Environmental Monitoring

Edited by

G. Bruce Wiersma

CRC PRESS

Environmental Monitoring

Environmental Monitoring

Edited by

G. Bruce Wiersma

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Environmental monitoring / edited by G. Bruce Wiersma.
p. cm.
Includes bibliographical references and index.
ISBN 1-56670-641-6 (alk. paper)
1. Environmental monitoring. I. Wiersma, G. B.

QH541.15.M64E584 2004
363.73'63—dc22 2003065879

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

Visit the CRC Press Web site at www.crcpress.com

Lewis Publishers is an imprint of CRC Press LLC

International Standard Book Number 1-56670-641-6
Library of Congress Card Number 2003065879
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

When I first entered the field of environmental monitoring 33 years ago as a new employee of a then very new U.S. federal agency called the Environmental Protection Agency, our efforts were concentrated on primarily chasing pollutant residues in the environment. Eight years later when I founded the international journal *Environmental Monitoring and Assessment* that was still certainly the case.

However, over the intervening years, while the importance of tracking and assessing chemical residues in the environment still remains, the concept of environmental monitoring has broadened to monitoring and assessment of the endpoints of environmental pollution. Environmental monitoring systems now look far beyond only measuring chemical residues in the environment to identifying and measuring the biological endpoints that more directly reflect the effect of human action rather than just the signature of human action.

The scope of environmental monitoring systems now encompasses landscape-scale monitoring networks, multimedia approaches, and far more biological indicators of environmental impact than were ever employed 20 or more years ago. In my opinion all these trends and changes are for the good and in the right direction.

Techniques and approaches are rapidly changing as well as the conceptual thinking used to design monitoring networks. For example, geostatistics were not widely applied 20 years ago, but they are commonly used today. Single media sampling programs used to be the norm 20 or more years ago, but today it is far easier to find monitoring programs that are multimedia in nature than are single media—as witnessed by the makeup of this book. I found it much easier to recruit authors dealing with ecological monitoring indicators, geostatistics, multimedia assessment programs, etc. than to identify authors who were working in the more traditional areas of air-, soil-, and water-sampling programs.

It was my intent, while thinking about the development of this book, to try to pull together a collection of articles that would represent the latest thinking in the rapidly changing field of environmental monitoring. I reviewed the current literature (within the last 5 years) for papers that I thought represented the latest thinking in monitoring technology. I then contacted these authors and asked them if they would be interested in writing a new paper based upon their current research and thinking. I also believed that the book needed a few chapters on major monitoring networks to show both the practical application aspects under field conditions as well as to provide some description of how current environmental monitoring systems are designed and operated.

I have been extremely gratified by the positive and enthusiastic response that I have received from the authors I contacted. My original letters of inquiry went out to over 50 authors, and 45 of them responded positively. Eventually that number was pared down to the 32 chapters that make up this book. I want to thank all the authors for their contributions.

About the Editor

Dr. Wiersma has been involved with environmental monitoring activities for almost 35 years. He began his career with the U.S. Environmental Protection Agency where he managed all the agency's national pesticides monitoring programs for 4 years. He then transferred to the Environmental Monitoring Systems Laboratory of the USEPA in Las Vegas, Nevada where he worked on the development of advanced monitoring techniques for the next 6 years.

In 1980 Dr. Wiersma transferred to the Idaho National Engineering Laboratory. There he helped set up their environmental sciences, geosciences, and biotechnology groups, eventually establishing and directing the Center for Environmental Monitoring and Assessment. In 1990 Dr. Wiersma became Dean of the College of Forest Resources at the University of Maine and currently is Dean of the College of Natural Sciences, Forestry and Agriculture and Professor of Forest Resources. His current research interest is focused on studying the impacts of atmospheric deposition on northern forests. One recent Ph.D. study was on the efficacy of the U.S. Forest Service's forest health monitoring indicators.

He was a member of the National Academy of Sciences/National Research Council Committee on a Systems Assessment of Marine Environmental Monitoring that resulted in the publication in 1990 of the book *Managing Troubled Waters: The Role of Marine Environmental Monitoring,* and was Chair of the National Academy of Sciences/ National Research Council Committee Study on Environmental Database Interfaces that resulted in the publication in 1995 of the book *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data*.

Dr. Wiersma has written more than 130 scientific papers and has been the managing editor of the international peer-reviewed journal *Environmental Monitoring and Assessment* for 25 years.

Contributors

Debra Bailey
agroscope
FAL Reckenholz
Swiss Federal Research Station for Agroecology and Agriculture
Zurich, Switzerland

Roger Blair
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Corvallis, Oregon

David Bolgrien
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Duluth, Minnesota

M. Patricia Bradley
U.S. Environmental Protection Agency
Office of Research and Development
Environmental Science Center
Meade, Maryland

Barbara Brown
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Norragansett, Rhode Island
Thamas Brydges
Brampton, Ontario
Canada

Giorgio Brunialti
DIPTERIS
University of Genova
Genova, Italy

Dale A. Bruns
Pennsylvania GIS Consortium
College of Science and Engineering
Wilkes University
Wilkes-Barre, Pennsylvania

Joanna Burger
Environmental and Occupational Health Sciences Institute
Consortium for Risk Evaluation with Stakeholder Participation,
and Division of Life Sciences
Rutgers University
Piscataway, New Jersey

Janet M. Carey
School of Botany
University of Melbourne
Victoria, Australia

Vincent Carignan
Institut des sciences de l'environnement
Université due Québec à Montréal
Montréal, Québec, Canada

Charissa J. Chou
Pacific Northwest National Laboratory
Richland, Washington

Mary C. Christman
Biometrics Program
Department of Animal and Avian Sciences
University of Maryland
College Park, Maryland

William D. Constant
Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, Louisiana

Susan M. Cormier
U.S. Environmental Protection Agency
National Exposure Research Laboratory
Cincinnati, Ohio

Joseph Dlugosz
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Duluth, Minnesota

Janet D. Eckhoff
National Park Service
Wilson's Creek National Battlefield
Republic, Missouri

J. Alexander Elvir
College of Natural Science, Forestry and Agriculture
University of Maine, Orono

Marco Ferretti
LINNAEA
Firenze, Italy

Paolo Giordani
DIPTERIS
University of Genova
Genova, Italy

Michael Gochfeld
Environmental and Occupational Health Sciences Institute
Consortium for Risk Evaluation with Stakeholder Participation,
and Division of Life Sciences
Rutgers University
Piscataway, New Jersey

James T. Gunter
University of Oklahoma
Norman, Oklahoma

Richard Haeuber
U.S. Environmental Protection Agency
Washington, D.C.

Stephen Hale
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Narragansett, Rhode Island

David Michael Hamby
Department of Nuclear Engineering and Radiation Health Physics
Oregon State University
Corvallis, Oregon

Steven Hedtke
National Health and Environmental Effects Research Laboratory
Research Triangle Park, North Carolina

Daniel Heggem
U.S. Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Las Vegas, Nevada

Felix Herzog
agroscope
FAL Reckenholz
Swiss Federal Research Station for Agroecology and Agriculture
Zurich, Switzerland

Paul F. Hudak
Department of Geography and Environmental Science Program
University of North Texas
Denton, Texas

Laura Jackson
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Research Triangle Park, North Carolina

K. Bruce Jones
U.S. Environmental Protection Agency
National Exposure Research Laboratory
Research Triangle Park, North Carolina

Romualdas Juknys
Department of Environmental Sciences
Vytautas Magnum University
Kaunas, Lithuania

I. Kalikhman
Yigal Allon Kinneret Limnological Laboratory
Israel Oceanographic and Limnological Research Ltd.
Haifa, Israel

Albert Köhler
Worms, Germany

Michael Kolian
U.S. Environmental Protection Agency
Clean Air Markets
Washington, D.C.

Frederick W. Kutz
Consultant in Environmental Science
Columbia, Maryland

Mandy M.J. Lane
Center for Ecology & Hydrology
Natural Environmental Research Council
Cumbria, United Kingdom

Barbara Levinson
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Research
Washington, D.C.

Yu-Pin Lin
Department of Landscape Architecture
Chinese Culture University
Taipei, Taiwan

Rick Linthurst
U.S. Environmental Protection Agency
Office of Research and Development
Office of Inspector General
Washington, D.C.

Michael E. McDonald
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Research Triangle Park, North Carolina

Jay J. Messer
U.S. Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Research Triangle Park, North Carolina

Jaroslav Mohapl
Waterloo
Canada

Karen R. Obenshain
Keller and Heckman LLP
Washington, D.C.

Anthony Olsen
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Corvallis, Oregon

Robert V. O'Neill
TN and Associates
Oak Ridge, Tennessee

Sharon L. Osowski
U.S. Environmental Protection Agency
Compliance Assurance and Enforcement Division
Office of Planning and Coordination
Dallas, Texas

John Paul
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Research Triangle Park, North Carolina

Steven Paulsen
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Corvallis, Oregon

James L. Regens
University of Oklahoma
Institute for Science and Public Policy
Norman, Oklahoma

Susannah Clare Rennie
Center for Ecology and Hydrology
CEH Merlewood
Cumbria, United Kingdom

Kurt Riitters
Forest Health Monitoring
USDA Forest Service
Southern Research Station
Research Triangle Park, North Carolina

Dirk J. Roux
CSIR Environmentek
Pretoria, South Africa

Estelle Russek-Cohen
Biometrics Program
Department of Animal and Avian Sciences
University of Maryland
College Park, Maryland

Elizabeth R. Smith
U.S. Environmental Protection Agency
Office of Research and Development
National Health Research Laboratory
Research Triangle Park, North Carolina

John Stoddard
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Corvallis, Oregon

Kevin Summers
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Gulf Breeze, Florida

Borys M. Tkacz
Forest Health Monitoring
USDA Forest Service
Washington, D.C.

Curtis Travis
Quest
Knoxville, Tennessee

Azamet K. Tynybekov
International Science Center
Bishkek, Kyrghyzy Republic

Kalliat T. Valsaraj
Gordon A. and Mary Cain Department of Chemical Engineering
Louisiana State University
Baton Rouge, Louisiana

Gilman Vieth
International QSAR Foundation for Reducing Animal Testing
Duluth, Minnesota

Tona G. Verburg
University of Technology
Interfaculty Reactor Institute
Department of Radiochemistry
Delft, The Netherlands

Marc-André Villard
Department de biologie
Universite de Moncton
Moncton, Nouveau-Brunswick, Canada

John William Watkins
Center for Ecology and Hydrology
Natural Environmental Research Council
Cumbria, United Kingdom

Chris Whipple
Environ Corp.
Emeryville, California

Gregory J. White
Ecological and Cultural Resources
Idaho National Engineering and Environmental Laboratory
Idaho Falls, Idaho

James D. Wickham
U.S. Environmental Protection Agency
National Health Research Laboratory
Las Vegas, Nevada

G. Bruce Wiersma
College of Natural Sciences, Forestry and Agriculture
University of Maine Orono, Maine
Director, Maine Agricultural and Forest Experiment Station
Orono, Maine

Hubert Th. Wolterbeek
University of Technology
Interfaculty Reactor Institute
Department of Radiochemistry
Delft, The Netherlands

Søren Wulff
Department of Forest Resource Management and Geomatics
Swedish University of Agricultural Sciences
Umeå, Sweden

Table of Contents

1 Conceptual Basis of Environmental Monitoring Systems: A Geospatial Perspective

D.A. Bruns and G.B. Wiersma

CONTENTS

1-56670-641-6/04/$0.00+$1.50

1.1 INTRODUCTION

The importance of and need for integrated environmental monitoring systems is well established. The U.S. National Science Foundation's (NSF) Long-Term Ecological Research (LTER) Program has 18 sites in the U.S., each with a study area which generally collects long-term descriptive measurements of air, water, soil, and biota, including data on forest or grassland stands, population and community inventories, and watershed–stream channel characteristics and habitats (e.g., Franklin et al.[1]). Originally, these observational data were intended to serve as the environmental context for basic ecosystem research conducted on an experimental basis to address the pattern and control of primary production and organic matter accumulation, nutrient cycling, population dynamics, and the pattern and frequency of site disturbance.

In a somewhat similar fashion, but focused on atmospheric pollutants,[2] the U.S. National Acid Precipitation Assessment Program (NAPAP) established a national network for long-term monitoring of wet and dry deposition of sulfates, nitrates, and "acid rain." In addition, NAPAP-sponsored ecological surveys (e.g., fish, invertebrates, forest conditions, and stream and lake water chemistry) were often collected for critically "sensitive" regions and ecosystems but on a much more geographically limited scope than for atmospheric components. Another more recent program for integrated environmental monitoring, building in part on past and ongoing LTER- and NAPAP-related activities, is the NSF's currently proposed National Ecological Observatory Network (NEON; see www.nsf.gov/bio/neon/start.htm).

The U.S. EPA also maintains an integrated monitoring network with a research agenda focused on developing tools to monitor and assess the status and trends of national ecological resources. This program, known as the Environmental Monitoring and Assessment Program (EMAP), encompasses a comprehensive scope of ecosystems (forests, streams, lakes, arid lands, etc.[3]) and spatial scales (from local populations of plants and animals to watersheds and landscapes[4]). EMAP usually holds an annual technical symposium on ecological research on environmental monitoring. For example, in 1997, EMAP addressed "Monitoring Ecological Condition at Regional Scales" and published the symposium proceedings in Volume 51 (Numbers 1 and 2, 1998) of the international journal *Environmental Monitoring and Assessment*.

The broadest and perhaps most compelling need for better and more integrated design principles for monitoring is based on the numerous and complex problems associated with global environmental change. This includes worldwide concern with climate change,[5,6] loss of biotic diversity,[7] nutrient (especially nitrogen via atmospheric deposition) enrichment to natural ecosystems,[8] and the rapid pace and impact of land-use change on a global basis.[9,10]

The necessity for a comprehensive global monitoring system was recognized in early publications of the International Geosphere–Biosphere Program (IGBP) and in later global change program proposals and overviews.[11–14] In particular, "geo-biosphere observatories" were proposed for representative biomes worldwide and would be the focus of coordinated physical, chemical and biological monitoring.[12,15,16] Bruns et al.[17–19] reviewed the concept of "biosphere observatories" and evaluated various aspects of monitoring programs for remote wilderness ecosystems and a geospatial watershed site for a designated American Heritage River in the context of global environmental change. These sites represent a broad spectrum of ecological conditions originally identified in the IGBP. Remote sites, especially at higher elevations, may be very sensitive to global factors like climate change, while the Heritage River watershed site is heavily impacted by regional scale "industrial metabolism."[12] The latter may provide an important "test-bed" for evaluation of geospatial technologies (see text below and References 19 and 20) and related spatial scales of land use change that might be applied later to more remote monitoring sites as part of long-term networks.

A conceptual basis for the design of integrated monitoring systems and associated networks has received growing attention in the last two decades as part of scientific research to address these environmental problems from a local to global perspective. Early and ongoing efforts include those of Wiersma, Bruns, and colleagues[17–27]—most of whom focused on conceptual design issues or monitoring approaches employed and exemplified at specific sites. Others have conducted similar work in relation to global environmental monitoring and research programs.[14,28,29] In addition, two major reports[30,31] sponsored by the National Research Council (NRC) cover a broad range of environmental monitoring issues, including consideration of comprehensive design principles. The former deals with marine monitoring and the latter report is focused on case studies to address the challenge of combining diverse, multimedia environmental data; this latter report reviewed aspects of the LTER program (at the H.J. Andrews Experimental Forest site), the NAPAP (Aquatic Processes and Effects), the Department of Energy's (DOE) CO_2 Program, and the first International Satellite Land Surface Climatology Project (ISLSCP) among others.

In this context of national and international global change programs, and the range of complex environmental problems from a global perspective, our objective in this chapter is to delineate and develop basic components of a conceptual approach to designing integrated environmental monitoring systems. First, general concepts from the National Research Council reports are reviewed to illustrate a broad perspective on monitoring design. Second, we highlight aspects of our previous and ongoing research on environmental monitoring and assessment with a particular focus on six components in the design of a systems approach to environmental monitoring. These are more specific but have evolved in the context of general ideas that emerge from the NRC reports. In particular, we use examples from our remote (wilderness) site research in Wyoming and Chile contrasted with an ongoing GIS watershed assessment of an American Heritage River in northeastern Pennsylvania. These examples are intended to facilitate illustration of design concepts and data fusion methods as exemplified in the NRC reports.[30,31] Third, we provide a general synthesis and overview of current general ideas and future directions and issues in environmental monitoring design. Finally, we wish to acknowledge the varied agencies and sponsors (see end of chapter) of our past and ongoing environmental research projects on which these conceptual design components are based.

1.2 GENERAL MONITORING DESIGN CONCEPTS FROM NRC REPORTS

The design of an integrated environmental monitoring strategy starts with identifying resources as risk in order to initiate development of a conceptual model.[30] This process of strategic planning is an iterative process whereby the model may be refined, elaborated, or enhanced based on practical and technical considerations, available resources, and defined monitoring objectives. This broad strategic approach (Figure 1.1a) usually will culminate in the development of testable questions that feed into the specifics of a detailed sampling and measurement design with a focus on parameter selection, quantifying data variability, and setting up a sampling scheme. This is also an iterative process (Figure 1.1b) with feedback to reframe questions and refine technical components of monitoring design. Data quality and statistical models for analyses also are identified as key components of this strategy.

Boesch et al. provide important insight into the use of conceptual models in monitoring design[30] and indicate that the term is sometimes misunderstood. A "conceptual model" typically begins as a qualitative description of causal links in the system, based on best available technical knowledge. Such a model may refer to descriptions of causes and effects that define how environmental changes may occur. For example, in monitoring toxic effects of point sources of pollutants, a conceptual model would identify critical sources of contamination inputs to the ecosystem and define which ecological receptors or endpoints (e.g., a particular species, a physical ecosystem compartment, or a target organ system) are likely to be impacted, modified, or changed. As a monitoring system is better defined, a more quantitative model or a suite of models based on different approaches (e.g., kinetic vs. numerical vs. statistical) may be used effectively to address complementary aspects of monitoring objectives.

Defining boundaries, addressing predictions and uncertainty, and evaluating the degree of natural variability are also broad concerns in the development of a monitoring strategy and sampling design.[30] For example, in monitoring pollutant impacts to streams and rivers, watershed boundaries may need to be established since upstream sources of contamination may be transported downstream during storm events, which may add uncertainty in the timing and movement of materials within the natural seasonal or annual patterns in the hydrologic cycle. For these reasons, a monitoring program should be flexible and maintain a continuous process of evaluating and refining the sampling scheme on an iterative basis.

Both NRC reports[30,31] highlight the need to address issues of spatial and temporal scales. Most monitoring parameters will vary on space and time scales, and no one set of boundaries will be adequate for all parameters. Also, it is expected that events that occur over large areas will most likely happen over long time periods, and both will contribute to natural variability in monitoring parameters—a condition confounding data interpretation.[30] Wiersma et al. identify spatial and temporal scales as one of the most apparent barriers to effective integration and analysis of monitoring data.[31] For example, geophysical and ecological processes may vary at different scales, and both can be examined from a variety of scales. No simple solution to scale effects has yet to emerge for monitoring design although a hierarchical approach to ecosystems and the use of appropriate information technologies like geographic information systems (GIS) and satellite remote sensing appear to be making progress on these issues.[31–33] Rosswall et al. and Quattorchi and Goodchild cover various ecological scaling issues for terrestrial ecosystems and biomes,[34,35] and Boesch et al. summarize a range of space and time scales[30] for various marine

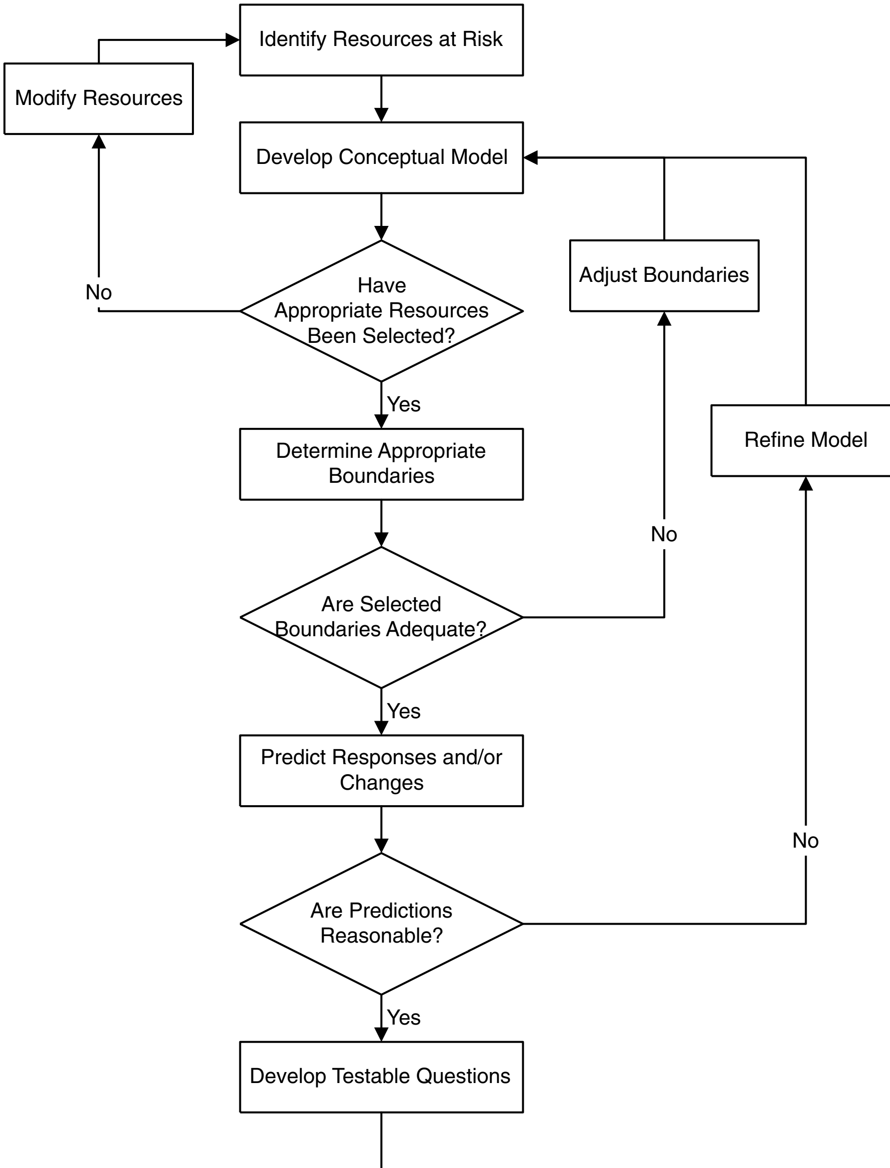


FIGURE 1.1A Designing and implementing monitoring programs: iterative flow diagram for defining a monitoring study strategy. (From Boesch, D.F. et al., *Managing Troubled Waters: The Role of Marine Environmental Monitoring,* National Academies Press, Washington, D.C., 1990. Reprinted with permission from the National Academy of Sciences. Courtesy of the National Academies Press, Washington, D.C.)

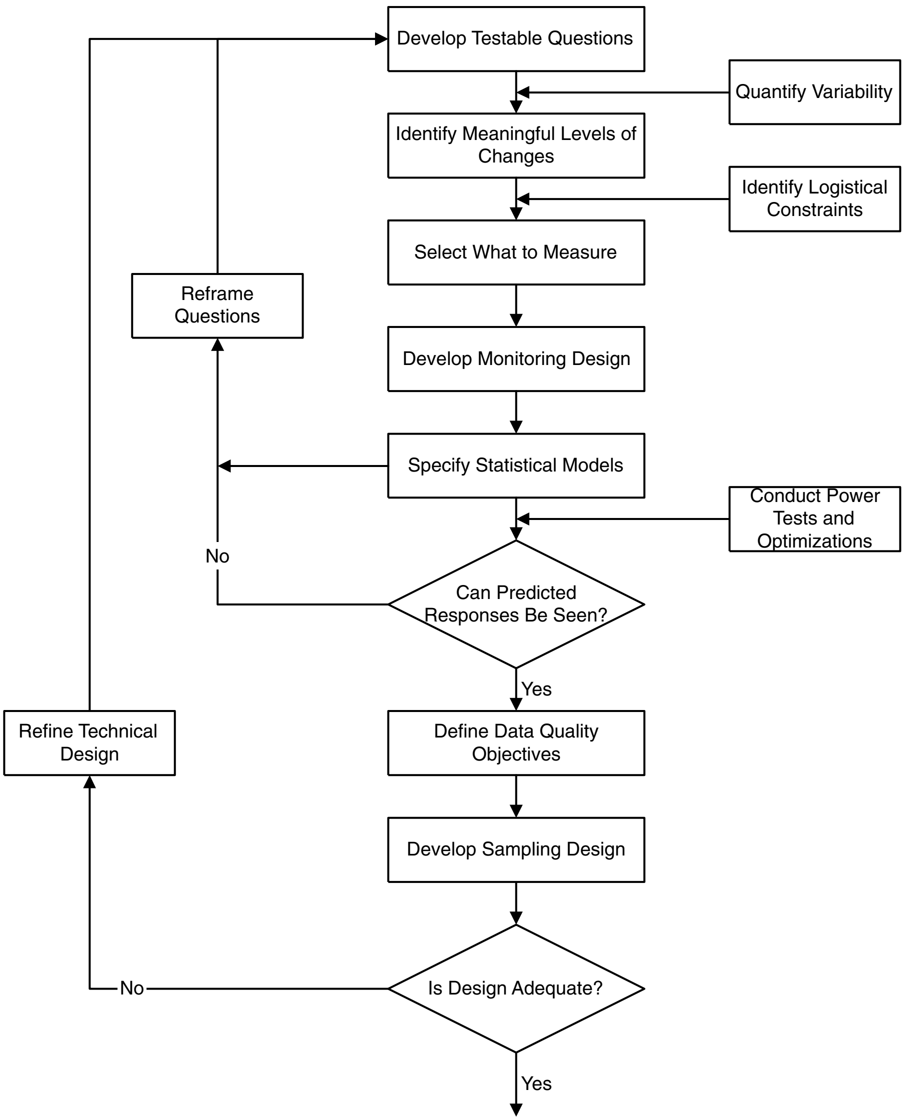


FIGURE 1.1B Designing and implementing monitoring programs: iterative flow diagram for developing an environmental measurement design. (From Boesch, D.F. et al., *Managing Troubled Waters: The Role of Marine Environmental Monitoring,* National Academies Press, Washington, D.C., 1990. Reprinted with permission from the National Academy of Sciences. Courtesy of the National Academies Press, Washington, D.C.)

impacts ranging from power plants, outfalls, and marinas to fishing, dredging, and natural events like storms and El Niño events.

Another general but key aspect in the overall planning for monitoring design relates to data quality assurance.[30,31] Boesch et al. highlight two aspects of quality assurance in the design of monitoring programs:[30] quality control (QC) and quality assurance (QA). QC might be viewed as strategic in nature since it is intended to "ensure that the data collected are of adequate quality given study objectives and the specific hypotheses to be tested."[30] QA is somewhat more "tactical" and deals with the everyday aspects of documenting sample analysis quality by repetitive measurements, internal test samples, use of standards and reference materials, and audits; specifically, sample accuracy and precision needs to be assessed, applied to data analysis and interpretation, and documented for reference. Standard Methods for the Examination of Water and Wastewater Analysis[36] is a well-known reference source for QA and QC procedures in microbiological and chemical laboratory analyses. In addition, QA/QC concepts and procedures are well addressed and documented in Keith[37] for a variety of multimedia environmental sampling methods. And finally, metadata (i.e., data about data) has emerged as a QC/QA component to monitoring programs during the last decade, given the emergence of relational databases and GIS for regular applications in environmental monitoring and assessment.[31] Later chapters in this book deal with detailed aspects of QA/QC and metadata and related data management tools are briefly addressed subsequently in this chapter under Data Integration.

1.3 OVERVIEW OF SPECIFIC CONCEPTUAL MONITORING DESIGN COMPONENTS

Conceptual components of our approach to environmental monitoring design (and application) have been detailed in papers by Wiersma and colleagues.[21,23,27] These components at that time included (1) application of a conceptual framework as a heuristic tool, (2) evaluation of source-receptor relationships, (3) multimedia sampling of air, water, soil, and biota as key component pathways through environmental systems, and (4) use of key ecosystem indicators to detect anthropogenic impacts and influences. This conceptual approach was intended to help identify critical environmental compartments (e.g., air, water, soil) of primary concern, to delineate potential pollutant pathways, and to focus on key ecosystem receptors sensitive to general or specific contaminant or anthropogenic affects. Also implicit in this monitoring design is a watershed or drainage basin perspective[17,18,38] that emphasizes close coupling of terrestrial–aquatic linkages within ecosystems.

Figure 1.2 summarizes these overall components of our approach,[27] especially at our remote monitoring sites in Chile, Wyoming, and the Arctic Circle (Noatak site). Remote sites were utilized for baseline monitoring and testing of design criteria and parameters. These sites were less impacted by local or regional sources of pollution or land use change and were expected to be more indicative of baseline conditions (in the context of natural variation and cycles) that might best serve as an "early warning" signal of background global environmental change.[18] In addition, field logistics were pronounced and rigorous at these remote sites for any type of permanent or portable monitoring devices and instrumentation. These conditions served as a good test of the practical limits and expectations of our monitoring design components.

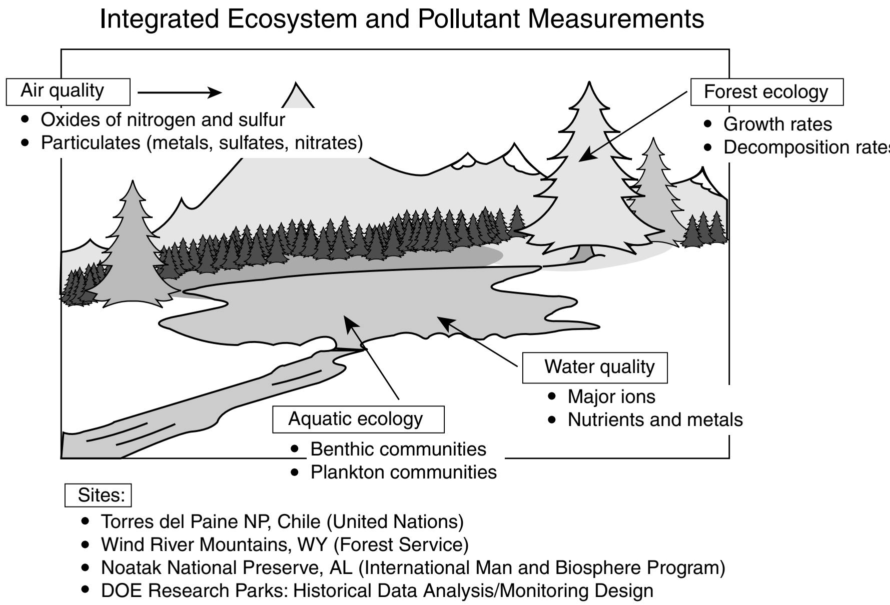


FIGURE 1.2 Conceptual design for global baseline monitoring of remote, wilderness ecosystems. (From Bruns, D.A., Wiersma, G.B., and Rykiel, E.J., Jr., Ecosystem monitoring at global baseline sites, *Environ. Monit. Assess.*, 17, 3, 1991. With permission from Kluwer Academic Publishers.)

Our overall monitoring design concept (Figure 1.2) also served as a basis for evaluating historical monitoring data from seven DOE National Environmental Research Parks. This conceptual assessment highlighted the need and opportunity inherent in geospatial technologies and data like Geographic Information Systems (GIS), satellite remote sensing imagery (RS), and digital aerial photography. In conjunction with the report by Wiersma et al.,[31] this DOE monitoring design assessment[27] facilitated start up of the GIS watershed research program and GIS Center in the GeoEnvironmental Sciences and Engineering Department at Wilkes University.[20] In addition, this general conceptual monitoring approach was used for: a regional land use plan for 16,000 acres of abandoned mine lands,[19,20,39] a successful community-based proposal to designate a regional watershed as an American Heritage River (AHR, see www.epa.gov/rivers/98rivers/), a National Spatial Data Infrastructure Community Demonstration Project (www.fgdc.gov/nsdi) and recipient of a U.S. government Vice Presidential "Hammer Award" (www.pagis.org/CurrentWatershedHammer.htm), and a GIS Environmental Master Plan for the Upper Susquehanna–Lackawanna River.[40]

Figure 1.3 provides additional overall conceptualization of our monitoring design for remote wilderness ecosystem study sites. This heuristic tool[22,41] highlights the atmospheric pathway for anthropogenic impacts to remote ecosystems and indicates the multimedia nature of our monitoring efforts based on field tested protocols evaluated in our remote site research program.[25,27] Details of this conceptual component of our monitoring design are provided below.

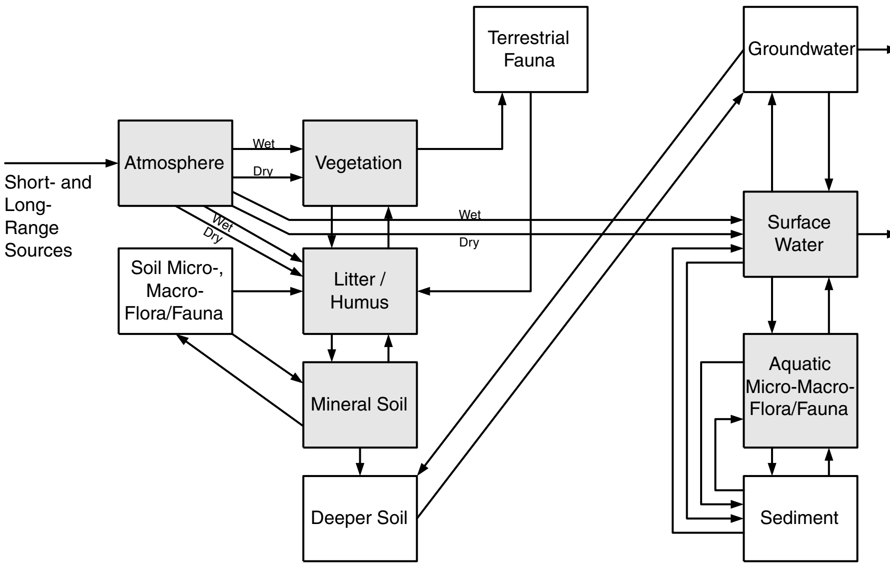


FIGURE 1.3 Systems diagram and heuristic tool for conceptualization of monitoring design for sensitive wilderness ecosystems. (From Bruns, D.A., Wiersma, G.B., and Rykiel, E.J., Jr., Ecosystem monitoring at global baseline sites, *Environ. Monit. Assess.*, 17, 3, 1991. With permission from Kluwer Academic Publishers.)

The NRC report by Wiersma et al.[31] significantly broadened our general conceptual approach to integrated environmental monitoring systems. This report's focus on combining diverse (including multimedia) environmental data sets and the extensive geographic spatial extent of two of the case studies (the ISLSCP example noted above, and use of remote sensing for drought early warning in the Sahel region of Africa) resulted in adding two additional components[19] to our design concepts: data integration with geospatial tools like GIS and remote sensing, and a landscape spatial scaling component, based in part again on GIS, but especially in the context of a hierarchical approach to ecosystems.[42]

1.4 CONCEPTUAL MONITORING DESIGN COMPONENTS

We have tested and evaluated different aspects of our monitoring design concepts depending on a range of criteria, including study site location and proximity, degree of local and regional pollutant sources and land use perturbations, funding agency and mission, duration and scope of the study (funding limitations), and issues of degree of spatial and temporal scaling. Our work at the Wyoming and Chile sites has been profiled and described in several contexts: global baseline monitoring,[25,27] freshwater ecosystems and global warming,[18] and testing and evaluation of agency (U.S. Forest Service) wilderness monitoring protocols for energy development assessment.[17,19] These are both remote, wilderness monitoring sites with the Torres del Paine Biosphere Reserve in southern Chile being one of the "cleanest" (and least disturbed globally), remote study areas from an atmospheric pathway; in contrast, the Wyoming site was "downwind" of significant ongoing and potential atmospheric emissions (oxides of sulfur and nitrogen) from regional energy development.

Numerous multimedia parameters were measured and evaluated at the Wyoming site, especially from a sampling protocol perspective. Details of this work and descriptions of the study site are provided in Bruns et al.[17,19] who focus on five evaluation criteria for monitoring design and implementation: ecosystem conceptual basis, data variability, uncertainty, usability, and cost-effectiveness. This study site gives the best perspective and detail on a wide range of monitoring parameters in our remote site work and serves as one of three (there is another remote site in southern Chile; for details on site conditions, see References 18, 25, and 27) examples of our conceptual approach to monitoring design.

The third study site for a basis of contrast and comparison to our remote sites is in northeastern Pennsylvania. This represents a 2000-square-mile portion of a watershed designated in 1998 by President Clinton as one of 14 American Heritage Rivers. A GIS watershed approach was employed for research in monitoring and assessment with geospatial tools to address environmental impacts from urban stormwater runoff, combined sewer overflows, acid mine drainage, impacts from abandoned mining lands, and regional suburbanization and land use change. Cleanup and reclamation costs for mining alone approach $2 billion, based on Congressional hearings in 2000.[40] As noted above, this site provides more perspective on geospatial tools and scaling issues vs. our earlier monitoring work at remote sites.

1.4.1 Conceptual Framework as Heuristic Tool

This component is generally considered as the starting point in monitoring design. It is not intended as a static or stand-alone element in the monitoring program. As a simple "box-and-arrow" diagram, it serves as an interdisciplinary approach to examine and identify key aspects of the monitoring program being designed. Basically, principal investigators and their technical teams, along with responsible program managers and agencies, often across disciplines and/or institutions, can take this simple approach to focus discussion and design on answering key questions: Is the study area of concern being potentially impacted by air or water pollution sources? What are the relative contributions of point vs. nonpoint sources of water pollution? What are the primary pollutant pathways and critical ecosystem components at risk? How are critical linkages between ecosystem components addressed and measured? What is the relative importance of general impacts like land use change vs. media specific impacts like air, water, or subsurface (e.g., landfills) contamination sources?

Figure 1.2 and Figure 1.3 as noted above illustrate the atmospheric route as primary disturbance and pollutant pathway to remote ecosystems such as those at our Wyoming and Chile study sites. In these cases, "wilderness" or "national park" status prevent immediate land use perturbations but atmospheric pollutants, either as a potential global background signal (e.g., particulates associated with "arctic haze") or from regional point sources like coal-fired power plants,[27] might be transported long distances and may affect remote ecosystems via wet and dry deposition processes.[43,44]

As expected at remote monitoring sites, field logistics and/or regulatory restrictions on available sources of electricity or weather protection may prohibit a number of instrument approaches to monitoring methods and techniques. Figure 1.3 shows shaded components (soil, water, vegetation, and aquatic community) of the monitoring program that are easily measured with simple field sampling devices and procedures. Forest survey methods, soil sampling trowels, and aquatic kick nets allow for rapid field assessments and sampling as field restrictions or time limits may dictate. We have also used this approach successfully in even more remote sites like the Noatak Biosphere Reserve in the Arctic Circle of Alaska[18,27] and in a mountainous "cloud forest" ecosystem of Fan Jing Shan Biosphere Reserve in south central China.[19] At the Wyoming remote monitoring site, metals in vegetation (terrestrial mosses), aquatic macroinvertebrates, and stream (water chemistry) alkalinity all scored highest across our five evaluation criteria noted above.[19]

Figure 1.4 shows a similar "systems diagram" developed for the northeastern Pennsylvania study site with a major focus on regional mining impacts. The eastern anthracite (coal) fields of Pennsylvania cover a general area of about 3600 mi^2, with about 2000 mi^2 directly within the Susquehanna–Lackawanna (US-L) watershed study area.[40] The watershed covers about an 11-county area with over 190 local forms of government or agencies and supports a regional population base of over 500,000 people. Due to the broad spatial extent of these impacts and complex set

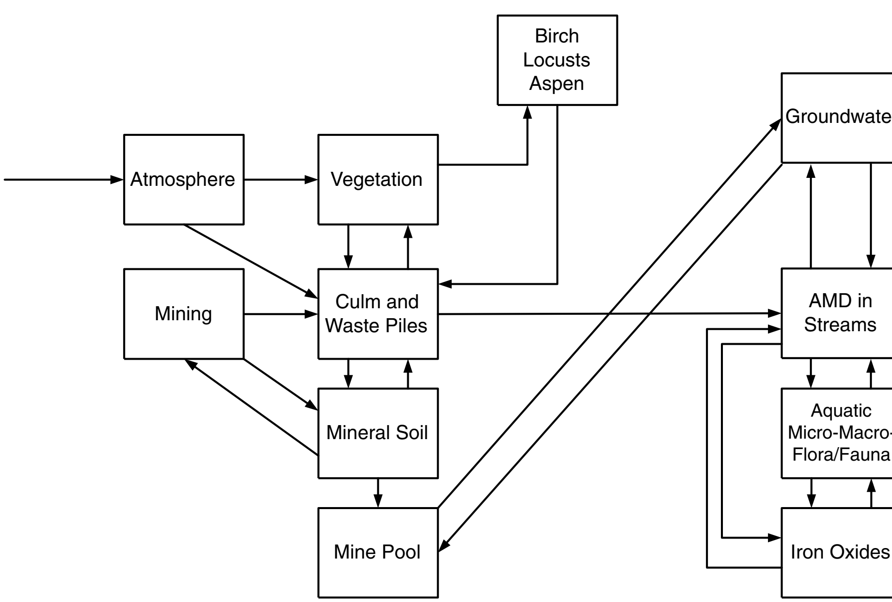


FIGURE 1.4 A GIS watershed systems approach to monitoring and assessment of regional mining impacts in Northeastern Pennsylvania. (From Bruns, D.A., Sweet, T., and Toothill, B., *Upper Susquehanna–Lackawanna River Watershed, Section 206, Ecosystem Restoration Report*, *Phase I GIS Environmental Master Plan*, Final Report to U.S. Army Corps of Engineers, Baltimore District, MD, 2001.

of environmental conditions, we have employed a major geospatial (GIS based) technological approach to monitoring and assessment at this site.[40,45] However, even here, the basic box-and-arrow diagram served a number of useful applications.

First, we assembled an interdisciplinary team of almost 20 members from various institutions and state and federal agencies. Hydrologists, geochemists, river ecologists, GIS technicians, plant ecologists, soil scientists, and engineers were represented for a one-day workshop on which these concepts and components were proposed, discussed, evaluated and agreed upon as a GIS watershed approach to regional monitoring and assessment.

Second, the general elements implicit in this conceptual framework allowed for scaling from local, site-specific and stream-reach applications (e.g., well-suited to local watershed groups) to broader watershed and landscape spatial scales (e.g., see the U.S. Environmental Protection Agency's (EPA) GIS Mid-Atlantic Integrated Assessment over a five-state region[4]). Our watershed monitoring research with federal sponsorship (e.g., EPA and U.S. Department of Agriculture [USDA]) facilitated our use of GIS, RS, aerial photography, and the Global Positioning System (GPS)—all of which are not generally available to local watershed groups or local branches of relevant agencies. Therefore, we avoided duplication of field measurements at a local level and instead focused on a watershed (and sub-catchment) approach with GIS. We were able to coordinate with local groups in public meetings and technical approaches due to a common conceptual design of the environmental system.

Third, a systems diagram of this nature also facilitated data analyses among key components, the pollutant sources, and the affected elements of the watershed and landscape. For example, the diagram in Figure 1.4 was used in setting out our statistical approach for prioritizing watershed indicators of potential use and identifying stream monitoring parameters for ranking of damaged ecosystems.[40] This also allowed us to incorporate land use and land cover databases derived from satellite imagery and integrate it with point samples of water (chemistry) quality and stream community biodiversity via statistical analysis (Figure 1.5).

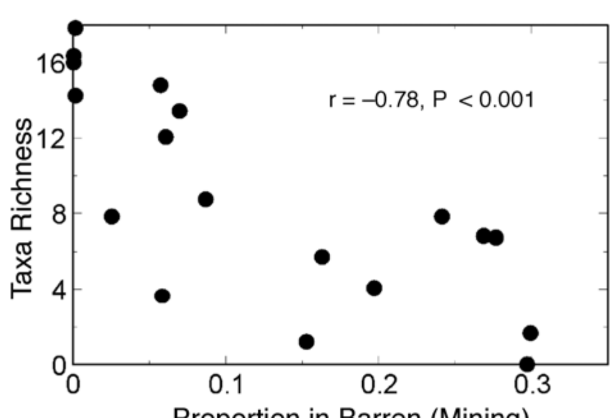


FIGURE 1.5 Statistical analysis of stream biodiversity vs. watershed area in mining land use.

And fourth, both our EPA- and USDA-sponsored GIS research projects maintained a public outreach and environmental education component. The basic systems diagram shown in Figure 1.4 successfully enhanced our educational component in this regard, both with other technical participants, and, especially, with the public. These outreach and education activities involved various public meetings for discussion, input, and coordination, in addition to posting of information, data, GIS analysis, and environmental concepts at our Website (www.pagis.org) for GIS research on the US-L watershed. Also, we participated in environmental education activities as part of public community riverfront park activities and high school student visits to campus for briefings on GIS, watershed analyses, and geospatial data applications.

1.4.2 Evaluation of Source–Receptor Relationships

This element in the conceptual design of a monitoring program is also implicit in the diagrams of Figure 1.2 to Figure 1.4. However, this component also mandates an interdisciplinary approach to environmental monitoring since soil scientists, forest ecologists, and stream ecologists may not have typical expertise in various techniques of water, air, and soil pollution monitoring. Likewise, environmental engineers involved with the design of air and water pollution control and monitoring technologies may lack the needed ecological expertise for identifying and measuring the response of critically sensitive ecosystem receptors or endpoints. Our work at remote sites in Wyoming and Chile benefited from a technical team approach since our research was sponsored during our employment with a DOE national laboratory where the necessary interdisciplinary mix of expertise was readily available to support work on remote site monitoring (e.g., see References 25 and 46.) For example, we had ready access to various scientists and staff with expertise in soils, forestry, river ecology, geology, air pollution, analytical chemistry, and general environmental science and engineering through various technical programs and organizations at the Idaho National Engineering Laboratory.

For the Pennsylvania GIS watershed project, similar concerns and issues were addressed. In this case, expertise in mining, engineering, geochemistry, hydrology, GIS, and stream ecology was derived through public and state and federal agency outreach during the public sector portion of the long-term project. However, the ultimate selection of key pollutant sources and critical ecosystem receptors needs to be well focused, since both remote site and watershed approaches often end with long lists of parameters for potential implementation in a monitoring program. In this situation, logistics, financial resources, and funding limitations require a subset of measurements that will allow assessment of the more important relationships. In some situations, peer-review by an outside panel,[47] case study reports,[31] or actual testing and evaluation of a range of parameters[17,19] may help to resolve differences or professional preferences and result in a more cost-effective but focused set of monitoring parameters. More examples are discussed below for the other design components.

