. (2002) found that patch size also was a good predictor of the occurrence of Lahontan cutthroat trout in the Lahontan basin of Oregon and Nevada.

Based on island biogeographic theory, isolated patches should have fewer species than would patches near sources of colonists. Dunham and Rieman (1999) explored this phenomenon for bull trout in Idaho where patches were stream catchments containing thermally suitable habitat. Using logistic regression models, they found that patch isolation, measured as the stream distance to the nearest occupied patch, interacted with patch size to determine the probability that a patch would contain bull trout. As expected, bull trout were less likely to occur in isolated patches, especially if the patches were small. Eros and Grossman (2005) used path analysis to show that patch location in the landscape (measured as the distance from downstream tributaries) and patch size (measured as pool volume) were important predictors of fish species richness in pools in a Hungarian stream.

For fish adapted to pool habitat in streams, shallow, fast-flowing reaches associated with road crossings or natural riffles can be an isolating mechanism. Lonzarich et al. (2000) found that pools separated by long riffles experienced less fish movement than did pools separated by short riffles. Management implications were that natural riffles or shallow, fast-flowing reaches created by human activities such as road culverts could be isolating mechanisms for some species and could slow fish recolonization after disturbances (Warren and Pardew 1998). On a larger scale, dams and their associated reservoirs fragment river systems into disjointed reaches of flowing water (Dynesius and Nilsson 1994). This can lead to the extirpation of obligate riverine fishes from isolated patches of river habitat (Winston et al. 1991; Quist et al. 2004b).

In lakes, distance to the nearest lake has been used as a measure of isolation. Olden et al. (2001) explored various measures of lake distance in their study of how lake isolation influenced fish assemblage patterns across a large watershed in Ontario. An isolation measure based on the straight-line distance between two lakes did not appear to be as insightful as measures that incorporated waterway distance and stream gradients between lakes. Understanding lake isolation helped in interpreting lake assemblage patterns. For example, some lakes contained coldwater species not predicted based on the lake's small area and shallow depth. These species appeared to be sustained because the lakes were in close proximity to deep lakes that could support these species. Thus, the small lakes appeared to be sink populations for coldwater species originating from source populations in other lakes. Measures of lake isolation that incorporate waterway distance and stream gradient also are important when modeling the dispersion of exotic species across a lake district (Hrabik and Magnuson 1999).

In addition to patch size and isolation, the juxtaposition of a habitat patch relative to other patches and landscape features can influence the biological characteristics of aquatic systems. For example, the position of a lake in a watershed can influence limnological properties. Riera et al. (2000) developed the concept of lake order to quantify a lake's position along a drainage network. Low-order lakes are isolated water bodies at the headwaters of drainage systems whereas high-order lakes are lower in the drainage and have river connections with other lakes. For the lake district studied by Riera et al., lake size, ion concentrations, crayfish abundance, and fish species richness all increased with lake order. Snodgrass and Meffe (1998) also provided evidence that the location of a water body within a watershed could influence biological properties. They found the composition of fish assemblages in South Carolina beaver ponds depended on landscape position. In particular, piscivorous fishes appeared to eliminate small-bodied prey species in upstream beaver ponds but not in downstream beaver ponds because the latter were closer to sources of colonists.

For streams, the spatial position of tributary streams within a drainage can influence fish assemblage composition. Adventitious first-order streams are small streams that originate low in the drainage basin and flow directly into large rivers. These streams naturally have more diverse fish assemblages than do first-order streams higher in the watershed (Osborne and Wiley 1992). Although stream

order is a strong predictor of fish assemblage characteristics, it fails to distinguish between the two types of first-order streams. This has important implications for attempts to quantify fish assemblage well-being through indices of biotic integrity. With such indices, stream order is used to calculate expected species richness at a site. However, this process would unfairly characterize first-order streams high in the drainage network as having degraded fish assemblages relative to adventitious first-order streams (Osborne et al. 1992).