Finally, it should be noted that data provided in Figure 1.5 represent one example of the successful selection of a key source of pollution with a sensitive ecological endpoint. In this case, the extent of mining disturbed lands within a watershed (or subwatersheds or "sampling" catchments) represented a major source of pollutant (and land use) impact—but derived from satellite imagery and ground-truthed with GPS.[45] Disturbed mining lands are devoid of natural vegetation and soil horizons and are susceptible to extreme amounts of sediment loading to streams and rivers where aquatic habitats are destroyed due to sedimentation processes (see literature review[40]). In addition, atmospheric exposure of pyritic mining wastes can generate a considerable amount of acid mine drainage to streams via the hydrologic cycle, so additional geochemical impacts are evident due to the eco-toxicity of high acidity and mobilized heavy metals like Cu and Zn from waste materials. Part of our research in this GIS watershed study was to determine what water chemistry and stream biotic variables would be best associated with regional mining impacts. The biodiversity of stream macroinvertebrates[40] was one of the better indicators in this regard as shown in Figure 1.5. As noted previously, stream macroinvertebrate parameters scored well over the five monitoring evaluation criteria employed for the Wyoming remote study site.[17,19]

1.4.3 Multimedia Monitoring

The rationale for monitoring various environmental media encompassing air, water, soil, and biota is based on several factors. First, the physical and chemical properties of pollutants demonstrate a wide range of fate and transport mechanisms with different pathways and effects upon ecological receptors. This is supported both by multimedia modeling approaches[48] and general estimation methods in ecotoxicology and environmental chemistry.[49,50] Second, focused population and community studies on the fate of metals and organic contaminants relative to bioaccumulation and trophic-food web transfer pathways[49,50] also indicate a need to approach monitoring design from a multimedia perspective. And third, larger-scale watershed and regional landscape investigations of particular pollutants like acid rain effects on freshwater ecosystems[51] and air pollution impacts to forests[52] should reinforce this design component if resource managers are to understand the fate and effects of pollutants in a holistic ecosystem framework.

Methods of sampling and analysis on a multimedia basis are well established[53] and detailed elsewhere in this volume. Our research on multimedia monitoring design has emphasized the testing and evaluation of methods for use in remote, wilderness ecosystems.[25,27] Table 1.1 lists monitoring parameters of various physical, chemical, and biological characteristics of a high-elevation ecosystem in Wyoming from a multimedia standpoint, and includes a cataloging of appropriate methods for use under potentially harsh field conditions. As indicated above, we have developed evaluation criteria for assessing the overall utility of these methods and the reader is referred to other reports and publications for more detailed consideration.[17,19,46]

1.4.4 Ecosystem Endpoints

The search for key ecosystem parameters for environmental monitoring and assessment has received considerable attention over the past two decades. Earlier studies, more aligned with environmental toxicology research or assessment of sewage pollution in streams, focused on population inventories or surveys of "indicator" species. Indicator

TABLE 1.1
Integrated Multimedia Monitoring Parameters at the Wind Rivers Study Site

Measurement	Method (Previously Evaluated)
Abiotic Measurements	
SO_4, NO_3, HNO_3, NH_3, NO_2, SO_2 (atm)	Transition flow reactor (filter pack) — EPA[54]
Source-term particle analysis	Scanning electron microscopy with energy dispersive x-ray analysis
Ozone	UV Photometry
Meteorological parameters	Standard sensors; plus dry depostion methods of Bruce Hicks/Oak Ridge National Laboratory
Trace metals (atm)	Low and high volume sampling
Trace metals (in water, litter, soil, vegetation)	Ecological sampling at study site[25,55]
Trace metals in snow	Snow cores before runoff (later analysis with Standard Methods[36])
Soil (organic matter, exchange bases, and acidity, pH, extractable sulfate)	U.S. Forest Response Program[47,56]
NO_3, PO_4, SO_4 (water)	National Surface Water Survey[57]
Lake/streams water chemistry (cations and anions)	National Surface Water Survey[57]
Biotic Measurements	
Lake chlorophyll *a*, zooplankton, benthic algae, fishes, benthic macroinvertebrates	U.S. Forest Service Wilderness Guidelines[47] and Standard Methods[36]
Stream ecosystem analysis (macroinvertebrate functional feeding groups, periphyton, decomposition, benthic organic matter)	River Continuum Concept[38,58–60]
Terrestrial (forest) ecosystem ˇ (productivity, needle retention, needle populations, litter decomposition, litterfall, foliage elemental composition, community structure)	Dr. Jerry Franklin; U.S. Department of Agriculture, Forest Service methods[61]

Source: From Bruns, D.A., Wiersma, G.B., and Rykiel, E.J., Jr., Ecosystem monitoring at global baseline sites, *Environ. Monit. Assess.*, 17, 3, 1991. With permission from Kluwer Academic Publishers.

species may include those that are either tolerant (e.g., tubificids or "bloodworms" thrive at low levels of oxygen due to organic loading of aquatic systems) or intolerant (e.g., various species of mayflies and stoneflies that require more pristine conditions of stream habitat and associated chemical constituencies) of pollutant concentrations or habitat disruptions.[50] However, Cairns[62] has cautioned against reliance on single indicator species since their known response is often in relation to very particular kinds of pollutants and may not warrant objective assessment of general or varied contaminant impacts. In this context, Schindler[63] has suggested that some individual species, like the crustacean *Mysis relicta*, may represent unique keystone species within aquatic food webs (i.e., occupying specialized niches) and are susceptible to a variety of stresses. In this case, a monitoring program would be more effective with the inclusion of such a species, assuming resource managers have access to previous knowledge and available data for decision support in the design process.

At present, ecological measurement and assessment methods encompass a hierarchical framework to ecosystem management. This seems due to the maturation of ecotoxicology as a science[64] along with further developments of environmental monitoring principles. For example, ecotoxicology texts[50] generally cover pollutant effects on individual organisms, populations, and communities, and ecosystem structure and function, and in some cases use an explicit hierarchical treatment to basic and applied concepts. In addition, resource managers are now responsible for and concerned about whole ecosystems—and the monitoring programs need to support these assessment objectives. Thus, scientists are being called upon to address multiple stresses on ecosystems.[65] In this context, a hierarchical approach represents the best available conceptual framework for dealing with complexities of both ecosystems and associated impacts of pollutants and physical disturbance to environmental systems such as changing land use.

Allen and Starr[66] and O'Neill et al.[67] originally developed the basic concepts of a hierarchical approach to understanding ecosystems. In its basic form, different levels of biological organization were recognized in a hierarchical fashion, with increasing degrees of ecological complexity. This hierarchy for either aquatic or terrestrial systems, from lowest to highest, included the following levels of biological organization: individual organisms (e.g., plants or wildlife), populations, communities, and ecosystems. The widespread use and availability of geospatial tools and data, like geographic information systems (GIS) and satellite remote sensing imagery, has facilitated further development of the hierarchical ecosystem concept that extends the scope of watershed and landscape spatial (and temporal) scales.[19,31,33] However, these aspects are covered below in the next two components in the conceptual design of monitoring systems.

A review of the ecological assessment literature (Table 1.2) indicated the use of a range of parameters across these various levels of ecological complexity.[68] Such measurements encompass biomass for a population (e.g., trout), biodiversity (e.g., stream macroinvertebrates) as an indicator of community structure, and nutrient cycling (e.g., water chemistry) as an integrator of ecosystem function. In general, the most common parameters have included trophic relationships, species diversity, succession (temporal changes in composition), energy flow, and nutrient cycling. Some investigators[51,63] have indicated that functional responses of ecosystems may be more robust than structural changes due to "functional redundancy" and variation in pollutant sensitivity among species; for this reason, individual species and community level monitoring has been recommended for detecting ecological impacts.

Our approach to ecosystem monitoring[17,19,25,27] has been to include both structural and functional parameters for terrestrial and aquatic habitats and environments on a watershed basis (see Table 1.1 and Figure 1.2, Figure 1.3, and Figure 1.4 above). At present, few studies and monitoring programs have produced long-term data on both structure and function,[63] and more research is needed before definitive guidelines can be set. Also, we have observed extreme impacts from land use change, like regional mining and urban stormwater runoff; while structural changes in these cases are more easily measured in the early stages of impact, we expect that in these

TABLE 1.2
Ecological Parameters: Recommendations for Monitoring and Assessment of Baseline Conditions and Human Impacts

	Author(s)												
	69	70	71	72	73	74	75	67	76	77	78	79	80
Parameter													
Abundance (biomass)					+	+	+	+	+(a)	+			
Reproduction				+	+	+	+		+(a)				
Behavior				+	+	+							
Community Structure													
Trophic relationships		+	+	+	+	+	–	+	+(a)	+	+	+	+
Species diversity	+	+	+	+	+	+	+	+				+	
Succession/ change in composition	+	+	+		+		+	+	+(a)	+	+	+	
Size relationships	+	+											
Ecosystem Function													
Energy flow	+	+	+	+	+	+	–	+	–(a) +(t)	+	+		
Nutrient cycling		+	+	+	+	+	+	+(a)	–(a)	+	+		+
Decomposition/ respiration		+	+	+	+	+		+(a)	–(a)	+	+		+
Biomass/nutrient pools	+	+	+			+							+

Notes: (a) = aquatic ecosystem; (t) = terrestrial ecosystem; + = good potential for monitoring and assessment; and – = robust; not indicative of *early* impacts or stresses. Plus signs reflect generally positive view of a particular parameter as a key indicator of impact or at least its potential utility to detect anthropogenic perturbations. A negative sign means that an author has found this parameter to be a poor indicator of ecological impact; these parameters were found to be too robust and were not very sensitive to impacts.

Source: Bruns, D.A. et al., An ecosystem approach to ecological characterization in the NEPA process, in *Environmental Analysis: The NEPA Experience*, Hildebrand, S.G. and Cannon, J.B., Eds., Lewis Publishers, Boca Raton, FL, 1993, 103. With permission.

extreme cases, functional changes are needed to define the total system collapse that warrants immediate attention to ecosystem restoration and pollution mitigation. Also, our experiences in remote site monitoring concurrently for aquatic (streams and lakes) vs. terrestrial (forests) systems suggests that functional measures like forest productivity, litterfall, and leaf decay rates will better reflect short-term impacts of atmospheric than compositional changes, given the long life cycle of most tree species vs. short-lived aquatic species.

General reviews of the monitoring literature[63,68,71,74,76] indicate that a number of important ecological impacts can be measured and assessed only at the ecosystem level. In addition, measuring many different parameters is not necessarily the optimal strategy for designing and implementing a monitoring program. In most cases, a selected subset of parameters can be defined from a conceptual basis and principles as outlined above, viewed in conjunction with knowledge of the published literature.

1.4.5 Data Integration

This is one of the major challenges to implementing a well-designed monitoring program and cuts across all of the other components of our systems approach. Generally, this aspect of a monitoring program and its practical utility in the "real world" will be limited to the extent that these other components are ignored, relegated to a minor role, or inadequately developed or addressed. A conceptual model or framework with clearly identified sources of pollution, their pathways, and likely environmental endpoints provides the broad overview and context within which data sets will be processed, summarized, and evaluated (i.e., "data fusion"; see Wiersma et al.[31]). This framework will provide an initial set of testable hypotheses for trend analysis and the inference of potential effects on ecosystems from point pollution sources and/or more diffuse impacts from nonpoint sources that may include changing land use over larger environmental extents and spatial scales. Actually collecting multimedia data and measuring key sensitive ecosystem endpoints are needed if resource managers are to manage, protect, and sustain environmental systems in a holistic fashion.

The NRC report by Wiersma et al.[31] provides a comprehensive review of issues associated with fusing diverse sets of environmental data. The authors review a series of case studies that encompass predicting droughts in the Sahel, atmospheric deposition in the U.S., the U.S. CO_2 program, ISLSCP (noted above), and marine fisheries. Numerous recommendations are provided, based on practical problems encountered from specific case study programs. These include organizational, data characteristics, and technological impediments to data fusion efforts. In this chapter, we focus on the challenge of data integration with a selected view toward aspects of data characteristics and geospatial technologies identified in Wiersma et al.[31] Organizational challenges, like agency mission, infrastructure, and coordination, are equally important but beyond the technical scope of this publication. The reader is referred to the original NRC report for more detailed information and insight into organizational factors in environmental monitoring design.

Geospatial technologies like GIS are emerging as the major approach to data fusion efforts, ranging from "enterprise GIS" in the business world to the "geodatabase" model in environmental management systems (www.esri.com and see Reference 81). The NRC report recommended using GIS and related technologies, like the GPS and satellite remote sensing imagery, in environmental monitoring and management programs to facilitate data acquisition (at various scales, see text below), data processing and analysis, and data dissemination to resource managers, political leaders, and the public. Concurrent with the NRC panel proceedings and publication, an applied GIS watershed research program[19,20,45,82] was being planned and implemented for the US-L—nationally designated in 1998 as 1 of 14 American Heritage Rivers. Four geospatial technologies were incorporated into this evolving program based in part on recommendations of the NRC report and geospatial data inventories at DOE national laboratories.[27] These aspects are highlighted here to demonstrate one approach to data fusion efforts in the spirit of the NRC report.

Figure 1.6 showcases how GIS is used to organize and integrate diverse environmental data sets to help solve environmental problems associated with past and on-going practices in land use. In particular, there is a $2-billion land reclamation and ecosystem restoration problem from over 100 years of regional coal mining. In addition, this watershed of 2000 square miles covers 196 local governments where urban stormwater runoff and combined sewer overflows (CSOs) have resulted in a $200 to 400 million aquatic pollution cleanup issue.[19,20,40]

Sampling of stream and river sites was of high priority given the nature of these environmental impacts to aquatic chemistry, habitats, and ecological communities (Figure 1.4 and Figure 1.6). GPS was used to locate each site and delineate point source features (mine water outfalls or CSOs) of pollution. Although field sampling techniques were "low-tech" based on standard methods, all ecological data of this type were easily integrated into a relational data base as part of the GIS for the watershed. In addition, sampling sites integrated into the GIS allowed for delineation and digitization of sampling site subcatchments for data analysis and integration from a comparative watershed perspective. For example, Figure 1.6 also shows the utility of GIS in visualizing data on a comparative basis between two watersheds (GIS graphic charts in lower left of figure). The subwatershed in a rural setting with no mining had high-stream macroinvertebrate biodiversity (clean water species), low acidity, and land cover mostly in forests and grassland meadows and minimal development vs. a mining watershed with more than 30% of land cover as mining disturbed areas, and with only pollution-tolerant aquatic species and high acidity in surface water streams.

In our Heritage River study area and region, satellite imagery (the Mid-Resolution Land Characteristics or MRLC, e.g., see Reference 83) was processed for land cover to facilitate watershed characterization for relating land use practices and problems to ecological conditions along environmental gradients within the watersheds.[40] Figure 1.5 and Figure 1.6 demonstrate one approach we used to data fusion by statistically relating stream biodiversity measures to mining land use (SPOT imagery shown in middle inset of Figure 1.6) within 18 delineated subwatersheds: (1) we used GPS to identify and locate point sampling sites on stream segments, (2) we digitized subwatersheds above each sampling point with a GIS data set of elevation contour lines, and (3) we processed SPOT imagery for land cover and land use[84,85] and conducted extensive ground-truthing of classified imagery with GPS.[20,45]

Land use impacts to ecological systems are generally viewed to be as widespread and prevalent worldwide to warrant a higher risk to ecosystems than global warming.[8–10] Satellite imagery also allows for a range of landscape[4] and watershed indicators[33] to be calculated for environmental monitoring and assessment at a broader spatial scale (see next section). Vogelmann et al.[86] surveyed data users of Landsat Thematic Mapper data (known as the National Land Cover Data set, or NLCD) from the early 1990s and found 19 different categories of application including land

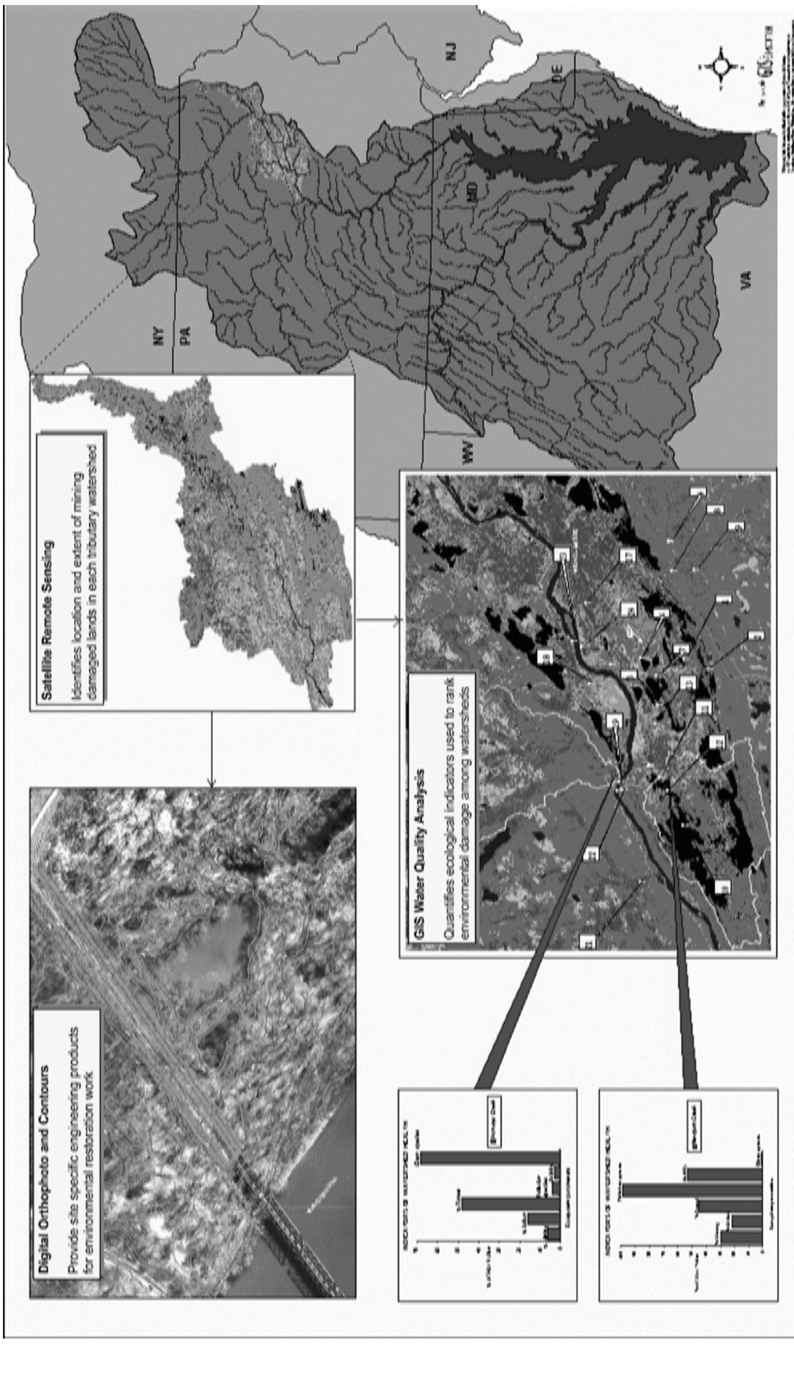


FIGURE 1.6 Integrated use of geospatial technologies for environmental monitoring on the heritage river watershed: GIS, GPS, remote sensing imagery, and orthoimagery.

cover change assessment, hydrologic-watershed modeling, environmental impact statements, water quality and runoff studies, and wildlife habitat assessments.

In addition to the integrated use of GIS, GPS, and remote sensing imagery in our PA Heritage River watershed research project (Figure 1.6), we have employed digital aerial photography as the fourth geospatial data source and technology.[20,40,84] Also known as orthoimagery, these data have been identified by the U.S. Federal Geographic Data Committee (FGDC)[87] as one the fundamental "framework" geodata sets for the National Spatial Data Infrastructure in the U.S. In this context, we surveyed 196 different local governments and regional state and federal agency offices within the 2000-square-mile watershed of the US–L River and found 10 of 11 counties lacking in local scale orthoimagery needed for tax assessment, land use and planning, emergency management, environmental cleanup, land and deeds records, ecological protection and monitoring, and floodplain management (see GIS watershed plan[40]). Falkner[88] provides an overview to methods and applications of aerial mapping from orthoimagery. Applications include mapping of geographically extensive wetlands,[89] cartographic support to management of state aquatic resources,[90] and floodplain management.[40]

A final element to data integration is the importance of QA and QC for the data sources themselves, along with metadata on all aspects of data development, processing and integration, and analysis.[31] Methods of multimedia field sampling and laboratory analysis (see references in Table 1.1) generally deal with adequate and established standard procedures of accuracy and precision (see also Reference 37 for general QA/QC issues). In contrast, geospatial metadata methods are still in various stages of development. GPS is generally accepted for most environmental applications in field mapping and is now commonly used for on-board aerial photography[88] and later aerotriangulation and accuracy calculations that require positional data as a replacement to conventional ground control surveys. In turn, either GPS[91] or accurate, georeferenced orthoimagery[83,86] may be used in accuracy assessments of remote sensed data classified for land use and land cover. Bruns and Yang[45] used GPS to conduct regional accuracy assessments on four such databases used in landscape–watershed analyses and reviewed general methods of accuracy assessment.[92,93]

1.4.6 Landscape and Watershed Spatial Scaling

The scope and extent of environmental contaminants in ecosystems, their potential for long-range transport through complex pathways, and their impact beyond simply local conditions, all dictate that environmental monitoring programs address pollution sources and effects from geographically extensive landscape and watershed perspectives. Although we have only recently added this final component to our conceptual design for monitoring systems,[19] there has been well over a decade of ecological research that serves as a foundation for successful inclusion of this element in monitoring programs. The success of this approach is supported from several standpoints.

As indicated above, landscape ecology has been well developed and investigated as part of a hierarchical perspective to ecosystem analysis.[42,66,67,94,95] Landscape parameters and indicators include dominance and diversity indices, shape metrics, fragmentation indices, and scale metrics, and are routinely incorporated into natural resource management texts on GIS and the emerging field of landscape ecotoxicology.[96] In a similar fashion, a hierarchical approach to spatial scales in environmental analysis[42] has been developed for both terrestrial and aquatic ecosystems, usually on an integrated basis relative to either a landscape or watershed context. Hunsaker and Levine[97] used GIS and remote sensing of land use in a hierarchy of 47 watersheds to assess water quality in the Wabash River System in Illinois. In this study, water quality monitoring sites were linked to their respective watershed segment in the hierarchy to address issues of terrestrial processes in the landscape and evaluate their relevance to environmental management practices. This GIS and hierarchical approach facilitated identification of water quality conditions at several spatial scales and provided resource managers with tools to enhance decision support and data maintenance.

O'Neill et al.[33] recommended the use of GIS and remote sensing data, along with recent developments in landscape ecology, to assess biotic diversity, watershed integrity, and landscape stability. These authors presented GIS watershed integrity results for the lower 48 states on the basis of 16 U.S. Geological Survey Water Resource Regions. In general, GIS and remote sensing imagery have strongly facilitated a hierarchical approach to spatial scale and watershed analysis. This has been due, in part, to the better availability of geospatial data and technology, but this also is based on the relevancy of these regional environmental assessments for broad geographic extents.[33,97]

A spatial hierarchy to watersheds has been employed in four other examples relevant to design principles for environmental monitoring. In the first example, Preston and Brakebill[98] developed a spatially referenced regression model of watershed attributes to assess nitrogen loading in the entire Chesapeake Bay watershed. These investigators used the EPA River Reach File to generate a spatial network composed of 1408 stream reaches and watershed segments for their regional analysis. From their GIS visual maps of the watershed, point sources of high nitrogen loading could be associated with specific urbanized areas of the Bay watershed and allowed the authors to acknowledge and identify large sewage-treatment plants as discharge points to stream reaches.

In the second example, an Interagency Stream Restoration Working Group (15 federal agencies of the U.S.) recently developed a guidance manual for use in stream restoration[99] based on a hierarchical approach to watersheds at multiple scales. This team recognized ecosystems at five different spatial scales from regional landscape to local stream reach, and stated that watershed units can be delineated at each of these scales—depending on the focus of the analysis and availability of data. Spatial scales of watershed ecosystems from the stream restoration guidance manual (Figure 1.7) were used in the GIS watershed master plan[40] and served as the basis for our regional heritage river designation and approach to the first steps of assessing environmental conditions in the US-L watershed. The illustration of spatial scale shown in Figure 1.7 is based on an example of ecosystem hierarchy in the overall Chesapeake Bay watershed. This hierarchy was employed in our study design and tributary analysis of the US-L watershed[40] that ranged from a regional landscape watershed to a local stream reach along a linear segment of stream or river corridor (see text below).

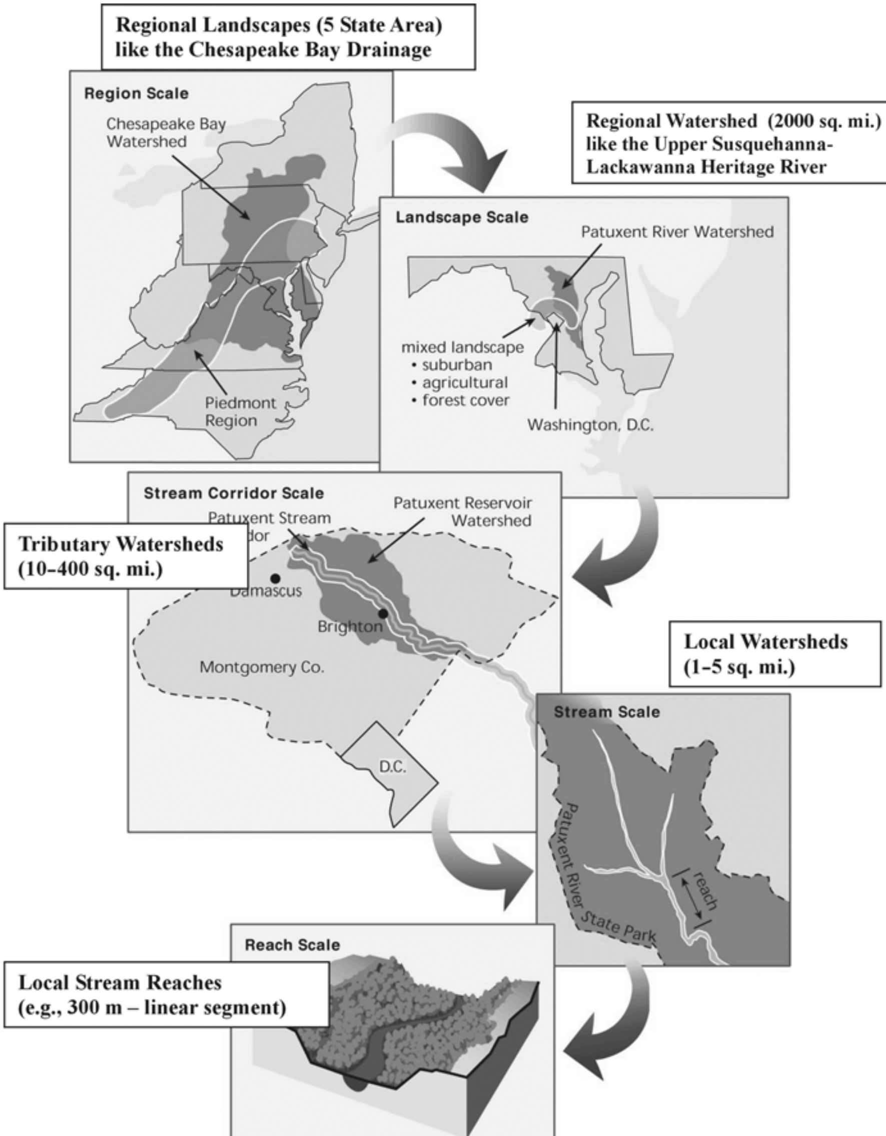


FIGURE 1.7 Ecosystems at multiple scales and used as the basis for regional to local GIS watershed analysis[40] for the Upper Susquehanna–Lackawanna American Heritage River. (Modified from the Interagency Stream Restoration Manual, Interagency Team, *Stream Restoration Guidance Manual, Stream Corridor Restoration: Principles, Processes, and Practices*, Federal Interagency Stream Restoration Working Group (FISRWG) (15 federal agencies of the U.S. government), GPO Item No. 0120-A; SuDocs ISBN-0-934213-59-3, Washington, D.C., 1998.)

The third example is the EPA EMAP ecological assessment of the U.S. Mid-Atlantic Region, including the Chesapeake Bay watershed plus portions of other watersheds like the Ohio and Delaware rivers over a five-state area.[4] These investigators developed 33 different watershed indicators of ecological conditions in 123 watershed units throughout the study area. This watershed assessment methodology is based on concepts developed earlier by O'Neill et al.[33] and relies strongly on land use and land cover data derived from satellite imagery, along with a number of complementary geospatial datasets like population density, roads, and hydrography.

The US-L watershed is one of the 123 watershed units from the EPA Mid-Atlantic regional study and we used their information to provide a regional context for our more detailed spatial analyses.[40] Thus, our fourth example is from the Heritage River watershed study and comprehensive GIS environmental master plan[40] where we conducted a GIS tributary analysis (on ten watershed indicators) and focused on 42 ecologically defined tributaries to the Susquehanna River (within the regional US-L watershed) including the river corridor segment of the mainstem. We highlighted results from five land cover classes (Multi-Resolution Land Characteristics or MRLC) that were evaluated in the watershed analysis. Of these, the mining land cover class was most effective (high) in detecting statistically significant impact differences between four (reference) rural watersheds without mining from another 12 tributaries with visibly observable mining/urban impacts.[40] Agricultural land cover was rated "medium" in its effectiveness in this regard. It was generally difficult to use forest, urban, and wetland cover classes to differentiate between mining/urban and "non-mining" rural reference tributaries. Other watershed parameters developed specially to address our regional environmental problems included the number of CSOs in a tributary subwatershed, the number of acid mining outfalls in a subwatershed, iron loading rates (from outfalls), and hydrogen ion (acidity) loading (also from mining outfalls). Statistical tests indicated significant differences in ranks between watershed conditions for CSOs, hydrogen ion loading, iron (Fe) loading, AMD outfalls, and an index based on the average of all seven watershed indicators.[40]

1.5 SYNTHESIS AND FUTURE DIRECTIONS IN MONITORING DESIGN

Our approach to monitoring design has been in continuous evolution, dependent on the growing awareness, concern, and the expanding scientific research and literature on regional environmental change, climate warming, and pollutants and land use change on a global basis.[14] The first four design components generally apply to our work at the southern Chile Biosphere Reserve, the Wind River Mountains high-elevation monitoring site, and the Pennsylvania Heritage River watershed. These components are based on long-founded interdisciplinary principles in the environmental sciences and ecology. For example, the conceptual framework of terrestrial-aquatic linkages within a river drainage basin[38,58] has been a key feature in river ecology research since the mid-1970s. Likewise, the use of heuristic ecosystem diagrams focused on source-receptor pathways of pollutant contaminants has been applied to biosphere research for the United Nations Global Environmental Monitoring System (GEMS) since the early and mid-1980s.[22,23]

The third and fourth design components, multimedia monitoring and ecosystem endpoints, are based on well-established methods in pollution measurements and field ecology but require an interdisciplinary teamwork approach. Long-standing monitoring programs like the LTER program at NSF, the NAPAP program for acid rain in the U.S., and EPA's EMAP program all exemplify these principles. In many ways, our work at the Wind Rivers site incorporated aspects of all of these programs[17–19] but pushed the limits of standardized instrumentation and methodology since most of our field measurements occurred above 10,000 feet elevation in alpine and subalpine forests, streams, lakes, and watersheds.

Technological innovations[100] and data availability,[4,87] NRC panel reports,[30,31] and the general spatial extent and complexity of regional and global environmental change, especially global land use change,[9,10,14] have all had a direct effect on our current focus on the last two recent components of an integrated conceptual approach to environmental monitoring. These encompass data integration with GIS and image processing software,[84,85,101] and landscape-watershed spatial scaling based on remote sensing imagery (SPOT, Landsat) and a landscape sampling design.[29,100–102] These have been more applicable to our research and work at the Chile and Pennsylvania sites since these have been ongoing, evolving, and have benefited from the availability of GIS, GPS, and remote sensing technologies. And finally, from an overall systems viewpoint, our approach has been influenced by and consistent with the landscape and watershed methods of O'Neill et al.[32,33] and the watershed analyses of the EPA Mid-Atlantic Landscape Atlas.[4]

Given the opportunities inherent in geospatial technologies and GIS data for environmental monitoring,[101] we are currently examining a number of watershed and landscape tools that would seem helpful to resource managers designing integrated monitoring programs. We predict that these and related GIS watershed and landscape software tools will facilitate ongoing and future efforts to monitor and assess environmental systems on a geographically extensive basis. These are briefly described below and highlighted in terms of their potential applicability to monitoring design considerations.

1.5.1 EPA BASINS

Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) is a multipurpose environmental analysis system for use by natural resource agencies to perform watershed- and water-quality-based studies. It was developed by EPA to address three objectives:

- To facilitate examination of environmental information
- To support analysis of environmental systems, and
- To provide a framework for examining management alternatives

BASINS[103] is intended to support a watershed-based approach to environmental and ecological studies in a watershed context. As such, the system has been designed to be flexible with a capability to support analysis at a variety of scales, using tools that range from simple to sophisticated. Comprehensive multimedia data sets compiled by EPA are provided in BASINS and one can access and query the datasets through a GIS data mining tool; other assessment tools include TARGET and ASSESS which allow one to evaluate water quality concerns and impacts on either a watershed or sampling point basis, respectively. BASINS requires a GIS platform as a separate software component but makes ample use of GIS utilities for data integration, analyses, graphic visualizations, and modeling. A number of models are also available, mostly for point sources of pollution regarding industrial discharges to surface waters. BASINS has been used in various capacities as part of a comprehensive GIS environmental plan for the US-L watershed and has helped contribute to all design components of a community water quality monitoring program (see Reference 40 and EMPACT Website at wilkes.edu/~gisriver/).

1.5.2 SWAT

The Soil Water Assessment Tool (SWAT) is now part of BASINS[103] but was originally developed by the U.S. Department of Agriculture[104] for management and assessment of agricultural and rangelands. SWAT is a GIS model that incorporates geospatial data on climate, rainfall, soils, slopes, and land cover and provides for a procedure to address changing land use (land cover) conditions in a watershed. Runoff and sediment yield from different patterns of land cover in a watershed are provided as model output[105] and as such give the resource manager a better idea of source-receptor issues in the design of a monitoring program. In addition, the hydrologic model is useful for evaluating conceptual relationships on a watershed basis, facilitating data integration, and predicting impacts to ecological endpoints.

1.5.3 CITYgreen Regional Analysis

CITYgreen is a GIS software tool for regional, local, and watershed–landscape analysis on the environmental function and economic value of trees and forests, especially in urban areas.[106] This is an environmental planning tool capable of using detailed forest (and tree) stand data locally or regional satellite imagery classified for land cover. The software uses readily available data on soils (STATSGO), slopes (STATSGO), and regional rainfall zones and precipitation—all of which are provided with the software program and user's manual. Model output includes runoff, carbon sequestration rates and storage (due to forests), and air pollution (sulfur dioxide, carbon monoxide, ozone, and nitrogen dioxide) removal potential due to deposition to forest vegetative surfaces. The software allows managers and planners to predict the outcome of various development scenarios by easily modifying (via user menus or software "wizards") the amount and type of land use change for a delineated study area.[107] CITYgreen appears to be a useful community planning tool that allows for data integration and especially predictive modeling of multimedia pathways and affects — both of which are useful to resource managers developing monitoring programs. The relative ease of operation for planning would seem to facilitate the iterative process necessary to monitoring design processes. We are currently starting analyses with CITYgreen to evaluate the potential for carbon sequestration if extensive re-forestation and ecosystem restoration efforts were to be implemented on regional areas of barren mining lands.

1.5.4 ATtILA

ATtILA refers to an EPA-developed GIS software component known as "Analytical Tools Interface for Landscape Assessments." This is a GIS software extension that allows users to easily calculate many common landscape metrics regardless of their level of GIS knowledge. Four metric groups are currently included in ATtILA[108]: landscape characteristics (e.g., percentage in forests or grasslands, landscape diversity, forest patch size), riparian characteristics (e.g., percentage of stream length adjacent to forests or urban land cover) human stresses (e.g., population numbers and change), and physical characteristics (e.g., average elevation, precipitation, total stream length). At present, ATtILA is available for R&D efforts mostly within the agency but we have examined the potential of ATtILA for application of landscape and watershed indicators in environmental monitoring and assessment. Again, due to the nature of a GIS approach, data integration, conceptual relationships, and assessment of landscape and watershed ecological endpoints would seem to be appropriate benefits to resource managers and researchers who need to address broad scale impacts from regional land use change and watershed modification.

1.5.5 Metadata Tools and Web-Based GIS

As a final note to future directions, we point out the importance of GIS metadata tools and the potential of Web-based GIS for data integration and showcasing environmental relationships at different spatial scales — local to global. The latter will also be extremely important for dissemination of data to resource managers, government leaders, scientists, industry and business leaders, and the public at large. The importance of metadata for GIS and environmental monitoring in general has been identified by the NRC report of Wiersma et al.[31] and briefly addressed above. ESRI's (Environmental Systems Research Institute) ArcGIS software now provides a metadata tool consistent with FGDC geospatial data standards. Metadata are crucial to QA and QC issues with environmental data of all types but are necessary also for accessing and querying databases on the Web for incorporation into a monitoring program or for a GIS decision support system for monitoring design. ESRI's ArcIMS (Internet Map Server) software and its Geography Network provide mechanisms for scientists, researchers, and resource managers to access a variety of environmental data for assessment, monitoring, and the design of a field measurement program. The reader is referred to the ESRI general Web page (www.esri.com) for additional information on software applications in this regard. Reports for the US-L American Heritage River watershed plan[40] provide diagrams and software architecture design components for how a GIS Web-based decentralized data distribution system can be used for community environmental programs, including monitoring of water quality and the status of watersheds in the region.

1.5.6 Homeland Security

The events of September 11, 2001 have focused detailed attention to the vulnerability of the nation's resources. An NRC report has recently addressed this issue[109] that has become the top priority to federal, state, and local government leaders and agencies, and the public at large. The NRC report highlights the need for a national program aimed at making the nation safer and outlines the technical elements necessary for homeland security. Geographic information systems (GIS) are identified as a critical component to this program within the context of information technology (IT) in general. Other reports by the National Center for Environmental Assessment (EPA web documents at http://www.epa.gov/nheerl/wtc/index.html) address the human and environmental risks associated with toxic chemical releases to the atmosphere associated with the destruction, combustion, and collapse of the New York World Trade Center towers. In the future, geospatial data and technologies are expected to be crucial analytical tools in both the collection of relevant environmental data and in making data available to emergency management personnel, government leaders, cleanup crews, environmental and public health managers and organizations, and the public at large.

Many of the nation's resources that are at high risk as identified in the NAS report on homeland security occur in rural communities, watersheds, and landscapes. These include energy, nuclear plants, groundwater resources, surface waters, food, agricultural soils, utilities and corridors, and rural transportation arteries between major cities and urban centers. In addition, EPA has a "Strategic Plan for Homeland Security" and the agency is especially responsible for security issues[110] regarding water resources, water supply, wastewater treatment facilities, and facilities of the chemical and energy industries.

Although "Homeland Security" in its many complex facets is beyond the scope of our objectives in the paper, it should be indicated that the ecological consequences of intentional releases of chemical and biological agents warrant a comprehensive systems approach to monitoring the fate and effects of these contaminants in the environment. In some cases, monitoring methods and measurements have direct applicable from more conventional programs of environmental management but the conceptual basis and six components outlined here are still intended to be useful starting points to support these newer efforts as part of national security.

1.6 CONCLUSION

The principles and concepts of environmental monitoring design are dynamic and iterative in nature. We have attempted to outline the key components of our approach to these concepts within the context of recommendations and reports of subcommittees on integrated monitoring sponsored by the NRC. The need has never been greater for integrated programs that collect, analyze, evaluate, and disseminate relevant environmental measurements at various scales on an ecosystem and multimedia basis. Local communities now grapple with issues of land use planning and change since most economic and environmental decisions are made at this local level of government. Such communities are often ill-equipped with the data and decision tools to make informed decisions on environmental health and the quality of life in general. State and federal agencies, research universities, and international organizations all maintain numerous environmental networks, but data may not be available to the general public or their use is beyond their resources for interpretation and decision making. Land use change is now viewed from a worldwide basis as more prevalent and critical (at least short-term) than global climate warming.[9,10] Other aspects of global environmental change[14] have important implications for natural and managed ecosystems worldwide and provide critical challenges to those professionals charged with designing and implementing integrated environmental monitoring programs. Global dispersal and deposition of air contaminants, toxics in the global environment, loss of biodiversity, and climate warming all warrant close attention to coordinated efforts and networks of monitoring programs.

The proposal of a network of "biosphere observatories" as originally described in an IGBP report[12] is still a viable concept to global environmental monitoring activities. The LTER program of NSF, EPA's EMAP, and the proposed and emerging NEON program represent select examples of critical, ongoing monitoring efforts. In addition, the Ecological Society of America has proposed the Sustainable Biosphere Initiative[111] with broad applicability and relevance to integrated monitoring programs along with recommendations for an aggressive research program to assess ecological responses to stress and loss of biodiversity both on a regional and global basis.

Our conceptual approach to these issues is intended to help scientists, researchers, decision makers, and other leaders develop, implement, and maintain integrated monitoring on a comprehensive basis but relevant to local, regional, national, and global perspectives. Our work at remote sites provides insight to the challenges of monitoring and detecting early-warning indicators against a background of natural variability in ecosystems. In contrast, the Heritage River watershed project has been developed with the intended benefit of geospatial technologies and the earlier design components from our remote site research. These sites provide a broad spectrum of applications ranging from baseline monitoring in relative un-impacted regions to regional watersheds where land use change and extensive mining has resulted in opportunities to assess and monitor the impacts of "industrial metabolism" (see References 12 and 19) from a landscape-watershed perspective with geospatial data and technologies. This collective work is intended to promote the biosphere observatory concept and to provide practical, real world "test-beds" for evaluating methods and approaches to integrated environmental monitoring.

ACKNOWLEDGMENTS

We would like to thank our numerous colleagues for their many contributions to our research at remote study sites, and especially Greg White who headed up the forest ecosystem measurements component at the Wind Rivers study area. We also appreciate the support of our Chilean colleagues including Guillermo Santana, who hosted our work as Manager of Torres del Paine National Park. Xiaoming Yang, Sid Halsor, Bill Toothill, Brian Oram, Bill Feher, Dave Skoronski, and Jim Thomas all contributed to various aspects of data collecting and processing for the Heritage River site; Kristopher Smith refined graphical materials for the chapter. This paper is part of the ongoing GIS environmental master plan for the US-L Watershed and has been supported by various people and agencies. Alex Rogers, the AHR Navigator, provided support and coordination for the community outreach aspects of the master plan cited throughout this chapter. Dave Catlin of the EPA's Office of Environmental

Information enthusiastically provided support and collaborative energy to our Heritage River site regarding application of geospatial data and technologies. Tom Sweet of the Pennsylvania GIS Consortium has provided insight to the use of local scale aerial photography data and how to incorporate that component into our conceptual design approach. Chris Cappelli of ESRI provided many hours of technical and conceptual support for the use of GIS, especially Web-based GIS, in data integration, analysis, and dissemination. Writing of aspects of this paper was funded and supported by the USDA (Cooperative State Research, Education, and Extension Service) Rural GIS program (some data analysis and manuscript writing), and EPA's Office of Environmental Information (writing and data analysis phases). And finally, we would like to thank Congressman Paul E. Kanjorski (U.S. 11th Congressional District in Pennsylvania) and his staff for their long-term support of GIS technology to solve environmental problems; the Congressman was tireless in his leadership to facilitate designation of the US-L watershed as an American Heritage River.

REFERENCES

1. Franklin, J.F., Bledsoe, C.S., and Callahan, J.T., Contributions of the long-term ecological research program, *BioScience*, 40, 509, 1990.
2. National Acid Precipitation Assessment Program (U.S. NAPAP), *1990 Integrated Assessment Report*, NAPAP Office of the Director, Washington, D.C., 1991.
3. Summers, J.K. and Tonnessen, K.E., Linking monitoring and effects research: EMAP's intensive site network program, *Environ. Monit. Assess.*, 51, 369, 1998.
4. Jones, K.B. et al., *An Ecological Assessment of the United States Mid-Atlantic Region: A Landscape Atlas*, U.S. EPA, Office of Research and Development, Washington, D.C., EPA/600/R-97/130, 1997.
5. Houghton, R.A. and Woodwell, G.M., Global climate change, *Sci. Am.*, 260, 36, 1989.
6. Houghton, R.A. et al., *Climate Change 1995: The Science of Climate Change*, Cambridge University Press, Cambridge, U.K., 1996.
7. Pimm, S.L. et al., The future of biodiversity, *Science*, 269, 347, 1995.
8. Vitousek, P.M. et al., Human domination of earth's ecosystems, *Science*, 277, 494, 1997.
9. Carpenter, S.R. et al., Global change and freshwater ecosystems, *Annu. Rev. Ecol. Syst*. 23, 19, 1992.
10. Chapin, F.S., III et al., Biotic control over the functioning of ecosystems, *Science*, 277, 500, 1997.
11. National Research Council, *Global Change in the Geosphere–Biosphere: Initial Priorities for an IGBP*, National Academies Press, Washington, D.C., 1986.
12. National Research Council, *Toward an Understanding of Global Change: Initial Priorities for U.S. Contributions to the International Geosphere–Biosphere Program*, National Academies Press, Washington, D.C., 1988.
13. Long-Term Ecological Research Program (NSF-LTER Network Office), *1990's Global Change Action Plan*, Long-Term Ecological Research Office, Seattle, WA, 1989.
14. Walker, B. et al., Eds., *The Terrestrial Biosphere and Global Change: Implications for Natural and Managed Ecosystems*, Cambridge University Press, Cambridge, U.K., 1999.
15. Scientific Committee on Problems of the Environment/Man and the Biosphere Programme (SCOPE/MAB), *Definition and Description of Biosphere Observatories for Studying Global Change*, Draft Report, SCOPE/MAB Workshop, Paris, France, 1987.

16. University Corporation for Atmospheric Research, *Opportunities for Research at the Atmosphere/Biosphere Interface*, Report of a Workshop, Boulder, CO, 1985.
17. Bruns, D.A., Wiersma, G.B., and Minshall, G.W., Evaluation of community and ecosystem monitoring parameters at a high-elevation Rocky Mountain study site, *Environ. Toxicol. Chem.*, 11, 359, 1992.
18. Bruns, D.A., Wiersma, G.B., and Minshall, G.W., Lotic ecosystems and long-term monitoring for global change, in *Global Warming and Freshwater Ecosystems*, Firth, P. and Fisher, S.G., Eds., Springer-Verlag, New York, 1992, chap. 14.
19. Bruns, D.A., Wiersma, G.B., and White, G.J., Testing and application of ecosystem monitoring parameters, *Toxicol. Environ. Chem.*, 62, 169, 1997.
20. Bruns, D.A. et al., *System for Environmental Survey and Land Use (GIS Analysis of watershed impacts from regional coal mining)*, Final Report to Earth Conservancy (Ashley, PA) and Department of Defense, Advanced Research Program Agency, Washington, D.C., 1997.
21. Wiersma, G.B., Conceptual basis for environmental monitoring programs, *Toxicol. Environ. Chem.*, 27, 241, 1990.
22. Wiersma, G.B. et al., The use of simple kinetic models to help design environmental monitoring systems, *Environ. Monit. Assess.*, 4, 233, 1984.
23. Wiersma, G.B. et al., Integrated global background monitoring network, in *Monitoring and Managing Environmental Impact: American and Soviet Perspectives*, Schweitzer, G.E. and Philips, A.S., Eds., Proceedings of the Fifth United States/Union of Soviet Socialist Republic Symposium on Comprehensive Analysis of the Environment, National Academies Press, Washington, D.C., 1986, 246.
24. Wiersma, G.B. et al., *Reconnaissance of Noatak National Preserve and Biosphere Reserve as a potential site for inclusion in the Integrated Global Background Monitoring Network*, Report for U.S. Man and the Biosphere Program, NTIS PB 88-100037, 1986.
25. Wiersma, G.B. et al., *Integrated Monitoring Project at Torres del Paine National Park, Chile, Methodology and Data Report—1984 to 1986*, EG&G Idaho, Inc., Informal Report, EGG-EES-7966, 1988. Idaho Falls, ID.
26. Wiersma, G.B. and Bruns, D.A., Monitoring for ecological assessment, in *North American Workshop on Monitoring for Ecological Assessment of Terrestrial and Aquatic Ecosystems*, Bravo, C.A., Ed., USDA Technical Report RM-GTR-284, 1996, 31.
27. Bruns, D.A., Wiersma, G.B., and Rykiel, E.J., Jr., Ecosystem monitoring at global baseline sites, *Environ. Monit. Assess.*, 17, 3, 1991.
28. Nilhgard, B. and Pylvanainen, M., Eds., *Evaluation of Integrated Monitoring Programme in Terrestrial Reference Areas in Europe and North America: The Pilot Programme 1989–91*, Environmental Data Centre, National Board of Waters and the Environment, Helsinki, 1992.
29. Heal, O.W., Menaut, J., and Steffen, W.L., Eds., *Towards a Global Terrestrial Observing System (GTOS): Detecting and Monitoring Change in Terrestrial Ecosystems*, MAB Digest 14 and IGBP Global Change Report 26, UNESCO, Paris, and IGBP, Stockholm, 1993.
30. Boesch, D.F. et al., *Managing Troubled Waters: The Role of Marine Environmental Monitoring*, National Academies Press, Washington, D.C., 1990.
31. Wiersma, G.B. et al., *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data*, National Academies Press, Washington, D.C., 1995.
32. O'Neill, R.V. et al., Indices of landscape pattern, *Landscape Ecol.*, 1, 153, 1988.
33. O'Neill, R.V. et al., Monitoring environmental quality at the landscape scale, *BioScience*, 47, 513, 1997.

34. Rosswall, T., Woodmansee, R.G., and Risser, P.G., Eds., *Scales and Global Change: Spatial and Temporal Variability in Biospheric and Geospheric Processes*, SCOPE 35, John Wiley & Sons, New York, 1988.
35. Quattrochi, D.A. and Goodchild, M.F., Eds., *Scale in Remote Sensing and GIS*, Lewis Publishers, Boca Raton, FL, 1997.
36. Clesceri, L.S., Greenberg, A.E., and Eaton, A.D., Eds., *Standard Methods for the Examination of Water and Wastewater*, American Public Health Association (APHA), Washington, D.C., 1998.
37. Keith, L.H., Ed., *Principles of Environmental Sampling*, American Chemical Society, Washington, D.C., 1988.
38. Minshall, G.W. et al., Developments in stream ecosystem theory, *Can. J. Fish. Aquat. Sci.*, 42, 1045, 1985.
39. EDAW, Inc., *Earth Conservancy Land Use Plan, Luzerne County, Pennsylvania*, Final Report to Earth Conservancy, Ashley, PA, 1996.
40. Bruns, D.A., Sweet, T., and Toothill, B., *Upper Susquehanna–Lackawanna River Watershed, Section 206, Ecosystem Restoration Report*, *Phase I GIS Environmental Master Plan*, Final Report to U.S. Army Corps of Engineers, Baltimore District, MD, 2001.
41. Wiersma, G.B. and Otis, M.D., Multimedia design principles applied to the development of the global integrated monitoring network, in *Pollutants in a Multimedia Environment*, Cohen, Y., Ed., Plenum Publishing, New York, 1986, 317.
42. O'Neill, R.V., Hierarchy theory and global change, in *Scales and Global Change: Spatial and Temporal Variability in Biospheric and Geospheric Processes*, Rosswall, T., Woodmansee, R.G., and Risser, P.G., Eds., SCOPE 35, John Wiley & Sons, New York, 1988, 29.
43. Olson, R.K. and Lefohn, A.S., Eds., *Effects of Air Pollution on Western Forests*, Transactions Series No. 16, ISSN 1040-8177, 1989.
44. Bohm, M. and Vandetta, T., *Atlas of Air Quality and Deposition in or near Forests of the Western United States*, U.S. Environmental Protection Agency, EPA/600/3-90/081, 1990.
45. Bruns, D.A. and Yang, X., An accuracy assessment of satellite imagery used in landscape–watershed assessments: a comparison of four databases, in *Papers and Proceedings of the Applied Geography Conferences*, Vol. 25, Montz, B.E. and Tobin, G.A., Eds., Applied Geography Conferences, Binghamton, NY, 2002, 230.
46. White, G.J. et al., *Evaluation of U.S. Forest Service Document: Guidelines for Measuring the Physical, Chemical, and Biological Condition of Wilderness Areas*, *Final Report*, EG&G Idaho, Inc., Informal Report, EGG-BE-9929, DOE National Engineering Lab., Idaho Falls, ID, 1991.
47. Fox, D.G., Bernabo, J.C., and Hood, B., *Guidelines for Measuring the Physical, Chemical, and Biological Condition of Wilderness Areas*, General Technical Report RM-146, U.S. Department of Agriculture, Forest Service Rocky Mountain Forest and Range Experimental Station, Fort Collins, CO, 1987.
48. Mackay, D., *Multimedia Environmental Models, the Fugacity Approach*, Lewis Publishers, Boca Raton, FL, 1991.
49. Landrum, P.F., Harkey, G.A., and Kukkonen, J., Evaluation of organic contaminant exposure in aquatic organisms: the significance of bioconcentration and bioaccumulation, in *Ecotoxicology: A Hierarchical Treatment*, Newman, M.C. and Jagoe, C.H., Eds., CRC Press, Boca Raton, FL, 1996, chap. 4.
50. Newman, M.C., *Fundamentals of Ecotoxicology*, Ann Arbor Press, Chelsea, MI, 1998.
51. Schindler, D.W., Effects of acid rain on freshwater ecosystems, *Science*, 239, 149, 1988.

52. Taylor, G.E., Jr., Forest ecosystems and air pollution: the importance of multiple stress interactions on a regional and global scale, in *Multiple Stresses in Ecosystems*, Cech, J.J., Jr., Wilson, B.W., and Crosby, D.G., Eds., Lewis Publishers, Boca Raton, FL, 1998, chap. 4.
53. Shaw, I.C. and Chadwick, J., *Principles of Environmental Toxicology*, Taylor and Francis, Bristol, PA, 1998.
54. Environmental Protection Agency (EPA), *Protocol for the Transition-Flow Reactor Concentration Monitor: Determination of Atmospheric Concentrations of Gaseous* HNO_3*,* SO_2*,* NO_2*, and* NH_3 *and of Fine Particulate Nitrate, Sulfate, and Ammonium Ion*, Atmospheric Sciences Research Laboratory, Office of Research and Development, Research Triangle Park, NC, 1987.
55. Wiersma, G.B., *Recommended Integrated Monitoring System for Pollutants in U.S. National Parks Designated as Biosphere Reserves*, EG&G Idaho Informal Report, EGG-PBS-6721, Idaho Falls, ID, 1985.
56. Robarge, W.P. and Fernandez, I., *Quality Assurance Methods Manual for Laboratory Analytical Techniques*, Draft EPA Report, Corvallis Environmental Research Laboratory, Corvallis, OR, 1986.
57. National Surface Water Survey (NSWS), *Analytical Methods Manual for the National Surface Water Survey*, Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV, 1986.
58. Vannote, R.L. et al., The river continuum concept, *Can. J. Fish. Aquat. Sci.*, 37, 130, 1980.
59. Bruns, D.A. et al., Ordination of functional groups and organic matter parameters from the middle fork of the Salmon River, Idaho, *Freshw. Invertebr. Biol.*, 1, 2, 1982.
60. Bruns, D.A. et al., Tributaries as modifiers of the river continuum concept: analysis by polar ordination and regression models, *Arch. Hydrobiol.*, 99, 208, 1984.
61. Baker, G.A., Harmon, M.E., and Greene, S.E., A study of selected ecosystem processes potentially sensitive to airborne pollutants, in *Proceedings of a Conference on Science in the National Parks: Physical Processes and Water Resources*, Flug, M., Ed., G. Wright Society and Colorado State University, Fort Collins, CO, 1986, 119.
62. Cairns, J., Jr., The myth of the most sensitive species, *BioScience*, 36, 670, 1986.
63. Schindler, D.W., Ecosystems and ecotoxicology: a personal perspective, in *Ecotoxicology: A Hierarchical Treatment*, Newman, M.C. and Jagoe, C.H., Eds., Lewis Publishers, Boca Raton, FL, 1996, chap. 13.
64. Newman, M.C. and Jagoe, C.H., Eds., *Ecotoxicology: A Hierarchical Treatment*, Lewis Publishers, Boca Raton, FL, 1996.
65. Cech, J.J., Jr., Wilson, B.W., and Crosby, D.G., Eds., *Multiple Stresses in Ecosystems*, Lewis Publishers, Boca Raton, FL, 1998.
66. Allen, T.F.H. and Starr, T.B., *Hierarchy: Perspective for Ecological Complexity*, University of Chicago Press, Chicago, IL, 1982.
67. O'Neill, R.V. et al., *A Hierarchical Concept of Ecosystems*, Princeton University Press, Princeton, NJ, 1986.
68. Bruns, D.A. et al., An ecosystem approach to ecological characterization in the NEPA process, in *Environmental Analysis: The NEPA Experience*, Hildebrand, S.G. and Cannon, J.B., Eds., Lewis Publishers, Boca Raton, FL, 1993, 103.
69. Rapport, D.J., Regier, H.A., and Hutchinson, T.C., Ecosystem behavior under stress, *Am. Nat.*, 125, 617, 1985.
70. Odum, E.P., Trends expected in stressed ecosystems, *BioScience*, 35, 419, 1985.
71. Schaeffer, D.J., Herricks, E.E., and Kerster, H.W., Ecosystem health: I. Measuring ecosystem health, *Environ. Manage.*, 12, 445, 1988.

72. National Research Council, *Testing for the Effects of Chemicals on Ecosystems*, National Academy Press, Washington, D.C., 1981.
73. Sheehan, P.J. et al., Eds., *Effects of Pollutants at the Ecosystem Level*, SCOPE 22, John Wiley & Sons, New York, 1984.
74. Taub, G.B., Indicators of change in natural and human-impacted ecosystems: status, in *Preserving Ecological Systems: The Agenda for Long-Term Research and Development*, Draggan, S., Cohrssen, J.J., and Morrison, R.E., Eds., Praeger, NY, 1987, 115.
75. Sigal, L.L. and Suter, G.W., II, Evaluation of methods for determining adverse impacts of air pollution on terrestrial ecosystems, *Environ. Manage*., 11, 675, 1987.
76. Schindler, D.W., Detecting ecosystem responses to anthropogenic stress, *Can. J. Fish. Aquat. Sci*., 44, 6, 1987.
77. Hinds, W.T., Towards monitoring of long-term trends in terrestrial ecosystems, *Environ. Conserv.*, 11, 11, 1984.
78. Beanlands, G.E. and Duinker, P.N., An ecological framework for environmental impact assessment, *J. Environ. Manage*., 18, 267, 1984.
79. Karr, J.R., Biological monitoring and environmental assessment: a conceptual framework, *Environ. Manage.*, 11, 249, 1987.
80. Ausmus, B.S., An argument for ecosystem level monitoring, *Environ. Monit. Assess.*, 4, 275, 1984.
81. Zeiler, M., *Modeling our World: The ESRI Guide to Geodatabase Design*, ESRI Press, Redlands, CA, 1999.
82. Bruns, D.A. and Sweet, T., *Tackling Environmental Clean-up with GIS: Regionally Coordinated Geographic Information Systems Provide Solutions to Susquehanna–Lackawanna Watershed Pollution Problems*, University of Wisconsin, Land Information Bulletin, USDA Rural GIS, Madison, WI, 2001.
83. Smith, J.H. et al., Impacts of patch size and land cover heterogeneity on thematic image classification accuracy, *Photogr. Eng. Rem. Sens.*, 68, 65, 2002.
84. Avery, T.E. and Berlin, G.L., *Fundamentals of Remote Sensing and Airphoto Interpretation*, Macmillan, NY, 1992.
85. Lunetta, R.S. et al., Remote sensing and geographic information systems data integration: error sources and research issues, *Photogr. Eng. Rem. Sens.*, 57, 677, 1991.
86. Vogelmann, J.E. et al., Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogr. Eng. Rem. Sens.*, 67, 650, 2001.
87. Federal Geographic Data Committee, *Framework Introduction and Guide*, Federal Geographic Data Committee, Washington, D.C., 1997.
88. Falkner, E., *Aerial Mapping Methods and Applications*, Lewis Publishers, Boca Raton, FL, 1995.
89. Lyon, J.G., *Wetland Landscape Characterization: Techniques and Applications for GIS, Mapping, Remote Sensing, and Image Analysis*, Ann Arbor Press, Chelsea, MI, 2001.
90. Dahlman, B.N. and Lanzer, E.L., Cartographic support for managing Washington State's aquatic resources, in *GIS Solutions in Natural Resource Management: Balancing the Technical–Political Equation*, Morain, S., Ed., OnWord Press, Santa Fe, NM, 1999, 115.
91. Brondizio, E. et al., Land cover in the Amazon estuary: linking of the Thematic Mapper with botanical and historical data, *Photogr. Eng. Rem. Sens.*, 62, 921, 1996.
92. Congalton, R.G., A review of assessing the accuracy of classifications of remotely sensed data, *Rem. Sens. Environ.*, 37, 35, 1991.

93. Congalton, R.G. and Green, K., *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, FL, 1999.
94. Turner, M.G., Landscape ecology: the effect of pattern on process, *Annu. Rev. Ecol. Syst.*, 20, 171, 1989.
95. Turner, M.G. and Gardner, R.H., Eds., *Quantitative Methods in Landscape Ecology*, Springer-Verlag, New York, 1991.
96. Holl, K.D. and Cairns, J., Jr., Landscape indicators in ecotoxicology, in *Handbook of Ecotoxicology*, Hoffman, D.J., Eds., Lewis Publishers, Boca Raton, FL, 1994, 185.
97. Hunsaker, C.T. and Levine, D.A., Hierarchical approaches to the study of water quality in rivers, *BioScience*, 45, 193, 1995.
98. Preston, S.D. and Brakebill, J.W., *Application of Spatially Referenced Regression Modeling for the Evaluation of Total Nitrogen Loading in the Chesapeake Bay Watershed*, USGS Water-Resources Investigations Report 99-4054, U.S. Department of the Interior/U.S. Geological Survey, Reston, VA, 1999.
99. Interagency Team, *Stream Restoration Guidance Manual, Stream Corridor Restoration: Principles, Processes, and Practices*, Federal Interagency Stream Restoration Working Group (FISRWG) (15 federal agencies of the U.S. government), GPO Item No. 0120-A; SuDocs ISBN-0-934213-59-3, Washington, D.C., 1998.
100. Lillesand, T.M. and Kieffer, R.W., *Remote Sensing and Image Interpretation*, John Wiley & Sons, New York, 1987.
101. Lunetta, R.S. and Elvidge, C.D., Eds., *Remote Sensing Change Detection: Environmental Methods and Applications*, Ann Arbor Press, Chelsea, MI, 1998.
102. Wessman, C.A., Spatial scales and global change: bridging the gap from plots to GCM grid cells, *Annu. Rev. Ecol. Syst.*, 23, 175, 1992.
103. Environmental Protection Agency (EPA), *Better Assessment Science Integrating Point and Nonpoint Sources (BASINS), User's Manual*, Version 3.0, U.S. EPA Office of Water (4305), Washington, D.C., EPA-823-H-01-001, 2001.
104. Arnold, J.G. et al., *SWAT — Soil Water Assessment Tool*, USDA, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX, 1994.
105. Hernandez, M. et al., Modeling runoff response to land cover and rainfall spatial variability in semi-arid watersheds, *Environ. Monit. Assess.*, 64, 285, 2000.
106. American Forests, *CITYgreen: Calculating the Value of Nature*, Version 5.0 User's Manual, American Forests, Washington, D.C., 2002.
107. Carl, J.R., DeAngelo, M.M., and Ruskey, D.C., Sustainable development plan and runoff study for a proposed development in the Hick's Creek watershed, in *Papers and Proceedings of the Applied Geography Conferences*, Vol. 25, Montz, B.E. and Tobin, G.A., Eds., Applied Geography Conferences, Inc., Binghamton, NY, 2002, 9.
108. Ebert, D.W. et al., *Analytical Tools Interface for Landscape Assessments (ATtILA)*, Quick Start Guide, Version 3.0, Draft, Environmental Protection Agency, Office of Research and Development, Las Vegas, NV, 2001.
109. National Research Council, *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism*, National Academies Press, Washington, D.C., 2002.
110. Environmental Protection Agency, *Strategic Plan for Homeland Security*, Web site at http://www.epa.gov/epahome/downloads/epa_homeland_security_strategic_plan.pdf, 2002.
111. Lubchenco, J. et al., The sustainable biosphere initiative: an ecological research agenda, *Ecology*, 72, 371, 1991.

2 Integrated Data Management for Environmental Monitoring Programs

A.M.J. Lane, S.C. Rennie, and J.W. Watkins

CONTENTS

1-56670-641-6/04/$0.00+$1.50

2.1 INTRODUCTION

The forces driving environmental change and the processes that govern response to it are inherently complex, multisectoral, multidimensional, and multiscale. Data managers face the considerable challenge of providing integrated information systems for managing and presenting a diversity of environmental monitoring data and associated knowledge in a flexible and accessible format to a broad user community. User requirements within global change research are not easily defined; they are diverse, they evolve, and the information system must evolve with them.

Recognition of the lack of reliable long-term runs of multidisciplinary data for environmental change research led in the late 1980s to the revival of environmental monitoring as a valuable scientific activity.[1,2] However, this time, more emphasis was placed on links with models and policy needs and, crucially, integrated databases were recognized as a key component. The U.S. Long-Term Ecological Research Network (LTER) had already recognized the importance of good data management in bringing together existing site-based data sets for multisite research.[3] New environmental monitoring programs like the U.K. Environmental Change Network (ECN)[4,5] and the UNECE Integrated Monitoring Programme[6] began to understand that integrated information management was central to realizing the potential of an expensive data gathering network. Even so, the scale of this requirement was underestimated. The role of data managers has since ballooned into full-scale information systems developers including building data capture, and query and analysis systems for the Internet. The emerging field of environmental informatics broadens this scope still further, viewing data management as an integrated component of the system of transforming raw environmental data into higher-grade knowledge suitable for decision making.

This chapter presents an overview of the main issues and pitfalls concerned with delivering information systems to support environmental monitoring. The scope and depth of the topic are considerable and are developing rapidly in line with new technology and computational techniques. References to key papers and current websites are provided for more detail.

2.2 INFORMATION SYSTEM DEVELOPMENT

2.2.1 User and System Requirements

"The nature of interdisciplinary global change research makes it impossible to clearly define a detailed and stable set of user requirements."[7] Environmental monitoring systems deliver information to a diverse community of scientific, policy, and public users. Scientific users in particular cannot be constrained by a set of predefined queries and reports, but require free rein to explore patterns in the data through *ad hoc* queries and flexible data exploration tools. There is a demand for remote, easy, and timely access to integrated data, with more control over query and styles of reporting, which means that interpretation relies less and less on personal contact with data providers. The database has become a resource relied upon to provide a comprehensive representation of all information gathered for the duration of the environmental monitoring program. The role of the meta-database has grown into that of a knowledge base, providing not only details of which data are where in what form, but also data quality warnings and guidance on appropriate analysis and interpretation.

To support these requirements, environmental monitoring information systems should ideally have the following characteristics;[8–10] they should be:

- Reliable: stable, holding data of good and known quality
- Secure: maintained in perpetuity with suitable access controls and backup systems
- Integrated: enabling integration of multidisciplinary data across spatial and temporal scales
- Comprehensive: incorporating integral data, metadata, and knowledge
- Flexible: designed to respond to changing requirements, for example, new data types and different data transformations
- Accessible: timely access to data through direct, easy-to-use Internet-based interfaces
- Analytical: support for spatio-temporal analyses and modeling at different scales, data interpretation, and decision-support tools
- Compatible: enable links to be made with other information sources and software systems

Reliability and security are clearly essential features and are embraced by most current practice. However, the degree to which data quality is addressed tends to depend on how resources are directed; data validation is time consuming, does not lend itself to performance measures, and so tends to be squeezed when resources are tight. The demand for accessibility and the rapid development in Web technology has driven a surge in data discovery, data delivery, and data exploration interfaces.

2.2.2 System Design Approaches

Basic steps in the process of acquiring, managing, and delivering environmental information can be summarized as follows (based on a diagram by Harmancioglu et al.[11]):

- Objectives and constraints
- Network design
- Data capture (measurement, sample collection, laboratory analyses)
- Data handling (data transfer and data transformation protocols)
- Database development
- Data access
- Data analysis and modeling
- Information interpretation and utilization
- Decision making

Data managers and system designers need to address all these elements of the data-into-knowledge process from objective definition through analysis, modeling, and decision-support. An integrated management strategy should be adopted that considers these elements as part of a unified information system. Each step needs to be considered in relation to the other components of the system, with the necessary requirements, activities, and information flows identified at each stage. This ensures a more user-oriented perspective, enables weaknesses and bottlenecks to be identified, promotes timely access to data, and aids the quality management process. Quality assurance (QA) issues apply throughout, including feedback mechanisms to keep quality objectives on track (see Section 2.3).

A functional partnership between data managers and scientists[12] is important to ensure that the information system is designed to adapt to changing environmental research objectives. The system needs to be sufficiently versatile to generate different user views, transformations and visualizations of data, and to accommodate new types of data. This has implications for system design approaches, which must build in this flexibility and adaptability. In a research environment, a rigid specification of the final user requirements may not be possible[7] because requirements evolve as research progresses. The systems design approach must be capable of incorporating this evolution through periodic review and reevaluation of design objectives. This form of iterative development has been used to manage the risk of projects failing to meet requirements. Boehm's spiral model[13,14] and rapid application development tools are approaches that enable the system to undergo evolutionary development with constant refinement of the overall aim.

The degree to which new information systems are standardized or harmonized and centralized or distributed is becoming less technology limited and more a function of scientific and policy objectives and resources. Where standardization is possible and appropriate, it is clearly desirable in order to aid data comparability. New integrated monitoring programs like the U.K. ECN have been in a position to implement standard formats and data dictionaries for key variables (e.g., space and time dimensions, species coding systems, chemical determinants) from the outset to build an integrated database. This, together with the centralization of ECN data, facilitated the early development of web-accessible dynamic database interfaces.[15–17] Figure 2.1 shows an overview of the ECN data management system currently in place. However, in order to participate in global change research, programs like ECN must still interact with other existing thematic, historic, national, and international databases, which will have captured data in different ways according to different criteria. Technology for dynamic linking of distributed heterogeneous databases is only just developing. Environmental monitoring and research programs are now initiating partnerships and finding funding opportunities to develop "data-grid" services and semantic mediation systems for environmental data (see Section 2.6). To create a distributed networked information system requires a strong coordinating body plus the right mix of IT, data management, computer science, and scientific expertise, plus the financial resources to develop the common framework on which an integrated resource is built. The U.S. LTER network, focused on linking existing databases, illustrates this approach. LTER has a well-established infrastructure necessary to build a network information resource and has begun to develop an interoperability framework across its ecological data holdings.[18–20]

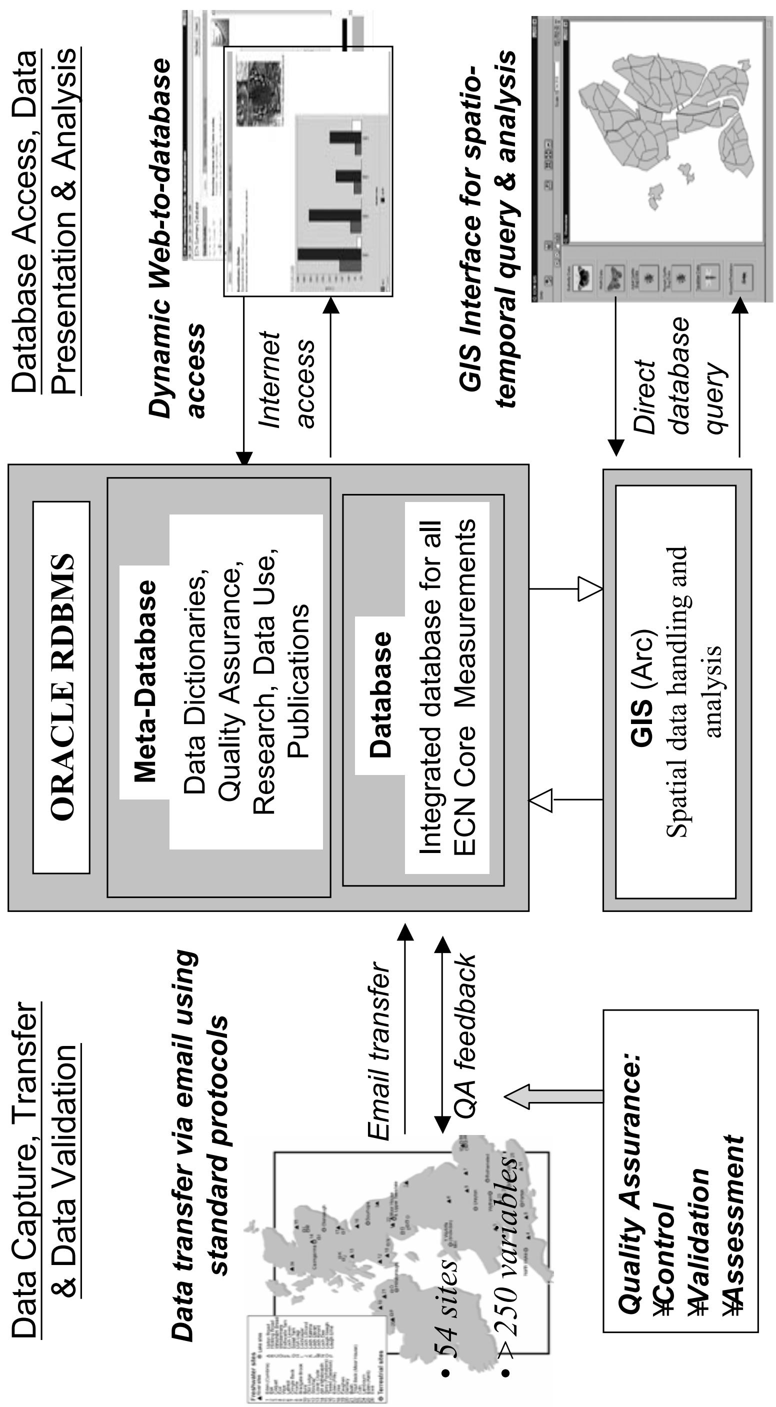


FIGURE 2.1 The U.K. Environmental Change Network (ECN) data management system.

2.2.3 Database Development

2.2.3.1 Data Definition and Specification

The choice of variables on which to base the monitoring of environmental change, and the methodology for sampling these variables, are subjects of other chapters in this book, but the data manager has an important role to play in these discussions. Baker et al.[19] describe how the involvement of data managers at an early stage in a U.S. LTER project ensured proper specification of data requirements. This is particularly true of environmental monitoring programs, where data management is often the catalyst that forces the proper specification of data objectives and highlights vagaries in sampling designs and data capture protocols. These protocols must include unambiguous and repeatable standard operating procedures and be designed to capture and deliver the required data, whether by manual or automatic means, according to stated objectives. They should include precise specifications for monitoring variables and guidelines on data management and data transfer. Standard formats for data handling and transfer should be designed to minimize error; Section 2.3 covers this in more detail. The U.K. ECN program's published protocols provide an example of this approach.[4,5]

The challenge for environmental monitoring data management is not the quantity of data but the diversity of data types and formats involved, from vertebrate populations to water quality and from field data to satellite imagery, across space and time dimensions. In devising specifications and data dictionaries for an integrated database, efforts should focus on a key subset of common parameters whose standardization will facilitate data interfacing,[7] for example, standards for expressing space and time dimensions, and standard code-lists for taxonomies and chemical determinants. The problems of integrating and analyzing data relating to different spatial and temporal scales is well known and much discussed among geographers and ecologists, and this issue has come to the fore with global change research. Information systems can at least facilitate scale transitions by supporting coding systems and data structures that permit easy movement through data hierarchies or changes in resolution through simple computation.

The Data Quality Objectives (DQO) process [21] is a planning approach to developing sampling designs for data collection activities that support decision making. The formal process is aimed at developing decision rules for selecting between two alternative conditions (e.g., compliance or noncompliance to a standard). Although some aspects are therefore difficult to apply to long-term environmental monitoring where the emphasis is on discovering patterns in the data, the DQO process embodies important principles for sampling design and data specifications. Examples include definition of the target population, minimizing bias, identifying practical constraints, defining data and metadata requirements, spatial and temporal dimensions, units of measurement, reporting precision, appropriate analytical methods, and detection limits.

2.2.3.2 Database Design

The requirement for greater functionality and complexity in data manipulation and analysis has increased the need for greater capability in database systems. Traditional relational database management systems (RDBMS) offer a predefined set of robust relational operations on a limited number of data types. To support more complex operations and data types within the database model along with reduced applications effort has required extension of the entity-relationship (ER) model[22] to include concepts from object-oriented design. Object-relational database management systems (ORDBMS) allow user-defined data types, with associated methods such as inheritance and encapsulation, to be constructed within the database model. Because object definitions plus the knowledge of their characteristics and functions are embedded within the database itself, the associated functionality is immediately accessible through any database interface. Ramakrishnan and Gehrke[23] provide a useful overview of object-oriented and object-relational database systems. The ANSI/ISO/IEC SQL:1999 standard[24] extended SQL to incorporate support for the object-relational model of data. These developments have lead to increased use of object-oriented modeling languages in database design, for example the Unified Modeling Language (UML).[25] This is not to say that an RDBMS implementation using ER modeling techniques is no longer appropriate as a database design method. ER and conventional normalization techniques still provide a powerful insight into the nature and issues involved in the data model implementation. However, the RDBMS implementation should be seen as a basic form of an ORDBMS implementation as the data model may need to be extended to include further object functionality at a future date.

The database design should be as generic as possible to provide sufficient flexibility for handling a wide range of data types and their relationships, plus later inclusion of new data types. At the same time, the design needs to be easily understood (particularly where scientists are using the database directly) and the structures easily interfaced with data access and analysis software. The meta-database will play a key role and is often the most important design to get right, as it provides the information, knowledge, and structures which form the hub of an integrated database.

It is important to distinguish between the logical data model produced through a process of normalization and the physical model of data tables to be implemented in the database system. The process of normalization should produce a clear understanding of the structure of the data and the issues involved in their dynamics.[22] This understanding should then be used to implement a physical data model that delivers the functionality the users require. If normal form is slavishly implemented, many complicated and time-consuming relational operations may be needed to construct the type of denormalized, derived data sets most commonly used in data analyses. These operations will have to be carried out *correctly* either manually by users or through the construction of a tailor-made interface. If the data in the database are of the "write-once, read-many" type commonly found in data archiving systems, then relaxation of the normal form to bring the physical implementation closer to the denormalized form required by the user might be desirable. Although this is an attractive strategy from the user's point of view, the database manager must recognize the dangers inherent in this approach. First, if the data are liable to change, all duplicate data in the physical data model must be kept synchronized through carefully constructed update procedures. These procedures will become more complicated as the data model evolves. Second, moving the physical data model toward one particular denormalized form will make the production of other derived data sets more difficult. For these reasons, denormalization issues tend to be handled through the creation of views of data tables, through applications code, or OLAP techniques (see Section 2.5.3.1). Either way it is important to recognize that the endpoint of the data modeling process is efficient use and maintenance of the data, not a perfect physical implementation of a logical model.

Specialized data models and techniques are required for spatial data to cater for multidimensional object definitions, surface representations, and relationships based on both distance and topology. Information systems for environmental monitoring will almost certainly need to provide for the handling of both vector (e.g., digital map data) and tessellation (e.g., satellite imagery) models of spatial data and related operations to provide for both object-oriented and location-oriented analyses.[26] Geographical information systems (GIS) and image-processing systems provide this functionality, but isolated development has inhibited system integration (see Section 2.2.3.3). Environmental change research requires the ability to integrate analyses across space and time, for which suitable data structures and functional systems are only just beginning to develop.[27–29]

An enormous body of literature exists on database and spatial database design and management; some key texts are referenced for this chapter.[22,23,30]

2.2.3.3 Software

Key considerations in choosing database management software for long-term environmental databases are reliability, security, flexibility, compatibility, and support. Cheaper open source software may be tempting, but there may be little support when things go wrong and no guarantee of continuity and upgrades in the longer term. Where data are sensitive, the system must be capable of handling variable access controls on database objects for different users. Rollback and automatic auditing are useful features. Multiuser access is clearly essential for a network information resource. The system should be capable of easy integration with applications software such as statistical packages, spatial analysis systems (GIS), and Web-based data access and display tools. The rather separate and isolated evolution of GIS resulted in an unhelpful divide between conventional database and spatial analysis systems. Environmental database managers and developers were faced with a choice of either using a database or a GIS, or duplicating data across both. To some extent, the difficulty was caused because conventional relational data structures were inefficient at supporting spatial object definitions. Better, more dynamically linked systems are now developing through development of standards for spatial geometries and object definitions (e.g., OpenGIS Consortium;[31] SQL/MM spatial) and their implementation by database software vendors (e.g., Oracle Spatial), coupled with a more open systems approach from GIS vendors.

2.2.3.4 Database Security and Recovery

The system needs to be secure against illegal or unintended access, deletions, or updates, while at the same time easily accessible to those authorized to use it. The two principle issues here are prevention through security schemas and recovery through back-up systems. Database and network security have expanded enormously in importance alongside the rapid spread of Internet access, and there is much literature and Web information on these topics. Security issues with respect to the Oracle RDBMS are discussed by Theriault and Heney.[32] Designing the database security schema is a priority as this forms the core element of the security system, with network-layer security software such as firewalls providing front-line protection. Monitoring accesses and maintaining audit trails are recommended to track access and to check for loopholes in the schema. Some simple precautions for environmental databases are provided by Nottrott.[33] Data managers should remember that the system needs to be protected from internal data management errors as well as the outside world; it is worth setting up different database management access modes to minimize write-enabled contact.

Perhaps the worst type of security breach for long-term databases is one which lies undiscovered and is only recognized during data analysis some months or even years later, after which several published papers have been written (the stuff of data manager nightmares!). Data verification rules (see Section 2.3.3) have a strong part to play here but, of course, they need to be run periodically on all the data in the database, not just on new data.

Back-up systems and audit trails provide the means of recovery from unscheduled changes. To keep storage space requirements to a minimum, back-ups are often overwritten after a specified time period, for example, daily backups overwritten monthly. However, it may be important to maintain periodic "snap-shots" of the database in perpetuity. Finally, original datasets including field sheets should be maintained wherever possible, preferably in a protected environment as the ultimate backup.

2.2.4 Institutional Issues

Any collaborative program that involves data collection or data exchange needs to establish an agreed data policy as early as possible to avoid later conflicts of interest or misinterpretations. A data policy should address:

- Data ownership and intellectual property rights
- Access protocols for data at different levels (raw/summarized/value-added):
  - Licensing requirements
  - Conditions of use
  - Waivers
  - Authorization: Who grants this and on what basis?
  - Security schemas and user authentication
  - Quality assurance — e.g., minimum requirements for data release
  - Charging policy
  - Time-period for exclusive use before data release, if appropriate
- Publication procedures: authorship, acknowledgment, etc.
- Data management procedures and standards, including back-up, security, archiving, and exchange formats
- Monitoring access to data and production of published material

International programs will need to accommodate local constraints and access legislation within their data policies. The design of data access systems should attempt to minimize the bureaucracy and bottlenecks which can develop during licensing and authorization of users, and maximize the degree to which the system can be automated. Requesting authorization can be a particular bottleneck, and it is important (1) to nominate an individual who is contactable on a regular basis (organizations are tempted to nominate a high-ranking official who is far too busy), and (2) to agree to a time limit for response. It may be possible to gain agreement for a certain level of summary data above which data can be made open access. ECN has adopted this strategy and was able to implement its Web-to-database access system without the need for specific security and authentication protocols (apart, of course, from the database level). Conditions of use clauses can be incorporated into the gateway for these systems, and the need for email to download data helps monitor use.[17]

Considering the urgent requirement for access to integrated data to understand complex environmental problems, it is surprising how many institutional philosophies still work in the opposite direction. Scientific reward systems can actually inhibit data sharing because they encourage competition through exclusive publication records. Programs intended to be collaborative have faltered because of competitive assessment mechanisms and the lack of incentive to share data and results. The U.S. National Committee for CODATA listed recommendations to address institutional barriers to data integration for global change research.[7] In summary, these are (1) reassessment of reward systems to encourage interdisciplinary research, (2) closer policy relevance, (3) giving data managers equal status with principle investigators and the opportunity to be involved from the start of a program, (4) early development of an agreed data policy and access protocols, (5) better communication among participating organizations, and (6) establishment of data and information analysis centers to derive added-value datasets and facilitate data exchange.

In the U.K., there have been additional barriers created through the encouragement of government departments and agencies to recover significant revenue through data licensing (for example, digital topographic data),[34,35] even by charging other government-funded organizations including universities and research institutes. This has been a considerable setback to the dissemination of information across public bodies and, indirectly, has been a factor that has inhibited integration of information and management practices. Revised and more enlightened policies, coupled with new legislation on freedom of information and the advent of the Web, are helping to reverse this process.

2.3 DATA QUALITY

The concept of quality is most useful when expressed as "fitness for purpose": an evaluation of a "product" (in this case data) in relation to a specified need. For environmental change research, which is concerned with the ability to distinguish signal from noise or real effects from measured artefacts, quality assurance (QA) procedures are critical. This means setting quality criteria and objectives at the outset of a data-gathering exercise, monitoring the process to assess how far these are being met, qualifying the data gathered, and providing feedback to enable procedures to be adjusted to keep the system on track.

Data of unknown quality are in effect as unreliable as poor quality data; little confidence can be placed in their analysis and interpretation. Users are typically unaware of the flaws that can arise in environmental data, the effects of compounding those errors through combining datasets,[36] and the limits to which datasets can be applied and interpreted. Spatial data are particularly vulnerable to misinterpretation, particularly where mapped interpolations and classifications are concerned. Metadatabases play an important role in providing information about the quality of data and how datasets have been derived; such metadata should always accompany data downloads.

The management of data quality should be an integral component of the quality system that applies to the monitoring program as a whole, from data capture to data delivery and presentation. The ISO 9000 series of standards[37] focuses on standardizing procedures for quality management. Although originally geared towards physical manufacturing products, ISO 9000 principles form a useful framework for developing QA systems for environmental monitoring.

The following is some useful terminology:

Quality assurance: A term embracing all planned and systematic activities concerned with the attainment of quality

Quality objectives: The specification of target values for quality criteria (such as accuracy and completeness)

Quality control: Operational techniques and activities that are used to fulfill requirements of quality

Quality assessment: Procedures for assessment of the degree to which quality objectives have been met after data capture — ideally providing evidence that they have been met. This should be a system of monitoring, which feeds back into the quality control process to maintain quality targets.

2.3.1 QA Procedures for Data Capture and Handling

The flow of data from capture to database, whether via manual or automated methods, should be designed to follow the standard operating procedures and support the quality criteria expressed in monitoring protocols, including the meta-information required to describe and qualify the data. Protocols should include data quality objectives and the quality control procedures that need to be applied in order to meet them, for example, guidance for handling of samples, and calibration and maintenance of automatic sensors. Guidelines for handling the data that derive from the field survey, sensor, or laboratory analysis must be clear for the less computer literate and designed to minimize error. Structures within the database should be designed to preserve the quality of the data and integrity rules used to prevent illegal entries, for example, duplicate records, null values in keys, or codes not present in a master list.

One of the principal sources of error for manual data capture is data entry. Field recording procedures and data entry systems should be designed together (they may be one and the same where field computers are used). The layout of field forms (whether on paper or field computer) should make recording as simple and straightforward as possible. They should reflect the method of survey to avoid wrestling with ten paper sheets in a gale where one would have done or unnecessary scrolling and jumping between screens. Transfer of data from field sheets to data entry screen in the laboratory should also be a simple read-across process; the system rather than the operator should perform the necessary transformation of data into suitable database structures. Transcription should be kept to a minimum. "Double punching" (a term from punch-card days), in which data are keyed in twice by independent operators and then computer-verified for agreement, is an excellent practice when resources allow, as subtle within-range typing errors can be detected this way. Standard data entry and transfer templates with in-built validation checks help prevent the input of "illegal" values and inconsistencies at an early stage, and ensure that data arrive at the data center in the expected format for further validation and database import.

Networked data import systems, for example, via the web, enable personnel at monitoring sites to submit and import data to the database directly, which speeds up the flow of data. However, as well as data verification checks, such systems must provide adequate explanation to users for failed data and proper indication of what should be done about such values. Automatic data capture systems (e.g., weather stations) can be linked via telemetry to provide data in near-real time to the database. However, the lack of manual intervention means that validation issues may not be resolved and the data should carry a "health warning."

Field forms, data entry systems and database design should not constrain the provision of information on factors which might affect data quality. Details of any deviations from standard procedures, faulty instrumentation, and problems during the sampling period should accompany the data transferred to the database and be stored as metadata, for example, through predefined quality codes attached to data records or, where these are insufficient, free text.

2.3.2 Missing Data and Uncertainty

Procedures for handling missing data values and uncertainty should be built into the data management system structures and documentation from the outset to avoid any ambiguity during data capture and interpretation. One recurring theme is the difference between missing values and zero, for example, the difference between "no catch" in the context of an insect trap that was not set (missing data) and "no catch" in the context of an insect trap set but no individuals found (zero). Protocols and data-handling guidelines should explain this important distinction in the context of the particular survey or sampling method. Missing data values may be recorded in the database using null values (sometimes frowned upon by relational database purists because of the "special" way the software may treat nulls) or through missing value codes. Codes enable direct look-up to explanations in the meta-database and work well with categorical data, but can be risky in numeric data fields if users are accessing data tables directly to generate summary statistics.

Laboratory analytical methods normally carry a limit of detection (LOD) below which a result cannot be accurately determined. Methods for both storing and analyzing these "less-than" values is another recurring theme, and the two should not be confused. Data managers need to ensure that this uncertainty is explicit within the database structures; what the researcher decides to do for their analysis is a separate issue which may depend on the statistical method. Whatever value is stored within the numeric field (often the detection limit itself) must be accompanied by a metadata flag identifying it as an LOD, and data downloads must include this metadata flag. Species identification is another common area where results can be uncertain. The taxonomic hierarchy can be used to handle different resolutions of identification, for example, recording at genus level where the species is not determined. However, there are often situations where the individual is known to belong to one or other species among many within the genus, and this information is lost if only the genus is recorded. At the very least, this additional information needs to be recorded within the metadata; some taxonomic databases may handle this information through explicit data structures.

2.3.3 Data Verification

Data verification is a key step in the quality assessment process. Data verification is sometimes called data validation, although a distinction between the two can be made as follows: data verification is the process of evaluating the compliance of data against the requirements of the method, and data validation attempts to assess the analytical quality of the data beyond the method as well as the impacts of its use.[38]

Data verification is concerned with screening data for contamination which may have occurred at any stage during data capture and handling. Edwards[39] explains that this will involve checking for (1) illegal values—values which are literally impossible, given the actual phenomenon observed, or (2) outliers—values which are unusually extreme, given the statistical model in use. It is important not to confuse the two. Simply because one has never observed a concentration below a given threshold and can't imagine it ever happening does not make it illegal.[39] Trapping illegal values may, for example, involve checking that count data include only integer numbers or comparing values against a master list, for example, a list of permitted categorical codes. Numeric range checks are often used as a "first-pass" procedure to detect possible outliers, which are then followed up through statistical investigation. Setting these ranges should also be the subject of statistical investigation: too broad and they miss important problems, too narrow and they throw up most of the data. Verification procedures might be designed to be site-specific, for example, to be sensitive to the variability in polluted as opposed to pristine sites or season-specific, based on expected seasonal patterns in the data. Multivariate checks can be made on variables which are known to be closely correlated. Visualization of data often shows up problems which are not apparent through range checking, for example, an instrument error which causes the same "valid" value to be recorded exclusively over a period of time. Visualization should ideally be part of the automated data verification process.[40] Other data verification procedures include formatting, logical integrity, and completeness checks.

An important issue is what to do with the outlying data values once detected and investigated. Long-term monitoring programs should adopt a very cautious approach to discarding data to avoid repeating the ozone hole mistake.[41] Data values should never be discarded unless there is a known and understood cause of contamination which has rendered the data meaningless, for example, a verified instrumentation fault. Otherwise, values identified as outlying should be kept and flagged up, with details provided in the meta-database. Where suspect values can be back-estimated, for example, for systematic errors like instrument drift, then both the original and the estimated data should be kept with both flagged up and described appropriately. Quality assessment procedures often involve performing an independent second survey or analysis as a check on correct application of the data capture methodology or on operator skills, for example, resurvey of vegetation plots or reanalysis of laboratory samples and interlaboratory trials.[8] Note that all sets of results may be equally valid; in any case, all should be maintained in the database with appropriate explanatory metadata.