18.6 CONCLUSIONS

Watershed-scale analyses will play an increasing role in fisheries management for three reasons. First, local habitat features often are the result of landscape level influences on stream catchments (Nelson et al. 1992; Isaak and Hubert 2001a). Thus, understanding fish–habitat relations often requires an understanding of the large-scale habitat features that determine fish distribution patterns (Wiley et al. 1997). Second, some processes that influence fish occurrence or abundance operate mainly at the landscape level, particularly phenomena such as habitat complementarity or metapopulation dynamics related to patch size, isolation, and juxtaposition (Schlosser 1995; Dunham and Rieman 1999; Fausch et al. 2002; Scheurer et al. 2003). Understanding the importance of these factors requires looking beyond site level habitat features to a landscape level perspective. Third, humans continue to cause large-scale alterations across the landscape. For example, the status of fish populations often is related to land use at the watershed scale (Allan et al. 1997; Moyle and Randall 1998; Schrank et al. 2001). Furthermore, human alterations increasingly occur at large spatial scales due to such phenomena as acid rain, climate warming, and watershed fragmentation.

An important reason for the growing interest in watershed-level analyses is the emergence of GIS technology that makes it relatively easy to quantify watershed features from existing maps or GIS coverages without the need for expensive and labor-intensive fieldwork (Fisher and Rahel 2004). For example, fish biomass (Lanka et al. 1987), maximum stream temperature (Isaak and Hubert 2001b), mean annual stream discharge (O'Shea 1995), trout occurrence (Dunham and Rieman 1999), and the number of nonnative fish species (Marchetti et al. 2004) can be predicted from characteristics such as watershed area, mean basin elevation, soil types, and the proportion of urban development in a watershed. This provides a way for fisheries scientists to gain insight about fish abundance or habitat conditions when extensive field measurements are not available or too costly to collect. With the increasing emphasis on ecosystem management, watershed-level approaches will play an important role in the management and conservation of fish populations.

18.7 REFERENCES

Abell, R. A., and ten coauthors. 2000. Freshwater ecoregions of North America: a conservation assessment. Island Press, Washington, D.C.

Allan, J. D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution and Systematics 35:257–284.

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37:149–161.

Anderson, D. R., and K. P. Burnham. 2001. Commentary on models in ecology. Bulletin of the Ecological Society of America 82:160–161.

Barton, D. R., W. D. Taylor, and R. M. Biette. 1985. Dimensions of riparian buffer required to maintain trout habitat in southern Ontario streams. North American Journal of Fisheries Management 5:364–378.

Baxter, C. V., C. A. Frissell, and F. R. Hauer. 1999. Geomorphology, logging roads, and the distribution of bull trout spawning in a forested river basin: implications for management and conservation. Transactions of the American Fisheries Society 128: 854–867.

Braaten, P. J., and C. S. Guy. 1999. Relations between physicochemical factors and abundance of fishes in tributary confluences of the lower channelized Missouri River. Transactions of the American Fisheries Society 128:1213–1221.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. G. Stone. 1984. Classification and regression trees. Wadsworth International Group, Belmont, California.

Brosse, S., J. F. Guegan, J. N. Tourenq, and S. Lek. 1999. The use of neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake. Ecological Modelling 120:299–311.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd edition. Springer-Verlag, New York.

Cattaneo, F., B. Hugueny, and N. Lamouroux. 2003. Synchrony in brown trout, *Salmo trutta,* population dynamics: a "Moran effect" on early-life stages. Oikos 100:43–54.

Chick, J. H., and C. C. McIvor. 1994. Patterns in the abundance and composition of fishes among beds of different macrophytes: viewing the littoral zone as a landscape. Canadian Journal of Fisheries and Aquatic Sciences 51:2873–2882.

Cumming, G. S. 2004. The impact of low-head dams on fish species richness in Wisconsin, USA. Ecological Applications 14:1495–1506.

De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192.

Diamond, J. M., and V. B. Serveiss. 2001. Identifying sources of stress to native aquatic fauna using a watershed ecological risk assessment framework. Environmental Science and Technology 35:4711–4718.

Digby, P. G. N., and R. A. Kempton. 1987. Multivariate analysis of ecological communities. Chapman and Hall, London.

Dolloff, C. A., D. G. Hankin, and G. H. Reeves. 1993. Basinwide estimation of habitat and fish populations in streams. U.S. Forest Service General Technical Report SE-83.

Dolloff, C. A., H. E. Jennings, and M. D. Owen. 1997. A comparison of basinwide and representative reach habitat survey techniques in three southern Appalachian watersheds. North American Journal of Fisheries Management 17:339–347.