Following up suspect data is a time-consuming business, especially when running a network of monitoring sites making frequent submissions of a broad range of data types. Even if verification procedures are well automated, intervention is often needed to follow up queries with site-based personnel. In addition, feedback to sites from the quality assessment phase should be provided to ensure systematic errors and deviations from the data capture and handling methodology are corrected. Logs or audit trails should be maintained on data processing verification and transformation procedures and on any changes which may subsequently be made when errors are detected at a later stage. These can be built within database itself for administrative purposes and linked to relevant data and quality information tables.

2.4 METADATA

The importance of metadata is now received wisdom. Datasets can be rendered useless unless accompanied by meta-information that at least describes their attributes, derivation, quality, and structure. Content standards for metadata have largely emerged through the development of spatial metadata and exchange standards since the early 1990s.[42–45] More recent versions of these standards have begun to cater explicitly to the time dimension, clearly important in environmental monitoring. The U.S. Federal Geographic Data Committee standard has been widely used internationally, and efforts are underway to harmonize this with the recently published ISO 19115 standard for geographic metadata.[46]

The U.S. Spatial Data Transfer Standard (SDTS) has set the basic framework for data quality that has been incorporated into U.S. and international efforts.[47] Data quality principles (ISO 19113[45]) are embedded in the developing international standards for geographic information. The principle quality criteria recognized by these standards are accuracy, completeness, lineage (derivation), and logical consistency (the fidelity of the logical relationships encoded in the database). Beard and Buttenfield[48] promoted a view which seemed particularly appropriate for environmental monitoring: that location, theme, and time should be seen as dimensions across which quality criteria can be applied; thus accuracy might apply to a coordinate (“positional accuracy”) or to an observation (theme) at that position (“attribute accuracy”) or to the timing of an event (“temporal accuracy”). Other related parameters which would normally accompany these quality criteria are resolution (the smallest quantity that can be recorded by a measurement device or method) and precision (the exactness with which a value is expressed).

2.4.1 Metadata Systems

The development of metadata systems has become an important and expanding part of environmental data management, especially since the increased visibility of data sources and direct access mechanisms through the Internet. There has been an explosion of metadata systems operating at the "data discovery" level that stop short at data center contact details; the data content and structures to support this are fairly generic. However, to provide direct access to data means providing the "technical" level metadata and knowledge necessary to use and interpret the dataset. Whereas meta-databases used to be regarded as a kind of data annex, they are now recognized as forming the hub of knowledge for an information system.

The description of the physical implementation of a database, the terminology used, and the logical structure of the data held are important features of a database which can often be quite specific to the nature of the monitoring program. The requirement placed on this metadata functionality by the use of the Internet is not only that this data must exist and be available to a global user community but also that the nature of this information must be sufficiently generic that it can be read and understood by a wide range of users, both human and machine. Metadata systems may be required to service automated requests for information in several formats as monitoring programs form global networks to address particular environmental issues. These requirements are by no means peculiar to environmental data management, and many Internet developments from other sectors are now and will continue to be relevant to this area. It is important that data managers seek out these developments (for example, through collaboration with the computer science and bibliographic communities) and do not attempt to reinvent the wheel.

Within the Internet environment, one particular database system is likely to be a member of a wider, federated network of peers. The federation allows peers to maintain their own internal data structures and protocols and take part in other networks. However, they must conform to common standards for communication of data requests and the servicing of these requests in a peer-to-peer model. A common protocol used in the peer-to-peer model is the ISO Z39.50 protocol for information search and retrieval. The protocol allows requests to be propagated across the network via Z39.50 client and server software at each node. This allows a user at any one node (or outside the network) to request information from catalogs of data entries held and maintained at all other nodes. This system is consistent with the federation ideas mentioned previously.

2.4.2 Extensible Markup Language (XML)

Having created a catalogue of data holdings that can be accessed over a network, a method is required to inform potential users of content and structure of those holdings. This method can then also be used to enable the transfer of data to and from other applications. Standards for the description and manipulation of document content are being rapidly transformed by the use of XML and associated Web-based standards.

XML provides a mechanism for describing the content of documents using a tagging system similar to HTML. It has been designed to exchange documents over the Internet and is increasingly used in other application domains. XML complements rather than replaces HTML. HTML uses predefined tags which define how information is to be formatted and displayed. XML tags are unlimited and user defined, and are used to describe the structure and contextual meaning of information. XML can be used to represent any kind of structured or semistructured document, e.g., papers, Web pages, database schemas, etc. This means that the structure of documents can be mapped, enabling the rapid and accurate location of relevant information. XML is more than just a markup language; it is a metalanguage which can be used to define new markup languages. It enables users to create a language designed specifically for their own application or domain. One example in the environmental domain is Ecological Metadata Language (EML).[49]

Database systems have adopted XML technologies to enable more flexible exchange of data between applications. The power of XML to describe the intended content and structure of data makes it ideal for enabling the transfer of metadata and data between applications when used in combination with search and retrieval protocols such the Z59.50 example described above. Examples of such networks can be seen via the Australia New Zealand Land Information Council (ANZLIC)[50] and the Australian Spatial Data Directory (ASDD),[51] and via the Climate and Environmental Data Retrieval and Archiving System (CERA)[52] at the Potsdam Institute for Climate Impact Research.

Protocols for XML descriptions are evolving rapidly via standards such as the World-Wide-Web Consortium (W3C)[53] XML schema standard and the concept of Web services. The idea of the Web service extends the previously described network model to include a services registry of nodes offering data operations. The services in the registry are a well-defined set of callable procedures. Web service descriptions and remote procedure calls between applications of this type are supported by W3Cs Web Services Description Language (WSDL) and Simple Object Access Protocol (SOAP), respectively. Using these developments, a metadata system can be developed that offers publication of data services to the global community via an Internet-based services registry and remote invocations of those services, not only by human users but also by automated applications. This is particularly relevant to the idea of a Web portal where a user can access many different services from a single Web page.

2.5 DATA ACCESS, EXPLORATION, AND ANALYSIS

Environmental information systems should allow researchers and decision makers access to heterogeneous data sources in a form suitable to address their particular questions. They should provide convenient navigation aids and analytical tools that help users to take advantage of information without requiring them to understand the sometimes-complex structures and relationships within a database.[54] Ideally, they should be able to integrate datasets from distributed sources, supporting the interoperability of these autonomous databases without requiring their physical centralization, adding value and allowing a more complete picture of the ecosystem to be developed.

2.5.1 Users of Environmental Information

Although routine collection of data from environmental monitoring programs has the potential to provide a wealth of information on the state of the environment, data from these programs can often be under-used because of (1) the difficulty of accessing the data, (2) the difficulty in assessing whether the data contain anything of interest or relevance in the time available to the user, and (3) the delay between collecting the data and its publication in a conventional form.[55]

The challenge facing managers of environmental databases is to provide data access systems to suit the often-different requirements of members of their user community. These could include, for example:

1. Scientific researchers who require access to raw data for detailed analysis of spatial and temporal patterns in the data.
2. Information brokers (e.g., sponsors of research programs or policy-makers) who may not necessarily require access to the high-resolution data, but who need summaries and interpretations of scientific research on which to base policy decisions. They may require guided access to summary data for inclusion in reports or interpreted information about trends in the data, for example, whether a pattern indicates a long-term trend that is likely to continue.
3. General public and schools that are likely to want easy-to-understand, interpreted information on environmental change, such as environmental indicators which enable easy assessment of how aspects of the environment are changing and how these changes could affect their lives (for example, see a set of climate indicators in the U.K.[56]). They may also want guided access to summary data if they wish to explore the data further.

2.5.2 Web-Based Data Access

The growth of the Internet has dramatically changed the way in which information can be managed and accessed. Users can retrieve information stored in thousands of interconnected networks across the world. The World Wide Web has become the obvious development platform for data discovery, query, and delivery systems, and is a highly accessible broadcast medium where up-to-date information in many different formats from different sources can be explored at low cost for the user. It is highly versatile: text, graphics, data, sounds, video, and other multimedia components can all be incorporated within a single system.

Web technology can be used to develop dynamic interfaces to environmental databases. Custom-built interfaces can be created either by using web integration tools provided with many database management, GIS and analytical software systems, or through the use of programming languages such as Java, ASP.NET, and other scripting languages. For example, the ECN Web site[57] includes an interface which enables users to build their own database queries by selecting any combination of ECN sites, core measurement variables and date ranges for dynamic generation of graphs and cross-tabulations, or for data download via electronic mail to a local machine.[17]

Other features on this site include near real-time access to data from automatic weather stations at ECN's sites and a dynamic set of climate change indicators for the U.K.

Many other monitoring programs provide access to their data over the Internet, for example, the U.S. Long-Term Environmental Research Program.[58] Others use the Internet to collect and display data from community-based monitoring programs, for example, the Canadian Ecological Monitoring and Assessment Network[59] or the GLOBE program.[60]

2.5.3 Data Analysis

Technologies such as online analytical processing and data mining have been developed over the last decade, originating in the business community. These are tools designed for information analysis and could have considerable potential for use in environmental information systems, for which they are only now beginning to be explored and applied.

2.5.3.1 Online Analytical Programming (OLAP)

OLAP is a term coined by Codd.[61] It is designed for the dynamic synthesis, analysis, and consolidation of large volumes of multidimensional data. OLAP systems use multidimensional structures either physically to store data or temporarily to hold data from databases or data warehouses. Multidimensional structures are best visualized as cubes of data with each side of the cube as a dimension; for example, a simple cube could include species, location, and time as dimensions (Figure 2.2). Cubes can be expanded to include other dimensions. These provide a compact and easy-to-understand way of representing data.

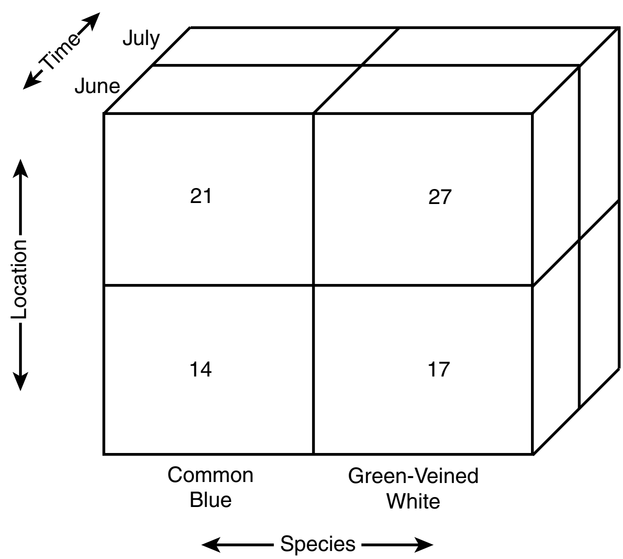
OLAP software uses these multidimensional views to provide quick access to information and allow users to analyze the relationships between data items, looking for patterns and trends.[62] The software aims to insulate users from complex query syntax so they do not have to understand the database structure or construct complex table joins. There are a wide range of OLAP functions such as navigation through levels of aggregation in the data; for example, data for a time dimension could be aggregated in years, quarters, months, weeks, and days (this is often called "drill up/down"). Another common OLAP function is termed "pivoting," which refers to the ability to look at data from different viewpoints. For example, one slice may look at all the butterfly species found at a particular location. Another slice could display data for a particular species over a year. This type of analysis is often performed along a time axis to analyze trends or find patterns in the data.

2.5.3.2 Data Mining

Data mining is the process of extracting valid, previously unknown information from large databases using a variety of techniques.[63] It is concerned with the analysis of data and the use of software techniques, including statistical, mathematical, and artificial

Location	Butterfly Species	Month	Count
A	Common blue	June	21
A	Common blue	July	32
A	Green-veined white	June	27
A	Green-veined white	July	37
B	Common blue	June	14
B	Common blue	July	23
B	Green-veined white	June	17
B	Green-veined white	July	28

(a) relational form



(b) a three-dimensional cube

FIGURE 2.2 Multidimensional data views.

intelligence technologies, to find patterns in sets of data.[64] Common statistical methods used include regression (linear, logistic, and nonlinear), discriminant analysis, and cluster analysis. However, as data mining becomes more common, machine learning techniques such as neural networks and rule induction are also getting more consideration.[65] Although complex in their own way, these methods require less statistical sophistication on the part of the users.

Neural networks are an approach to computing that involves developing mathematical structures with the ability to learn. They can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Rule induction is a complementary technique. Working either from the complete dataset or a subset, a "decision tree" is created representing a series of rules that lead to a class or value. Each class has a unique pattern of values which forms the class description and makes it possible to predict the class of unseen objects.

These rules provide a useful, comprehensible "audit trail" for nonspecialists to understand how the results were obtained.

Data visualization helps understand and explore information in large, complex datasets. It can be a powerful way to sift through information and spot patterns and trends or identify outliers, and can be useful for quality assurance of data as well as data analysis. Visualizing data is not a new idea, but the rapid development of computer graphics techniques and analytical tools, together with the growth in computing power, can now support real-time interactive environments, which have spawned an explosion of user interest from all sectors. Visualization is increasingly seen as an integral component of an information system. Advanced techniques are being applied in environmental data analysis and data mining, and in environmental modeling to allow users to test scenarios of change, for example, adjusting model driving variables and viewing the impacts of policy or customizing landscapes. Virtual reality approaches enable immersion in virtual worlds supporting not only visual awareness but other sensory interactions.

2.6 DEVELOPING TECHNOLOGIES FOR NETWORKED INFORMATION SYSTEMS

A growing requirement in global change research is the ability to integrate data from distributed sources in order to allow multidisciplinary, multiscale, and geographically extensive analyses. Research should theoretically be unconstrained by the way data are organized in regional, national, and thematic databases. Science is now being undertaken by "virtual organizations," whereby teams who may belong to several different institutions located around the world can rapidly combine resources to work on a specific project. The mechanisms enabling this also allow groups to reform in different combinations to work on other projects. Each project will dynamically call on distributed computing and information resources, depending on the nature of the question being addressed.

Developments in networking technology mean that Internet access to a virtual information resource composed of linked and harmonized distributed databases is becoming possible. The new challenge facing data managers and scientists is how databases can be dynamically linked, not just technically but (probably more difficult) in terms of the data themselves to enable comparative analyses across theme, time, and space. The solution lies in devising methods for describing terms and their relationships between databases which can be used to translate and transform data into comparable integrated forms. Whereas this activity has traditionally been carried out by scientists and statisticians according to need, to achieve truly "interoperable" databases these descriptions must be embedded in and used dynamically by the system in response to user query. This implies the use of intermediate software agents that can take descriptive information, such as taxonomies of data items and inference rules, to build the required derived data set from separate data resources. Development of these semantic capabilities is described below. Using such facilities allows accesses to and transitions between data sources, providing a unified data discovery, query, and delivery interface across distributed data sources.

2.6.1 Grid-Distributed Computing

The grid is an emerging infrastructure that enables computers and networks to be linked together into a seamless common resource. The initial impetus came from the supercomputing sector and was focused on sharing computer power. However, the potential for linking distributed heterogeneous information resources via data grids, and linking these with appropriate application software has far-reaching implications for global research, and this is now a rapidly developing area of technology. The grid will be a secure, flexible infrastructure for sharing geographically distributed resources, such as computing, data and software services among groups of individuals,[66] and will enable the development of powerful tools to access and combine data from distributed data sources, providing toolkits for analyses on-the-fly.[67,68] The term "E-science" refers to the collaborative science that will be enabled by this new technological infrastructure.

Users should not necessarily need to know what resources they are using or where they are located, and they should be able to access distributed data in a uniform fashion. From the user perspective, the operation of the grid is as follows:

- A user submits a request via a web portal specifying the requirements of the query.
- Grid services seek appropriate data and software tools, and allocate resources, such as computing power and storage space, in order to satisfy the request.
- Security and authentication checks are carried out for access to data and resources.
- The progress of the request is monitored.
- Results (information, analyses, dynamic visualizations, etc.) are presented to the user.

A range of "middleware" (software that connects applications) platforms are being developed, together with emerging standards such as Open Grid Services Architecture (OGSA). Among these platforms are Globus[69] and Unicore,[70] which include resource allocation and authentication tools. Currently, most data-grid development activities are limited to data stored as flat files. However, an exception is the San Diego Supercomputer Center (SDSC) Storage Resource Broker (SRB)[71] which enables scientists to create, manage, and access unified virtual data collections located on distributed heterogeneous data resources, including relational databases. The Global Grid Forum working group on Database Access and Integration Services (DAIS-WG)[72] seeks to promote standards for the development of grid database services, focusing principally on providing consistent access to existing, autonomously managed databases. However, there is still some way to go before these developments are able to support the type of seamless and flexible access to resources envisaged by E-science.

2.6.2 The Semantic Web and Knowledge-Based Data Systems

Combining distributed data resources across the Internet requires embedded knowledge of the terms and relationships used within and across those resources. A formal description of the meaning of data content not only provides better quality of information for the human reader but also provides the material for software applications to automatically interpret and draw inference of how data can or cannot be integrated in answer to a user request. A key area of development is work on standards for the Semantic Web through the working groups of the World Wide Web Consortium (W3C).[53] The idea is to add knowledge representation and semantic descriptions to Web resources, allowing logic questions, choices, and inferences to be made about their content. The use of ontological languages such as the W3C's Ontological Web Language (OWL) standard is one such development for knowledge representation.

The idea of adding meaning and logic to data resources has been extended to the grid environment, and it is likely that the developments from the Semantic Web will feed into grid technologies. Such grid-based facilities would begin to provide the tools for researchers to combine different resources and services dynamically across the Internet to answer complex questions requiring access to many different resources. The job of the data manager is then extended to include the representation of knowledge on how to use the data in semantic descriptions and ontologies. This is an extension of metadata toward more formal structures that will enable the type of interoperability that is required by Internet-based researchers. At this time systems capable of such dynamic integration of environmental data over the Internet are developmental. One new initiative in this area is the Science Environment for Ecological Knowledge (SEEK) program.[20]

2.7 CONCLUDING REMARKS

As data and information management is promoted towards center stage, huge demands are placed on data managers and information system developers. They must not only design systems to capture, manage, and provide access to data within one particular database but also enable that resource to be employed in combination with other resources using Internet-based technologies. In environmental monitoring in particular, they must understand enough about each topic area to build appropriate data models and structures and cater to the increasingly complex needs of scientists. The raised profile of quality assurance requires them to be responsible for the management of data quality throughout the information system.

These demands and developments need resources and a change of culture. If environmental science is to reap the benefits of new information technology, then it needs to regard information management as an integral part of its science strategy and make appropriate investments for the future. This is slowly beginning to happen as scientists see the benefits of new information technology and as the role of IT professionals is expanding. Scientific programs are increasingly urged to involve IT and data management specialists at the outset to define information management needs and to allocate resources accordingly. The recent U.K. E-science program was launched in recognition of the need to encourage links between science and IT. There are now obligations for scientific organizations to manage and provide access to their data resources to approved standards and to present research results via easy-to-use Web-enabled interfaces.

The complexity and scope of environmental change research demands a comprehensive and integrated approach to science and information management that breaks down the barriers between disciplines, roles, and technologies. For example, to resolve issues of sustainable development, researchers will wish to answer questions that cut across sociological, economic, and environmental fields, combine data from different sources, and use a range of analytical and visualization tools. A meeting of minds between environmental scientists, data managers, data analysts, and technologists is required to promote a more fluid system of transforming data into information and knowledge for research and policy. Devising appropriate methodologies that take advantage of new technology is the focus of the emerging disciplines of environmental and ecological informatics. Together with grid-based computing, these developments offer a huge step forward for science in providing flexible and dynamic access to integrated resources for exploration and analysis of large and complex distributed data sources.

REFERENCES

1. Burt, T.R., Long-term study of the natural environment — perceptive science or mindless monitoring? *Prog. Phys. Geogr.*, 18(4), 475, 1994.
2. Tinker, P.B., Monitoring environmental change through networks, in *Long-Term Experiments in Agriculture and Ecological Sciences*, Leigh, R.A. and Johnston, A.E., Eds., CAB International, Wallingford, U.K., 1994, 407.
3. Michener, W., Data management and long-term ecological research, in *Research and Data Management in the Ecological Sciences*, Michener, W.K., Ed., Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina Press, Columbia, 1986, 1.
4. Sykes, J.M. and Lane, A.M.J., Eds., *The United Kingdom Environmental Change Network: Protocols for Standard Measurements at Terrestrial Sites*, Her Majesty's Stationery Office, London, 1996.
5. Sykes, J.M., Lane, A.M.J., and George, D.G., Eds., *The United Kingdom Environmental Change Network: Protocols for Standard Measurements at Freshwater Sites*, Centre for Ecology and Hydrology, Huntingdon, U.K., 1999.
6. Integrated Monitoring Programme Centre, *Manual for Integrated Monitoring,* UN ECE Convention on long-range trans-boundary air pollution, International Co-operative Programme on Integrated Monitoring on Air Pollution Effects, Finnish Environment Institute, Helsinki, 1998.
7. U.S. National Committee for CODATA, *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data.* National Academies Press, Washington, D.C., 1995.
8. Lane, A.M.J., The U.K. Environmental Change Network database: an integrated information resource for long-term monitoring and research, *J. Environ. Manage.*, 51, 87, 1997.
9. Gurtz, M.E., Development of a research data management system: factors to consider, in *Research and Data Management in the Ecological Sciences*, Michener, W., Ed., Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina Press, Columbia, 1986, 23.
10. Stafford, S.G., Brunt, J.W., and Michener, W.K., Integration of scientific information management and environmental research, in *Environmental Information Management*

and Analysis: Ecosystem to Global Scales, Michener, W.K., Brunt, J.W. and Stafford, S.G., Eds., Taylor & Francis, London, 1994, 3.
11. Harmancioglu, N.B., Alpaslan, M.N., and Singh, V.P., Needs for Environmental Data Management, in *Environmental Data Management*, Harmancioglu, N.B., Singh, V.P., and Alpaslan, M.N., Eds., Kluwer Academic, Netherlands, 1998, 1.
12. Strebel, D.E., Blanche, W.M., and Nelson, A.K., Scientific information systems: A conceptual framework, in *Environmental Information Management and Analysis: Ecosystem to Global Scales*, Michener W.K., Brunt J.W., and Stafford, S.G., Eds., Taylor & Francis, London, 1994, 59.
13. Boehm, B., A Spiral Model of Software Development and Enhancement, ACM SIGSOFT Software Engineering Notes, 1986.
14. Boehm, B. et al., Using the WinWin spiral model: a case study, *IEE Computer*, 18, 33, 1998.
15. Brocklebank, M., Lane, A.M.J, Watkins, J.W., and Adams, S., Access to the Environmental Change Network Summary Database via the World Wide Web, ECN Technical Report 96/1, Grange-over-Sands, Cumbria, U.K., 1996.
16. Lane, A.M.J. and Parr, T.W., Providing information on environmental change: data management strategies and Internet access approaches within the UK Environmental Change Network, in *Proceedings of the 10th International Conference on Scientific and Statistical Database Management*, Rafanelli, M. and Jarke, M., Eds., IEEE Computer Society, Los Almitos, CA, 1998, 244.
17. Rennie, S.C., Lane, A.M.J., and Wilson, M., Web access to environmental databases: a database query and presentation system for the UK Environmental Change Network, *Proceedings of ACM Symposium on Applied Computing*, Como, Italy, 2000, 894.
18. Brunt, J.W., The LTER network information system: a framework for ecological information management, in *North American Science Symposium: Toward a Unified Framework for Inventorying and Monitoring Forest Ecosystem Resources*, Aguirre-Bravo, C. and Franco, C.R., Eds., Guadalajara, Mexico, 1999. Proceedings RMRS-P-12. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 1999, 435–440.
19. Baker, K.S. et al., Evolution of a multisite network information system: the LTER information management paradigm, *BioScience*, 50(11), 963, 2000.
20. Partnership for Biodiversity Informatics: Enabling the Science Environment for Ecological Knowledge (SEEK), Proposal to the National Science Foundation, 2002; information available online at http://seek.ecoinformatics.org.
21. U.S. Environmental Protection Agency, Guidance for the Data Quality Objectives Process, EPA QA/G-4, Office of Environmental Information, Washington, D.C., 2000.
22. Date, C.J., *An Introduction to Database Systems*, 7th ed., Addison-Wesley, Reading, MA, 2000.
23. Ramakrishnan, R. and Gehrke, J., *Database Management Systems*, 3rd ed., McGraw-Hill, New York, 2003.
24. Melton, J. and Simon, A.R., *SQL:1999: Understanding Relational Language Components*, Morgan Kaufman, San Francisco, 2002.
25. Booch, G., Jacobson, I., and Rumbaugh, J., *The Unified Modelling Language User Guide*, Addison-Wesley, Reading, MA, 1998.
26. Peuquet, D.J., Issues involved in selecting appropriate data models for global databases, in *Building Databases for Global Science*, Mounsey, H. and Tomlinson, R.F., Eds., Taylor & Francis, London, 1988, 66.
27. Snodgrass, R.T. and Ahn, I., A taxonomy of time in databases, *Proceedings ACM SIGMOD International Conference on Management of Data*, Austin, TX, 236, 1985.

28. Jenson, C.S. et al., A glossary of temporal database concepts, *SIGMOD Rec.*, 21(3), 35, 1992.
29. Langran, G., *Time in Geographical Information Systems*, Taylor & Francis, London, 1992.
30. Burrough, P.A., *Principles of Geographical Information Systems for Land Resources Assessment*, Monographs on Soil and Resources Survey, Clarendon Press, Oxford, 1987.
31. OpenGIS Consortium; information available online at http://www.opengis.org.
32. Theriault, M. and Heney, W., *Oracle Security*, O'Reilly & Associates, Sebastopol, CA, 1998.
33. Nottrott, R., Communications and Networking, in *Data and Information Management in the Ecological Sciences: A Resource Guide*, Michener, W.K., Porter, J.H., and Stafford, S.G., Eds., LTER Network Office, University of New Mexico, Albuquerque, NM, 1998.
34. Department of the Environment, Handling Geographic Information, Report to the Secretary of State for the Environment of the Committee of Enquiry into the Handling of Geographic Information, HMSO, London, 1987.
35. Rhind, D.W., Data access, charging and copyright and their implications for geographical information systems, *Int. J. Geogr. Inf. Syst.*, 6(1), 13, 1992.
36. Goodchild, M.F. and Gopal, S., *The Accuracy of Spatial Databases*, Taylor & Francis, London, 1989.
37. International Standards Organisation (ISO), ISO 9000 series: Quality Management, ISO, 2000.
38. U.S. Environmental Protection Agency, Guidance on Environmental Data Verification and Data Validation, EPA QA/G-8, Office of Environmental Information, Washington, D.C., 2002.
39. Edwards, D., Data quality assurance, in *Ecological Data: Design, Management and Processing*, Michener, W.K. and Brunt, J.W., Eds., Blackwell, Oxford, U.K., 2000, 70.
40. Chapel, S.E. and Edwards, D., Automated smoothing techniques for visualisation and quality control of long-term environmental data, in *Environmental Information Management and Analysis: Ecosystem to Global Scales*, Michener W.K., Brunt J.W., and Stafford, S.G., Eds., Taylor & Francis, London, 1994, 141.
41. Stolarski, R.S. et al., Nimbus 7 satellite measurements of the springtime Antarctic ozone decrease, *Nature*, 322, 808, 1986.
42. National Committee for Information Technology Standards (NCITS), Spatial Data Transfer Standard (SDTS), ANSI NCITS 320-1998, American National Standards Institute (ANSI)/U.S. Geological Survey, Rolla, MO, 1998.
43. Federal Geographic Data Committee (FGDC), FDGC-STD-001-1998 Content Standard for Digital Geospatial Meta-data (revised June 1998), FDGC, Washington, D.C., 1998.
44. Association for Geographic Information (AGI), Guidelines for Geographic Information Content and Quality, AGI, London, 1996.
45. International Standards Organisation (ISO), ISO 19113:2002 Geographic Information — Quality Principles, ISO, 2002.
46. International Standards Organisation (ISO), ISO 19115 Geographic Information — Metadata, ISO, 2003.
47. Chrisman, N.R., Metadata required to determine the fitness of spatial data for use in environmental analysis, in *Environmental Information Management and Analysis: Ecosystem to Global Scales*, Michener W.K., Brunt J.W., and Stafford, S.G., Eds., Taylor & Francis, London, 1994, 177.

48. Beard, M.K. and Buttenfield, B.P., Spatial, statistical and graphical dimensions of data quality, *Comput. Sci. Stat.*, 24, 408, 1992.
49. Ecological Metadata Language; information available online at http://knb.ecoinformatics.org/software/eml
50. Australia and New Zealand Spatial Information Council (ANZLIC); information available online at http://www.anzlic.org.au/
51. Australian Spatial Data Directory (ASDD); information available online at http://www.auslig.gov.au/asdd/
52. Climate and Environmental Data Retrieval and Archive (CERA); information available online at http://www.dkrz.de/forschung/project/cera.html
53. The World Wide Web Consortium; information available online at http://www.w3.org.
54. Hale, S.S., Bahner, L.H., and Paul, J.F., Finding common ground in managing data used for regional environmental assessments, *Environ. Monit. Assess.*, 63(1), 143, 2000.
55. Parr, T.W. and Hirst, D.J., The UK Environmental Change Network and the Internet: their role in detecting and interpreting environmental change, in *Advances in Sustainable Development. Environmental Indices: Systems Analysis Approach*, EOLSS Publ., Oxford, UK, 1999, 223–236.
56. Cannel, M. et al., Indicators of Climate Change in the UK, Report to the Department of the Environment, Roads and Transport, 1999; information available online at http://www.nbu.ac.uk/iccuk/
57 United Kingdom Environmental Change Network (ECN); information available online at http://www.ecn.ac.uk
58. United States Long-Term Ecological Research Network (LTER); information available online at http://lternet.edu
59. Canadian Ecological Monitoring and Assessment Network (EMAN); information available online at http://www.eman-rese.ca/eman/
60. The GLOBE Programme; information available online at http://www.globe.gov
61. Codd, E.F., Codd, S.B., and Salley, C.T., Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate, Arbor Software, 1993; available online at http://www.hyperion.com/products/whitepapers/
62. Forsman, S., OLAP Council White Paper, 1997; available online at http://www.olapcouncil.org/research/whtpaply.htm.
63. Simoudis E., Reality check for data mining, *IEEE Expert*, October 26, 1996.
64. Hand, D.J., Statistics and data mining: intersecting disciplines, *SIGKDD Explorations*, 1(1), 16, 1999.
65. Chen, M., Han, J., and Yu, P.S., Data mining: an overview from a database perspective, *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866, 1996.
66. Chervenak, A., Foster, I., Kesselman, C., Salisbury, C., and Tuecke, S., The data grid: towards an architecture for the distributed management and analysis of large scientific datasets, *J. Netw. Comput. Appl.*, 1999.
67. Foster, I., Kesselman, C., and Tuecke, S., The anatomy of the grid: enabling scalable virtual organizations, *Int. J. Supercomput. Appl.*, 2001.
68. Foster, I., Kesselman, C., Nick, J.M., and Tuecke, S., The physiology of the grid: an open grid services architecture for distributed systems integration, Open Grid Service Infrastructure WG, Global Grid Forum, 2002.
69. The Globus Project; information available online at http://www.globus.org
70. The Unicore Forum; information available online at http://www.unicore.org
71. San Diego Supercomputer Center (SDSC) Storage Resource Broker (SRB); information available online at http://www.npaci.edu/DICE/SRB/
72. The Global Grid Forum; information available online at http://www.ggf.org/

3 Using Multimedia Risk Models in Environmental Monitoring

C. Travis, K.R. Obenshain, J.T. Gunter, J.L. Regens, and C. Whipple

CONTENTS

3.1 INTRODUCTION

Taking environmental measurements is an essential component of any environmental program. The motivational force behind taking environmental samples is to obtain the knowledge necessary to limit current or future health or environmental impacts resulting from contamination in the environment. Thus, environmental monitoring is closely linked to risk assessment. Risk assessment is a methodological approach frequently used to evaluate the health and environmental impacts of environmental pollutants. The National Research Council (NRC) defines risk assessment as an "evaluation of

1-56670-641-6/04/$0.00+$1.50

information on the hazardous properties of substances, on the extent of human exposure to them, and on the characterization of the resulting risk." Risk assessment is not a single, fixed method of analysis. Rather, "it is a systematic approach to organizing and analyzing scientific knowledge and information of potentially hazardous activities or substances that might pose risks under specified conditions."[1]

A cornerstone of risk analysis is the frequent use of computerized analytical tools, such as multimedia environmental transport models, to organize and evaluate the large quantities of information necessary to understand the present and future implications of contamination at a site. These models are considered key analytical tools for assessing potential, current, and future human and ecological exposure to environmental contaminants.[2,3] The multimedia approach involves tracking contaminants from sources through multiple environmental media (e.g., air, water, soil, and biota) to points of human and ecological exposure.

Characterization costs tend to be significant cost-drivers in most environmental remediation or decommissioning programs. Historically, characterization costs may run as high as 40% of the total project cost.[4] Deciding on the appropriate amount of environmental sampling involves balancing the need for sound environmental data upon which to base decisions and the high cost of obtaining the samples. Multimedia models are increasingly being used as an aid in this balancing act. Multimedia models provide a vehicle for organizing and evaluating existing and new environmental data in a manner that allows understanding of complex situations and contributes to deciding what additional data are needed for the decision at hand. Drivers for the increased use of multimedia models are (1) the need to ensure that the type, quality, and quantity of the data collected are appropriate for the decision-making process and (2) the need for increased flexibility in the data-gathering process to decrease the time and cost involved in performing site characterization.

3.2 MULTIMEDIA MODELS

Multimedia models are analytical tools that employ complex systems of equations to assessing potential human and ecological risks from current and future exposures to environmental contaminants.[2,3] They incorporate both empirical and numerical techniques to estimate risks from potential exposure to a contaminant. Typically, multimedia models simulate contaminant dispersion, transport, and fate from sources through multiple environmental media (e.g., air, water, soil, biota) to estimate potential exposure, dose, and risk to onsite or offsite targets through exposure pathways. They typically include the following components:[5]

- Source term component that predicts the rate and quantity of contaminate released
- Fate and transport component that simulates contaminant movement and final disposition within and among media
- Exposure component that estimates the dose received by a target population in contact with the contaminant
- Effects component that predicts the impact of health of the target population based on the dose of the contaminants received

Multimedia models offer some important advantages.[5,6] They require risk assessors to:

1. Conceptualize the relationships between environmental contaminants and potential exposures through many media
2. Provide a framework for resolving conflicts
3. Allow stakeholder values and beliefs to be incorporated into the risk assessment process
4. Clearly express relationships in a simplified way so that the processes of contaminant dispersion, transport, fate, exposure, and uptake can be simulated in parallel for many media
5. Enhance understanding of how exposures may occur
6. Consolidate engineering and scientific knowledge in a way that can be replicated, can be used to test theories, and provide insights
7. Incorporate interactions within and among media
8. Quantify the sensitivity of estimated outputs to simplifying assumptions as well as measured and estimated model parameters

These advantages make multimedia models valuable tools for risk assessment and for planning environmental monitoring programs.

Multimedia models were originated to meet the assessment requirements and needs for specific sites. During the risks assessment process, risks assessors typically want to use multimedia models to gain insight into what are the most significant sources of contamination, which contaminants contribute most risks, and what are the most important pathways of exposure.[7] Many multimedia models evolved from their original purpose to provide these insights for a broader range of hazard classes (e.g., radionuclides, industrial chemicals, landfill contaminants, power plant emissions, etc.) and sites. It is important to know the origins of the models in order to select the models that will appropriately simulate contaminant dispersion, fate, and transport processes. The models must consider the targets of concern (e.g., onsite remediation workers, offsite residents, sensitive ecological species, etc.) and the pathways most likely to expose the target (groundwater, surface water, soil ingestion, dust inhalation, accumulation in foodstuffs, dermal absorption, "shine," etc.). Model flexibility is also an important consideration. The models selected must be able to consider complex geology, terrain, exposure pathways, and regulatory requirements. Finally, models must conform to guidelines set forth by regulatory oversight agencies.[7]

3.3 MULTIMEDIA MODEL SELECTION

Since their introduction in the late 1970s, multimedia models have gained growing use as tools to assess potential current and future human and ecological exposure to contamination at a site.[8] The multimedia approach involves using a combination of empirical expressions and analytical numerical solutions to differential equations to estimate relationships among contaminant release rates, pollutant transport, intermedia fluxes, and subsequent exposure. Multimedia models typically have been designed to combine these estimates of human exposure via ingestion, inhalation, external radiation, and dermal absorption with measures of potency for cancer and noncancer endpoints to assess risks.[9,10]

Numerous multimedia models are utilized to estimate exposure and risk from the transport and fate of chemical and radiological contaminants at sites in the U.S. Department of Energy (DOE) nuclear weapons complex, current and former U.S. Department of Defense (DoD) installations, and private sector facilities.[11] Because of their ability to consider multiple pathways and to compute environmental transport across variable timeframes, three multimedia models have come to be widely used at complex sites: Multimedia Environment Pollutant Assessment System (MEPAS), Residual Radiation (RESRAD), and Multimedia Contaminant Fate, Transport, and Exposure Model (MMSOILS). MEPAS, developed by Pacific Northwest National Laboratory, MMSOILS, developed by the U.S. Environmental Protection Agency (EPA), and RESRAD, developed by Argonne National Laboratory, are multimedia environmental transport models that consider transport of contaminants in air, surface water, and groundwater, and bioaccumulation in plants and animals. Table 3.1, derived from the benchmarking report by Cheng et al.,[12] summarizes the major attributes of these three multimedia models. The reader who is interested in the technical formulations and detailed assumptions of each model is referred to the manuals for MEPAS, MMSOILS, and RESRAD.[13–15]

3.4 LIMITATIONS OF MULTIMEDIA MODELS

The perceived complexity and lack of transparency of multimedia transport models represent significant barriers to their acceptance and widespread use. For this reason, most risk assessments still rely on hand-calculations and spreadsheet methods. However, an even greater barrier is the view that results obtained using computer models are highly dependent on user input and, therefore, subject to manipulation. It is widely recognized that for decisions to be both credible and implementable, the public must have confidence in both the scientific basis for judgments involved and the decision processes employed.[3] Most of the limitations associated with multimedia models fall into three categories: model uncertainty, parameter uncertainty, and scenario uncertainty.

3.4.1 MODEL UNCERTAINTY

Model uncertainty stems from gaps in the theories required to accurately predict contaminant transport, and target exposures and doses.[16] Because models are mathematical approximations of nature, they do not provide absolute estimates of risk. They provide conditional estimates based on multiple assumptions about source term, environmental settings, transport characteristics, exposure scenarios, toxicity, and other variables. Model uncertainties are introduced from the simplifying assumptions made during the modeling process. These uncertainties can have significant impacts on the accuracy or reliability of risk estimates. Numerical uncertainties, which arise from descretization or programming-related errors, are also a component of model uncertainty. Applying models to situations beyond their capabilities can create additional uncertainties.[7,16]

TABLE 3.1
Major Attributes of Three Multimedia Models

Multimedia Model Component	RESRAD	MMSOILS	MEPAS
Contaminants			
Organic chemicals	*	*	*
Metals	*	*	*
Radionuclides (parent)	*	—	*
Radionuclides (progeny)	*	—	*
Nonaqueous phase liquids	—	—	—
Source Types			
Contaminated soil	*	*	*
Landfill	*	*	*
Surface impoundment	—	*	*
Injection well	—	*	*
Underground storage tank	—	*	*
Waste pile	*	*	*
Trench with cap	*	—	*
Stack without plume rise	—	—	*
Stack with plume rise	—	—	*
Source Term Characteristics			
Mass balance	*	*	*
Multimedia partitioning	*	*	*
Source decay	*	*	*
Source ingrowth (radionuclides)	*	—	*
Multiple contaminants per simulation	*	*	*
Multiple sources per simulation	—	—	*
Source Release Mechanisms			
Erosion	*	*	*
Volatilization	*	*	*
Runoff	*	*	*
Leaching	*	*	*
Resuspension	*	*	*
Direct release to Vadose zone	*	—	*
Groundwater	*	—	*
Surface water	—	—	*
Air	*	—	*
Overland	—	—	*
Medium-Specific Flow			
Air			
Box model (O-D, complete mixing)	*	*	*
Steady state joint-frequency-of-occurrence (wind speed, stability class, direction)	*	*	*

(*continued*)

TABLE 3.1 (Continued)
Major Attributes of Three Multimedia Models

Multimedia Model Component	RESRAD	MMSOILS	MEPAS
Channeling of winds	—	—	*
Surface Hydrology			
Precipitation — Runoff, infiltrate, etc.	*	*	*
Surface Water			
Surface water (1-D steady state)	—	*	*
Surface water (n-D dynamic)	—	—	—
Wetlands	—	—	*
Groundwater			
Vadose zone (SS infiltration-soil moisture)	*	*	*
Vadose zone (1-D dynamic)	—	—	*
Vadose zone multiphase flow	—	—	—
Groundwater (1-D steady state)	*	*	*
Groundwater (n-D dynamic)	—	—	—
Medium-Specific Contaminant Transport			
Atmosphere			
Near field (0-D, complete mixing)	*	*	—
Far field–simple terrain (Gaussian plume)	*	*	*
Far field–complex terrain	—	—	*
Surface Water			
Overland (advection only)	—	*	*
Simple water bodies (0-D, complete mixing)	*	*	*
Streams & Rivers			
1-D advection & dispersion	—	—	*
1-D advection, 2-D dispersion	—	—	—
n-D advection & dispersion	—	—	—
Lakes & Reservoirs			
1-D advection & dispersion	—	—	—
n-D advection & dispersion	—	—	—
Groundwater			
Vadose zone (1-D advection only)	*	*	*
Vadose zone (1-D advection & dispersion)	—	*	*
Homogeneous aquifer (1-D advection only)	*	*	*
Homogeneous aquifer (1-D advection & n-D dispersion)	—	*	*
Homogeneous aquifer (n-D advection & dispersion)	—	—	—
Radionuclide progeny-specific retardation	*	—	—

(*continued*)

TABLE 3.1 (Continued)
Major Attributes of Three Multimedia Models

Multimedia Model Component	RESRAD	MMSOILS	MEPAS
Contaminant Transportation & Path Process			
1st order decay (no decay products)	*	*	*
1st order decay (with chained decay products)	*	—	*
Non-1st order decay	—	—	—
Linear partitioning (water/air, water/soil)	*	*	*
Nonlinear partitioning (water/soil)	—	—	—
Chemical reaction/speciation	—	—	—
Intermedia Contaminant Fluxes			
Surface Soil			
Air (volatilization, resuspension)	*	*	*
Vadose zone (leaching)	*	*	*
Overland (erosion, runoff)	*	*	*
Offsite surface soil	—	*	*
Overland			
Surface water (erosion, runoff)	—	*	*
Aquatic organisms (uptake)	*	*	*
Surface Water			
Sediment (sedimentation)	—	—	—
Groundwater (percolation)	*	*	*
Vadose Zone			
Air (volatilization)	*	*	*
Surface water	*	*	*
Groundwater			
Surface soil (deposition)	*	*	*
Air			
Surface water (deposition)	—	*	*
Vegetation (deposition)	*	*	*
Vegetation (uptake, deposition)	*	*	*
Soil, Irrigation			
Vegetation (uptake, disposition)			
Vegetation, Soil, Water			
Dairy animals (uptake)	*	*	*
Exposure pathways			
Ingestion (plant, meat, milk, aquatic food, water, soil)	*	*	*
Inhalation (gases, particulates)	*	*	*
External radiation	*	—	*
Dermal	*	*	*

(*continued*)

TABLE 3.1 (Continued)
Major Attributes of Three Multimedia Models

Multimedia Model Component	RESRAD	MMSOILS	MEPAS
Human Health Endpoints			
Radiological dose	*	—	*
Cancer (risk)	*	*	*
Non-cancer (hazard quotient)	*	*	*
Individual (mean, EMI)	*	*	*
Population (distribution, cumulative)	—	—	*

Note: * indicates component is included; — indicates component is not included.

Source: From J.J. Cheng, J.G. Droppo, E.R. Faillace, E.K. Gnanapragasam, R. Johns, G. Laniak, C. Lew, W. Mills, L. Owens, D.L. Stenge, J.F. Sutherland, G. Whelan, and C. Yu, *Benchmarking Analysis of Three Multimedia Models: RESRAD, MMSOILS, and MEPAS* DOE/ORO-2033, Oak Ridge National Laboratory, Oak Ridge, TN, 1995.

Since models are simplified approximations of empirical reality, they often omit several important considerations. Regens et al. (1999)[7] found that the following technical issues are not typically addressed by the existing generation of multimedia models:

- **Complex terrain:** Terrains that have complex topographies (i.e., complicated drainage patterns, extreme slopes) present a challenge when using mathematical models. To facilitate model computations, a generalized terrain which simplifies actual complexities usually is specified. Such generalization can affect the reliability and accuracy of risk estimates, especially in localized areas.
- **Nonaqueous phase liquids (NAPLs):** NAPLs are classified as being lighter than water (LNAPLs) or denser than water (DNAPLs). NAPLs can serve as a continuing source of dissolved contamination as well as a mechanism for transport of contaminants independent of the direction of water flow. Multimedia models typically do not have the capability to handle contaminant transport in the presence of NAPLs.
- **Fractured and karst media:** Fractured and karst media are distinct hydrological and geomorphological terrain that allow for rapid dispersion of contaminants through fast flow pathways (i.e., voids in the strata). Multimedia models commonly assume that groundwater flow moves through a semisolid rock matrix, in which retardation is typically more effective. This simplification can lead to conclusions that are not reflective of actual site conditions.
- **Mass balance between pathways:** Multimedia models examine the partitioning, fate, and transport of contaminants in different environments (i.e., soil, groundwater, surface water, air, and biota). Multimedia models

often do not ensure that total contaminant mass, accounting for degradation products, summed over all environmental media remains constant. Without this constant, multimedia models cannot give an accurate accounting of the relative importance of each exposure pathway.

Model uncertainty can contribute the most uncertainty to an analysis.[17] Employing multiple models and assessment techniques can aid with defining the boundaries of model uncertainty, but model uncertainty is one of the most difficult uncertainties to reduce in an assessment.

3.4.2 Parameter Uncertainty

Parameter uncertainty is defined as the lack of knowledge regarding the parameter values required to assess risks.[16] Parameters can be grouped into two categories. The first set consists of nonsite-specific parameters which include chemical properties (e.g., chemical-specific cancer slope factor and unit risk factor), target physical characteristics (e.g., age, weight, and location), and reference risk assessment parameters (e.g., inhalation and ingestion rates). The second category includes parameters that are site specific. Site-specific parameters are typically measured at the site. They may employ averaged or typical values from other locations when site-specific data are unavailable. These values are typically determined from sampling the site's properties that are required for modeling risks. Besides inherent variation, uncertainty originates from several sources. These sources include measurement error, systematic error (bias), and sampling error, and error from using surrogate data.[2,17] Measurement error typically arises from instrumentation and tools, and is customarily controlled through calibration and quality control checks. Systematic error results from procedures or data collection techniques that fail to wholly characterize the desired sampled parameter characteristic. Sampling error tends to occur when the number of samples used to infer a desired population characteristic is inadequate. Parameter uncertainty is the best understood source of uncertainty and typically does not cause large variation in assessment results if carefully considered.[17]

3.4.3 Scenario Uncertainty

Scenario uncertainty stems from incomplete or missing information required for a risk assessment that may result in the incorrect conceptualization of past, present, or future site conditions.[7,16] It involves the introduction of uncertainties resulting from generalized assumptions about the nature and amount of contaminants released at a site, the spatial and temporal distribution of potential targets, and the nature of the exposure pathways. The major causes of scenario uncertainty are (1) incomplete knowledge of historical events, and (2) incorrect assumptions about future exposure scenarios.[7] Because future land use and resultant future exposures are often unknown, assumptions must be made in terms of which scenarios should be considered. Incomplete analysis is an important source of scenario uncertainty.[2] Failing to assess an important exposure pathway can produce flawed results. Controlling scenario uncertainty resulting from incomplete analysis requires assessors to carefully rationalize, obtain feedback, and document every exposure scenario that will be included and excluded in an analysis.[2] Worst-case scenarios often are used to ensure conservative, worse-than-expected estimates of risk.

3.5 DATA QUALITY OBJECTIVES

In order to streamline and increase the efficiency of environmental field data collection programs, the Environmental Protection Agency[18,19] encourages the development of Data Quality Objectives (DQOs) prior to initiation of sampling. The DQO process is a planning tool that identifies an environmental problem, defines the data collection process, and ensures that the type quality, and quantity of the data collected are appropriate for the decision-making process. The basic idea is to identify the question to be answered by the data and then design the data collection and analysis process to provide an answer to the question.[19]

The DQO planning process consists of seven key steps[19]:

1. State the problem: Stakeholders work together to define their concerns and issues based on descriptions of the site, waste stream, issue, etc., and agree on the question or problem to be studied.
2. Identify the decision: Stakeholders design the answer or result that will answer the question or solve the problem, including the threshold level for action.
3. Identify inputs to the decision: Stakeholders define the measurements needed to answer the question.
4. Define the boundaries: Stakeholders define the time and space circumstances covered by the decision.
5. Develop a decision rule: Technical staff and stakeholders develop the formulation to obtain the needed data (quality and quantity) and to identify acceptability or confidence in the ultimate decision.
6. Specify acceptable limits on decision errors: In concert with Step 5, stakeholders define the tolerance for making incorrect decisions.
7. Optimize data design: Technical staff identifies the most resource-effective data collection design.

Implementation of the DQO process forces data suppliers and data users to consider the following questions:

- What decision has to be made?
- What type and quality of data are required to support the decision?
- Why are new data needed for the decision?
- How will new data be used to make the decision?

The DQO planning process has several notable strengths. It brings together the right players (stakeholders and technical staff) at the right time to gain consensus and commitment about the scope of the project. This interaction results in a clear understanding of the problem, the actions needed to address that problem, and the level of uncertainty that is acceptable for making decisions. Through this process, data collection and analysis are optimized so only those data needed to address the appropriate questions are collected.[19]

When taking environmental measurements, we need to measure things that add value to our decisions. Multimedia models can help identify precisely those measurements that will most reduce decision uncertainty. Multimedia models can identify the contaminants and pathways that make the largest contributions to overall risk and, thus, are most in need of monitoring. They can identify the location at a site that is the largest contributor to risk, and that is most important to monitor. Conversely, they can identify areas that are not important contributors to risk and do not need additional or increased monitoring. A sensitivity analysis of the model can identify variables in the model that lead to the most uncertainty, and, thus, the variables most in need of monitoring. They can be used to sum the impact of all pollutants released from a facility and to develop a new single metric such as pounds of toxics released per unit product built.[20]

3.6 CONCEPTUAL SITE MODEL DEVELOPMENT

The construction of a conceptual site model is an essential first step in conducting risk assessments for environmental remediation. The conceptual site model provides a blueprint outlining the problem to be analyzed in a site risk assessment by summarizing what is known about present site conditions at a level of detail that is appropriate for the scope, analytical capabilities, and resources available. The conceptual site model defines the assumptions the risk assessor makes about contamination concentrations and locations as well as the spatial and temporal distribution of potential receptors. A conceptual site model represents (1) chemicals and radionuclides of potential concern including, to the extent known, contaminant locations, concentrations, and chemical forms; (2) ongoing or anticipated processes that will produce changes in these concentrations such as chemical transformation or radioactive decay; (3) ongoing or anticipated processes that will result in the transport of those substances, including movement from one medium to another; and (4) receptor locations and exposure scenarios appropriate to the site-specific conditions, including anticipated future land use scenarios (e.g., agricultural, recreational, residential) to be assessed. Typically, this information is used in a conceptual site model to provide a qualitative characterization of the contamination situation and delineate the exposure pathways to be considered by the multimedia models. Simple quantitative calculations can be used during the development of a conceptual site model to assure that omitted pathways do not contribute significantly to risk.

The computational methods or multimedia models selected for use in risk assessment should be capable of characterizing the important properties and processes of the principal contaminants of concern (e.g., volatilization from soil to air, radioactive decay). The method or multimedia model selected also should support analysis of the pollutant transport and exposure pathways specified in the conceptual site model. The conceptual site model defines what processes are to be analyzed and what calculations are to be done. As a result, the conceptual site model must be specified with analytical capabilities and limitations in mind. Moreover, because the conceptual

TABLE 3.2
Judgments in Development of Conceptual Site Models

Source Term Properties

- Spatial distribution of contaminants
- Barriers preventing or restricting release of contaminants
- Degradation rate of barriers
- Release rate of contaminants through barriers
- Fraction of source term contributing to transport pathways of interest
- Flux of contaminants to transport pathways of interest

Site Geometry

- Topography of site
- Dimensions of contaminated area
- Location of waste in contaminated area
- Dimension of unsaturated layers
- Depth of saturated layer

Transport Properties

- Homogeneity of subsurface geology
- Physical and chemical characteristics of media (e.g., porosities, organic matter contents, bulk densities, pHs, hydraulic conductivities)
- Effect of seasonal conditions (storm events, intense cold)
- Fraction of precipitation reaching groundwater vs. surface water
- Direction and velocity of groundwater flow
- Contribution of karst conditions to groundwater flow
- Vertical and horizontal components of groundwater flow

Receptor Properties

- Future land use
- Location of future exposure
- Pathways of future exposure

Exposure Factors

- Future population distribution
- Exposure duration
- Ingestion rates
- Bioaccumulation factors
- Toxicity factors

site model is a conceptualization of past, present, and future conditions at the hazardous waste site, its construction requires that the risk analyst to make numerous subjective judgments and simplifying assumptions. These subjective judgments can greatly influence the validity of risk estimates. Table 3.2 illustrates the types of judgments typically made by the analyst.

Results of the Regens et al.[21] study imply that the most important factor to consider when applying multimedia models as an aid in developing environmental monitoring plans is to utilize a multimedia model that is capable of analyzing the conceptual model developed for the site. The linkage between site conceptual models and multimedia models is critical to obtaining realistic results. Not all multimedia models are capable of representing the conceptual model of a site.

3.7 EXPEDITED SITE CHARACTERIZATION

The DOE has developed a methodological approach to site characterization called Expedited Site Characterization (ESC) that combines DQOs with an iterative approach to gathering data that shortens the length of the assessment period and can reduce costs at many sites. At its core is a reliance on an initial conceptual site model that is iteratively modified with the use of multimedia models (if necessary) to integrate field measurements into a refined conceptual site model. The ESC method has been successfully applied at many DOE sites. Based on initial experience with the method, the American Society for Testing and Materials has developed standards for application of the methodology.[22]

A fundamental aspect of the ESC methodology is the use of a preliminary conceptual hydrogeologic model (together with a computer implementation of this conceptualization, if necessary, at the more complex sites) to guide site investigations.[23] The preliminary conceptual model is iteratively refined as additional site data yield a more detailed understanding of the subsurface environment. As the conceptual model evolves, the characterization program is modified to address the specific data needs identified at each stage of the site investigation. Essentially, the computer model (conceptual site model) dictates, at each step of the process, what additional data are needed to reduce uncertainty to acceptable levels.

The principal elements of ESC are:

Step I: Review existing site characterization data for quality and completeness. Develop an initial conceptual model and use computerized groundwater transport models to integrate existing data to obtain a preliminary understanding of the extent of contamination and its movement. The model is often a set of hypotheses describing the essential features of the site that control the pathways between the likely source of contamination and all potential receptors.

Step II: Use output from the subsurface models to identify and obtain necessary additional monitoring data necessary to reduce uncertainty in current understanding of the contamination distribution in groundwater and soils.

Step III: Use data collected in Step II, along with computerized models to predict the contamination's long-term fate and migration, and evaluate risk at site. The preliminary mathematical models of groundwater flow and contaminant transport are recalibrated, incorporating new data. Predictive flow and transport runs are then used to simulate the effects on long-term contaminant migration. The output is used to decide on the most appropriate actions at the site.

Noninvasive measurements like electromagnetic and geophysical surveys precede invasive techniques like soil borings or well installations. Team members analyze data and immediately integrate results into field work. The team also develops a conceptual model of the site to predict measurements formally. Analysis ensures the plan stays within regulator-approved bounds. The result is a fast, cost-efficient, and scientifically thorough investigation that focuses on those issues critical to DOE and its stakeholders.

3.8 CASE STUDIES

We present three case studies that illustrate the possible benefits of utilizing multimedia models in the site characterization process to significantly reduce the cost or time involved.

3.8.1 Los Alamos National Laboratory: Using Models to Reduce the Number of Wells

The Los Alamos National Laboratory (LANL) was founded in 1943 as part of the U.S. effort to develop atomic weapons. A variety of contaminants, including radionuclides, metals, and organic compounds, have been released from LANL activities into canyons surrounding the facility in the intervening years since 1943. Most of these contaminants tend to adsorb to sediments near the release point, but they have been redistributed into nearby canyons by floods. For sampling and analysis purposes, the site was divided into eight watersheds, each containing a mesa top, a connecting canyon system, and all subsequent pathways to the Rio Grande. Environmental transport pathways considered were surface soils, surface water, stream sediments, and groundwater.

Working with the State of New Mexico, the DOE developed an initial Groundwater Protection Program Plan that called for 87 wells at a cost of $53 million. Both the State of New Mexico and the DOE have reservations about the initial plan. The DOE was concerned with the cost and the fact that there was not a technically defensible justification for the number and location of the wells. The state was concerned about the paucity of current knowledge regarding the laboratory's subsurface geology and hydrological setting. They were both concerned that there was insufficient understanding of the hydrogeologic setting to be able to design an effective monitoring system.

The state and DOE decided to develop a conceptual model of the site, together with a computer model of each watershed. Using the computer models, they were able to organize what was known and what was not known about contaminate movement at the site. This knowledge allowed development of a preliminary conceptual model of the site and a preliminary data collection design to evaluate the appropriateness of the conceptual model assumptions.

Based on the computer models and the available data, the DOE installed 11 wells on the LANL site to improve understanding of groundwater movement. The location of the wells was based on computer models and designed to reduce uncertainty in the final decisions. The refined understanding of the subsurface environment obtained through the initial characterization effort was to provide a better foundation for future monitoring efforts. In 2002, it was decided to install six additional groundwater characterization and monitoring wells. Using this iterative process of combining data collection with computer modeling, the final cost of the monitoring system was lowered to $43 million, a savings of $11 million.[24] In addition, both sides have more confidence in the final decisions that will be made.

3.8.2 Pantex: Minimization of Characterization Costs

The Pantex Plant is a DOE-owned facility, encompassing approximately 16,000 acres located in Carson County, TX, approximately 17 mi northeast of Amarillo. Slightly more than 2,000 acres of the DOE-owned property are used for industrial operations. The remaining DOE-owned land serves for DOE safety and security purposes. DOE operated the Pantex Plant to assemble and disassemble nuclear weapons, including fabrication, assembly, testing, and disassembly of nuclear ammunition and weapons. Past waste management operations included burning chemical wastes in unlined pits, burying wastes in unlined landfills, and discharging plant wastewaters into onsite surface waters. The contaminant plume resulting from these operations has migrated past the plant boundaries and onto adjacent landowners' property to the southeast. The Ogallala Aquifer, a major source of water for the region, lies beneath the Pantex Plant. Chemicals discharged from plant operations have been detected in parts of the perched aquifer, but not in the Ogallala. Pantex has an extensive groundwater-monitoring program consisting of over 70 wells, which are monitored quarterly or semi-annually for an extensive list of contaminants.

Contaminant migration at Pantex is complicated by a perched aquifer at a depth of 250 ft and the 400-ft thick Ogallala aquifer at 450 ft. The U.S. Army Corps of Engineers[25] spent 2 years and $2.1 million completing the first of a proposed four-phase comprehensive traditional characterization effort that was scheduled to last 5 years and cost $11.9 million. The initial phase of the COE effort included the drilling of 11 wells. Based on this work, an additional 43 wells were called for. At that point, DOE began an ESC effort[4] which resulted in only four soil borings, one of which was later turned into a well, and nine cone penetrometer pushes. After the DOE team left the site, an additional 10 borings were made and 5 monitoring wells were installed to determine precisely what remediation efforts would be required to treat the contaminant plume.

Table 3.3 compares the number of penetrations estimated for completion of site characterization at Pantex. This table does not include the 11 monitoring wells drilled by the COE during the initial phase. If the 11 initial phase wells were included, total penetrations for both methods would be increased by 11. This table demonstrates that the ESC approach resulted in installation of significantly fewer monitoring wells.

It is estimated that the Expedited Site Characterization effort saved the DOE $6.8 million, which was a 57% cost savings and a 40% reduction in time.[4] The Agency for Toxic Substances and Disease Registry has reviewed the site[26] and concluded that Pantex Plant in Amarillo, TX, did not pose a threat to public health at that time based on available environmental sampling data.

TABLE 3.3
Post-Phase I Penetrations at the Pantex Plant

Type of Penetration	Expedited Site Characterization Actual Number	Traditional Planned Number
Borings	13	0
CPT pushes	9	0
Wells (perched aquifer)	5	39
Wells (Ogallala)	1	4

3.8.3 Department of Energy Hanford Site: Using Models to Identify Data Needs

The DOE owns and operates a 586-mi^2 Hanford Site in the southeastern part of Washington State. The site was the major nuclear weapons plutonium production complex for over 40 years. Production of plutonium at Hanford stopped in 1987 with the closure of the last plutonium reactor. Currently, Hanford is engaged in the world's largest environmental cleanup project. More than 1,500 waste sites have been identified at Hanford, ranging from small areas of surface contamination to 177 underground storage tanks that hold about 54 million gallons of highly radioactive waste. Another major problem at Hanford is nearly 2000 tons of spent nuclear fuel stored in two water-filled basins just a quarter mile from the Columbia River. The bottoms of these basins are covered with as much as 2 to 3 ft of radioactive sludge. Cleanup will take decades and cost tens of billions of dollars. Even then, it will not be technically possible to completely clean the Hanford site and the residual contamination must be kept away from people and the environment for the thousands of years that it will remain hazardous.

Consider Hanford's composite analysis as an illustration of the use of models in designing environmental monitoring programs. This example is not for identifying where samples should be taken, but shows what type of information (sampling data) would be most useful in reducing overall remediation decision uncertainty. The composite analysis was performed at Hanford to determine the impact of interacting source terms on the groundwater transport of radionuclides away from waste disposal facilities. The concern was that even though previous analyses had indicated that material left at individual waste sites after cleanup would produce no public harm, the cumulative impact of all remaining sources had not been investigated.

To probe this issue, the Hanford Groundwater Team developed a site-wide groundwater flow and transport model that was used to simulate future groundwater flow conditions and to predict the migration of existing and future contaminants in the unconfined aquifer system at the Hanford Site. The groundwater model was a three-dimensional numerical model developed at Hanford between 1992 and 1997.

This project was as an application of the DQO process to model prediction data rather than to plan for field sample collection. That is, how much confidence does one need in model results to be able to make decisions about future events that will be protective of human health? In this application, stakeholders were asked to set acceptable decision error limits when the inputs to the decision are the output from release, fate, and transport, and risk assessment models.

The composite analysis identified lack of knowledge concerning the inventory and release rates of contaminates at the waste sites as the major factor contributing to uncertainty in the multimedia models, and consequently in the decision-making process. The rate of release from waste sites, now and in the future, is needed to conduct site-wide assessments of cumulative impact to the Columbia River and surrounding communities. Limited data are available on actual release rates. Hence, the primary technical gap is the need for data and models to describe release rates from all waste that will reside at the Hanford Site in the postclosure period. These modeling results provide temporal and spatial information for system assessments.

An example of data needs identified is the release rates from a soil-waste matrix near previous tank releases. The soils will have been modified due to contact with high heat and high pH tank wastes, effectively producing a new waste form. The rate of release from these modified areas to underlying soils needs to be determined through environmental monitoring. The environmental monitoring results will serve as a basis for developing computer models to predict releases from such modified soils over future times.

In addition, future tank releases will have an opportunity to travel the same path within the vadose zone as past tank releases. Thus, the influence of the initial leak on the original soil/sediment profile is important to understand as the basis for predicting the migration and fate of future releases. Thus, it is necessary to take measurements of the soil/sediment profile in areas where past releases have occurred.

Information obtained by addressing these needs will provide an improved technical basis for making site regulatory decisions and therefore reduce the uncertainty associated with the basis for these decisions. The estimated life cycle cost savings was estimated at $200 million.[27]

3.9 CONCLUSIONS

Multimedia models are increasingly being used as an aid in the development of environmental monitoring plans. Drivers for the increased use of multimedia models are (1) the need to ensure that the type, quality, and quantity of the data collected are appropriate for the decision-making process and (2) the need for increased flexibility in the data-gathering process to decrease the time and cost involved in performing site characterization.

Multimedia models provide the capability to organize and summarize large amounts of information to estimate exposures and assess risks. As reliance on multimedia models for risk assessment has increased, it has become important to delineate their advantages and limitations. General conclusions that can be drawn regarding the use of multimedia models in the design of environmental monitoring plans include the following:

- Linking the application of the multimedia models to the site conceptual model using real data for inputs requires skill on the part of the risk assessor, coupled with understanding of the environmental setting and parameterization of the models to avoid introducing substantial estimation error.

- The Regens et al.[21] multimedia model comparison study demonstrates that when MEPAS, MMSOILS, and RESRAD were applied to real DOE sites using actual data, those models were used successfully to identify exposure pathways and estimate risk. The models achieved agreement on major exposure pathways, although contaminant concentrations could differ by two orders of magnitude due largely to differences in transport and fate. This points to the need for order-of-magnitude analysis to avoid model misuse.
- The primary benefit of the screening-level capability inherent in multimedia models is the identification of contaminants and pathways that make the largest contributions to overall risk. To do this well, a multimedia model needs the capacity to reflect realistically site conditions. This involves the ability to accept generalized assumptions concerning the amount and location of contaminants, spatial and temporal distributions of potential receptors, and exposure pathways. In essence, because multimedia models necessarily are simplified approximations of empirical reality, the flexibility of a multimedia model *vis a vis* the site conceptual model is critical. This underscores the importance of utilizing a multimedia model that is capable of analyzing the conceptual model developed for the site.

The Regens et al. (2002)[21] multimedia model comparison study demonstrated that multimedia models can be applied to derive theoretically plausible estimates of risks by major pathways using actual site data if the multimedia model is compatible with the conceptual site model. The degree to which a conceptual site model can be defined that is consistent with the multimedia models' capabilities is key to the successful use of multimedia media models in designing environmental sampling plans. If the conceptual site model is poorly chosen, the results of the analysis are likely to be incorrect, no matter how well the multimedia model is able to characterize future exposures.

REFERENCES

1. National Research Council (NRC). 1994. *Science and Judgment in Risk Assessment*, National Academies Press, Washington, D.C.
2. Environmental Protection Agency (EPA). 1992. Guidelines for Exposure Assessment, *Fed. Reg.* 57(104): 22888–22938.
3. National Research Council (NRC). 1983. *Risk Assessment in the Federal Government: Managing the Process,* National Academies Press, Washington, D.C.
4. Department of Energy (DOE). 1998. Innovative Technology Summary Report: Expedited Site Characterization (DOE/EM-0420).
5. Cox., L.A. 2000. *Risk Analysis: Foundations, Models, Methods*, Kluwer Academic, Boston.
6. Van De Meent, D. De Bruijn, J.H.M., De Leeuw, F.A.A.M., De Nijs, A.C.M., Jager, D.T., and Vermeire, T.G., 1995. Exposure modeling, in *Risk Assessment of Chemicals: An Introduction*, Van Leeuwen, C.J. and Hermens, J.L.M., Eds., Kluwer Academic, Boston, pp. 103–145.

7. Regens, J.L., Travis, C., Obenshain, K.R., Whipple, C., Gunter, J.T., Miller, V., Hoel, D., Chieruzzi, G., Clauberg, M., and Willis, P.D., 1999. *Multimedia Modeling and Risk Assessment*, Medical University of South Carolina Press, Charleston, SC.
8. Brenner, R., 1995. Predicting chemical risks with multimedia fate models, *Environ. Sci. Technol.* 29, 556A–559A.
9. Mills, W.B., Cheng, J.J., Droppo, J.G. Jr., Faillace, E.R., Gnanapragasam, E.K., Johns, R.A., Laniak, G.F., Lew, C.S., Strenge, D.L., Sutherland, J.F., Whelan, G., and Yu, C., 1997, Multimedia benchmarking analysis for three risk assessment models: RESRAD, MMSOILS, and MEPAS, *Risk Anal.* 17, 187–201.
10. Laniak, G.F., Droppo, J.G., Faillace, E.R., Gnanapragasam, E.K., Mills, W.B., Strenge, D.L., Whelan, G., and Yu, C., 1997. An overview of a multimedia benchmarking analysis for three risk assessment models: RESRAD, MMSOILS, and MEPAS, *Risk Anal.* 17, 203–214.
11. Moskowitz, P.D., Pardi, R., Fthenakis, V.M., Holtzman, S., Sun, L.C., and Irla, B., 1996. An evaluation of three representative multimedia models used to support cleanup decision-making at hazardous, mixed, and radioactive waste sites, Risk Anal. 16, 279–287.
12. Cheng, J.J., Droppo, J.G., Faillace, E.R., Gnanapragasam, E.K., Johns, R., Laniak, G., Lew, C., Mills, W., Owens, L., Stenge, D.L., Sutherland, J.F., Whelan, G., and Yu, C., 1995. *Benchmarking Analysis of Three Multimedia Models: RESRAD, MMSOILS, and MEPAS* DOE/ORO-2033, Oak Ridge National Laboratory, Oak Ridge, TN.
13. Buck, J.W., Whelan, G., Droppo, J.G. Jr., Stenge, D.L., Castleton, K.J., McDonald, J.P., Sato, C., and Streile, G.P., 1995. *Multimedia Environmental Pollutant Assessment System (MEPAS) Application Guidance, Guidelines for Evaluating MEPAS Input Parameters for Version 3.1-PNL-10395*, Pacific Northwest Laboratory, Richland, WA.
14. U.S. Environmental Protection Agency, 1996. *MMSOILS: Multimedia Contaminant Fate, Transport, and Exposure Model, Documentation and User's Manual, Version 4.0*, Office of Research and Development, Washington, D.C.
15. Yu, C., Zielen, A.J., Cheng, J.-J., LePoire, D.J., Gnanapragasam, E., Kamboj, S., Arnish, J., Wallo, A., Williams, III, W.A., and Peterson, H., 2001. *User's Manual for RESRAD Version 6, ANL/EAD-4*, Argonne National Laboratory, Argonne, IL.
16. Baker, S.R., Driver, J., and McCallum D. (Eds.), 2001. *Residential Exposure Assessment: A Sourcebook*, Kluwer Academic/Plenum Publishers, New York.
17. Haimes, Y.Y., 1998. *Risk Modeling, Assessment, and Management*, John Wiley & Sons, New York.
18. U.S. Environmental Protection Agency, 1993. Guidance for Planning for Data Collection in Support of Environmental Decision Making Using the Data Quality Objectives Process (EPA QA/G-4).
19. U.S. Environmental Protection Agency, 2000. Data Quality Objectives Process for Hazardous Waste Site Investigations (EPA/600/R-00/007).
20. Eaton, R.J., 1999. Getting the Most Out of Environmental Metrics http://www.nae.edu/nae/naehome.nsf/weblinks/NAEW-4NHM88?opendocument
21. Regens J.L., et al., 2002. Conceptual site models and multimedia modeling: comparing MEPAS, MMSOILS, and RESRAD, *Hum. Ecol. Risk Assess.* 8(2), 391–403.
22. American Society for Testing and Materials, 1998. Standard Practice for Expedited Site Characterization of Vadose Zone and Groundwater Contamination at Hazardous Waste Contaminated Sites (ASTM Standard D6235).
23. Sedivy, R.A., 1998. The role of the conceptual hydrogeological model in site characterization and dynamic risk assessment, *Proceedings of the 1st International Symposium on Integrated Technical Approaches to Site Characterization*, Chicago, IL, pp. 187–201.

24. U.S. Department of Energy, 1996. Los Alamos National Laboratory Groundwater Protection Program Plan, http://dqo.pnl.gov/casestudies/cases/lanl.htm
25. Corps of Engineers (COE), 1992. Hydrogeologic Assessment for the Pantex Plant, prepared by the U.S. Army Corps of Engineers, Tulsa District, Tulsa, OK, for the U.S. Department of Energy, Pantex Plant, Amarillo, TX.
26. Agency for Toxic Substances and Disease Registry, 1998. Public Health Assessment for the U.S. Department of Energy, Pantex Plant, Amarillo, TX.
27. U.S. Department of Energy, 2001. Technology Needs/Opportunities Statement: Provide Method for More Accurate Estimates of Waste Constituent Release Rates and Modes from Waste, http://www.hanford.gov/boards/stcg/pdfs/fy02needs/subsurf/ss42.pdf

4 Basic Concepts and Applications of Environmental Monitoring

T. Brydges

CONTENTS

1-56670-641-6/04/$0.00+$1.50

4.1 INTRODUCTION

One of the definitions for the verb "monitor" given by the *Oxford Dictionary* (1995) is to "maintain regular surveillance over." It is essential in all environmental monitoring programs to make measurements at regular intervals over a substantial length of time.

There are two fundamental reasons for monitoring the natural environment: first, to establish baselines representing the current status of ecosystem components and, second, to detect changes over time—particularly, any changes that are outside of the natural variation in these baselines (Hicks and Brydges 1994). Hence, the importance of having long-term data records is clear.

Closely associated with these reasons is the desire to define why the observed changes are occurring. This chapter provides a detailed description of different approaches to monitoring, a listing of the problems confronting current monitoring networks, and a description of some representative networks.

Ecological monitoring programs fall into four broad categories.

4.1.1 Simple Monitoring

Simple monitoring records the values of a single variable at one geographical point over time. In practice, however, such single parameter monitoring is frequently expanded to include measurements of the parameter at many geographical locations. Air temperature is one example of the application of simple monitoring. Data from around the world are used to calculate the average global air temperature which is one of the keystone measurements in the global warming/climate change issue. The graphs in Figure 4.1 represent the longest temperature records that can be constructed from direct air temperature measurements and are frequently used to show the increase in average global air temperature over the past 150 years.

Another example of simple monitoring is the measurement of atmospheric concentrations of carbon dioxide, the dominant greenhouse gas. The long-term monitoring record from the observatory at Mauna Loa, HI, is shown in Figure 4.2. It clearly shows the increasing concentrations, and this observation has been very influential in the global warming issue. This simple monitoring approach has been extended to the point where atmospheric carbon concentrations are now measured at many sites around the world.

4.1.2 Survey Monitoring

In many cases, environmental problems have become obvious without any historical monitoring record of the changes that led up to the problem. The absence of an historical record can be replaced by a survey of current conditions over a geographical area. The monitoring survey is designed to include areas that are affected and areas not affected by the observed stress. The affected areas are assumed to have

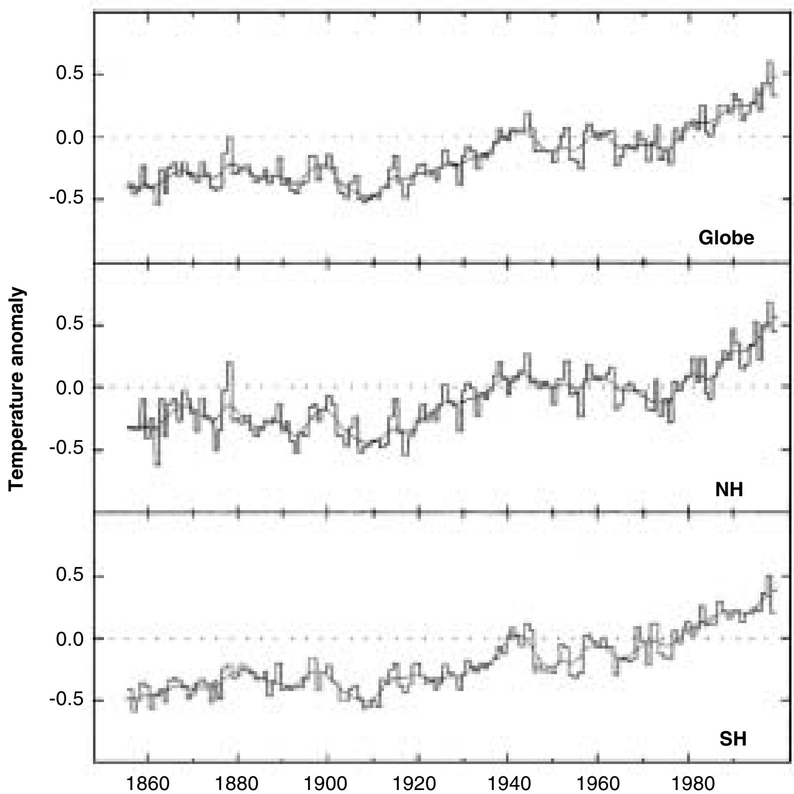


FIGURE 4.1 Global and hemispheric annual temperature anomalies 1856–1999.

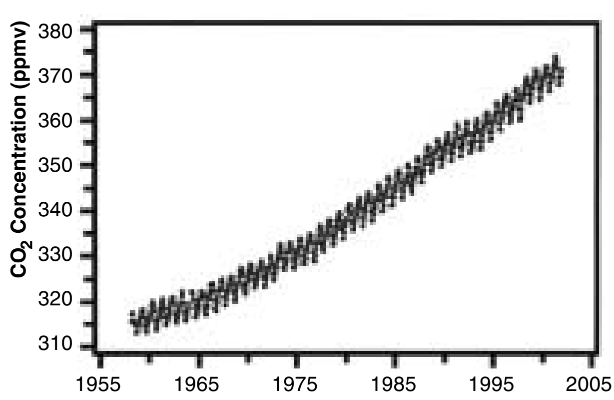


FIGURE 4.2 Atmospheric monthly CO_2 measurements at Mauna Loa, Hawaii. Source: Dave Keeling and Tim Whorf (Scripps Institution of Oceanography).

had similar environmental characteristics as the unaffected areas at some unknown time in history. To take as an example, eutrophication (excessive algae growth) of Lake Erie and Lake Ontario became obvious in the 1960s. An historical monitoring record might have shown concentrations of phosphorus and algae simultaneously increasing. However, no such monitoring was undertaken. Vollenweider (1968) overcame the lack of historical data by comparing current survey data on algae growth with nutrient levels for many lakes showing a range of eutrophic states. He developed a relationship between total phosphorus loading and algae growth in both affected and unaffected lakes. He applied this relationship to Lake Erie and Lake Ontario, making the assumption that their historic nutrient and algae concentrations were similar to currently unaffected lakes. This reasoning, based on survey monitoring data from the lakes and supported by laboratory and field experimental data, was applied to establishing scientifically defensible phosphorus control programs for these lakes (International Joint Commission 1969). These control programs have been successful in reducing phosphorus concentrations and the corresponding algae growth (Dobson 1994).

4.1.3 Surrogate or Proxy Monitoring

Another way to compensate for the lack of long-term monitoring records involves using surrogate or proxy information to infer historical conditions in the absence of actual measurements of the desired variable. In this approach, data are obtained from information "stored" in the environment that relates to the desired variable. For example, to evaluate global warming trends it would be ideal to have temperature records from the beginning of time. This being impossible, several surrogates for temperature have been used to construct long historical records.

Information stored in arctic ice cores has been used to infer air temperatures over very long periods of time. Dansgaard et al. (1993) reported on results for two ice cores drilled in central Greenland that represented the accumulation of ice for some 250,000 years. The ice cores were 3028.8 m long. The ratio of oxygen 18 to oxygen 16 stable isotopes of the air trapped in snow particles is a function of the air temperature at the time of deposition of the snow. The deposited snow becomes packed into ice and the oxygen isotope ratios in trapped air bubbles preserve a "temperature" record in the built-up ice layers over thousands of years. Measurements of this ratio from the ice cores were used to compare the extent and rates of change of the inferred temperature over this very long period of time. The authors concluded that the temperature has been remarkably stable during the past 10,000 years but, prior to that, instability dominated the North American climate.

In the 1970s, lakes and rivers in eastern Canada were observed to have unexpectedly low pH and high sulfate concentrations. Acid rain was postulated as the cause. However, there was a serious lack of data to show that the historical pH values had in fact been higher. Critics of the acid rain theory postulated that the observed conditions were natural and did not represent recent acidification. It was known that the fossil remains of some diatoms and chrysophyte algae species accumulated in the sediment layers at the bottom of lakes. It was further shown

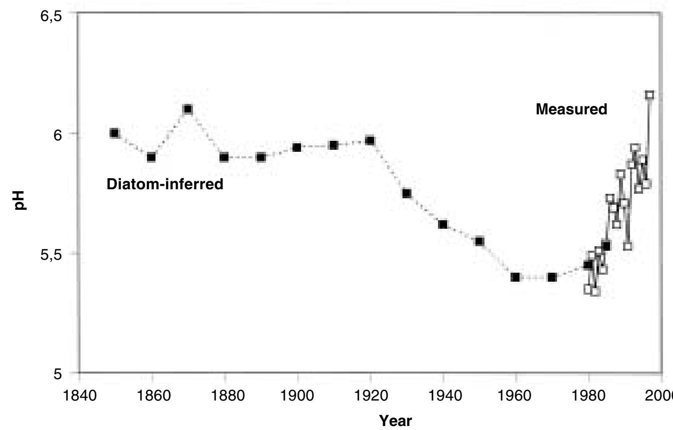


FIGURE 4.3 pH of George Lake. Diatom-inferred values (Dixit et al. 1992) are given for the period 1850 to 1985 (closed squares). Measured values (OMNR/OMOE unpublished data) are given for the period 1981 to 1998 (open squares).

that the relative numbers of these species changed with lake pH. Combined with the ability to determine when the sediment layers were deposited, a method to infer the historical lake pH was developed using the algae species composition in the sediments. The technique has been applied extensively in eastern Canada (Jeffries 1997) where large geographical areas containing thousands of lakes have been damaged by acid rain. Specifically, it has been shown that acidity and metal concentrations began to increase about 1920 in the lakes near the large sulfur dioxide and metal sources at Sudbury, Ontario (Dixit et al. 1992). In 54 lakes in south-central Ontario, all lakes with present measured pH less than 6.0 had acidified since approximately 1850 (Hall and Smol 1996), according to the pH inferred from the diatom assemblages. Their historical proxy measured pH values are plotted in Figure 4.3 along with recent measured values (E. Snucins, Department of Biology, Laurentian University, Sudbury Ontario, Canada, personal communication 1999). The results show the proxy-measured decline of pH consistent with the known high local emissions of sulfur dioxide, followed by the measured recovery of pH after the emissions were reduced. These proxy measures of historical pH have been very important in establishing the reality and the long-term nature of surface water acidification caused by acid rain.

4.1.4 Integrated Monitoring

While long-term records from simple monitoring, surveys, and surrogate data have provided substantial information on what has been changing in the environment, they are generally unable, by themselves, to answer the important question of why these changes are occurring. A much more detailed set of ecological information is needed to establish cause-and-effect relationships. The concept of integrated monitoring has been developed with the overall objectives of recording changes in the environment and of understanding and defining the reasons for those changes; in other words, to define what is changing and why.

Integrated monitoring has four specific objectives: to establish cause-and-effect relationships; to derive scientifically defensible pollution control or resource management programs; to measure the environmental response to the control measures; and to provide early warnings of new problems. For example, much of the preliminary information on the ecological effects of acid rain on surface waters came from the data sets being gathered to study eutrophication of lakes. In recent years, substantial information on the ecological effects of climate variability/change has, in turn, been derived from integrated monitoring being conducted to measure the effect of acid rain.

The integrated monitoring sites are characterized by long term (i.e., indefinite) multidisciplinary monitoring, i.e., meteorology, precipitation chemistry, runoff chemistry, and a full suite of biological factors. A centerpiece of integrated monitoring sites is frequently a "calibrated watershed" where the monitoring strives to develop a detailed balance of the inputs and outputs of water and chemicals along with intensive biological monitoring of the terrestrial and aquatic components of the watershed. The resulting information is often sufficient to answer both the questions of what changes are occurring and why they are happening. The integrated monitoring is usually carried out in conjunction with detailed research projects and often in conjunction with ecological manipulation experiments. These experiments involve the deliberate alteration of the environment under highly controlled and monitored conditions. For example, whole-lake additions of nutrients, combined with integrated monitoring of the lakes, helped to resolve the question of whether to reduce nitrogen or phosphorus in order to control eutrophication (Schindler 1974). Whole-lake acidification experiments (Schindler 1980) have provided key information on the sequence of biological changes as lake pH decreased, and on the relative importance of nitric and sulfuric acid (Rudd et al. 1990).

Results of experimental flooding of a boreal forest wetland have been reported (Kelly and Rudd 1997). These studies simulate the effect of flooding behind dams and the effects of possible water level changes due to climate change/variability. Flooding caused the wetland to change from a carbon sink to a carbon source with respect to the atmosphere and also caused an increase in the microbial conversion of inorganic mercury to methyl mercury. The changes in fluxes are summarized in Figure 4.4.

In addition to these four categories of monitoring, there is a large field of technology that applies to within-plant industrial compliance monitoring. This monitoring measures the content of effluents to determine total emissions and their value relative to legally established limits. This is important information, particularly with respect to determining stresses applied to the environment. However, this chapter is restricted to monitoring programs that are carried out within the natural environment.

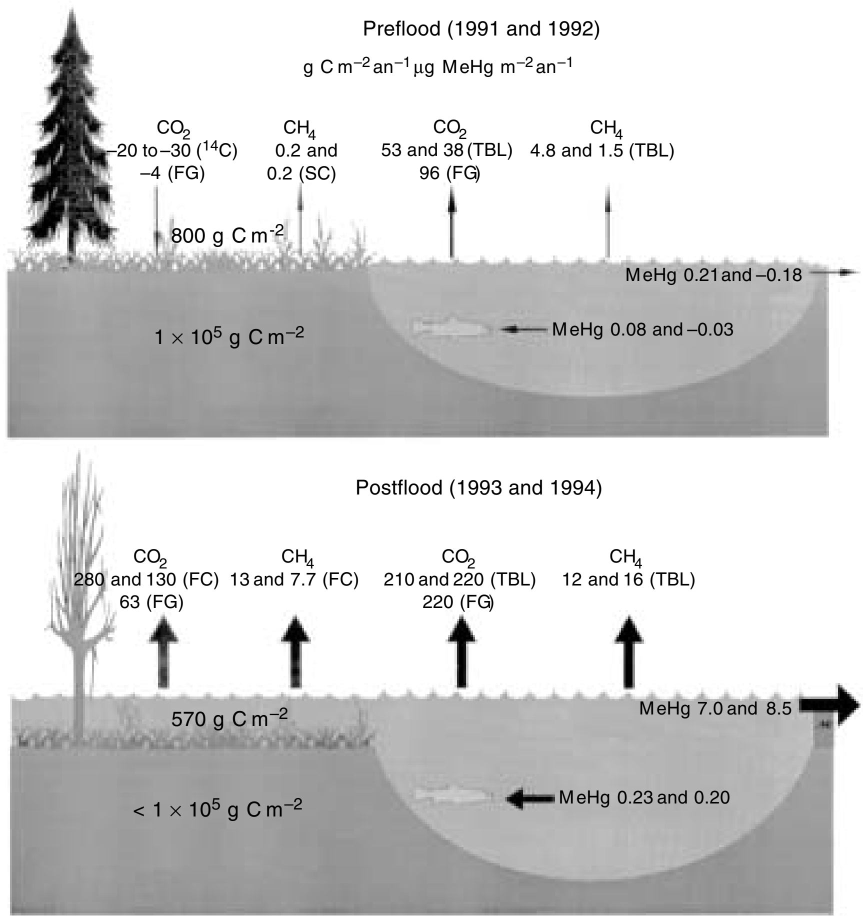


FIGURE 4.4 Annual downstream loss (mass output–mass input) of MeHg from the experimental wetlands annual fluxes of MeHg into fish for 2 years prior to and 2 years after fleeding, and pre- and postflood fluxes of CH_4 and CO_2 from the pestland and pond surfaces of the wetland to the atmosphere. Methods used for measurement of the gas fluxes are shown in brackets: TBL, thin boundary layer method; FG, Flux-gradient (1992 and 1993 only); ^{14}C, carbon accumulation rate by ^{11}C dating; FC, floating chambers; BC, static chambers on the pestland surface.

4.2 THE NATURE OF SOME CURRENT ENVIRONMENTAL ISSUES

The nature of the ecological responses to the current suite of stresses on ecosystems poses major challenges in meeting the overall and specific objectives of integrated monitoring.

4.2.1 Natural Elements and Compounds Causing Problems

Essential elements such as phosphorus, sulfur, nitrogen, and carbon can lead to significant environmental change and damage if present in the environment in excessive amounts or if in a particular form. Examples illustrate these points. Excessive amounts of phosphorus in lakes can lead to inordinate algae growth (eutrophication) (Vollenweider 1968; Schindler and Fee 1974). Excessive sulfur in the form of sulfuric acid in precipitation can cause acidification of sensitive soils and surface waters, which leads to environmental degradation (National Research Council of Canada 1981; Environment Canada 1997). Ammonium and nitrates also contribute to acidification of terrestrial and aquatic systems (Reuss and Johnson 1986; Environment Canada 1997). Increasing amounts of carbon in the form of carbon dioxide in the atmosphere may lead to changes in global climate (Houghton et al. 1990). Consequently, these elements are subjected to controls in order to eliminate their negative effects.

Since these elements are essential building blocks of life and are naturally present in the environment, it is neither possible nor desirable to reduce their concentrations to zero. Therefore, in setting objectives for these compounds in the environment, it has become the practice to apply the concept of a critical load or level which is not zero but rather is defined as "the highest load/level that will not cause chemical changes leading to long-term harmful effects of the most sensitive ecological systems" (Nilsson 1986). Very detailed ecological information (largely derived from integrated monitoring data) is needed to establish these critical loads. The challenge to the monitoring programs is further complicated by the fact that some environmental changes that occur might be deemed beneficial. For example, increasing the amount of nitrogen deposition to forests may increase their growth rate (Nilsson 1988). However, nitrogen deposition may also lead to a terrestrial ecosystem that is made up of plant species different from the natural ecosystem (Nilsson 1986). The new plant community could be perceived by some to be less desirable or, by others, as more desirable than the original condition of the system.

4.2.2 Long-Term Ecological Effects

The ecological responses to these elements and other factors such as the invasion of exotic species are long-term (i.e., decades). Phosphorus contained in raw and partially treated sewage has been discharged into Lake Erie and Lake Ontario since the beginning of the 20th century, yet the dense algal mats, fish kills, and oxygen depletion in the bottom waters of the lakes did not become obvious until the 1960s.

Atlantic salmon began to decrease in numbers from several Nova Scotian rivers in the 1950s and completely disappeared by 1980 (Bargay and Riorden 1983). Monitoring of the Atran-Hogvadsan River in Sweden documented the nearly complete loss of salmon (*Salmo salar*) between 1952 and 1978 (Edman and Fleischer 1980). These decreases have been explained by the slow acidification of the runoff water in response to the acid deposition that began in the early part of this century. Global average temperature (Figure 4.1) has shown several trend characteristics. There was an increase prior to 1940, little further change or even a decrease until about 1980, and then a sharp increase until about 1984, followed by a slower increase. Long-term data are needed to define these changes and to establish that there is an overall increasing trend for the period. Data sets of several decades can give misleading information on the long-term trends.

Ryan et al. (1994) have analyzed tree ring growth for sugar maples in Ontario on a regional basis. After removing the effects of aging and weather variations, they found decreases in growth rates that they associated with environmental pollution. These growth decreases had been occurring for over 30 years in trees that appeared healthy. Watmough et al. (1999) have shown that sugar maples currently observed to be in decline have experienced decreased growth rates since the 1940s.

The invasion of zebra mussel (*Dreissena polymorpha*) in Lake St. Clair has brought about gradual decreases to the point of extinction in the native mussel population as measured over an 8-year study period (Nalepa et al. 1996). The increased water filtering capacity of the zebra mussel population has increased water clarity twofold (Griffiths 1993).

Zebra mussel have increased rapidly in the St. Lawrence River, from almost nonexistent in 1991 to densities of 20,000 mussels/m^2 at the Soulanges Canal (Ricciardi et al. 1996), and at Beauharnois, Becancour, and Île d'Orléans (De LaFontaine, et al. 2000). While there is concern for ecological damage by this invasion, there are some scientific benefits. Since the mussels filter so much water, they accumulate contaminants. Researchers at the St. Lawrence Centre are taking advantage of this feature by monitoring physiological change in the mussels such as the rupture of genetic material and hormonal changes in reproductive function—"biomarkers" of the long-term effects of the contaminants.

4.2.3 Effects Occurring at the Ecosystem Level

The biological response to stresses is often manifested at the ecosystem level. The sulfate component of acid rain leads to leaching of nutrients, notably calcium and magnesium, from sensitive soils, which in turn decreases tree growth rates (Hall et al. 1997). This may further lead to weakening of the trees, making them more vulnerable to attack by insects and diseases, the observed cause of death. There is a long chain of ecological responses beginning with acid rain as the applied stress and ending with the observed result, death by insects or disease.

4.2.4 Overlapping Ecological Effects of Various Stresses

The responses to one stress can overlap with and aggravate the responses to another stress. For example, when watersheds affected by acid rain experience dry weather conditions, sulfuric acid can be formed by the reoxidation of stored sulfur compounds, leading to pulses of acidic runoff and reacidification of the receiving lakes (Keller et al. 1992). Thus, the warmer climate conditions accentuated the effects of another stress—acid rain, in this case—with severe results.