Dunham, J. B., and B. E. Rieman. 1999. Metapopulation structure of bull trout: influences of physical, biotic, and geometrical landscape characteristics. Ecological Applications 9:642–655.

Dunham, J. B., B. E. Rieman, and J. T. Peterson. 2002. Patch-based models to predict species occurrence: lessons from salmonid fishes in streams. Pages 327–334 *in* J. M. Scott, and six coeditors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.

Dunham, J. B., and G. L. Vinyard. 1997. Incorporating stream level variability into analyses of site level fish habitat relationships: some cautionary examples. Transactions of the American Fisheries Society 126:323–329.

Dutilleul, P. 1993. Modifying the *t* test for assessing the correlation between two spatial processes. Biometrics 49:305–314.

Dynesius, M., and C. Nilsson. 1994. Fragmentation and flow regulation of river systems in the northern third of the world. Science 266:753–762.

Ellison, A. M., and A. A. Agrawal. 2005. The statistics of rarity. Ecology 86:1079–1080.

Eros, T., and G. D. Grossman. 2005. Fish biodiversity in two Hungarian streams: a landscape-based approach. Archiv fur Hydrobiologie 162:53–71.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52: 483–498.

Fisher, W. . L., and F. J. Rahel. 2004. Geographic information systems in fisheries. American Fisheries Society, Bethesda, Maryland.

Fortin, M.-J., and J. Gurevitch. 2001. Mantel tests: spatial structure in field experiments. Pages 308–326 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments, 2nd edition. Oxford University Press, New York.

Frimpong, E. A., and seven coauthors. 2005. Determination of optimal riparian forest buffer dimensions for stream biota–landscape association models using multimetric and multivariate responses. Canadian Journal of Fisheries and Aquatic Sciences 62:1–6.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10:199–214.

Greenfield, B. K., T. R. Hrabik, C. J. Harvey, and S. R. Carpenter. 2001. Predicting mercury levels in yellow perch: use of water chemistry, trophic ecology, and spatial traits. Canadian Journal of Fisheries and Aquatic Sciences 58:1419–1429.

Grenouillet, G., D. Pont, and C. Herisse. 2004. Within-basin fish assemblage structure: the relative influence of habitat versus stream spatial position on local species richness. Canadian Journal of Fisheries and Aquatic Sciences 61:93–102.

Gresswell, R. E., W. J. Liss, G. L. Larson, and P. J. Bartlein. 1997. Influence of basin-scale physical variables on life history characteristics of cutthroat trout in Yellowstone Lake. North American Journal of Fisheries Management 17:1046–1064.

Griffith, D. A. 1987. Spatial autocorrelation—a primer. Association of American Geographers, Washington, D.C.

Griffith, G. E., J. M. Omernik, and A. J. Woods. 1999. Ecoregions, watersheds, basins, and HUCs: how state and federal agencies frame water quality. Journal of Soil and Water Conservation 54:666–677.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Harig, A. L., and K. D. Fausch. 2002. Minimum habitat requirements for establishing translocated cutthroat trout populations. Ecological Applications 12:535–551.

Hilderbrand, R. H., and J. L. Kershner. 2000. Conserving inland cutthroat trout in small streams: how much stream is enough? North American Journal of Fisheries Management 20:513–520.

Hinch, S. G., K. M. Somers, and N. C. Collins. 1994. Spatial autocorrelation and assessment of habitat–abundance relationships in littoral zone fish. Canadian Journal of Fisheries and Aquatic Sciences 51:701–712.

Hosmer, D. W. and S. Lemeshow. 2000. Applied logistic regression, 2nd edition. Wiley, New York.

Hrabik, T. R., and J. Magnuson. 1999. Simulated dispersal of exotic rainbow smelt *(Osmerus mordax)* in a northern Wisconsin lake district and implications for management. Canadian Journal of Fisheries and Aquatic Sciences 56 (Supplement 1):35–42.

Hynes, H. B. 1975. A stream and its valley. Internationale Vereinigung fur Theoretische und Angewandte Limnologie 19:1–15.

Isaak, D. J., and W. A. Hubert. 2001a. Production of stream habitat gradients by montane watersheds: hypotheses tests based on spatially explicit path analyses. Canadian Journal of Fisheries and Aquatic Sciences 58:1089–1103.

Isaak, D. J., and W. A. Hubert. 2001b. A hypothesis about factors that affect maximum summer stream temperatures across montane landscapes. Journal of the American Water Resources Association 37:351–366.

Jackson, D. A. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology 74:2204–2214.

Jackson, D. A. 1997. Compositional data in community ecology: the paradigm or peril of proportions. Ecology 78:929–940.

Jackson, D. A., and H. H. Harvey. 1989. Biogeographic associations of fish communities—local versus regional processes. Ecology 70:1472–1485.

Jackson, D. A., and H. H. Harvey. 1993. Fish and benthic-invertebrates—community concordance and community–environment relationships Canadian Journal of Fisheries and Aquatic Sciences 50:2641–2651.

Jackson, D. A., and K. M. Somers. 1989. Are probability estimates from the permutation model of Mantel's test stable? Canadian Journal of Zoology 67:766–769.

Jennings, M. J., E. E. Emmons, G. R. Hatzenbeler, C. Edwards, and M. A. Bozek. 2003. Is littoral habitat affected by residential development and land use in watersheds of Wisconsin lakes? Lake and Reservoir Management 19:272–279.

Jensen, M. E., and P. S. Bourgeron, editors. 2001. A guidebook for integrated ecological assessments. Springer-Verlag, New York.

Jessup, B. K. 1998. A strategy for simulating brown trout population dynamics and habitat quality in an urbanizing watershed. Ecological Modelling 112:151–167.

Johnson, L. B., and S. H. Gage. 1997. Landscape approaches to the analysis of aquatic ecosystems. Freshwater Biology 37:113–132.

Kolar, C. S., and D. M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. Science 298:1233–1236.

Kruse, C. G., W. A. Hubert, and F. J. Rahel. 1997. Geomorphic influences on the distribution of Yellowstone cutthroat trout in the Absaroka Mountains, Wyoming. Transactions of the American Fisheries Society 126:418–427.

Kruse, C. G., W. A. Hubert, and F. J. Rahel. 2001. An assessment of headwater isolation as a conservation strategy for cutthroat trout in the Absaroka Mountains of Wyoming. Northwest Science 75:1–11.

Laë, R., S. Lek, and J. Moreau. 1999. Predicting fish yield of African lakes using neural networks. Ecological Modelling 120:325–335.

Lanka, R. P., W. A. Hubert, and T. A. Wesche. 1987. Relations of geomorphology to stream habitat and trout standing stocks in small Rocky Mountain streams. Transactions of the American Fisheries Society 116:21–28.

Legendre, P., and L. Legendre. 1998. Numerical Ecology, 2nd edition. Elsevier. Amsterdam.

Lek, S., and J. F. Guégan. 1999. Artificial neural networks as a tool in ecological modeling, an introduction. Ecological Modelling 120:65–73.

Lonzarich, D. G., M. R. Lonzarich, and M. L. Warren, Jr. 2000. Effects of riffle length on the short-term movement of fishes among stream pools. Canadian Journal of Fisheries and Aquatic Sciences 57:1508–1514.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacNally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between—and reconciliation of —'predictive' and 'explanatory' models. Biodiversity and Conservation 9:655–671.

Magnuson, J. J., W. M. Tonn, A. Banerjee, J. Toivonen, O. Sanchez, and M. Rask. 1998. Isolation vs. extinction in the assembly of fishes in small northern lakes. Ecology 79: 2941–2956.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology, 2nd edition. Chapman and Hall, London.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209–220.

Marchetti, M. P., T. Light, P. B. Moyle, and J. H. Viers. 2004. Fish invasions in California watersheds: testing hypotheses using landscape patterns. Ecological Applications 14:1507–1525.

Marsh-Matthews, E., and W. J. Matthews. 2000. Geographic, terrestrial and aquatic factors: which most influence the structure of stream fish assemblages in the Midwestern United States? Ecology of Freshwater Fish 9:9–21.

Maxwell, J. R., C. J. Edwards, M. E. Jensen, S. J. Paustian, H. Parrott, and D. M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). U.S. Forest Service General Technical Report NC-176.