4.2.5 Increased Size of Geographical Areas Affected by Environmental Stress

Environmental stresses have increased in geographical scope. Eutrophication of lakes was essentially a local problem, although it assumed binational proportions in the

Great Lakes region. Acid rain affected lakes and forests in a major part of eastern North America and also in most of Western Europe; thus, it became a multinational problem with respect to effects and controls. Stratospheric ozone depletion affects even larger geographic areas than acid rain, and, although the effects were still confined to the more southerly and northerly latitudes, emission controls affect nearly every country in the world. Climate change affects the entire globe and all countries are involved in the complex of effects and controls.

These characteristics of natural elements causing damage/change over long periods of time, at the ecosystem level and on a global scale, have made integrated monitoring essential and also very difficult.

4.3 GLOBAL RESPONSES TO ENVIRONMENTAL CHALLENGES

In recognition of the need for data to define ecological change and to defend expensive control measures, there has been a global response to the design and conduct of monitoring programs, representing all four categories of monitoring activity.

4.3.1 International Conventions to Protect the Environment

In order to deal with regional and global environmental change, it has become necessary to develop new scientific and political mechanisms that operate at the international level. The first step in the process is the establishment of a written agreement known as a Convention. Conventions usually deal with one issue and are frequently negotiated under the general guidance of the United Nations Environmental Programme. There are over 260 international Conventions and agreements that deal with environmental issues (Nelson 1998).

Many international Conventions require countries to establish monitoring programs in support of the particular issue covered by the Convention.

International Conventions usually have a common set of characteristics as follows:

- The Convention is intended to build international consensus that, in fact, a particular ecological, wildlife, or pollution problem exists and will, eventually, require international control action.
- The Convention is worded in general terms. Specific control action requirements usually are not included. This allows for countries to sign the Convention, thus agreeing that there is a problem but without agreeing on control measures which bring along corresponding difficult economic and social consequences.
- The Convention commits countries to conducting further research and monitoring on the issue, and frequently there is an agreement to prepare scientific assessments of the problem. This promotes more and better ecological information to be gathered, leading to improved consensus on

the issue through the assessment process. Governments are then in a better position to defend the need for control actions.

- The Convention commits countries to various reporting requirements, including submission of regular reports to the Conferences of the Parties regarding what they have done to enforce the Convention.
- The Convention sets up a secretariat to manage the overall process.
- The Conventions include agreement to negotiate protocols on specific control or other management actions needed to resolve the problem being addressed. This allows for a highly flexible approach to complex and multipollutant issues. Single compounds are addressed, leaving the more difficult or less important ones until later. Time frames for control action do vary for different compounds and are readily changed in light of new science and technology. Countries may choose not to sign a protocol without having to give up participation in the Convention activities.
- Control actions and schedules are frequently influenced by new ecological data as they are gathered from the monitoring networks.

The following are some examples of Conventions that require monitoring.

The **Convention on Long-Range Transboundary Air Pollution 1979**, signed by the countries of the United Nations Economic Commissions for Europe, sets out a reference to monitoring in its preamble. It says, "by means of… research and monitoring, to coordinate national action for combating air pollution." This objective is repeated in Article 6. Article 7 states that signatories shall initiate and cooperate in the conduct of research of the effects of sulfur compounds and other major pollutants on human health and the environment, including agriculture, forestry, materials, aquatic and other natural ecosystems and visibility with a view to establishing a scientific basis for dose/effect relationships designed to protect the environment.

Article 2.2(a) of the **Vienna Convention for the Protection of the Ozone Layer 1985** calls on signatories to "cooperate by means of systematic observations, research and information exchange in order to better understand and assess … the effects on human health and the environment from modification of the ozone layer."

Article 7(b) of the **Convention on Biological Diversity** requires signatories to "monitor, through sampling and other techniques … components of biological diversity; (d) Maintain and organise … data derived from … monitoring activities." Article 12(b) commits countries to promote and encourage research … [on] biological diversity. Annex 1 sets out the need for identification and monitoring of ecosystems and habitats, species and communities, and described genomes and genes.

The **UN Framework Convention on Climate Change** has two specific references to monitoring. Article 4(g) calls for countries to promote and cooperate in scientific, technological, technical, socio-economic and other research; systematic observation, and development of data archives related to the climate system … regarding the causes, effects … of climate change. Article 5(a) requires signatories to support and further develop … networks … aimed at defining, conducting, assessing and financing research, data collection and systematic observation; and (b) … strengthen systematic observation and national scientific research capabilities.

4.4 APPLICATION OF THE RESULTS OF MONITORING

Conducting monitoring programs, reporting results, and preparing scientific assessments are essential components of international environmental issues. Many of the Conventions include agreements by signatory countries to not only conduct monitoring and research, but also to report results and prepare scientific assessments.

The results of monitoring programs are frequently reported to the public in the form of indicators. An indicator is a statistic or parameter that is measured over time to provide information on trends in an environmental condition. Good indicators are sensitive to change, are supported by reliable, readily available monitoring data, are relevant to the issue, and are understood and accepted by intended users. Environmental indicators are selected key statistics that represent or summarize a significant aspect of the state of the environment, or of natural resource sustainability and the related human activities. To the extent possible, indicators provide information on what stresses are causing the trends. Environmental indicators are important tools for translating and delivering concise, scientifically credible information in a manner that can be readily understood and used by decision makers. They are usually accompanied by technical reports that discuss the manner in which the ecosystem and its components are responding to these changes and also the societal response undertaken to reduce or ameliorate these stresses.

Each year, The Worldwatch Institute in Washington, D.C. publishes a collection of monitoring data (Brown et al. 2002) that, as the book describes, indicates "the trends that are shaping our future." It reports on a wide range of environmental, social, economic, and business issues for which long-term monitoring information is available.

In Canada, a National Environmental Indicator Series has been produced for about 10 years by Environment Canada for a number of issues of national concern. In April 2003, National Indicators and Reporting Office published a report entitled *Environmental Signals: Canada's National Environmental Indicator Series 2003* that is comprised of 55 key indicators including most of the original National Environmental Indicator Series plus new issues such as solid waste and biodiversity. The report is available as a printed document and on The State of the Environment Infobase Website (http://www.ec.gc.ca/soer-ree/english/default.cfm). This report will be updated periodically as new data and information are acquired.

This larger report is accompanied by a shorter report entitled *Environmental Signals: Headline Indicators 2003*, which covers about a dozen of the main indicators (also available in print format and accessible at the same web site http://www.ec.gc.ca/soer-rree/English/headlines/toc.cfm).

Scientific assessments have become a major vehicle for assembling and reporting on the findings of monitoring programs. A notable example is the assessments prepared by the Intergovernmental Panel on Climate Change (IPCC). Three comprehensive assessments have been reported under the Framework Convention on Climate Change. The Convention commits signatories to conduct an assessment every 5 years. The latest assessment was released in 2001 (IPCC 2001b). These assessments make extensive use of monitoring results such as those shown in Figure 4.1 and Figure 4.2. The Executive Summaries for policy makers put the findings of the large, complex documents into a short report that is understandable to the public. These summaries are very influential in driving public response and action on climate change.

4.5 NATIONAL AND INTERNATIONAL MONITORING PROGRAMS

Thousands of environmental monitoring programs are being carried out around the world at local, national, and international levels. Following are descriptions of six examples that represent the type of work being done and that illustrate how the monitoring programs are being expanded to cope with new demands.

4.5.1 Canada's Ecological Monitoring and Assessment Network

Multidisciplinary environmental studies, particularly at the small watershed level, have been carried out in Canada for several decades. Studies were initiated by governments and academic institutions, usually to deal with environmental problems of interest to the specific location. For example, in the 1960s, the Federal Government initiated studies on lake eutrophication at the Experimental Lakes Area near Kenora, Ontario (Hecky, Rosenberg, and Campbell 1994), Laval University began the Centre for Arctic Studies at Kuujjaauapik which has focused on arctic and subarctic ecological processes, and studies on nutrient processes in surface waters began at Kejimkujik National Park. The Last Mountain Lake site in Saskatchewan was established as a National Wildlife Area in 1887. As new issues have emerged, sites such as Turkey Lakes in Ontario and Duschenay in Quebec were established in response to the need for more information on acid rain. These multiyear, integrated monitoring studies were very effective in resolving the site-specific scientific and policy questions set out by the supporting agencies.

In 1994, Environment Canada created the Ecological Monitoring and Assessment Network (EMAN) and a Coordinating Office (EMAN CO) to augment Canada's ability to describe ecosystem changes, provide timely information to decision makers, and help inform the Canadian public. EMAN links the many groups and individuals involved in ecological monitoring in Canada in order to better detect, describe, and report ecosystem changes. Essential elements of EMAN include various national and regional monitoring programs, over 80 long-term integrated ecosystem monitoring sites, and a diversity of ecological monitoring initiatives conducted by numerous partners at all levels of government, by nongovernment organizations, and by volunteers.

EMAN CO functions include:

- Coordinating within the Network the collection, access, integration, management, interpretation, and application of sound data on ecosystem status and trends; the timely identification of emerging environmental issues;

collaboration with related international networks; and the effective delivery of scientific information so as to improve decision making.
- Fostering the development, delivery, and promotion of best practices in ecological monitoring through protocol development and standardization, data collection and management, peer-reviewed journal publication, Web presence, National Science meetings, public education programs, and outreach.
- Developing and maintaining the existing Network and increasing its relevance to Environment Canada's program and priorities.

4.5.2 Monitoring of Stratospheric Ozone Depletion

Ozone is formed in the stratosphere by reactions between ultraviolet (UV) solar radiation and oxygen molecules. Ozone is not a very stable compound and it decomposes to molecular oxygen. The formation and decomposition are in a dynamic equilibrium with daily, seasonal and year to year variations in the amount of ozone present at any given time.

Stratospheric ozone is important to the Earth's climate as it interacts with solar radiation and with the thermal radiation emitted by the Earth's surface. Ozone also partially screens the Earth's surface from UV-B radiation that would otherwise be harmful to living organisms.

Scientists first postulated and then observed that anthropogenic chemicals such as chlorofluorocarbons, halons, and some chlorine containing solvents can react with ozone to accelerate the decomposition reactions, leading to a depletion of the ozone concentrations. This depletion would allow more of the high-energy UV-B radiation to reach the earth surface with possible biological effects. Less ozone would also affect the atmospheric radiation balance with implications for global weather.

Stratospheric ozone depletion was one of the first environmental problems of global proportions and control programs implemented under the Montreal Protocol and its Amendments involve most countries.

The amount of ozone in the atmosphere over a given point can be measured at ground level by spectrophotometric methods. The Canadian-invented Brewer spectrophotometer continuously measures the total ozone in the column of the atmosphere above the instrument and the UV-B radiation.

Canada has established a national network of 12 stations with Brewer instruments, and there is a global network that has over 150 Brewers. The Brewer instruments give the total ozone in the column of air above the site but not the height at which it occurs. Small instruments sent aloft by balloons measure the vertical concentration profiles. Measurements are made daily and monthly at about 260 locations around the world.

Ozone data from the monitoring networks are gathered at the World Ozone Data Centre (WODC) that was initiated in connection with the International Geophysical Year of 1957/8. It has been run since the early 1960s in Toronto by the Meteorological Service of Canada (MSC) for the World Meteorological Organization. In 1992 the MSC took on the additional task of operating the World Ultraviolet Radiation Centre (WURC). Together, the two data centers comprise the World Ozone and Ultraviolet

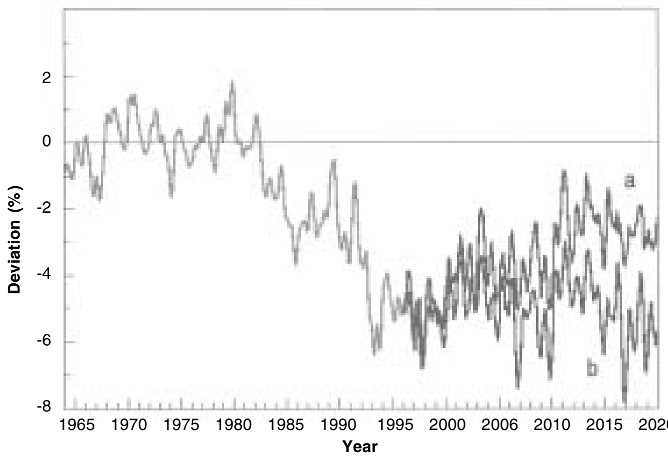


FIGURE 4.5 Statistical distribution of decadinal changes in biological oxygen demand (BOD) of rivers by continent and time period.

Radiation Data Centre (WOUDC) that includes participation by over 100 independent agencies from 75 countries.

In addition, the MSC operates the Brewer Data Management System (BDMS) that gathers data from over 80 Brewer spectrophotometers located in 15 countries. The WMO Global Atmospheric Watch (GAW) provides external advice to the WOUDC through the Advisory Group on UV radiation and ozone.

The annual MSC/WMO publication *Ozone Data for the World* (ODW), also called the "Red Book," was the main output of the WODC for many years but has been replaced since 2000 by an ODW CD-ROM. The ozone and UV data, the software collection, and descriptive information are available and can be found through the WOUDC Website (www.woudc.org).

Ozone depletion has been observed during the past two decades. Figure 4.5 shows the global average values along with two possible future scenarios that depend on the on-going success of control programs.

4.5.3 International Cooperative Programs Established under the Convention on Long-Range Transboundary Air Pollution

There has been extensive international pollution control and monitoring activity conducted under the Convention on Long-range Transboundary Air Pollution. The Convention established the European Monitoring of Atmospheric Pollution (EMAP) network that measures air quality and deposition of several pollutants in Europe. As pollution control Protocols were developed and ratified, there was an increasing interest in monitoring the effects of the pollutants and the ecological response to the control programs. This has lead to the establishment of five International Co-operative Monitoring Programs (ICPs) that include Forests, Rivers and Lakes, Crops, Building

Materials and Cultural Monuments, and Integrated Monitoring. Canada participates in all programs except Crops. The programs have generated considerable information about the environmental effects of air pollutants—notably sulfur, nitrogen, and tropospheric ozone.

The Integrated Monitoring Programme (IMP) has the four specific objectives listed above in its description of Integrated Monitoring. The program presently includes about 50 sites in 20 countries in Europe plus Canada (Kleemola and Forsius 2001).

The sites are divided into two categories (Kleemola and Laine 1997):

1. Intensive monitoring sites (A-sites) where samples are collected and observations made for many compartments in the ecosystem for the application of complex models. Intense investigations of dose–response relationships are also carried out. Strict siting criteria have been set for the A-sites. These sites are normally located in protected areas.
2. Biomonitoring sites (B-sites) have the objective of quantifying the variation between sites concerning some of the more important features like input–output mass balance models of elements and models for bioindicators on the spatial basis. Biomonitoring for detecting natural changes, effects of air pollutants, and climate change is a particular aim of these sites.

The 1997 Annual Report (Kleemola and Laine 1997) of the IM concluded that decreases in sulfur dioxide emissions have resulted in decreases in wet and dry deposition of sulfate. Correspondingly, the sulfate concentrations and the acidity in runoff water have generally decreased.

The results have also shown a complicated watershed response to nitrogen deposition. Some sites have decreasing nitrate concentrations in runoff even though the nitrate deposition has not changed. Other sites have recorded increased nitrate in runoff, indicating the possible onset of nitrogen saturation.

Multivariate statistical analysis of forest damage data from IM sites explained damage as being 18%, 42%, and 55% caused by the combined action of ozone and acidifying sulfur and nitrogen compounds in the air (Kleemola and Forsius 2001).

4.5.4 Monitoring the Effects of Acid Rain

The Canadian Forest Service (CFS) established the ARNEWS in 1984 in response to prevalent concerns about the effect of acid rain on the health of Canada's forests. Analysis of ARNEWS data strived to detect early signs of change or damage to forest trees and soils attributable to air pollution and not to damage associated with natural causes or management practices. ARNEWS was a good example of one of the main reasons for monitoring, i.e., simple monitoring applied to determine the status of the resource and to determine if any observed changes are outside of the normal variation. Long-term changes in vegetation and soils attributable to acid rain and other pollutants were monitored. Symptoms of damage from air pollution are not obvious and frequently resemble damage from natural causes. Experience of field professionals trained to distinguish these symptoms from abnormal climatic conditions, inherent nutrient deficiencies, and the effects of insects and diseases was crucial to ARNEWS.

Forest decline is defined as a continued and sustained deterioration of conditions ending in the death of trees. ARNEWS did not identify any extensive areas of decline other than that of white birches near the Bay of Fundy where acid fog has been implicated. However, a clear correlation has been determined between tree mortality and exceedances over critical loads, using data from a number of ARNEWS plots. Researchers at the University of New Brunswick found that tree mortality increased where the amount of acid deposited from the atmosphere exceeds the calculated acid tolerance level of the forest soil at the site. It was determined that in 1995, trees on 17 ARNEWS forest plots had visible damage. All 17 plots were in areas where acid deposition exceeded critical loads. On the other hand, none of the plots outside of the exceedance area had visible damage that could not be explained by other factors such as insects, weather, or stand dynamics (Moayeri and Arp 1997).

The New England Governors and Eastern Premiers (NEG/EP) Secretariat is currently engaged in a project to map the sensitivity of northeastern forests to acidic deposition (New England governors and Eastern Premiers 1998). The Forest Mapping Working Group of NEG/EP is developing critical loads and exceedances for these forests. Data from ARNEWS comprise the basis for the modeling exercise involving the Canadian component of the study. The project is also examining the impact of forest management practices on critical loads and exceedances. NEG/EP hopes to finalize the sensitivity maps by 2004 and produce a final report on the impact of continued acidic deposition on the productive forest land of northeastern North America.

On some ARNEWS plots in areas of higher sulfate and nitrate deposition, soil base cation depletion from the forest floor was observed and the prevailing thought is that this depletion will likely be sustained or increased at current deposition levels (Morrison et al. 1995).

In response to changing program objectives, this network of 151 plots across the country ceased to exist as an entity in 1998. Some ARNEWS plots were archived in a manner that would permit their use in the future if the need arose, while other plots were incorporated into continuing issue-based monitoring being carried out by CFS and its partners.

CFS has incorporated 18 ARNEWS plots into the Forest Indicators of Global Change Project (FIGCP) that incorporates a gradient of atmospheric pollution conditions from Ontario to the Maritimes. In total, 25 plots make up the gradient, also incorporating four former North American Maple Project (NAMP) plots as well as three additional sites. This initiative began in 1999 and is designed to develop new, early warning indicators of forest condition; to investigate interactions between air pollution, climate change, and forest productivity; and to establish an array of permanent research-monitoring (integrated monitoring) plots to conduct detailed studies of nutrient/carbon cycling in eastern Canada. ARNEWS plots also comprise part of a network of plots set up to study aspen decline in Alberta and Saskatchewan.

4.5.5 Global Environmental Monitoring System (GEMS)

Canada participates in the global freshwater quality agenda through its United Nations sponsorship of the GEMS/Water Programme. The program headquarters are located at the National Water Research Institute in Burlington, Ontario. GEMS/Water is a UN program on global water quality that was initiated in 1976 by United Nations Environment Programme (UNEP). It is the only international UN program strictly devoted to water quality and significantly contributes toward a global appreciation of current water quality status and trends while promoting sustainable freshwater quality management. The main activities of the GEMS/Water program include international cooperative data program and monitoring; data and information sharing; global and regional assessments; capacity building and technical expertise; advice to governments and international agencies; information products; and partnerships.

Data are submitted to GEMS/Water for entry into the global database (GLOWDAT) from approximately 800 stations in about 70 participating countries including South Africa, Japan, Thailand, Belgium, France, Senegal, New Zealand, and Russia. The database currently consists of approximately two million data points, representing six classes of water assessment variables. GEMS/Water has also recently inaugurated a new database on pathogenic organisms. Extracts from the database are used for the preparation of regional and global water assessments requested by UN agencies and in the course of various research projects carried out by public and private-sector organizations. Figure 4.6 is an illustration of data presentation.

The GEMS/Water Programme Office also contracts with outside sources to carry out special activities such as the evaluation and assessment of the strategic water quality monitoring program of the Panama Canal Authority.

A number of activities are being undertaken to solicit new countries to participate in the GEMS/Water program, with an emphasis upon developing countries, particularly in Africa. Since 1998, participation in GEMS/Water by national governments and agencies has significantly increased. This reflects the growing awareness and concern about freshwater quality and availability in the environment, and the impact that these resources have upon the lives of the people of the world. GEMS/Water provides a focal point for water quality data and information designed to provide input to scientific assessments undertaken by the UN on activities leading to sustainable development.

4.5.6 Global Climate Observing System

The increase in atmospheric concentrations of greenhouse gases, atmospheric emissions of sulfur dioxide, formation of tropospheric ozone, and depletion of stratospheric ozone creates a complex set of interactions affecting the radiation balance of the atmosphere (IPCC 2001a). These changes are likely to have already increased the average global temperature (Wigley et al. 1998; Crowley 2000). Scientists predict increases of 1.4 to 5.8°C over the period 1990 to 2100 (IPCC 2001a). In addition, there are variations in the global temperature brought about by natural fluctuations in the sun's energy, and these were also likely to have contributed to the observed increase in temperature over the past 150 years (Wigley et al. 1998). The El Niño phenomenon causes large variations in weather conditions around the globe on a decade time scale.

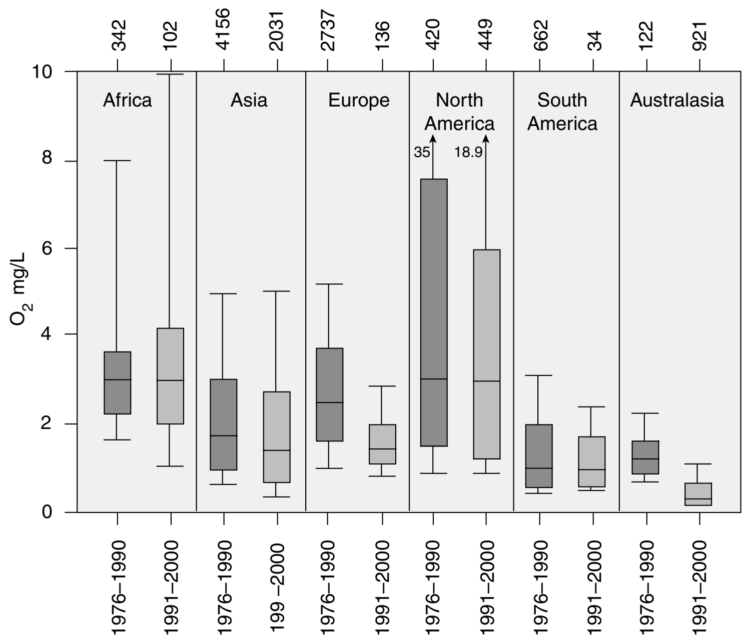


FIGURE 4.6 The global total ozone record from 1964 to 1996 and two hypothetical projections based on different control scenarios. (a) The best case assuming that the Montreal Protocol and its amendments will be fully implemented. (b) Based on the assumption that concentrations of all ozone-depleting substances remain at their 1997 levels.

The sheer geographical extent (i.e., global) and complexity of anthropogenically caused climate change requires a concerted international monitoring effort to measure changes and to determine their causes and ecological and sociological consequences. This international effort is being developed by the World Meteorological Organization and is coordinated by the Global Climate Observing System (GCOS).

While there are numerous monitoring programs already in place, GCOS recognized that new and reorganized programs are needed to enable nations to:

- Detect and quantify climate change at the earliest possible time
- Document natural climate variability and extreme climate events
- Model, understand, and predict climate variability and change
- Assess the potential impact on ecosystems and socio-economics
- Develop strategies to diminish potentially harmful effects
- Provide services and applications to climate-sensitive sectors
- Support sustainable development (Spence 1995)

The planning workshop in 1995 (Karl 1995) set out an array of monitoring needs in three main categories.

1. *Climate forcing and feedbacks:* This will include monitoring of the full suite of greenhouse gases: carbon dioxide (CO_2), methane (CH_4), carbon monoxide (CO), nitrous oxide (N_2O), chlorofluorocarbons (CFCs), ozone (O_3), and water vapor. This information will be combined with measurements of aerosols, solar radiation, and cloud cover to calculate energy budgets.
2. *Climate responses and feedbacks:* This program has 11 components ranging from detailed measurements of sea level, surface and sub-surface oceans processes, land surface air temperature, tropospheric and stratospheric temperatures, and precipitation and cryospheric changes to reanalyzes of climate model predictions and determining new variables that are sensitive to anthropogenic climate forcing.
3. *Climate impacts:* In the final analysis, it is the ecological effects of the changing atmosphere and changing climatic variables that are of most interest to the general public, resources managers, planners, and governments. Such questions — how is agricultural (food) production affected? what is happening to the forests? do extreme weather conditions threaten life and property? — are uppermost in the minds of the public.

GCOS will look beyond the climatological aspects of atmospheric change to assess changes in the global ecological condition. There will be continuing work on developing, measuring, and reporting on environmental indicators. Long-term monitoring of land surface characteristics such as vegetation, soil moisture, runoff, and surface temperatures are being considered.

The GCOS Planning Workshop concluded *inter alia* that "adequate long-term climate monitoring will continue to be critically dependent on developing a partnership among network operators, data managers, analysts, and modelers. Multi-purpose observing systems used for operations, research, and monitoring are likely to be the most practical means of achieving an economical long-term climate monitoring system."

4.5.7 Volunteer Monitoring Networks

In Canada and in other countries, there are large-scale, simple monitoring programs carried out by trained volunteers. These programs have great value by themselves in gathering and reporting on changes in the environment. However, they also augment the networks operated by scientific professionals. In an ideal case, the scientific explanations for change are derived at integrated monitoring sites. These explanations can then be applied to the observations from the volunteer networks in order to provide a complete picture of the geographical cause and extent of the environmental change.

In establishing volunteer monitoring programs it is important to select variables that are relatively easy and inexpensive to measure so that large numbers of people can participate. It is also essential to establish scientifically sound protocols and to provide training to participants via manuals, videos, and lectures.

The following are examples of volunteer networks operating in Canada.

4.5.8 Environment Canada's Volunteer Climate Network

Across Canada, approximately 1000 individuals or agencies record their local weather conditions twice daily and send this information to Environment Canada. The data include daily maximum and minimum temperature, rainfall, snowfall, and snow cover. The majority of the reports are mailed to regional Climate Centres for inclusion in the climate data record. However, there is an increasing number of observers who input their observations on-line for immediate use by forecasters and climatologists. This tradition of weather record-keeping dates back to 1840 and is a proven source of valuable data for climate inquiries and studies.

4.5.8.1 PlantWatch

PlantWatch is a phenology (study of the seasonal timing of life-cycle events) program which links students and other observers as the "eyes of science," tracking the green wave of spring moving north. The initial program, established in 1995, can be reached at the Website www.devonian.ualberta.ca/pwatch. Students develop scientific skills while observing springtime changes in plants and learning about biodiversity. Observers monitor flowering of native and cultivated plants and report the bloom times to central scientists over the Internet or electronic mail. A Teacher Guide describes the program and curriculum connections. Schools are encouraged to establish "PlantWatch Gardens," planting the key indicator species.

There has been an observed trend in western Canada to earlier flowering associated with strong El Niño events, warmer ocean temperatures, and warmer winter temperatures as show in Figure 4.7 (Beaubien and Freeland 2000). This valuable seasonal information helps decision making for farmers and foresters, i.e., to correctly time operations such as planting, fertilizing, and crop protection, and to predict harvest timing. It also is useful in wildlife management (e.g., in early springs, more deer fawns are successful); human health (pollen warnings for allergy sufferers); and tourism (best times to photograph flowers or animals or to go fly-fishing).

The national coordinator for PlantWatch is based at the University of Alberta's Devonian Botanic Garden, home of the Alberta Wildflower Survey (renamed Alberta PlantWatch). Observations on the common purple lilac have been carried out in Europe and Asia (Beaubien 2003).

In 2001 Environment Canada and the Canadian Nature Federation expanded PlantWatch by funding work to locate coordinators in each province and territory, and by producing a booklet in 2002 and Web page describing PlantWatch (see www.naturewatch.ca).

4.5.8.2 NatureWatch

In recent years, Environment Canada has expanded volunteer networks and Community Based Monitoring (CBM). CBMs often take the form of citizen-science monitoring programs that are coordinated and supported by governmental agencies. Examples of CBMs are the NatureWatch programs established by EMAN in partnership with the Canadian Nature Federation (see www.naturewatch.ca). NatureWatch programs are designed to collect reliable information that can contribute to

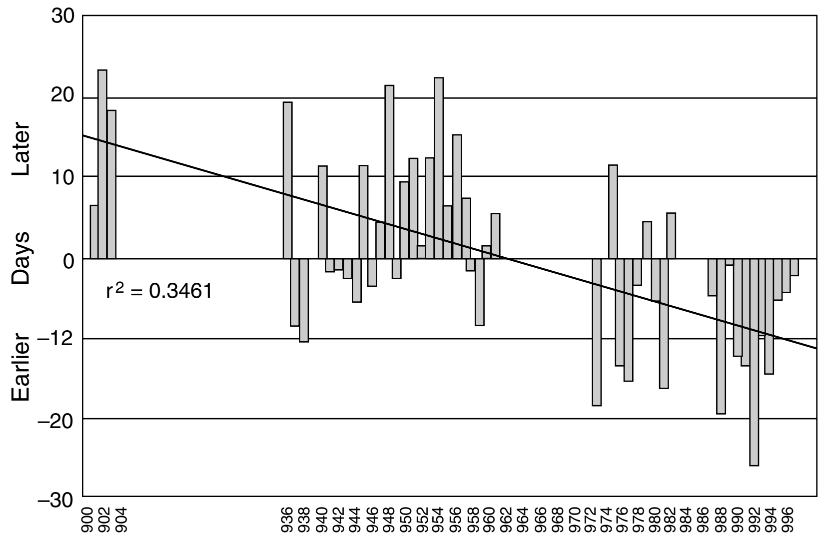


FIGURE 4.7 Long-term trend (1901–1997) in first-flowering dates of *Populus tremuloides* at Edmonton, Alberta. The Julian dates of flowering are shown as deviations from the mean bloom date for all data. Phenology data for 1901–1903, 1936–1961, 1973–1982 and 1987–1997 are plotted (no deviation values = zero).

local, regional, and national monitoring programs. The NatureWatch programs are internet-based but allow for hardcopy observation submissions. FrogWatch (www.frogwatch.ca) was launched in the spring of 2000 and collects information on the distribution and abundance of anurans across Canada. This program is supported by partnerships with anuran experts in each province and territory who check the submissions for accuracy and investigate outliers. WormWatch (www.wormwatch.ca), developed in partnership with the agriculture industry and AgriFood Canada, collects information on the distribution and abundance of earthworms in Canada and was launched in the fall of 2001. IceWatch (www.icewatch.ca), developed in partnership with the Meteorological Service of Canada and Laval University, collects information on lake and river ice phenology and was also launched in the fall of 2001.

4.6 THE FUTURE

Environmental monitoring programs have a good track record in meeting the key objectives of defining problems and their solutions, reporting on the effectiveness of control programs, and in identifying new issues. In spite of substantial success in solving, or at least taking action, on many environmental problems, we have lost ground in the overall objective of protecting the environment. Environmental disruption is occurring for many reasons, such as land use change and the deliberate and accidental introduction of exotic species.

In Lake Erie, for example, prior to 1992, reductions in phosphorus input resulted in the expected reductions in the growth of algae (Dobson 1994). However, since then, the accidental introduction of exotic species, notably zebra and quagga mussels, has resulted in a complex web of ecological changes affecting phosphorus concentrations, algal growth, hypolimnetic oxygen concentrations, and fish production (Charlton 1999; Nicholls 1999). This has lead Charlton (1999) to conclude that "Lake Erie in the 1990s is a lake in transition," a matter of consderable concern given the great economic and social importance of the lake as a municipal water supply and recreation and commercial fishery. In 1997, in an attempt to explain the variations in long-term trends in algal growth in Lake Erie, Nicholls (1997) concluded that "Effective ecosystem management of the Great Lakes depends on sound interpretation of long-term environmental data." Events are proving him correct as even the seemingly successful phosphorus management plan itself has been questioned regarding its ongoing effectiveness (Charlton 1999). Monitoring and understanding the changes in the Great Lakes is of critical importance for the effective and safe management of this vital water resource.

A very serious environmental issue facing the globe is the fact that anthropogenic activities are changing the chemical composition of the global atmosphere. For example, increasing concentrations of carbon dioxide, methane, nitrous oxide, and HFCs have been recorded by monitoring programs. In addition, sulfur and nitrogen compounds have changed the chemical character of precipitation (acid rain) on a regional scale in North America, Europe, and Asia. These changes, in turn, alter the physical properties of the atmosphere with responses such as stratospheric ozone depletion, ground-level ozone formation, and modifications to the radiation balance. Changes in radiation balance are expected to increase the average global temperature (global warming). The biosphere is affected by the chemistry and physics of the atmosphere, so we would expect it to respond to these new conditions, and it already has.

Emissions of sulfur dioxide and nitrogen oxides have altered the chemical characteristics of precipitation in areas near and downwind of large sources of these pollutants. Damage to lakes, forests, human health, building materials, and atmospheric visibility has been documented in eastern North America (Environment Canada 1997). That report also drew attention to the fact that present control programs are not strict enough to fully protect the environment. Further reductions in emissions, particularly of sulfur dioxide, are needed.

Keeling et al. (1996) have reported on changes in the amplitude and timing of the global carbon cycle. The amplitude of the yearly cycle has increased by 20% at the Mauna Loa, HI observatory over the past 30 years and by about 40% over arctic sites at Alert in Canada and Pt. Barrow in the U.S. In addition, the yearly minimum atmospheric concentration of carbon dioxide that occurs in July is now observed to occur about a week earlier than 30 years ago. These changes appear to be a response to increasing average temperatures. Keeling et al. (1996) noted "These striking increases over 30 years could represent unprecedented changes in the terrestrial biosphere, particularly in response to some of the highest global annual average temperatures since the beginning of modern records, and particularly in response to plant growth being stimulated by the highest concentrations of atmospheric CO_2 in the past 150,000 years."

Briffa et al. (1998) reported on tree growth from 300 locations at high latitudes in the Northern Hemisphere. Over the past 50 years, the expected patterns of growth related to temperature have not been observed. Instead, growth rates are less than expected. While the reasons for these observed changes are not known, they are an indication of wide-scale disruptions of normal processes in the biosphere. McLaughlin and Percy (1999) report that "in North America the regional patterns of the most frequent disease problems documented by forest surveys, are spatially consistent with the patterns of highest levels of ozone and acidic deposition." The physiological effects of air pollutants may be predisposing trees to other stresses or amplifying their effects. Such changes should be viewed, at least as an early warning, of further ecosystem responses to the changing atmosphere.

The IPCC 2001 Executive Summary of Working Group II (IPCC 2001c) drew attention to the complexities of plant growth by noting that previous attribution of increased terrestrial uptake of carbon dioxide to increased CO_2 concentrations and temperature and moisture changes were not confirmed by field observations. Changes in uptake of carbon may be more due to changes in land uses and management than climate. The need to understand terrestrial sinks of carbon dioxide is of vital importance as such sinks are accepted as part of countries overall carbon dioxide emission control actions (Kyoto Protocol 1997).

Wardle (1997) has reported on recent changes in stratospheric ozone depletion over Canada's Arctic. Low values of ozone in the spring have been increasing in both frequency and severity, due to low temperatures in the stratosphere. The lower temperatures in the stratosphere may in turn be caused by higher concentrations of carbon dioxide. Wardle postulates that increasing atmospheric concentrations of carbon dioxide will cause even lower temperatures to occur in the winter stratosphere, thus increasing the early spring ozone depletion. This may happen even though the concentrations of CFCs in the atmosphere have increased very little in past few years.

Overall, we can anticipate a wide range of complex ecological responses to the changing chemical and physical properties of the atmosphere. Land use change and exotic species will further add to the ecological stress. The natural resources that are being affected are the basis of large parts of the economies of North America. It is essential that we record and understand changes in these ecosystems in order to manage the associated industries in a sustainable way. This represents a major challenge for our monitoring programs and particularly for the integrated monitoring sites.

ACKNOWLEDGMENTS

Many people have provided essential reports, text, and graphs. Special thanks to E. Beaubien, Brian Craig, A. Fenech, E. Hare, H. Hirvonen, W. Hogg, Gary Ironside, J. Rudd, Bryan Smith, E. Snucins, H. Vaughan, D. Wardle, and the UNEP & WHO GEMS/Water Collaborating Centre. Helpful reviews of the text have been provided by Marilyn Brydges and Adam Fenech.

This chapter is a revision of Chapter 33, Survellance Environmentale, in *Environment et Santa Publique, Fondements et Praticues*, Gerin, M. et al., Eds., Edisem, 2003. I very much appreciate the editors' permission to print this revision and for their assistance in preparing some of the figures.

REFERENCES

Bangay, G.E. and C. Riordan (co-chairs). 1983. *Memorandum of Intent on Transboundary Air Pollution*, Impact Assessment Work Group 1, Final report.

Beaubien, E.G., and H.J. Freeland. 2000. Spring phenology trends in Alberta, Canada: links to ocean temperature. *Int. J. Biometeorol.* 44: 53–59.

Beaubien, E.G., 2003. Devonian Botanic Garden, University of Alberta, Edmonton, Alberta, Canada, T6G2E1. Personal communication.

Briffa, K.R., F.H. Schweingruber, P.D. Jones, T.J. Osborn, S.G. Shiyatov, and E.A. Vaganov. 1998. Reduced sensitivity of recent tree growth to temperature at high northern latitudes. *Nature* 391: 678–682.

Brown, L.R. et al. 2002. *Vital Signs.* W.W. Norton, New York.

Charlton, M.N., R. LeSage, and J.E. Milne, 1999. Lake Erie in transition: the 1990s. In *State of Lake Erie (SOLE) — Past*, *Present and Future*. M. Munawar, T. Edsall, and I.F. Munawar (Eds.). Ecovision World Monograph Series, pp. 97–123.

Crowley, T.J. 2000. Causes of climate change over the past 1000 years. *Science* 289 (July 14): 270–276.

Dansgaard, W., S.J. Johnsen, H.B. Clausen, D. Dahl-Jensen, N.S. Gundestrup, C.U. Hammer, C.S. Hvidberg, J.P. Steffensen, A.E. Sveinbjörnsdóttir, J. Jouzel, and G.C. Bond. 1993. Evidence for general instability of past climate from a 250-year ice-core record. *Nature* 264: 218–220.

De LaFontaine, Y., F. Gagne, C. Blaise, G. Costan, and P. Gagnon. 2000. Biomarkers in Zebra mussel (*Dreissena polymorpha*) for the assessment and monitoring of water quality of the St. Lawrence River (Canada). *Aquat. Toxicol.* (accepted for publication).

Dixit, S.S., J.P. Smol, J.C. Kingston, and D.F.I. Charles. 1992. Diatoms: powerful indicators of environmental change. *Environ. Sci. Technol.* 26: 23–33.

Dobson, H.F.H. 1994. *Lake Ontario water quality trends, 1969 to 1992: some observational nutrient-science for protecting a major and vulnerable source of drinking water.* National Water Research Institute Contribution No. 94-58.

Edman, G. and S. Fleischer. 1980. The River Hogvadsan liming project — a presentation. *Proceedings of the International Conference on the Impact of Acid Precipitation*, SNSF Project, Norway, pp. 300–301.

Environment Canada. 1997. *Canadian Acid Rain Assessment*, *Volume One*, *Summary of Results.*

Environment Canada. 2003. Environmental Signals: Canada's National Environmental Indicator Series 2003. http://www.ec.gc.ca/soer-ree/English/Indicator_series/default.cfm#pic.

Environment Canada. 2003. Environmental Signals: Headline Indicators 2003. http://www.ec.gc.ca/soer-ree/English/headlines/toc.cfm.

Griffiths, R.W. 1992. Effects of Zebra Mussels (*Dreissena polymorpha*) on the benthic fauna of Lake St. Clair. In *Zebra Mussels: Biology, Impacts and Control*. T.F. Nalepa and D.W. Schloesser (Eds.). Lewis Publishers/CRC Press, Boca Raton, FL, pp. 415–437.

Hall, P. et al. 1997. *Canadian Acid Rain Assessment, Volume Four, The Effects on Canada's Forests*, Environment Canada, Catalogue number En56-123/4-1997E.

Hall, R.I. and J.P. Smol. 1996. Paleolimnological assessment of long-term water quality changes in south central Ontario lakes affected by cottage development and acidification. *Can. J. Fish. Aquat. Sci.* 53: 1–17.

Hecky, R.E., D.M. Rosenberg, and P. Campbell. 1994. The 25th Anniversary of the experimental lakes area and the history of lake 227. *Can. J. Fish. Aquat. Sci.* 51: 2243–2246.

Hicks, B.B. and T.G. Brydges. 1994. A strategy for integrated monitoring. *Environ. Manage.* 18(I): 1–12.

Houghton, J.J., G. Jenkins, and J.J. Ephraums (Eds.). 1990. *Climate Change: The IPCC Scientific Assessment*. Cambridge University Press, Cambridge, U.K.

International Joint Commission, 1969. *Pollution of Lake Erie, Lake Ontario and the International Section of the St. Lawrence River. Volume 1 — Summary.* Report by the International Lake Erie Water Pollution Board and the International Lake Ontario — St. Lawrence River Water Pollution Board.

IPCC, 2001a. *Summary for Policy Makers, A Report of Working Group 1 of the IPCC.* World Meteorological Organization/United Nations Environment Programme.

IPCC, 2001b. *Climate Change 2001: The Third Assessment Report of the Intergovernmental Panel on Climate Change.* World Meteorological Organization/United Nations Environment Programme.

IPCC, 2001c. *Summary for Policy Makers: Climate Change 2001: Impacts, Adaptation and Vulnerability.* IPCC Working Group II World Meteorological Organization/United Nations Environment Programme.

Jeffries, D.S. 1997. *Canadian Acid Rain Assessment, Volume Three, The Effects on Canada's lakes, Rivers, and Wetlands.* Environment Canada.

Karl, T., F. Bretherton, W. Easterling, C. Miller, and K. Trendberth. 1995. Long-term climate monitoring by the Global Climate Observing System (GCOS): an editorial. *Climate Change* 31(2–4): 135–147.

Keeling, C.D., J.F.S. Chin, and T.P. Whorf. 1996. Increased activity of northern vegetation inferred from atmospheric CO_2 measurements. *Nature* 382(6587): 146–149.

Keller, W., J.R. Pitblado, and J. Carbone. 1992. Chemical responses of acidic lakes in the Sudbury, Ontario area to reduced smelter emissions. *Can. J. Fish. Aquat. Sci.* 49(Suppl. 1): 25–32.

Kelly, C.A. and W.M. Rudd. 1997. Increases in fluxes of greenhouse gases and mercury following flooding of an experimental reservoir. *Environ. Sci. Technol.* 31(5): 1334–1344.

Kleemola, S. and Y. Laine. 1997. In S. Kleemola and M. Forsius (Eds.). 6th Annual Report 1997, UN ECE Integrated Monitoring. *Finnish Environ.* 116: 6–11. Finnish Environment Institute, Helsinki, Finland.

Kleemola, S. and M. Forsius (Eds.) 2001. International Cooperative Programme on Integrated Monitoring of Air Pollution Effects on Ecosystems, 2001. 10th Annual Report. Finnish Environment Institute, Helsinki, Finland.

Kyoto Protocol, 1997. United Nations Environment Programme.

McLaughlin, S. and K. Percy. 1999. Forest health in North America: some perspectives on actual and potential roles of climate and air pollution. *Water Air Soil Pollut.* 116: 151–197.

Moayeri, M.H. and P.A. Arp. 1997. Unpublished data, assessing critical soil acidification load effects for ARNEWS sites; preliminary results. University of New Brunswick, Fredericton.

Morrison, I.K., R.E. Fournier, and A.A. Hopkins. 1995. Response of forest soil to acidic deposition: results of a five year re-sampling study in Canada. In *Air Pollution and Multiple Stresses.* R.M. Cox, K.E. Percy, K.F. Jensen, and C.J. Simpson (Eds.). *Proceedings of the 16th IUFRO International Meeting for Specialists in Air Pollutant Effects on Forest Ecosystems*, Fredericton, September 7–9, 1994. Nat. Res. Can., Can. For. Ser., Fredericton, 402 pp.

Nalepa, T.F., D.J. Harston, G.W. Gostenik, D.L. Fanslow, and G.A. Lang. 1996. Changes in the freshwater mussel community of Lake St. Clair: from Unionidae to Dreissena polymorpha in eight years, *J. Great Lakes Res.* 22(2): 354–369.

National Research Council of Canada. 1981. Acidification in the Canadian aquatic environment: scientific criteria for assessing the effects of acidic deposition on aquatic ecosystems. Report No. 18475, National Research Council of Canada, Ottawa.

Nelson, D.D. 1998. *International Environmental Auditing.* Government Institutes, Rockville, MD, U.S.A.

New England Governors and Eastern Premiers, 1998. *Acid Rain Action Plan*, prepared by The Committee on the Environment of The Conference of New England Governors and Eastern Canadian Premiers. Adopted by NEG/EP in Fredericton, New Brunswick, Canada, June 1998.

Nicholls, K.H. 1997. Planktonic green algae in Western Lake Erie: the importance of temporal scale in the interpretation of change. *Freshwater Biol.* 38: 419–425.

Nicholls, K.H., G.J. Hopkins, and S.J. Standke. 1999. Reduced chlorophyll to phosphorus ratios in nearshore Great Lakes waters coincide with the establishment of dreissenid mussels. *Can. J. Fish. Aquat. Sci.* 56: 153–161.

Nilsson, J. (Ed.). 1986. Critical Loads for Sulfur and Nitrogen, Nordic Council, Copenhagen.

Nilsson, J. (Ed.). 1988. Critical Loads for Sulfur and Nitrogen, Rep. NORD — 1988: 15, Workshop Skokloster, Sweden, Nordic Council of Ministers.

Reuss, J.O. and D.W. Johnson. 1986. *Acid Deposition and the Acidification of Soils and Waters*, Springer-Verlag, Berlin.

Ricciardi, A., F.G. Whoriskey, and J.B. Rasmussen. 1996. Impact of the (*Dreissena*) invasion on native unionid bivalves in the upper St. Lawrence River. *Can. J. Fish. Aquat. Sci.* 53: 685–695.

Ryan, D.A.J., O.B. Allen, D.L. McLaughlin, and A.M. Gordon. 1994. Interpretation of sugar maple (*Acer saccharum*) ring chronologies from central and southern Ontario using a mixed linear model. *Can. J. For. Res.* 24: 568–575.

Rudd, J.W.M., C.A. Kelly, D.W. Schindler, and M.A. Turner. 1990. A comparison of the acidification efficiencies of nitric and sulfuric acids by two whole-lake addition experiments. *Limnol. Oceanogr.* 35: 663–679.