Miles, J. N. V. 2000. Statistical software for microcomputers: Amos 4.0. British Journal of Mathematical and Statistical Psychology 53:141–144.

Modde, T., R. C. Ford, and M. G. Parsons. 1991. Use of a habitat-based stream classification system for categorizing trout biomass. North American Journal of Fisheries Management 11:305–311.

Montgomery, D. R., G. E. Grant, and K. Sullivan. 1995. Watershed analysis as a framework for implementing ecosystem management. Water Resources Bulletin 31:369–386.

Moyle, P. B., and P. J. Randall. 1998. Evaluating the biotic integrity of watersheds in the Sierra Nevada, California. Conservation Biology 12:1318–1326.

Nash, C. H., J. S. Richardson, and S. G. Hinch. 1999. Spatial autocorrelation and fish production in freshwaters: a comment on Randall et al. (1995). Canadian Journal of Fisheries and Aquatic Sciences 56:1696–1699.

Nelson, R. L., W. S. Platts, D. P. Larsen, and S. E. Jensen. 1992. Trout distribution and habitat in relation to geology and geomorphology in the North Fork Humboldt River drainage, northeastern Nevada. Transactions of the American Fisheries Society 121:405–426.

Odland, J. 1988. Spatial autocorrelation. Sage Publications, Newbury Park, California.

Olden, J. D., and D. A. Jackson. 2000. Torturing data for the sake of generality: how valid are our regression models? Ecoscience 7:501–510.

Olden, J. D., and D. A. Jackson. 2001. Fish–habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. Transactions of the American Fisheries Society 130:878–897.

Olden, J. D., and D. A. Jackson. 2002a. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling 154:135–150.

Olden, J. D., and D. A. Jackson. 2002b. A comparison of statistical models for modeling fish species distributions. Freshwater Biology 47:1976–1995.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2001. Spatial isolation and fish communities in drainage lakes. Oecologia 127:572–585.

Olsen, A. R., and seven coauthors. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54:1–45.

O'Shea, D. T. 1995. Estimating minimum instream flow requirements for Minnesota streams from hydrologic data and watershed characteristics. North American Journal of Fisheries Management 15:569–578.

Osborne, L. L., and six coauthors. 1992. Influence of stream location in a drainage network on the index of biotic integrity. Transactions of the American Fisheries Society 121:635–643.

Osborne, L. L., and M. J. Wiley. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences 49:671–681.

Peres-Neto, P. R. 2004. Patterns in the co-occurrence of fish species in streams: the role of site suitability, morphology, and phylogeny versus species interactions. Oecologia 140: 352–360.

Peres-Neto, P. R., and D. A. Jackson. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia 129:169–172.

Peres-Neto, P. R., D. A. Jackson, and K. M. Somers. 2003. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. Ecology 84:2347–2363.

Phillips, J. M., J. R. Jackson, and R. L. Noble. 1997. Spatial heterogeneity in abundance of age-0 largemouth bass among reservoir embayments. North American Journal of Fisheries Management 17:894–901.

Porter, M. S., J. Rosenfield, and E. A. Parkinson. 2000. Predictive models of fish species distribution in the Blackwater drainage, British Columbia. North American Journal of Fisheries Management 20:349–359.

Prepas, E. E. and twelve coauthors. 2001. Landscape variables influencing nutrients and phytoplankton communities in Boreal Plain lakes of northern Alberta: a comparison of wetland- and upland-dominated catchments. Canadian Journal of Fisheries and Aquatic Sciences 58:1286–1299.

Quist, M. C., W. A. Hubert, and D. J. Isaak. 2004a. Fish assemblage structure and relations with environmental conditions in a Rocky Mountain watershed. Canadian Journal of Zoology 82:1554–1565.

Quist, M. C., W. A. Hubert, and F. J. Rahel. 2004b. Relations among habitat characteristics, exotic species, and turbid-river cyprinids in the Missouri River drainage of Wyoming. Transactions of the American Fisheries Society 133:727–742.

Rahel, F. J. 1984. Factors structuring fish assemblages along a bog lake successional gradient. Ecology 65:1276–1289.

Rahel, F. J., C. J. Keleher, and J. L. Anderson. 1996. Potential habitat loss and population fragmentation for cold water fish in the North Platte River drainage of the Rocky Mountains: response to climate warming. Limnology and Oceanography 41:1116–1123.

Rahel, F. J., and J. J. Magnuson. 1983. Low pH and the absence of fish species in naturally acidic lakes: inferences for cultural acidification. Canadian Journal of Fisheries and Aquatic Sciences 40:3–9.

Rahel, F. J., and N. P. Nibbelink. 1999. Spatial patterns in relations among brown trout distribution, summer air temperature, and stream size in Rocky Mountain streams. Canadian Journal of Fisheries and Aquatic Sciences 56 (Supplement 1):43–51.

Randall, R. G., J. R. M. Kelso, and C. K. Minns. 1995. Fish production in freshwaters: are rivers more productive than lakes? Canadian Journal of Fisheries and Aquatic Sciences 52:631–643.

Rathert, D., D. White, J. C. Sifneos, and R. M. Hughes. 1999. Environmental correlates of species richness for native freshwater fish in Oregon, USA. Journal of Biogeography 26:257–273.

Reed, J. M. 1996. Using statistical probability to increase confidence of inferring species extinction. Conservation Biology 10:1283–1285.

Rieman, B. E., D. C. Lee, and R. F. Thurow. 1997. Distribution, status, and likely future trends of bull trout within the Columbia River and Klamath River basins. North American Journal of Fisheries Management 17:1111–1125.

Rieman, B. E., and J. D. McIntyre. 1995. Occurrence of bull trout in naturally fragmented habitat patches of varied size. Transactions of the American Fisheries Society 124:285–296.

Riera, J. L., J. J. Magnuson, T. K. Kratz, and K. E. Webster. 2000. A geomorphic template for the analysis of lake districts applied to the Northern Highland Lake District, Wisconsin, USA. Freshwater Biology 43:301–318.

Rodriguez, M. A., and W. M. Lewis. 1997. Structure of fish assemblages along environmental gradients in floodplain lakes of the Orinoco River. Ecological Monographs 67:109–128.

Rohlf, F. J. 1993. NTSYS-pc. Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket, New York.

Saunders, D. L., J. J. Meeuwig, and A. C. J. Vincent. 2002. Freshwater protected areas: strategies for conservation. Conservation Biology 16:30–41.

Seaber, P. R., F. P. Kapinos, and G. L. Knapp. 1987. Hydrologic unit maps. U.S. Geological Survey, Water-Supply Paper 2294, Corvallis, Oregon.

Scheuerell, M. D., and D. E. Schindler. 2004. Changes in the spatial distribution of fishes in lakes along a residential development gradient. Ecosystems 7:98–106.

Scheurer, J. A., K. D. Fausch, and K. R. Bestgen. 2003. Multiscale processes regulate brassy minnow persistence in a Great Plains river. Transactions of the American Fisheries Society 132:840–855.

Schlosser, I. J. 1995. Critical landscape attributes that influence fish population dynamics in headwater streams. Hydrobiologia 303:71–81.

Schlosser, I. J., and P. J. Angermeier. 1995. Spatial variation in demographic processes in lotic fishes: conceptual models, empirical evidence, and implications for conservation. Pages 360–370 *in* J. L. Nielsen, editor. Evolution and the aquatic ecosystem: defining unique units in population conservation. American Fisheries Society, Symposium 17, Bethesda, Maryland.

Schramm, H. L., Jr., and W. A. Hubert. 1999. Ecosystem. Pages 111–123 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Schrank, A. J., and F. J. Rahel. 2004. Movement patterns in inland cutthroat trout: management and conservation implications. Canadian Journal of Fisheries and Aquatic Sciences 61:1528–1537.

Schrank, S. J., C. S. Guy, M. R. Whiles, and B. L. Brock. 2001. Influence of instream and landscape-level factors on the distribution of Topeka shiners *Notropis topeka* in Kansas streams. Copeia 2001:413–421.

Sheldon, A. L., and G. K. Meffe. 1995. Path analysis of collective properties and habitat relationships of fish assemblages in coastal plain streams. Canadian Journal of Fisheries and Aquatic Sciences 52:23–33.