Schindler, D.W. 1980. Experimental Acidification of a Whole Lake — a Test of the Oligotrophication Hypothesis. In *Proceedings of the International Conference on the Ecological Impact of Acid Precipitation*. D. Drablos and A. Tollan (Eds.). SNSF — Project Sandefjord, Norway, pp. 370–374.

Schindler, D.W. and E.J. Fee. 1974. Experimental Lakes Area: whole-lake experiments in eutrophication. *J. Fish. Res. Board Can.* 31(5): 937–953.

Spence, T. and J. Townshend. 1995. The Global Climate Observing System (GCOS): an editorial. *Climate Change* 31(2–4), 130–134.

Thompson, D. (Ed.). 1995. The Concise Oxford Dictionary, 9th ed. Clarendon Press, Oxford, U.K., 879 pp.

Vollenweider, R.A. 1968. The Scientific Basis of Lake and Stream Eutrophication, with Particular Reference to Phosphorus and Nitrogen as Eutrophication Factors. Technical Report DAS/CSI/68, OECD, Paris, 27: 1–182.

Wardle, D.I. 1997. Trends in Ozone over Canada, Ozone Depleting Substances and the UV-B. Air Waste Management Association, Calgary, 19 September.

Wardle, D.I., J. Kerr, C.T. McElroy, and D.R. Francis (Eds.). 1997. *Ozone Science: A Canadian Perspective on the Changing Ozone Layer.* Environment Canada, Atmospheric Environment Service. CARD 97-3.

Watmough, S., T. Brydges, and T. Hutchinson. 1999. The tree-ring chemistry of declining sugar maple in central Ontario, Canada. *Ambio* 28(7): 613–618.

Wigley, T.M.L., R.L. Smith, and B.D. Santer. 1998. Anthropogenic influence on the autocorrelation structure of hemispheric-mean temperatures. *Science* 282 (Nov. 27): 1676–1679.

5 Assessment of Changes in Pollutant Concentrations

J. Mohapl

CONTENTS

1-56670-641-6/04/$0.00+$1.50

5.1 INTRODUCTION

International agreements, such as the Clean Air Act Amendments of 1990 and the Kyoto Protocol, mandate introduction and enforcement of policies leading to systematic emission reductions over a specific period of time. To maintain the acquaintance of politicians and general public with the efficiency of these policies, governments of Canada and the U.S. operate networks of monitoring stations providing scientific data for assessment of concominant changes exhibited by concentrations of specific chemicals such as sulfate and nitrate. The highly random nature of data supplied by the networks complicates diagnosis of systematic changes in concentrations of a particular substance, as well as important policy related decisions such as choice of the reduction magnitude to be achieved and the time frame in which it should be realized.

This chapter offers quantitative methods for answering some key questions arising in numerous policies. For example, how long must the monitoring last in order that a reduction can be detected given the precision of current measurements? Based on the recent annual rate of change, how many more years will it take to see the desired significant impact? Do the data, collected over a specific period of time, suggest an emission reduction at all? How do we extrapolate results from isolated spots to a whole region? How do we compare changes measured by different networks with specific sampling protocols and sampling frequencies? An accurate answer to these and other questions can avoid wasting of valuable resources and prevent formulation of goals, the achievement of which cannot be reasonably and reliably verified and therefore enforced in a timely manner.

The statistical method for assessment of changes in long-term air quality data described in the next section was designed and tested on samples by three major North American monitoring networks: CASTNet, run by the U.S. Environmental Protection Agency, CAPMoN, operated by the Canadian Federal Government, and APIOS-D, established by the Ontario Ministry of Environment and Energy. Despite that, the method is general enough to have a considerable range of application to a number of regularly sampled environmental measurements. It relies on an indicator of long-term change estimated from the observed concentrations and on statistical tests for decision about the significance of the estimated indicator value. The indicator is interpreted as the average long-term percentage change. Its structure eliminates short-term periodic changes in the data and is invariant towards systematic biases caused by differences in measurement techniques used by different networks. The latter feature allows us to carry out a unified quantitative assessment of change over all of North America. Since inference about the indicator values and procedures utilizing the indicator for answering policy-related questions outlined above require a reliable probabilistic description of the data, a lot of attention is devoted to CASTNet, CAPMoN, and APIOS-D case studies.

A basic knowledge of statistics will simplify understanding of the presented methods; nevertheless, conclusions of data analysis should be accessible to the broadest research community. The thorough, though not exhaustive, analysis of changes exhibited by the network data demonstrates the versatility of the percentage decline indicator, the possibilities offered by the indicator for inference and use in policy making, and a new interesting view of the long-term change in air quality over North America from 1980 to 1998.

5.1.1 Frequently Asked Questions about Statistical Assessment

Among practitioners, reputation of statistics as a scientific tool varies with the level of understanding of particular methods and the quality of experience with specific procedures. It is thus desirable to address explicitly some concerns related to air quality change assessment often occurring in the context of statistics. The following section contains the most frequent questions practitioners have about inference and tests used throughout this study.

Question 1: Why should statistics be involved? Cannot the reduction of pollutant concentrations caused by the policies be verified just visually? Why cannot we rely only on common sense?

Answer: A reduction clearly visible, say, from a simple plot of sulfur dioxide concentrations against time, would be a nearly ideal situation. Unfortunately, the variability of daily or weekly measurements is usually too high for such an assessment and the plots lack the intuitively desired pattern. Emission reductions require time to become noticeable, but if the policies have little or no effect, they should be modified as soon as possible. Hence, the failure of statistics to detect any change over a sufficiently long time, presumably shorter than the time an obvious change is expected to happen, can be a good reason for reviewing the current strategy. Conversely, an early detection of change may give us space to choose between more than one strategy and select and enforce the most efficient one.

Question 2: Inference about the long-term change is based on the probability distribution of the observed data. The distribution is selected using the goodness-of-fit test. However, such a test allows one only to show that the fit of some distribution to the data is not good, but lack of statistical significance does not show the fit is good. Can the goodness-of-fit information be thus useful?

Answer: In this life, nothing is certain except death and taxes (Benjamin Franklin), and scientific inference is no exception. Statistical analysis resembles largely a criminal investigation, in which the goodness-of-fit test allows us to eliminate probability distributions suspected as useless for further inference about the data. Distributions that are not rejected by the test are equally well admissible and can lead to different conclusions. This happens rarely though. Usually, investigators struggle to find at least one acceptable distribution describing the data. Although the risk of picking a wrong probability distribution resulting in wrong conclusions is always present, practice shows that it is worth assuming.

Question 3: Some people argue that inference about concentrations of chemical substances should rely mainly on the arithmetic mean because of the law of conservation of mass. Why should one work with logarithms of a set of measurements and other less obvious statistics?

Answer: A simple universal yes–no formula for long-term change assessment based on an indicator such as the arithmetic mean of observed concentrations is a dream of all policy makers and officials dealing with environment-related public affairs. In statistics, the significance of an indicator is often determined by the ratio of the indicator value and its standard deviation. To estimate the variability of the indicator correctly is thus the toughest part of the assessment problem and consumes the most space in this chapter.

Question 4: Series of chemical concentrations observed over time often carry a substantial autocorrelation that complicates estimation of variances of data sets. Is it thus possible to make correct decisions without determining the variance properly?

Answer: Probability distributions describing observed chemical concentrations must take autocorrelation into account. Neglect of autocorrelation leads to wrong conclusions. Observations that exhibit a strong autocorrelation often contain a trend that is not acknowledged by the model. Numerous methods for autocorrelation detection and evaluation are offered by the time-series theory and here they are utilized as well because it is impossible to conduct statistical inference without correctly evaluating the variability of the data.

5.1.2 Trend Analysis vs. Change Assessment

The high variability of air chemistry data supplied by networks such as CASTNet, CAPMoN, and APIOS-D and the complex real-world conditions generating them lead researchers to focus on what is today called trend detection and analysis. The application of this method to filter pack data from CASTNet can be found in Holland et al. (1999). A more recent summary of various trend-related methods frequently used for air and precipitation quality data analysis is found in Hess et al. (2001). The advantage of trend analysis is that it applies well to both dry and wet deposition data (Lynch et al. 1995; Mohapl 2001; Mohapl 2003b). Some drawbacks of trend analysis in the context of the U.S. network collected data are discussed by Civeroloa et al. (2001). Let us recall that the basic terminology and methods concerning air chemistry monitoring in network settings are described in Stensland (1998).

A trend with a significant, linearly decreasing component is commonly presented as a proof of decline of pollutant concentrations. Evidence of a systematic decline, however, is only a part of the assessment problem. The other part is quantification of the decline. One approach consists of estimation of the total depositions of a chemical over a longer time period, say per annum, and in the use of the estimated totals for calculation of the annual percentage decline (Husain et al. 1998; Dutkiewicz et al. 2000). A more advanced approach, applied to CASTNet data, uses modeling and fluxes (Clarke et al. 1997). There is no apparent relation between the analysis of trends, e.g., in sulfate or nitrate weekly measurements, and the flux-based method for the total deposition calculation. Trend analysis reports rarely specify the relation between the trend and the disclosed percentage declines. What do the significance of trend and confidence intervals for the percentage change, if provided, have in common is also not clear.

Besides the presence of change, there are other questions puzzling policy makers and not easily answered by trend analysis as persuasively and clearly as they deserve. If the change is not significant yet, how long do we have to monitor until it will prove as such? Is the time horizon for detection of a significant emission reduction feasible? Is the detected significant change a feature of the data or is it a consequence of the estimator used for the calculation?

Though analysis of time trends in the air chemistry data appears inevitable to get proper answers, this chapter argues that the nature of CAPMoN, CASTNet, and APIOS-D data permits drawing of conclusions using common elementary statistical formulas and methods. Since each site is exposed to particular atmospheric conditions, analysis of some samples may require more sophisticated procedures.

5.1.3 Organization of This Chapter

Section 5.2 introduces the annual percentage change and decline indicators. In the literature, formulas for calculation of percentage declines observed in data are rarely given explicitly. A positive example, describing calculation of the total percentage from a trend estimate, is Holland et al. (1999). The main idea here is that an indicator should be a well-defined theoretical quantity, independent of any particular data set and estimation procedure and admitting a reasonable interpretation. Various estimators of the quantity, differing in bias, variability, speed of convergence, etc., can be then designed and studied according to the features of the available data.

Introduction of the long-term percentage change and decline indicators, which are central to this study, does not require a specific probability distribution. Interpretation of the indicators in the context of stationary processes that are commonly used in large network data analysis is given in Section 5.2.1. Applicability of the indicators to the CASTNet data set is discussed in Section 5.3 and throughout the rest of this study.

Section 5.4 presents the elementary statistics for estimation and temporal inference about the change and decline indicators, including confidence regions. It shows how the estimators work on the CASTNet data set. The results are interesting in comparison to those in Holland et al. (1999), Husain et al. (1998), and Dutkiewicz et al. (2000). Section 5.4.3 utilizes the decline indicator to gain insight into the regional changes of the CASTNet data.

Section 5.5 develops methods for statistical inference about the percentage change in the simplest but fairly common case, occurring mainly in the context of small data sets when the data entering the indicators appear mutually independent and identically distributed. A set of policy-related problems concerning long-term change assessment is also solved.

Section 5.6 extends the results to data generated by stationary processes and applies them to the CASTNet observations. Problems concerning policies are reformulated for data generated by stationary processes and solutions to the problems are extended accordingly. Further generalization of the change indicator is discussed in Section 5.7.

Spatial distribution of air pollutants is frequently discussed in the context of concentration mapping (Ohlert 1993; Vyas and Christakos 1997), but rarely for the purpose of change assessment. Section 5.8 generalizes definition of the indicators from one to several stations. The spatial model for construction of significance tests and confidence intervals for the change indicators is built using a multivariate autoregressive process.

The CAPMoN data carry certain features that require further extension of the percentage change estimators in Section 5.9. Besides analysis of changes in time and space analogous to the CASTNet study, they offer the opportunity to use the change indicators for comparison of long-term changes estimated from the two networks (see Section 5.9.3). Both CAPMoN and CASTNet maintain common sampling sites at Egbert and Pennsylvania State University serving network calibration. Comparison of the annual rates of decline, quantities that essentially determine the long-term change indicators, is used to infer about similarities and differences in changes measured by the two networks.

Another example of how to apply change indicators to comparison of pollutant reductions reported by different networks is presented in Section 5.10. Data from three stations that hosted CAPMoN and APIOS-D devices during joint operation of the networks demonstrate that the indicator is indeed invariant towards biases caused by differences in measurement methods.

Most case studies in this chapter focus on dry deposition data in which pairs are natural with regard to the sampling procedure. Section 5.11 demonstrates its power on CASTNet precipitation samples, where the paired approach is not particularly optimal due to the irregular precipitation occurrence reducing the number of pairs. Still, the application shows the considerable potential of the method and motivates the need for its further generalization.

Since decisions about the trend parameter of a stationary $AR(p)$ process are essential for inference about the indicators, and the results of inference have a straight impact on quality and success of policies that will implement them, the plain average estimator vs. the least squares and maximum likelihood estimators are discussed in Section 5.12.1. The presented theory shows that the so-called average percentage decline estimator remains optimal even for correlated data, though the inference must accommodate the autocorrelation accordingly.

5.2 THE ASSESSMENT PROBLEM

This section presents the annual percentage decline indicator as a quantity describing the change exhibited by concentrations of a specific pollutant measured in the air over a 2-year observation period. It is derived for daily measurements, though weekly or monthly data would be equally useful. The only assumption the definition of the indicator needs is positiveness of the observed amounts. Practice requires assessment of change over longer periods than just 2 years. Introduction of the long-term percentage decline, central to our inference about the air quality changes, thus follows.

5.2.1 THE SPOT AND ANNUAL PERCENTAGE CHANGES

Let us consider concentrations of a chemical species in milligrams per liter (mg/l) sampled daily from a fixed location over two subsequent nonleap years, none of them missing and all positive. It is to decide if concentrations in the first year are in some sense systematically higher or lower than in the second year.

For the purpose of statistical analysis, each observed concentration is represented by a random variable c. Due to the positiveness of concentrations, the random variable c is also positive and admits the description

$$c = \exp\{m + \eta\}, \tag{5.1}$$

where m is a real number and η is a random variable with zero mean. In applications, m is not known, hence the value of η is not observable.

Let c describe a concentration in year one and let c' be the concentration observed the day exactly one year later. Then c' admits the representation

$$c' = \exp\{m' + \eta'\}, \tag{5.2}$$

and our task is to compare c to c'. This can be done either by assessing how far the difference $c - c'$ lies from zero or how much the ratio c/c' differs from one. While dealing with $c - c'$ appears more natural, the fraction c/c' turns out as much more operational. That is because c/c' is again a positive random variable with the representation

$$\frac{c}{c'} = \exp\{\mu + \zeta\}, \tag{5.3}$$

where $\mu = m - m'$ and $\zeta = \eta - \eta'$. Since the use of c/c' is not quite common, we focus on the quantity

$$pc_{\zeta} = 100\frac{c' - c}{c} = 100\left(\frac{c'}{c} - 1\right) = 100(\exp\{-\mu - \zeta\} - 1). \tag{5.4}$$

The quantity pc_{ζ} can be called the annual *spot percentage change*. If a systematic change occurred over the 2 years, then at least intuitively it is captured by the deterministic value of μ.

The sampling methods used by CASTNet and CAPMoN networks produce results that are systematically biased towards each other (Mohapl 2000b; Sickles and Shadwick 2002a). Other networks suffer systematic biases as well (Ohlert 1993). It is thus important to emphasize that the quantity pc_{ζ} is not affected by the bias. The bias means that, in theory, if the precision of CASTNet and CAPMoN were exactly the same up to the bias, then the CASTNet measurements would be $c_1' = \alpha c_2'$ and $c_1 = \alpha c_2$, respectively, where c_2' and c_2 are CAPMoN observations taken at the same time and location. The relations

$$\frac{c_1' - c_1}{c_1} = \frac{\alpha c_2' - \alpha c_2}{\alpha c_2} = \frac{c_2' - c_2}{c_2},$$

show that the percentage change is not affected by the bias.

Similarly, if two networks issue measurements in different units, then the spot percentage declines computed from those results are comparable due to the same argument. The annual percentage decline is thus unit invariant.

A random variable is not a particularly good indicator of a change. That is why we introduce the *annual percentage change* using the quantity

$$pc = 100(\exp\{-\mu\} - 1), \tag{5.5}$$

which arises from pc_{ζ} by suppression of the noise. Policy makers think usually in terms of an annual percentage decline to be achieved by their policies, and this decline is a positive number. Hence, we introduce the *annual percentage decline* indicator $pd = -pc$, or in more detail

$$pd = 100(1 - \exp\{-\mu\}). \tag{5.6}$$

It is rather clear that pd grows as μ increases. The parameter μ is called the annual rate of change or annual rate of decline.

At the moment, the annual percentage decline pd is a sensible indicator of the annual change only if μ is common for all spot changes obtained from the two compared years. Though this is a serious restriction expressing a belief that the decline proceeds in some sense uniformly and linearly, justification of this assumption for a broad class of concentration measurements will be given shortly.

5.2.2 The Long-Term Percentage Change

To explain the difference between the annual and long-term change, let us denote c_t and c'_t positive concentration amounts of a chemical sampled with the same frequency either daily, weekly or monthly over two equally long periods measured in years. Due to (5.1), (5.2), and (5.3),

$$c_t = \exp\{m_t + \eta_t\} \tag{5.7}$$

and

$$c'_t = \exp\{m'_t + \eta'_t\}, \tag{5.8}$$

respectively, where m_t and m'_t represent a trend, and η_t and η'_t capture irregularities in concentration amounts due to the randomness of weather conditions and inaccuracies of the measuring procedure. Recall that the only assumption for representations (5.7) and (5.8) is positiveness of the observed values. Depending on the situation, the time index t can denote the order number of the observation in the sample, e.g., t-th week, but it can also denote a time in a season measured in decimals. For example, under weekly sampling, $t = n/52$ is the n-th week of the year. Hopefully, the reader will not confuse t with the familiar t-test statistics.

Air quality monitoring networks are running over long time periods. Suppose we have two sets of data, each collected regularly over P years, with W observations in each year. For the moment, let c_t and c'_t be observations from the first and second periods, respectively. Then the spot percentage change (5.4), defined by pairs of observations from now and exactly P years later, has the form

$$pc_t = 100\frac{c'_t - c_t}{c_t} = 100\left(\frac{c'_t}{c_t} - 1\right) = 100(\exp\{-\mu - \zeta_t\} - 1). \tag{5.9}$$

From (5.9) we can arrive at the same indicators pc and pd as in (5.5) and (5.6), respectively. However, μ in (5.5) and (5.6) cannot be interpreted as an annual rate of decline anymore.

To illustrate why, let m_t and m'_t in (5.7) and (5.8), respectively, have at selected points $t_n = n/W$, $n = 1,\ldots, WP$, the form

$$m_{t_n} = s + rt_n + \pi_{t_n}$$

and

$$m'_{t_n} = s + r(t_n + P) + \pi_{t_n},$$

where π_t is an annual periodic component. Such a representation is quite frequent in air pollution modeling (Lynch et al. 1995; Holland et al. 1999), and determines μ as

$$\mu = m_{t_n} - m'_{t_n} = -rP, \qquad n = 1,\ldots, WP,$$

which means the more years the compared periods contain, the larger the absolute value of μ. Consequently, the *annual rate of change* or *annual rate of decline* ρ satisfies in this more general setting the equation

$$\rho = \frac{\mu}{P}, \tag{5.10}$$

where P is the number of years in each of the compared periods. The parameter μ will be simply called the long-term *rate of change* or *rate of decline*. If $P = 1$, then μ, the long-term rate of change, agrees with the annual rate of change. We can thus introduce the long-term *percentage change pc* and *percentage decline pd* indicators using relations (5.5) and (5.6), respectively, with μ defined as $\mu = \rho P$. We recall that P is always one half of the total observation period covering data available for analysis and the trend m_{t_n} declines only if $r < 0$.

A large part of the analysis in this chapter has to do with verification of the assumption that the parameter μ determining the change indicators *pc* and *pd* is the same for all pairs in the sample. If the rate μ is the same for all pairs, then due to (5.3),

$$\ln c_t - \ln c_t' = \mu + \zeta_t, \tag{5.11}$$

where μ is a constant parameter representing the magnitude of the systematic change in pollutant concentration over time and ζ_t is a series of zero-mean random variables. Assumption (5.11) expresses our belief that by subtracting observations with the same position in the compared periods we effectively subtract out all periodicities, and if a linear change in the concentrations prevails, the parameter μ will be significant. For justification see Figure 5.1 of the CASTNet data sampled at Woodstock.

Model (5.11) turns into a powerful tool for change assessment if the data do not contradict the hypothesis that $\zeta = \{\zeta_t, -\infty < t < \infty\}$, the noise-generating mechanism for our measurements, is a *stationary process*. Stationarity means the covariance between any two ζ_t and ζ_{t+h} depends on the lag h only. More formally,

$$cov(\zeta_t, \zeta_{t+h}) = R(h) \tag{5.12}$$

for some finite function $R(h)$, called the *covariance function* of the process ζ. If in addition to the stationarity condition (5.12)

$$\sum_{h=-\infty}^{\infty} |R(h)| < \infty, \tag{5.13}$$

then $R(h)$ has a spectral density function, and the law of large numbers and the central limit theorem are true (Brockwell and Davis 1987, Chapter 7). These large

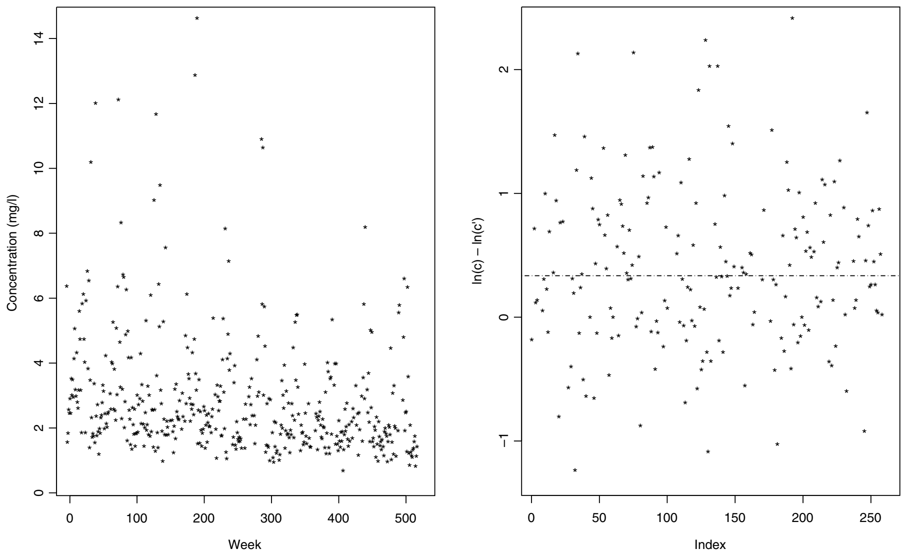


FIGURE 5.1 Observations of teflon filter SO_4 (mg/1), collected during 1989–1998 at the CASTNet station Woodstock, Vermont, USA, demonstrate appropriateness of assumption (5.11). The dashed line on the right is the estimate of μ.

sample properties result in accurate statistics for decision about significance of the observed percentage change, or more precisely, about significance of the observed rate of change (decline). For more details on the statistics of stationary time series see also Kendall and Stuart (1977) (Volume 3, Section 47.20).

It is emphasized that assumption (5.11) does *not* mean we impose any restrictions on the distribution of c and c', or equivalently, of η and η', respectively. In other words, if we are interested only in the percentage change, it suffices to concentrate on the distribution of the process ζ in (5.11).

Given the previous results, the problem of long-term change assessment essentially consists in determining μ and in deciding if it differs significantly from zero. Though this is a relatively narrow formulation from the practice point of view, its solution yields results applicable to a broad class of air quality data and provides enough space for better understanding more complicated problems.

5.3 CASE STUDY: ASSESSMENT OF DRY CHEMISTRY CHANGES AT CASTNET SITES 1989–1998

The Clean Air Status and Trends Monitoring Network (CASTNet) is operated by the U.S. Environmental Protection Agency (EPA). Dry chemistry sampling consists in sucking of a prescribed volume of air through a pack of filters collecting particles and gases at designated rural areas. The CASTNet filter contents are analyzed weekly in a central laboratory for amounts of sulfate and nitrate extracted from the teflon, nylon, and celulose filters. The extracted chemicals are called teflon filter

TABLE 5.1
Summary of Monitored Species and Their Interpretation. WNO_3 Is Usually Not Interpreted

Raw Chemical	TSO_4	TNO_3	TNH_4	$NHNO_3$	WNO_3	NSO_4	WSO_2
Interpretation	SO_4^-	NO_3^-	NH_4^+	HNO_3		$SO_2 = NSO_4 + WSO_2$	

TABLE 5.2
Summary of Monitoring Periods

Years of Monitoring	2	4	6	8	10
Number of Pairs	52	104	156	208	260
Number of Stations	5	17	2	2	40

sulfate, nitrate, and ammonium (TSO_4, TNO_3, and TNH_4, respectively), nylon filter sulfate and nitrate (NSO_4 and $NHNO_3$, respectively), and cellulose filter sulfur dioxide and nitrate (WSO_2 and WNO_3, respectively). Their interpretation is in Table 5.1. A more detailed overview of CASTNet operation and setting is given in Clarke et al. (1997) and in this handbook.

Sometimes $SO_2 = NSO_4 + WSO_2$ is called the total sulfur dioxide. Studies of the CASTNet data aiming to assess long-term changes, such as Holland et al. (1999), use total sulfur dioxide and nitrogene values calculated according to the formula

$$N = \frac{14}{62}\mathrm{TNO_3} + \frac{14}{63}\mathrm{NHNO_3}.$$

Access to the CASTNet data is provided to the public at http://www.epa.gov/castnet/data.html. A number of interesting details concerning the sampling procedures is available (Sickles and Shadwick 2002a). The CASTNet data analyzed here were collected at 66 stations. Data from all stations end in 1998, and 40 of them start 10 years earlier in 1989. The frequency Table 5.2 summarizes how many stations have been sampling 2, 4, etc., years prior to 1998. It also shows the possible maximal sample size of pairs available from the period.

Locations of the stations are indicated in Figure 5.2 and Figure 5.3. The stations are in two groups representing the western and eastern U.S. The set of pairs for spot change calculation contains a substantial amount of missing data. The main reason an observation is missing is the start or end of monitoring in the middle of the year, which can eliminate up to 50% and more of paired observations from stations with short history prior to 1998. A labor dispute interrupted sampling at about half of the stations from October 1995 to February 1996 or later and substantially contributed to the missing pair set. Causes for data missing through natural problems with air pumps, filter pack, etc., are listed on the CASTNet

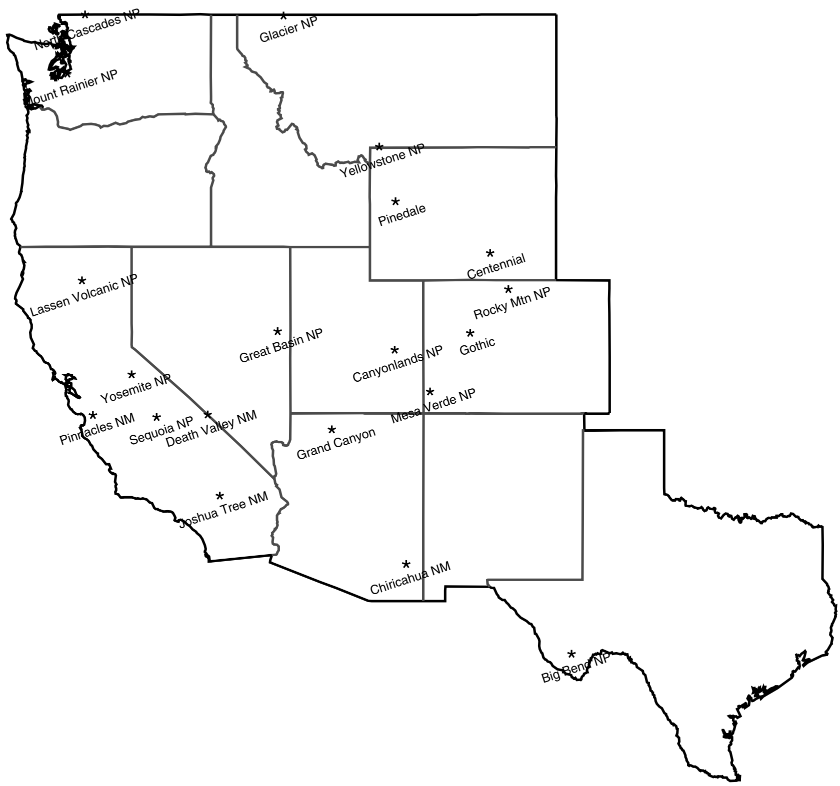


FIGURE 5.2 CASTNet stations in the western United States.

homepage. The actual percentages of pairs of data used for analysis are given in Table 5.3 through Table 5.5. The sometimes low percentage numbers show that the exclusive use of pairs can lead to a considerable loss of information. Due to the presence of seasonal trends, restriction to the pairs is important for comparison.

A theory for estimation and inference about the long-term percentage change indicator has to be developed before the CASTNet analysis can be approached. Because of the large extent of CASTNet dry deposition data, the basic theory is laid out next in Section 5.4 and the case study continues later as the theory evolves.

5.4 SOLUTION TO THE CHANGE ASSESSMENT PROBLEM

The formal solution to the change assessment problem is simple. If the data do not contradict the assumption that the rate of change μ is constant and the process ζ in (5.11) is stationary, then μ can be estimated by a simple sample average, variance of μ is determined by the spectral density of the process, and inference about μ, including confidence regions, can be carried out in a standard manner. Substitution

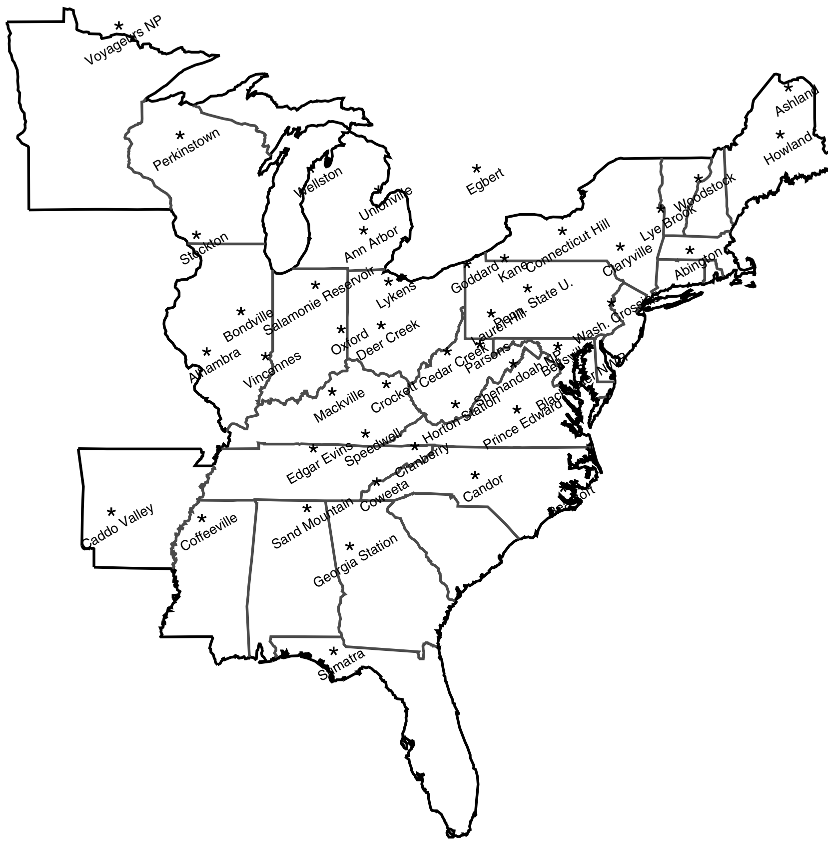


FIGURE 5.3 CASTNet stations in the eastern United States.

of the sample average for μ leads to an interesting interpretation of *pc* and *pd*, respectively, in terms of the spot percentage change. Since substitution of the average for the true rate turns *pc* and *pd* indicators into random variables, properties of these random variables must be determined to evaluate their bias and standard deviation.

5.4.1 Estimation of μ and Inference

Let us consider positive concentration amounts from two subsequent periods measured in years obeying the model (5.11) with noise ζ, that is, a stationary process with covariance function satisfying (5.13). The amounts admit representation (5.7) and (5.8), respectively. The most popular statistics for estimation of μ in (5.11) from a sample $\ln(c_t/c_t')$, $t = 1, \ldots, N$, is the arithmetic mean

$$\hat{\mu} = \frac{1}{N} \sum_{t=1}^{N} \ln(c_t/c_t'). \tag{5.14}$$

TABLE 5.3
Western US CASTNet Stations. The Percentage of Pairs Used for Analyses out of the Total Available Theoretically Given the Duration of Monitoring

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Big Bend NP	50	50	50	46	50	50	50	4
Canyonlands NP	70	70	70	53	70	70	70	4
Centennial	77	66	77	76	77	76	77	10
Chiricahua NM	82	82	82	81	82	82	82	10
Death Valley NM	77	77	76	26	77	76	75	4
Glacier NP	92	92	92	91	91	87	91	10
Gothic	84	77	83	68	83	75	82	10
Grand Canyon	81	80	81	67	81	80	81	10
Great Basin NP	70	70	70	32	70	68	69	4
Joshua Tree NM	68	68	68	46	67	67	67	4
Lassen Volcanic NP	50	50	50	29	50	47	50	4
Mesa Verde NP	64	63	64	56	64	64	64	4
Mount Rainier NP	89	81	75	53	60	60	64	2
North Cascades NP	70	66	70	38	55	49	68	2
Pinedale	79	73	77	69	78	76	77	10
Pinnacles NM	61	62	61	43	61	60	62	4
Rocky Mtn NP	77	76	77	67	78	76	75	4
Sequoia NP	43	43	43	32	43	42	43	2
Yellowstone NP	89	83	91	58	91	91	91	2
Yosemite NP	49	48	47	36	49	41	49	4

According to Brockwell and Davis (1987) (Section 7.1), the estimator $\hat{\mu}$ is unbiased and the stationarity assumption combined with (5.13) implies it is also consistent. In addition, it can be shown that for large samples, $\hat{\mu}$ is approximately Normal in the sense that

$$\frac{\hat{\mu}-\mu}{\sqrt{var(\hat{\mu})}} \sim \mathcal{N}(0,1). \tag{5.15}$$

In our notation, $\mathcal{N}(m,V)$ means the Normal distribution with mean m and variance V. The tilde denotes membership in a family of distributions. Without a consistent estimator of $var(\hat{\mu})$, the relation (5.15) is of little use. If we denote such an estimator $v\hat{a}r(\hat{\mu})$, then

$$Z_{\mu} = \frac{\hat{\mu}-\mu}{\sqrt{v\hat{a}r(\hat{\mu})}} \tag{5.16}$$

is also asymptotically Normal. The last statistics serves for construction of the test for decision if μ equals to a particular value specified, for example, by a policy, and it also provides an approximate confidence region for μ.

TABLE 5.4.
Eastern US CASTNet Stations. Part I. The Percentage of Pairs Used for Analysis out of the Total Available Theoretically Given the Duration of Monitoring

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Abington	67	67	67	67	66	66	66	4
Alhambra	87	87	87	87	86	86	86	10
Ann Arbor	72	72	72	72	71	71	71	10
Ashland	81	77	81	81	80	69	78	10
Beaufort	72	72	72	72	72	72	72	4
Beltsville	74	74	74	74	73	73	72	10
Blackwater NWR	29	29	28	28	28	27	27	4
Bondville	80	80	79	79	79	79	79	10
Caddo Valley	80	74	79	80	79	79	79	10
Candor	83	83	83	83	83	83	83	8
Cedar Creek	83	66	83	82	82	83	83	10
Claryville	88	81	88	89	89	88	88	4
Coffeeville	72	71	71	71	71	71	70	10
Connecticut Hill	85	84	85	85	84	83	82	10
Coweeta	94	65	94	93	93	87	92	10
Cranberry	81	69	81	81	81	80	81	10
Crockett	68	66	68	68	68	67	66	6
Deer Creek	80	80	80	80	79	79	77	10
Edgar Evins	82	73	82	82	82	82	81	10
Egbert	90	90	89	89	89	89	89	4
Georgia Station	79	78	79	79	79	79	79	10
Goddard	82	82	81	81	80	80	79	10
Horton Station	82	81	81	81	81	81	81	10

To test the hypothesis $\mu = \mu_0$ against $\mu \neq \mu_0$, we simply calculate Z_{μ_0} and reject the null hypothesis on the critical level α if $|Z_{\mu_0}| > u(\alpha)$. The symbol $u(\alpha)$ denotes the α quantile of the Normal distribution, a number exceeded by the absolute value of a standard Normal random variable with probability α, shortly $Prob(|U| > u(\alpha)) = \alpha$.

The 100 $(1 - \alpha)$% confidence region for μ has the form

$$\hat{\mu} - \sqrt{v\hat{a}r(\hat{\mu})}u(\alpha) < \mu < \hat{\mu} + \sqrt{v\hat{a}r(\hat{\mu})}u(\alpha). \tag{5.17}$$

Consistency means the tendency of an estimator to approach the true value of the parameter it estimates with growing sample size. In the case of μ, consistency agrees with the *law of large numbers* (LLN). The convergence of the statistics Z_μ to the standard Normal distribution is known as the *central limit theorem* (CLT). Both LLN and CLT are crucial in the probability theory and statistics of large samples (Feller 1970; Loéve 1977). Nonstationary processes obeying LLN and

TABLE 5.5
Eastern US EASTNet Stations. Part II. The Percentage of Pairs Used for Analysis out of the Total Available Theoretically Given the Duration of Monitoring

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Howland	61	50	61	61	61	59	61	6
Kane	79	69	79	79	78	78	78	10
Laurel Hill	82	69	82	82	82	82	81	10
Lye Brook	31	28	31	31	31	30	31	4
Lykens	68	68	68	68	68	68	68	10
Mackville	74	74	74	74	73	72	71	8
Oxford	95	95	95	95	95	95	95	10
Parsons	94	94	94	94	94	94	93	10
Penn. State U.	94	94	93	93	93	93	92	10
Perkinstown	95	95	95	95	95	92	94	10
Prince Edward	81	74	81	81	81	80	80	10
Salamonie Reservoir	80	80	79	79	78	78	77	10
Sand Mountain	82	82	81	81	81	81	81	10
Shenandoah NP	85	81	85	85	85	85	85	10
Speedwell	72	71	71	72	71	71	70	10
Stockton	51	51	50	50	50	50	50	4
Sumatra	87	84	87	87	87	84	85	10
Unionville	82	82	82	82	82	82	82	10
Vincennes	95	95	95	95	95	95	95	10
Voyageurs NP	94	94	92	83	92	92	92	2
Wash. Crossing	80	80	80	80	78	78	77	10
Wellston	78	77	78	78	78	77	77	10
Woodstock	76	60	76	74	75	70	74	10

CLT are often more difficult to work with, their features, such as distribution and correlation of variables, are more complicated to verify, and estimates of variance and perhaps other parameters are not easy to obtain. Hence, though the theoretical setting of the following considerations could be more general, stationary processes are the best choice for the intended application. It is known (Brockwell and Davis 1987, Section 7; or Grenander and Rosenblatt 1984, Section 3.7) that stationarity, (5.13) and finite fourth order moments of a process assure that LLN holds for the sample mean and variance and for the maximum likelihood estimators. The CLT is also true.

In the case of stationary processes, $var(\hat{\mu})$ can be described in terms of the covariance function $R(h)$, introduced in (5.12) and the *spectral density* function $f(\lambda)$, defined by the integral

$$f(\lambda) = \frac{1}{2\pi}\int_{-\pi}^{\pi} R(h)\exp\{-i\lambda h\}dh, \tag{5.18}$$

where $i = \sqrt{-1}$. It should be noted that convergence of the integral is a consequence of (5.13). For large samples,

$$var(\hat{\mu}) \approx \frac{1}{N} 2\pi f(0), \tag{5.19}$$

where N denotes the sample size (Brockwell and Davis 1987, Section 7.1, Remark 1).

It is useful to denote $v\hat{a}r(y)$ as the sample variance

$$v\hat{a}r(y) = \frac{1}{N-1} \sum_{t=1}^{N} (\ln(c_t/c_t') - \hat{\mu})^2 \tag{5.20}$$

and assume that for large samples, $v\hat{a}r(y)$ tends to the true variance of the process

$$y = \{y_t : y_t = \ln c_t - \ln c_t', \quad -\infty < t < \infty\}. \tag{5.21}$$

Then for large N, $v\hat{a}r(y) \approx R(0)$, and

$$\frac{\hat{\mu} - \mu}{\sqrt{var(\hat{\mu})}} \approx \frac{\hat{\mu} - \mu}{\sqrt{v\hat{a}r(y)}} \sqrt{N} \sqrt{\frac{R(0)}{2\pi f(0)}}. \tag{5.22}$$

Consequently, for large samples, Z_μ is approximately the familiar t-test statistics multiplied by the coefficient

$$v = \sqrt{\frac{R(0)}{2\pi f(0)}}. \tag{5.23}$$

The statistics

$$Z_\mu' = \frac{\hat{\mu} - \mu}{\sqrt{v\hat{a}r(y)}} \hat{v} \sqrt{N}, \tag{5.24}$$

where $\hat{v}$ is a suitable estimator of v, can be used for inference about μ and construction of the 100 $(1 - \alpha)$% confidence regions the same way as Z_μ:

$$\hat{\mu} - \sqrt{v\hat{a}r(y)} u(\alpha)/(\hat{v}\sqrt{N}) < \mu < \hat{\mu} + \sqrt{v\hat{a}r(y)} u(\alpha)/(\hat{v}\sqrt{N}). \tag{5.25}$$

Omission of v for inference about the data often leads to wrong conclusions!

5.4.2 The Average Percentage Decline in Air Pollution

Using the average $\hat{\mu}$, described by (5.14), we can estimate the long-term percentage change (5.5) and the percentage decline (5.6) as

$$p\hat{c} = 100(\exp\{-\hat{\mu}\} - 1) \tag{5.26}$$

and

$$p\hat{d} = 100(1 - \exp\{-\hat{\mu}\}), \tag{5.27}$$

respectively. Before analyzing the statistical features of these estimators, let us have a look at their meaning.

The estimator $p\hat{c}$, and therefore $p\hat{d}$, can also be interpreted as the average long-term percentage change and decline, respectively. To see why, let us write $p\hat{c}$ down explicitly in the form

$$p\hat{c} = 100\left(\exp\left\{\frac{1}{N}\sum_{t=1}^{N}\ln\left(\frac{c_t'}{c_t}\right)\right\} - 1\right). \tag{5.28}$$

The natural logarithm is nearly linear in the vicinity of one in the sense that $\ln(1 + x) \approx x$. Hence, if the spot percentage change pc_t, introduced in (5.9), is not too large, so that $pc_t/100$ is a small quantity, then

$$\frac{pc_t}{100} \approx \ln\left(1 + \frac{pc_t}{100}\right) = \ln\left(\frac{c_t'}{c_t}\right)$$

and consequently,

$$\frac{1}{N}\sum_{t=1}^{N}\frac{pc_t}{100} \approx \frac{1}{N}\sum_{t=1}^{N}\ln\left(\frac{c_t'}{c_t}\right).$$

If we apply a similar argument to $p\hat{c}$, we get

$$\frac{p\hat{c}}{100} \approx \ln\left(1 + \frac{p\hat{c}}{100}\right) = \frac{1}{N}\sum_{t=1}^{N}\ln\left(\frac{c_t'}{c_t}\right).$$

Comparison of the last two approximate equalities provides

$$p\hat{c} \approx \frac{1}{N}\sum_{t=1}^{N} pc_t$$

justifying (5.28) as a quantification of the average long-term percentage change. Similar arguments apply to $p\hat{d}$.

The last relation motivates introduction of the estimator

$$p\tilde{c} = \frac{1}{N}\sum_{t=1}^{N} pc_t. \tag{5.29}$$

TABLE 5.6
Western US CASTnet Stations. Estimates of the Long-Term Percentage Decline *pd*. The Asterisk Denotes a Significant Change Based on the Statistics Z_μ. Model for Z_μ Is in Table 5.13

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Big Bend NP	–5	81*	–9	39*	–3	–29*	–16	4
Canyonlands	–3	–3	–11	38*	4	1	–26*	4
Centennial	15*	0	11*	34*	–7*	2	–1	10
Chiricahua NM	–4	–23*	–5	44*	–17	–6	–10*	10
Death Valley NM	1	2	0	–5	4	–14*	–26*	4
Glacier NP	12*	17*	15*	33*	3	–5	–18	10
Gothic	9*	–7	7*	26*	–4	15*	–5	10
Grand Canyon	7	0	3	33*	–10*	9	–8	10
Great Basin NP	–14*	–4	–15*	7	–15	–73*	–19*	4
Joshua Tree NM	6	–3	0	37*	8	–1	–55*	4
Lassen Volcanic NP	6	12	13	30*	24*	–18	8	4
Mesa Verde NP	–5	12*	–4	47*	7	1	–20*	4
Mount Rainer NP	–12	–6	–53*	–15	–46*	–80*	–5	2
North Cascades NP	–8	–14	–15	–24	–86*	–69*	–19	2
Pinedale	11*	2	9*	30*	–10*	–5	2	10
Pinnacles NM	0	–2	0	46*	31*	–12	20*	4
Rocky Mtn NP	–15*	–32*	–22*	46*	–6	–64*	–38*	4
Sequoia NP	–33*	9	–9	–46*	12	5	27	2
Yellowstone NP	9	14	–3	–21	4	14	–2	2
Yosemite NP	–3	10	–3	36*	20*	–3	6	4

Example 3.5.1 and Example 4.5.1. in Section 5.5.1 show that (5.29) is a very poor estimator of *pc* and its use is strongly discouraged.

5.4.3 Long-Term Concentration Declines at CASTNet Stations

Statistics (5.27) was used for computation of the percentage decline observed for the 10-year period at the CASTNet sites; see Table 5.6 through Table 5.8. Since the magnitude of the long-term percentage change depends on the length of the observation period (see Section 5.2.2), the tables contain, besides the change computed from data of the species, the length of the observation period. The asterisk denotes a significant decline based on the statistics Z_μ introduced in (5.16) and discussed further in the following sections. Computation of Z_μ requires a probabilistic model describing the data and assisting in estimation of the variance of $\hat{\mu}$. A positive number in the table accompanied by an asterisk means a significant decline expressed in percentages over the period measured in years. A negative number with an asterisk is interpreted as an increase in concentrations of the species at the particular station.