Simon, T. P., editor. 1999. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Smith, K. L., and M. L. Jones. 2005. Watershed-level sampling effort requirements for determining riverine fish species composition. Canadian Journal of Fisheries and Aquatic Sciences 62:1580–1588.

Snodgrass, J. W., and G. K. Meffe. 1998. Influence of beavers on stream fish assemblages: effects of pond age and watershed position. Ecology 79:928–942.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry, 3rd edition. Freeman, New York.

Somers, K. M., and D. A. Jackson. 1993. Adjusting mercury concentration for fish-size covariation: a multivariate alternative to bivariate regression. Canadian Journal of Fisheries and Aquatic Sciences 50:2388–2396.

Steedman, R. J. 1988. Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. Canadian Journal of Fisheries and Aquatic Sciences 45:492–501.

Steedman, R. J. 2000. Effects of experimental clearcut logging on water quality in three small boreal forest lake trout *(Salvelinus namaycush)* lakes. Canadian Journal of Fisheries and Aquatic Sciences 57(Supplement 2):92–96.

Tiffan, K. F., R. D. Garland, and D. W. Rondorf. 2002. Quantifying flow-dependent changes in subyearling fall Chinook salmon rearing habitat using two-dimensional spatially explicit modeling. North American Journal of Fisheries Management 22:713–726.

Toepfer, C. S., W. L. Fisher, and W. D. Warde. 2000. A multistage approach to estimate fish

abundance in streams using geographic information systems. North American Journal of Fisheries Management 20:634–645.

Torgersen, C. E., D. M. Price, H.W. Li, and B. A. McIntosh. 1999. Multiscale thermal refugia and stream habitat associations of Chinook salmon in northeastern Oregon. Ecological Applications 9:301–319.

Van Kirk, R. W., and L. Benjamin. 2001. Status and conservation of salmonids in relation to hydrological integrity in the Greater Yellowstone Ecosystem. Western North American Naturalist 61:359–374.

Ver Hoef, J. M. and N. Cressie 2001. Spatial statistics: analysis of field experiments. Pages 289–307 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments, 2nd edition. Oxford University Press, New York.

Wall, S. S., C. R. Berry, Jr., C. M. Blausey, J. A. Jenks, and C. J. Kopplin. 2004. Fish-habitat modeling for gap analysis to conserve the endangered Topeka shiner *(Notropis topeka)*. Canadian Journal of Fisheries and Aquatic Sciences 61:954–973.

Wang, L., J. Lyons, and P. Kanehl. 2003. Impacts of urban land cover on trout streams in Wisconsin and Minnesota. Transactions of the American Fisheries Society 132: 825–839.

Wang, L., J. Lyons, J. P. Kanehl, and R. Gatt. 1997. Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries (Bethesda) 22(6):6–12.

Warren, M. L., Jr., and M. G. Pardew. 1998. Road crossings as barriers to small-stream fish movement. Transactions of the American Fisheries Society 127:637–644.

Wesche, T. A., and D. J. Isaak. 1999. Watershed management and land use practices. Pages 217–248 *in* C. C. Kohler and W. A. Hubert, editors. Inland fisheries management in North America, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Whittier, T. R., D. P. Larsen, S. A. Peterson, and T. M. Kincaid. 2002. A comparison of impoundments and natural drainage lakes in the Northeast USA. Hydrobiologia 470:157–171.

Wiley, D. J., R. P. Morgan II, R. H. Hilderbrand, R. L. Raesly, and D. L. Shumway. 2004. Relations between physical habitat and American eel abundance in five river basins in Maryland. Transactions of the American Fisheries Society 133:515–526.

Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshwater Biology 37: 133–148.

Williams, L. R., C. M. Taylor, M. L. Warren, and J. A. Clingenpeel. 2002. Large-scale effects of timber harvesting on stream systems in the Ouachita Mountains, Arkansas, USA. Environmental Management 29:76–87.

Winston, M. R., C. M. Taylor, and J. Pigg. 1991. Upstream extirpation of four minnow species due to damming of a prairie stream. Transactions of the American Fisheries Society 120:98–105.

Young, M. K., and P. M. Guenther-Gloss. 2004. Population characteristics of greenback cutthroat trout in streams: their relation to model predictions and recovery criteria. North American Journal of Fisheries Management 24:184–197.

Index