Regardless of any test outcome, researchers often want to know how the observed decline depends on the geography of the monitored region. A plot of the observed

TABLE 5.7
Eastern US CASTNet Stations. Part I. Estimates of the Long-Term Percentage Decline *pd*. The Asterisk Denotes a Significant Change Based on the Statistics Z_μ. Model for Z_μ Is in Table 5.14

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Abington	–3	5	–8	46*	–10*	–78*	–14	4
Alhambra	12*	–3	11*	40*	–4	26*	–26*	10
Ann Arbor	16*	15*	16*	46*	5	17*	–28*	10
Ashland	28*	–9	29*	46*	27*	36*	9	10
Beaufort	–15	0	–24*	52*	–1	–65*	–10	4
Beltsville	18*	17*	20*	36*	5	14*	–61*	10
Blackwater NWR	–25*	25	–24	50*	–19	–48*	–16*	4
Bondville	15*	–5	11*	43*	–3	15*	–32*	10
Caddo Valley	5	0	5	46*	2	13	–23*	10
Candor	5	–20*	–1	39*	–3	–6	–17*	8
Cedar Creek	16*	–15	10*	42*	–1	32*	–16	10
Claryville	–2	–9	0	57*	10*	–49*	–2	4
Coffeeville	11*	7	15*	38*	–26*	12	–16*	10
Connecticut Hill	16*	–17*	10*	44*	10*	31*	–9	10
Coweeta	9*	4	2	41*	–4	14	–14*	10
Cranberry	9*	–2	4	40*	–2	–1	–13*	10
Crockett	10*	12	3	56*	5	15*	10*	6
Deer Creek	13*	1	10*	43*	–1	20*	–36*	10
Edgar Evins	10*	26*	12*	44*	–7*	21*	–24*	10
Egbert	4	4	6	61*	7	–16	–1	4
Georgia Station	9*	–17	4	39*	–1	25*	–21*	10
Goddard	13*	–2	11*	38*	5	24*	–33*	10
Horton Station	12*	3	8*	38*	–5	4	–11	10

change against the longitude and latitude is the easiest way to find out. Let us suppose that the longitude is measured in degrees east of Greenwich, which leads to negative longitude values of locations in the U.S., and the latitude is measured in degrees north of the equator. The U.S. locations thus have a positive latitude. Plots of the annual percentage decline, computed using the annual rate (5.10), revealed nothing in particular. Some species, with percentage decline calculated from full 10 years of observation, show growing decline in the northeast direction, however. If we realize that the Ohio River Valley belongs traditionally to the most polluted areas, the northeast decline in concentration would be an anticipated positive news. Figure 5.4 to Figure 5.6 show that TSO_4, TNH_4, and $NHNO_3$ declines tend to grow when plotted against the latitude and longitude, respectively.

It is tempting to infer about the significance of the growth exhibited by the percentage declines in a particular direction using standard regression methods. However, those are designed only for independent Normal random variables (Draper and Smith 1981), and $\hat{pd}$ in not Normal. The data tend also to have a heavy spatial

TABLE 5.8
Eastern US CASTNet Stations. Part II. Estimates of the Long-Term Percentage Decline *pd*. The Asterisk Denotes a Significant Change Based on the Statistics Z_μ. Model for Z_μ Is in Table 5.15

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3	Years
Howland	10*	12	3	55*	17*	–12	8	6
Kane	16*	–1	10*	40*	7*	21*	–31*	10
Laurel Hill	14*	–5	10*	43*	6*	26*	–48*	10
Lye Brook	0	–66*	–7	38*	8	–113*	–18*	4
Lykens	14*	–7	11*	45*	–2	21*	–27*	10
Mackville	10*	–15*	4	46*	–6	20*	–15*	8
Oxford	19*	–2	15*	41*	7*	26*	–32*	10
Parsons	15*	10	10*	44*	7*	38*	–28*	10
Penn. State U.	15*	–2	10*	35*	8*	13*	–42*	10
Perkinstown	13*	2	11*	43*	1	14*	0	10
Prince Edward	16*	7	11*	41*	10*	10	–27*	10
Salamonie Reservior	13*	15*	13*	40*	–13*	15*	–17*	10
Sand Mountain	7*	–8	7	39*	2	14*	–29*	10
Shenandoah NP	13*	–32*	6*	34*	–1	26*	0	10
Speedwell	12*	–4	10*	43*	–1	8	–28*	10
Stockton	3	6	9	53*	7	–19	–3	4
Sumatra	7	–9	3	43*	4	15*	–9*	10
Unionville	18*	10*	19*	46*	–2	13*	–18*	10
Vincennes	16*	–2	9*	42*	6	32*	–31*	10
Voyageurs NP	9	27*	9	14	10	7	5	2
Wash. Crossing	16*	10	16*	39*	3	14*	–39*	10
Wellston	22*	12	23*	41*	7	24*	–3	10
Woodstock	24*	–2	19*	44*	16*	35*	–4	10

correlation (see Figure 5.10). The results of a common regression statistics could thus be misleading.

Let us disregard, for the moment, the possible spatial trends and assume that the rate of decline μ is common for all locations where sampling lasted, say, the full 10 years. If the model (5.11) is true, then averaging over all 10-year rates yields an estimate of the overall rate μ. Due to problems mentioned earlier, construction of a multivariate test for the hypothesis that data of a given species from each location have μ equal to the common average versus the hypothesis that at least one of the variables has μ different from the average is difficult.

A simple, single-variable approach consists of the calculation of the 95% confidence region for the rate of change of a particular species at a particular location and the calculation of overall averages estimating μ of 4- and 10-year samples, respectively, because most of the stations have been operating over these years. The construction of the confidence region requires a reasonable probabilistic model for the particular species and location reflecting autocorrelation detected in the data. The averages substituted for $\hat{\mu}$ in the percentage decline estimator (5.14) provide

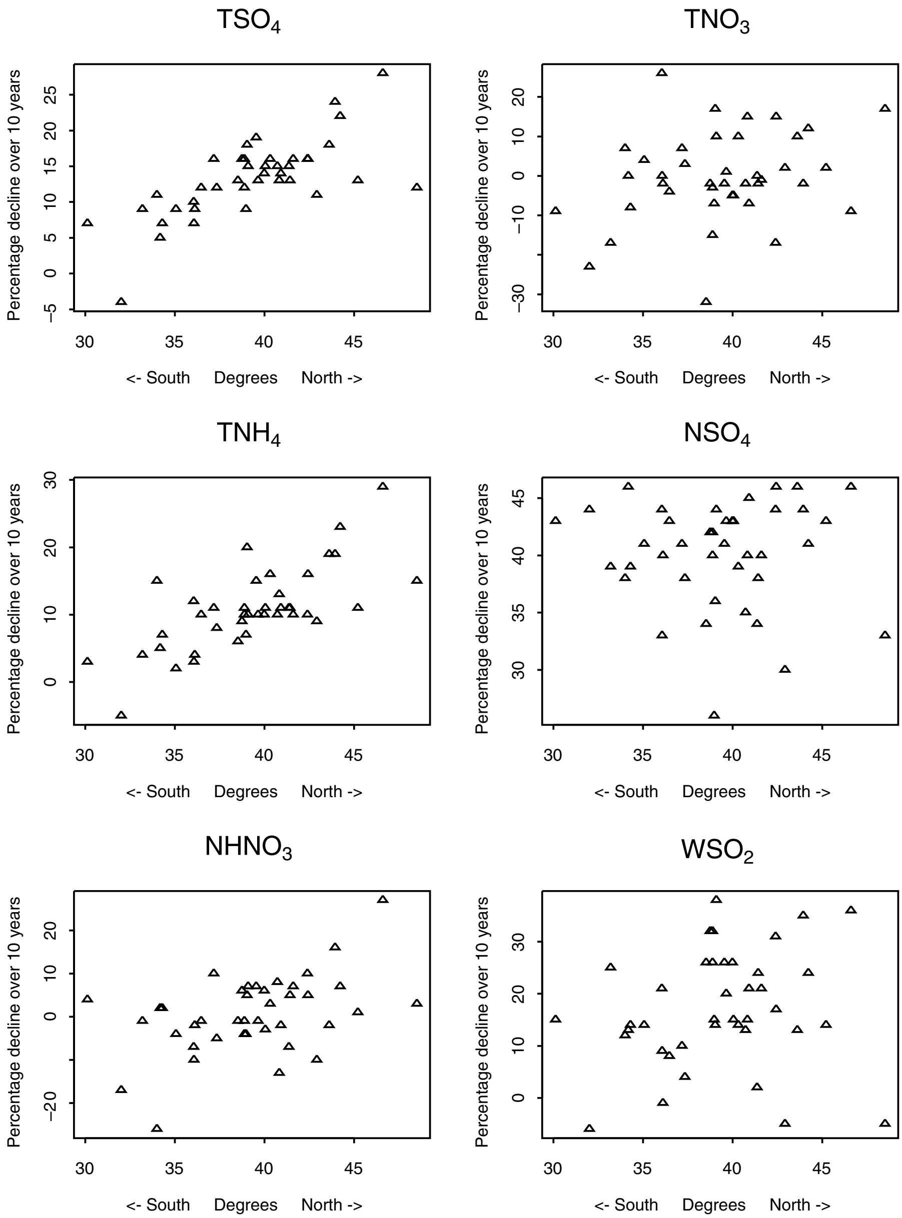


FIGURE 5.4 Percentage change over a 10-year period observed at the CASTNet stations in direction from south to north.

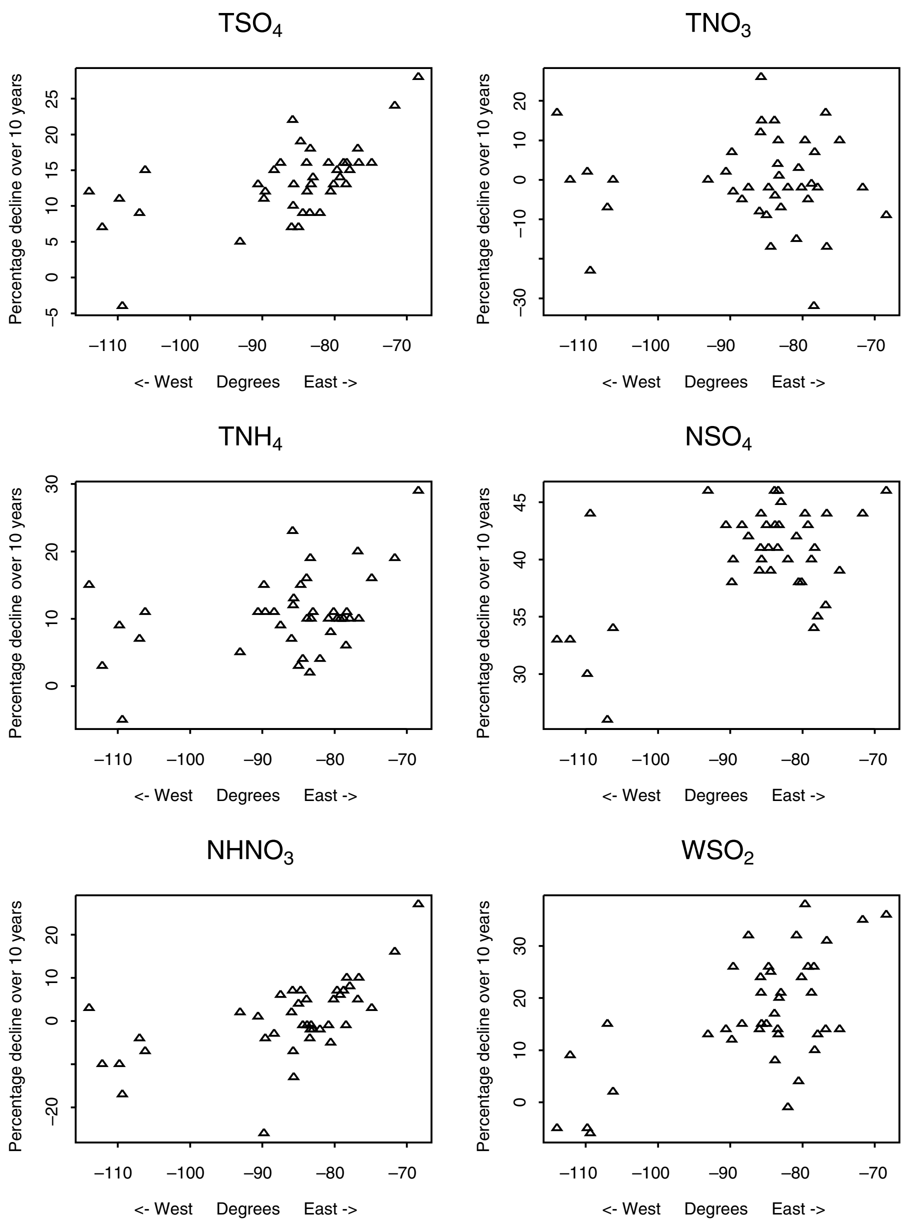


FIGURE 5.5 Percentage change over a 10-year period observed at the CASTNet stations in direction from east to west.

estimates of the overall declines. If the average is covered by the confidence region, a plus sign appears in Table 5.11 and Table 5.12, telling us which stations show a change significantly different from the overall one in Table 5.9 and Table 5.10. The failure of the interval to cover the overall average can mean an excessive increase or decline of concentration compared to the regional averages. Frequencies in

TABLE 5.9
The Average Percentage Decline Over Ten Years for Species Monitored in the Air by 40 of the CASTNet Stations Listed in Table 5.11 and the Number of Stations with Confidence Region Covering the Average. Confidence Regions for the Percentages Are in Table 5.22

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Percent	13	0	11	40	1	18	–19
No. of Stations	37	32	35	39	32	28	28

TABLE 5.10
The Average Percentage Decline Over Four Years for Species Monitored in the Air by 17 of the CASTNet Stations Listed in Table 5.11 and the Number of Stations with Confidence Region Covering the Average. Confidence Regions for the Percentages Are in Table 5.23

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Percent	–4	0	–5	46	5	–31	–13
No. of Stations	17	15	17	13	15	9	13

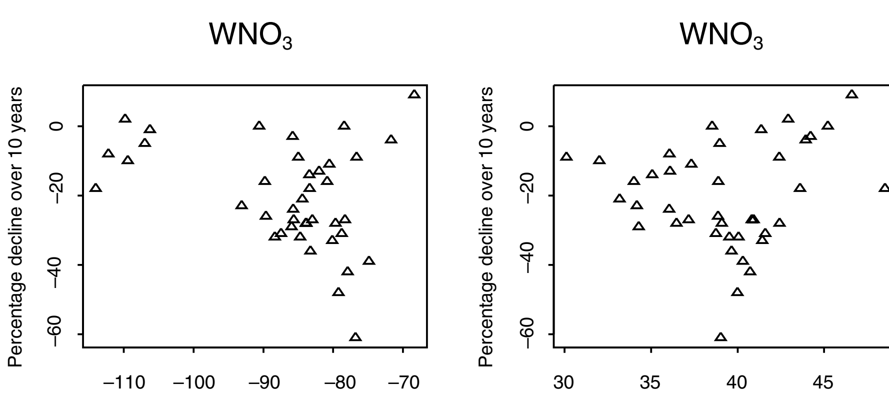


FIGURE 5.6 Percentage changes of cellulose filter nitrate concentrations observed over a 10-year period at the CASTNet stations.

TABLE 5.11
CASTNet Stations with Monitoring Period Ten Years. The + Sign Means the 95% Confidence Region for the μ at the Station and Particular Species Does Not Contain the Overall Average, – Sign Means the Average Is Covered by the Interval. Consequently, Decline (Growth) Observed at the Station Differs Significantly from the Overall Average

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Albambra	–	–	–	–	–	+	–
Ann Arbor	–	+	–	–	–	–	–
Ashland	+	–	+	–	+	+	+
Beltsville	–	+	+	–	–	–	+
Bondville	–	–	–	–	–	–	–
Caddo Valley	–	–	–	–	–	–	–
Cedar Creek	–	–	–	–	–	+	–
Centennial	–	–	–	–	–	–	+
Chiricahua NM	+	+	+	–	+	+	–
Coffeeville	–	–	–	–	+	–	–
Connecticut Hill	–	+	–	–	+	+	–
Coweeta	–	–	–	–	–	–	–
Cranberry	–	–	–	–	–	+	–
Deer Creek	–	–	–	–	–	–	–
Edgar Evins	–	+	–	–	–	–	–
Georgia Station	–	–	–	–	–	–	–
Glacier NP	–	+	–	–	–	+	–
Goddard	–	–	–	–	–	–	–
Gothic	–	–	–	+	–	–	+
Grand Canyon	–	–	–	–	+	–	–
Horton Station	–	–	–	–	–	+	–
Kane	–	–	–	–	–	–	–
Laurel Hill	–	–	–	–	–	–	+
Lykens	–	–	–	–	–	–	–
Oxford	–	–	–	–	–	–	–
Parsons	–	–	–	–	–	+	–
Penn. State U.	–	–	–	–	–	–	+
Perkinstown	–	–	–	–	–	–	+
Pinedale	–	–	–	–	–	+	+
Prince Edward	–	–	–	–	+	–	–
Salamonie Reservoir	–	+	–	–	+	–	–
Sand Mountain	–	–	–	–	–	–	–
Shenandoah NP	–	+	–	–	–	–	+
Speedwell	–	–	–	–	–	–	–
Sumatra	–	–	–	–	–	–	–
Unionville	–	–	–	–	–	–	–
Vincennes	–	–	–	–	–	+	–
Wash. Crossing	–	–	–	–	–	–	+
Wellston	–	–	+	–	–	–	+
Woodstock	+	–	+	–	+	+	+

TABLE 5.12
CASTNet Stations with Monitoring Period Four Years. The + Sign Means the 95% Confidence Region for the μ at the Station and Particular Species Does Not Contain the Overall Average, – Sign Means the Average is Covered by the Interval. Consequently, Decline (Growth) Observed at the Station Differs Significantly from the Overall Average

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Abington	–	–	–	–	+	+	–
Beaufort	–	–	–	–	–	–	–
Big Bend NP	–	–	–	–	–	–	–
Blackwater NWR	–	–	–	–	–	–	–
Canyonlands NP	–	–	–	–	–	+	–
Claryville	–	–	–	+	–	–	–
Death Valley NM	–	–	–	+	–	–	–
Egbert	–	–	–	+	–	–	–
Great Basin NP	–	–	–	+	–	+	–
Joshua Tree NM	–	–	–	–	–	–	+
Lessen Volcanic NP	–	–	–	–	–	–	+
Lye Brook	–	+	–	–	–	+	–
Mesa Verde NP	–	–	–	–	–	+	–
Pinnacles NM	–	–	–	–	+	–	+
Rocky Mtn NP	–	+	–	–	–	+	+
Stockton	–	–	–	–	–	–	–
Yosemite NP	–	–	–	–	–	–	–

Table 5.9 and Table 5.10 indicate that the overall averages are representative for most stations. A more complex and rigorous procedure follows in Section 5.8.

Compared to what we are used to seeing in the literature (Holland et al. 1999), except the impressive NSO_4 value, the overall percentage declines for each species are somewhat more moderate. In fact, the 4-year monitoring period does not exclude the possibility of growing TSO_4, TNH_4, and WSO_2 concentrations. We discuss this phenomenon more in detail at the end of Section 5.8.6.

Remarkably, the exceptionally high or low values of NSO_4 do not seem to be accompanied by excessive changes in other monitored species. In fact, all occurrences of the plus sign seem rather random and unrelated to each other. Also, nothing suggests an accumulation of plus signs at a particular geographic region.

5.4.4 Statistical Features of the Indicators $\hat{pc}$ and $\hat{pd}$

Since $\hat{\mu}$ is asymptotically Normal, $\exp\{-\hat{\mu}\}$ is asymptotically a Lognormal random variable. Any *Lognormal random variable* is described by (5.3), where ζ is assumed to be Normal. Under the additional normality assumption, the expectation of (5.3) is

$$E\left(\frac{c}{c'}\right) = \exp\{\mu + \nu^2/2\}, \tag{5.30}$$

where v denotes the standard deviation of ζ and the variance of c/c' is

$$var\left(\frac{c}{c'}\right) = \exp\{2\mu + v^2\}(\exp\{v^2\} - 1). \tag{5.31}$$

See Finney (1941) or Kendall and Stuart (1977), Volume 1, for details. Lognormal random variables are frequently used for chemistry data analysis (Atchison and Brown 1957).

In consequence of (5.30) and (5.31), the expectation and variance of $p\hat{d}$, for example, are

$$Ep\hat{d} = 100(1 - \exp\{-\mu + var(\hat{\mu})/2\}) \tag{5.32}$$

and

$$var(p\hat{d}) = 10000\exp\{-2\mu + var(\hat{\mu})\}(\exp\{var(\hat{\mu})\} - 1), \tag{5.33}$$

respectively. Expression (5.32) shows that $p\hat{d}$ is a biased estimator because it underestimates the true pd. The true variance of $\hat{\mu}$ is not available due to the lack of the true parameters and must be replaced in applications by the estimated variance $v\hat{a}r(\hat{\mu})$. Since $v\hat{a}r(\hat{\mu})$ tends to zero with growing sample size, the estimator is asymptotically unbiased.

Due to the monotony of the natural logarithm, the approximate 100(1 − α)% confidence region for pd is

$$\begin{aligned} &100(1 - \exp\{-\hat{\mu} + \sqrt{v\hat{a}r(\hat{\mu})}u(\alpha)\}) < pd \\ &< 100(1 - \exp\{-\hat{\mu} - \sqrt{v\hat{a}r(\hat{\mu})}u(\alpha)\}). \end{aligned} \tag{5.34}$$

Using the statistics Z'_μ we get

$$\begin{aligned} &100(1 - \exp\{-\hat{\mu} + \sqrt{v\hat{a}r(y)}u(\alpha)/(\hat{v}\sqrt{N})\}) < pd \\ &< 100(1 - \exp\{-\hat{\mu} - \sqrt{v\hat{a}r(y)}u(\alpha)/(\hat{v}\sqrt{N})\}). \end{aligned} \tag{5.35}$$

In a similar way we can obtain characteristics of $p\hat{c}$ and the confidence region covering pc.

A correction to the bias of the pd estimator can be easily derived from results in Finney (1941). However, calculation of the corrected estimator involves parameters which also must be estimated and that imports a new kind of bias in the corrected estimator. Hence, here we prefer to live with the bias and have a common straightforward quantity easy to calculate and well suited for comparison purposes.

If the length of each compared period is P years, since $\hat{\mu}$ is an unbiased estimate of μ, the quantity

$$\hat{\rho} = \frac{\hat{\mu}}{P} \tag{5.36}$$

is an unbiased estimator of the annual rate of decline ρ introduced in (5.10). The estimator is useful, e.g., for testing of agreement between declines estimated from two different data sets collected at the same location; see the end of Section 5.9.5. Let us recall that the $100(1-\alpha)\%$ confidence region for $\hat{\rho}$ is

$$\hat{\rho} - u(\alpha)\sqrt{v\hat{a}r(\hat{\mu})}/P < \rho < \hat{p} + u(\alpha)\sqrt{v\hat{a}r(\hat{\mu})}/P. \tag{5.37}$$

5.5 DECLINE ASSESSMENT FOR INDEPENDENT SPOT CHANGES

The long-term percentage decline, as introduced in formula (5.6), is a natural indicator of change using daily or weekly observations divided in two equally long seasons. Weekly observations, for example, gathered over two subsequent years, produce a set of 52 spot changes. Such small samples rarely exhibit autocorrelation and can be analyzed using the familiar *t*-test. Results of such an assessment can be interpreted in a straightforward manner and, in conjunction with the case study of 10 years of CASTNet data, provide an important insight in the nature of dry chemistry measurements.

Models formed by series of mutually independent, identically distributed (*iid*) Normal random variables describe numerous CASTNet air quality data. Table 5.13 to Table 5.15 list description of models considered in some sense optimal for the analysis. A chemical monitored by a station with the symbol −.0.. or +.0.. in the proper column is best described by *iid* Normal variables. Since the model occurs rather frequently, this section recalls the elements of statistical inference in the context of change assessment assuming the observed concentrations follow model (5.11) and ζ_t are *iid* and Normal. Independence of ζ_t in (5.11) is equivalent to independence of the spot percentage changes introduced by (5.4).

5.5.1 Estimation and Inference for Independent Spot Changes

Let us consider data from two subsequent periods measured in years, admitting representation (5.7) and (5.8), respectively. We are interested in estimation and inference about the parameter μ in (5.11) under the assumption that $\zeta_t = \sigma\varepsilon_t$, where the variables ε_t, $t = 1,\ldots, N$, are *iid* and Normal with zero mean and variance one. Its justification in the context of observed data is outlined in Section 5.5.2 on model validation. If a linear change in the concentrations prevails, the parameter μ is significantly different from zero. To get an idea about the actual value of μ, we must estimate it. The familiar estimator of μ for a sample of Normal mutually independent data is the sample average. Due to the Normal distribution of the logarithms, the average $\hat{\mu}$ is also a Normal random variable with mean μ and variance σ^2/N. This is not as large a sample result as the one presented in Section 5.4.1! The Normal distribution assumption about the data thus admits a bit more accurate conclusions recalled next.

The null hypothesis that the parameter μ equals the value required by the policy, versus the alternative that μ is different from the policy value, can be tested using

TABLE 5.13
The Model for Western US CASTNet Stations Is $\ln(c_t/c'_t) = \alpha + \beta t + \zeta_t$, Where ζ Is an $AR(p)$ Process. Each Column Contains: Sign of β,* if β Is Significant, the Order p, k if Null Hypothesis Rejected by KS Test and c if Rejected by χ^2 Test. A Dot Means a Non-Significant Result

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Big Bend NP	–.0..	–.0..	–.0..	–.10..	–.7..	–.0..	–.0..
Canyonlands NP	–.1..	+*0..	–.1..	+.1k.	–.1..	+.0..	+.1..
Centennial	–.1kc	–.0..	–*1..	+*5..	–.0..	–*6..	–.3..
Chiricahua NM	–.1..	–*6..	–*1..	+*7..	–.4k.	–.3..	+*2..
Death Valley NM	–.0..	+.0..	–.0..	–.0..	+.0..	–.2..	+.1..
Glacier NP	+.2..	+.0..	+.2..	+*4..	–*3..	–*7..	+.4..
Gothic	+.1kc	–.1..	–*1k.	+*4..	–.3k.	–*1..	–.2..
Grand Canyon	+.1..	–*5..	–.1..	+*1..	–.0..	–.0..	+.3..
Great Basin NP	+.0..	+*1..	+.0..	–.0..	+.1..	–.1..	–.1..
Joshua Tree NM	+.3..	+*0..	–.3..	–.0..	+.1..	+.3..	+*1..
Lassen Volcanic NP	–.0..	–.4..	–.0..	–*0..	–.0..	–.0..	..0..
Mesa Verde NP	+.0..	+*1..	–.0..	+.0..	+.0..	+.0..	–.4..
Mount Rainier NP	+.0..	+.0..	+.4..	–*0..	–.0..	+.0..	–*0k.
North Cascades NP	+.0..	–.0k.	–.0..	–*0k.	–.4..	–.0..	–*1..
Pinedale	+*1..	–.1k.	+.1..	+*4..	–.1..	–*0..	–*4..
Pinnacles NM	–*0..	+.0..	–*0..	–*0..	–*0..	–.0..	–.1k.
Rocky Mtn NP	+.0..	+.0..	+.0.c	+.1k.	–.0..	–.2..	–.1..
Sequoia NP	–*2..	–*2..	–.0..	–*0..	–.0..	–.1..	–.1..
Yellowstone NP	–.0..	–.0..	–.0..	–*1..	–.0..	+.0..	–.0..
Yosemite NP	–*4..	–.0..	–*2..	–*1..	–.2k.	–.0k.	–.1..

the likelihood ratio (Kendall and Stuart 1977, Volume II, Section 24.1), resulting in the statistics

$$t = \frac{\hat{\mu} - \mu}{\hat{\sigma}} \sqrt{N}, \tag{5.38}$$

with Student's T_{N-1}-distribution of $N - 1$ degrees of freedom. Let us recall that $\hat{\sigma}^2 = v\hat{a}r(y)$ is the sample variance defined by (5.20) and N is the number of pairs available for testing. The test rejects the null hypothesis in favor of the alternative if $|t| > T_{N-1}(\alpha)$, where α is the prescribed significance level and $t_{N-1}(\alpha)$, is the quantile for which $Prob(|T_{N-1}| > t_{N-1}(\alpha)) = \alpha$. The $100(1 - \alpha)$% confidence region for μ is (Kendall and Stuart 1977, Volume 2, Section 20.31)

$$\hat{\mu} - \hat{\sigma} t_{N-1}(\alpha)/\sqrt{N} < \mu < \hat{\mu} + \hat{\sigma} t_{N-1}(\alpha)/\sqrt{N}. \tag{5.39}$$

Mutually independent, identically distributed Normal random variables form a stationary process with $R(0) = \sigma^2$ and $R(h) = 0$ for any other $h \neq 0$. Results from

TABLE 5.14
The Model for Eastern US CASTNet Stations Is $\ln(c_t/c'_t) = \alpha + \beta t + \zeta_t$, Where ζ Is an $AR(p)$ Process. Each Column Contains: Sign of β,* If β Is Significant, the Order p, k if Null Hypothesis Rejected by KS Test and c if Rejected by χ^2 Test. A Dot Means a Non-Significant Result

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Abington	+.2..	−*0..	−.2..	+.2k.	+.6..	−.1..	+*1..
Alhambra	−.0k.	−*0..	−*0k.	+*9..	−.0..	−*0..	+*6k.
Ann Arbor	−.1..	−.2..	−*1..	+*3..	−.0k.	−*0..	+*3.c
Ashland	−.0..	−*2..	−*0..	+*3..	−*0..	−*3..	−.2..
Beaufort	−.5..	+.0..	−.8..	+.2..	−.1..	−.1..	−.0..
Beltsville	−*0..	−.0k.	−*0..	+*2k.	−.0kc	−*2..	+*3.c
Blackwater NWR	+.0..	−*0..	+.0..	−.0..	+*1..	+.0..	+.0..
Bondville	−.0..	−.0..	−*0..	+*3..	+.1..	−*0..	+*4..
Caddo Valley	−*1..	−*0..	−*1..	+*4..	−.1..	−*1..	+*3..
Candor	−*2..	+.0..	−*11..	+*4..	+.0..	−.3..	+.3..
Cedar Creek	−.1..	−.1..	−.2..	+*3..	−.1..	+.2..	+*3..
Claryville	−.9..	−*2..	−.6..	+.2..	+.6..	+.3..	−.0..
Coffeeville	−.1..	+*1..	+.2..	+*6..	−.2..	−.0..	+*3..
Connecticut Hill	−.1..	−.1..	−*10..	+*4..	+.0..	−.4..	+*3k.
Coweeta	−*1..	−*1..	−*1..	+*4.c	−*3..	−*2..	+.4..
Cranberry	−*0..	−.0..	−*1..	+*4..	−*1.c	−*3.c	+.3..
Crockett	−.0..	+*0..	+.1..	+*2..	+.1..	−*0..	−.0..
Deer Creek	−.0..	−*2..	−*0..	+*3..	+.1..	+.9kc	+*3..
Edgar Evins	−.2..	−.1..	−*3..	+*3..	−.1..	−.1..	+*3..
Egbert	+.2..	+.0..	−.0..	+.1..	+.0..	+.0..	−.1..
Georgia Station	−*0..	−.6..	−*3..	+*3..	−*1..	−*1..	+*1..
Goddard	−.0..	−.2..	−*0..	+*3..	+.0..	−.1..	+*8..
Horton Station	−.0..	−*3..	−*0..	+*5..	−*0..	−.1..	+.8..

Section 5.4.4 can thus be applied. The spectral density of the process ζ has for *iid* random variables form $f(\lambda) = \sigma^2/(2\pi)$ for all $\lambda \in (-\pi, \pi)$. Consequently, $\hat{\nu} = \nu = 1$ in (5.24) and $Z'_\mu = t$ in (5.38). For large N, the T_{N-1} distribution converges to the Normal, hence, for large samples the tests based on t and Z'_μ agree.

The long-term percentage change (5.5) and percentage decline (5.6) can be estimated using (5.26) and (5.27), respectively. The expectation of $\hat{pd}$ is

$$E\hat{pd} = 100(1 - \exp\{-\mu + \sigma^2/2N\}) \tag{5.40}$$

and the variance is

$$var(\hat{pd}) = 10000\exp\{-2\mu + \sigma^2/N\}(\exp\{\sigma^2/N\} - 1). \tag{5.41}$$

These results are now accurate, not asymptotic.

TABLE 5.15
The Model for Eastern US CASTNet Stations Is $\ln(c_t/c_t') = \alpha + \beta t + \zeta_t$, Where ζ Is an $AR(p)$ Process. Each Column Contains: Sign of β,* if β Is Significant, the Order p, k if Null Hypothesis Rejected by KS Test and c if Rejected by χ^2 Test. A Dot Means a Non-Significant Result

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Howland	−.5..	+.0..	−.1.c	+*3..	−.0..	−*3..	−*1..
Kane	−.0..	−*1..	−*0..	+*3..	+.0..	−.1..	+*4..
Laurel Hill	−.0..	−.2..	−.1k.	+*4..	−.1..	+.3..	+*4..
Lye Brook	+.0..	−*0..	−.0..	+*0..	+.0..	+.0..	+*0..
Lykens	−.0..	−*1..	−*0..	+*3..	+.0..	+.0..	+*4..
Mackville	−.1..	+.0..	−.0..	+*3k.	+.1k.	−.1..	+*2..
Oxford	−.0..	−.5..	−.0..	+*6..	−.1..	+.7..	+*10..
Parsons	−.0..	−*3..	−.0..	+*4kc	−.1k.	−.3..	+*5..
Penn. State U.	−.1..	−.5..	−.0..	+*3..	+.2..	−.3..	+*4..
Perkinstown	−.0..	−.0..	−*0..	+*3..	−.0..	−*0..	−.3..
Prince Edward	−.0..	−.6..	−*0..	+*4..	−.2..	−.5..	+*5..
Salamonie Reservoir	−*0..	−*3..	−*0.c	+*4..	−.1..	−*8..	+*2..
Sand Mountain	−*1..	−.2..	−*2..	+*3..	−.0.c	−.10..	+*3..
Shenandoah NP	−.0..	−*2..	−*0..	+*5..	−*0..	−*2..	+*7k.
Speedwell	−.1..	−.1..	−.2..	+*4..	−.0..	−*4k.	+*6.c
Stockton	−*3..	+.3k.	−.3..	+.2..	−*0..	−.0..	−.2..
Sumatra	−.2..	−.0..	−*2..	+*4k.	−.5..	−*1..	+*1..
Unionville	−.0..	−*0..	−*0..	+*5..	−.0..	−*0..	+.2..
Vincennes	−.0..	−*1..	−.0..	+*4..	−.1..	−.7..	+*6..
Voyageurs NP	−.0..	+.0..	+.0..	−*2..	−.0..	+.0..	−.0..
Wash. Crossing	−*2..	−*1..	−*0..	+*4k.	−.1..	−*1k.	+*2..
Wellston	−.0..	−*0..	−*0k.	+*4..	−.0..	−*0..	+.4..
Woodstock	+.0..	−.1..	−.7..	+*3..	−.0..	−*1..	+.2..

Expression (5.40) shows that $\hat{pd}$ is a biased estimator underestimating the true pd. Due to the monotony of the natural logarithm, the $100(1-\alpha)\%$ confidence region for pd is

$$\begin{aligned}&100(1-\exp\{-\hat{\mu}+\hat{\sigma}t_{N-1}(\alpha)/\sqrt{N}\}) < pd \\ &< 100(1-\exp\{-\hat{\mu}-\hat{\sigma}t_{N-1}(\alpha)/\sqrt{N}\}).\end{aligned} \tag{5.42}$$

Characteristics of $\hat{pc}$ and the confidence region for pc can be obtained similarly.

Example 1.5.1: To see the influence of the bias on the percentage decline estimate, let us consider a set of $N = 52$ *iid* Normal spot changes obtained from 2 years of weekly data. If $\mu = 0$ and $\sigma = 0.600$, then $E\hat{pd} = -0.347\%$, whereas the true value is zero percent. The standard deviation of $\hat{pd}$ is $\sqrt{var(\hat{pd})} = 8.364\%$. That compares reasonably to the measurement error reported for the CASTNet monitored species by Sickles and Shadwick (2002b).

Example 2.5.1: Suppose a year of weekly data yields *iid* Normal spot changes with $\hat{\mu} = 0.000$ and $\hat{\sigma} = .600$. Then $t_{51}(0.05) = 2.008$ and the 95% confidence region for $pd = 0$, i.e., no decline, is (−18.184, 15.386). Hence, the observed percentage decline can be somewhere in the interval (−19, 16) percent and still has to be considered as a rather random event with high occurrence frequency 95%.

Example 3.5.1: Here we compare the bias and variance of the estimator $p\hat{c}$ and the estimator $p\tilde{c}$, introduced in (5.29).

The expectation of $p\tilde{c}$ is, according to (5.30),

$$Ep\tilde{c} = 100(\exp\{-\mu + \sigma^2/2\} - 1), \tag{5.43}$$

which means the estimator is biased and the bias does not vanish with growing sample size.

For independent spot changes, the variance of $p\tilde{c}$ is, according to (5.31),

$$var(p\tilde{c}) = 10000\frac{1}{N}\exp\{-2\mu + \sigma^2\}(\exp\{\sigma^2\} - 1). \tag{5.44}$$

Therefore, the estimator is consistent and approaches the value $Ep\tilde{c}$ with growing sample size. The speed of convergence is similar to that of $p\hat{c}$. That is because for large N, $N(\exp\{\sigma^2/N\} - 1) \approx \sigma^2$, as we know from elementary calculus, and in consequence of (5.41),

$$var(p\hat{c}) = var(p\hat{d}) \approx 10000\frac{\sigma^2}{N}\exp\{-2\mu + \sigma^2/N\}. \tag{5.45}$$

Example 4.5.1: The bias of $p\tilde{c}$ is certainly not negligible. For example, if $\mu = 0$ and $\sigma = 0.6$, then substitution in (5.43) yields $Ep\tilde{c}$ = 19.72%! The estimator $p\tilde{c}$ in (5.29) is thus not particularly useful. The presence of σ in $Ep\tilde{c}$, unchanged even for large samples, is very unpleasant because σ reflects all artefacts arising during the sampling process, in particular the measurement error.

5.5.2 Model Validation

To verify the independence and Normal distribution of the data one can use the familiar procedures recalled next. The aim of the tests and diagnostic plots is to assure that the data exhibit no obvious conflict with the normality and independence hypothesis. For simplicity the data are considered standardized, which means the trends were removed and they are scaled to have variance equal to one.

If the data $e_1, \ldots, e_N$ are *iid* and Normal, then the plot of e_t against the lag t should exhibit a band of randomly scattered points with no apparent clusters and outliers. The edges of the band should be parallel and not wave or form other patterns (see Figure 5.3b).

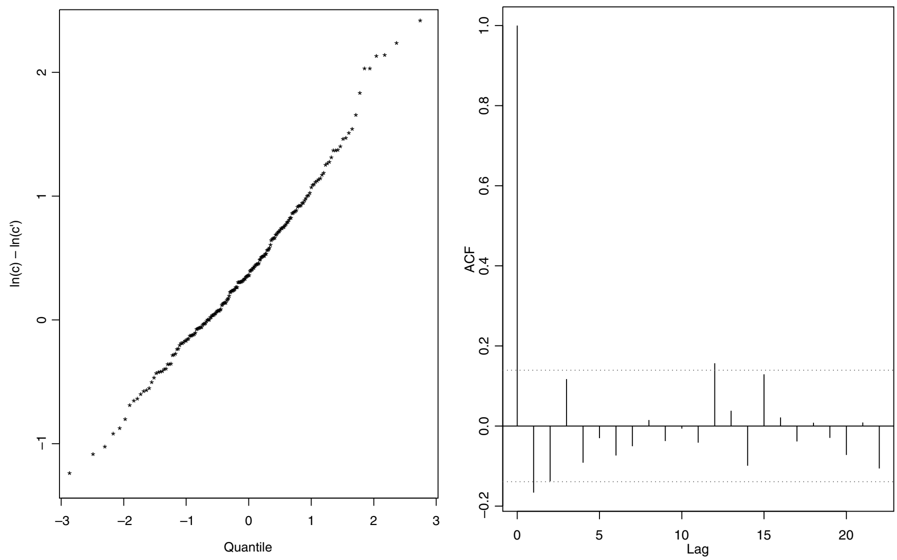


FIGURE 5.7 Diagnostic plots from Woodstock teflon filter SO_4 (mg/1) observations $\ln(c_t) - \ln(c_t')$ in Figure 5.3. The quantile plot supports the Normal distribution assumption. The ACF plot indicates second order autocorrelation.

A further step towards the Normality verification is the *quantile plot*. The plot arises by drawing the observed e_t against the quantiles $\Phi^{-1}(\hat{\Phi}(e_t))$. In our notation, $\Phi(x)$ is the standard Normal distribution, $\Phi^{-1}(x)$ is its inverse, and $\hat{\Phi}(x)$ is the empirical distribution function of the data. If the data come from $\Phi(x)$, then we see an almost straight line because $\hat{\Phi}(x) \approx \Phi(x)$ (see Figure 5.7a). Construction of $\hat{\Phi}$ is described in Kendall and Stuart (1977, Volume II, Section 30.46).

After the plot inspection, the testing proceeds using the Kolmogorov–Smirnov (KS) and χ^2 Normality tests described in Kendall and Stuart (1977) (Volume II, Section 30.49). The two tests are derived under the assumption that the data are *iid* Normal. The null hypothesis is that the distribution function of the data equals Φ. The alternative hypothesis is *not* that the distribution is not Normal. The examined data could be well generated by two different Normal distributions with different variances, for example, because the sampling procedure changed at some point in time. The inference utilizing KS statistics is based on the fiducial argument (Kendall and Stuart 1977, Chapter 21). The distribution of the KS statistics $\{x: |\Phi(x) - \hat{\Phi}(x)|, -\infty < x < \infty\}$, is derived under the null hypothesis, and if the observed value of the maximum is unlikely under this distribution, the hypothesis is simply rejected.

The use of the goodness-of-fit tests is recommended along with the quantile plots to get a more accurate idea about the reason for rejection. The KS test is sensitive to discrepancies in the center of the empirical distribution and towards outliers, and the χ^2 test is sensitive to violations at the tails of the distribution. The quantile plot seems less affected by presence of autocorrelation than the KS and χ^2 test results.

The independence of the data can be studied using the autocorrelation function plot, *ACF plot* for short. If $e_1,\ldots, e_N$ are *iid*, then

$$\hat{R}(h) = \frac{1}{N}\sum_{t=h+1}^{N} e_t e_{t-h} \tag{5.46}$$

should form a series of mutually independent random variables for $h > 1$. Hence, dividing each $\hat{R}(h)$ by $\hat{R}(0)$ and plotting the scaled values against h, we should see a randomly scattered series such as the one in Figure 5.7b. Since $\hat{R}(h)$ is an estimator of $R(h)$, large values, especially at the beginning of the plot, signal the presence of autocorrelation.

5.5.3 Policy-Related Assessment Problems

The relation between the annual decline statistics $\hat{pd}$ and the t-test (5.38) in Section 5.5.1 allows us to answer the basic quantitative questions relating to policies and their enforcement. Suppose we have two subsequent years of weekly concentrations.

Problem 1.5.3: What accuracy must the concentration measurements have should the t-test detect a 6% decline as significant on a 5% significance level?

Problem 2.5.3: What is the lowest percentage decline $\hat{pd}$ detectable as significant given a measurement accuracy?

The answers follow upon investigation of the critical level k that must be exceeded by $\hat{pd}$ to be recognized as significant. The inequality $\hat{pd} > k$ is true if and only if $\hat{\mu} > \ln(1 - k/100)^{-1}$. The quantity $\ln(1 - k/100)^{-1}$ should thus agree with the critical value of the one-sided t-test for the significance of $\hat{\mu}$. Consequently, k is chosen to satisfy

$$\ln\left(1-\frac{k}{100}\right)^{-1} = \frac{\hat{\sigma}}{\sqrt{N}}\, t_{N-1}(\alpha), \tag{5.47}$$

where $\hat{\sigma}$ is given in (5.20) and $t_{N-1}(\alpha)$ is the quantile of Student's T_{N-1} distribution with significance level α. For example, if $N = 52$ weeks and $\alpha = 0.05$, then $t_{51}(0.05) = 1.675$.

Solution to Problem 1.5.3: If k = 6% is on the edge between significant and nonsignificant, then $\hat{\sigma} = \ln(1 - 0.06)^{-1}\ \sqrt{52}/1.675 = 0.266$ is the largest admissible $\hat{\sigma}$. Hence, $\hat{pd} = 6\%$ will be assessed as a significant change on the 5% significance level if and only if the sample deviation (5.20) will not exceed 0.266. If c_t and c'_t are Lognormal and follow (5.7) and (5.8) with variance parameter $v^2 = var(\eta_t) = var(\eta'_t)$, then the deviation of the observed $\ln(c_t/c'_t)$ is $\hat{\sigma} \approx \sqrt{2}\, v$, which means deviation v of the logarithms should not exceed $0.266/\sqrt{2} = 0.188$. The expected relative measurement error, expressed in percentages, must be thus kept under 15.18% if a 6% change over 2 years of monitoring should be detectable. For example, CASTNet filter pack data have $\hat{\sigma}$ somewhere between 0.4 and 0.8. Notice that due to the long time gap between observing c_t and c'_t we consider η and η' mutually independent.

Note: If the observed concentrations follow model (5.1), and η is interpreted as randomness due to the measurement error, then the expected, or average, relative measurement error expressed in percentages is

$$e = 100E\left|\frac{\exp\{m+\eta\}-\exp\{m\}}{\exp\{m\}}\right| = 100E\,|\exp\{\eta\}-1|. \tag{5.48}$$

If η has zero mean, Normal distribution, and deviation v, then the expectation has the form

$$\begin{aligned} E\,|\exp\{\eta\}-1| &= \int_0^{\infty}(\exp\{x\}-1)\phi(x)dx - \int_{-\infty}^{0}(\exp\{x\}-1)\phi(x)dx \\ &= \int_0^{\infty}(\exp\{x\}-\exp\{-x\})\phi(x)dx, \end{aligned} \tag{5.49}$$

where $\phi(x)$ is the density function of the Normal distribution. Evaluation of the integral yields

$$e = 100\exp\{v^2/2\}(\Phi(v)-\Phi(-v)), \tag{5.50}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left\{-\frac{u^2}{2}\right\}du \tag{5.51}$$

is the standard Normal distribution function.

Solution to Problem 2.5.3: The right site of (5.50) is a monotone-growing function of v. If e is known from the design of the network, then v can be determined uniquely from (5.50). If the assumptions of the model are correct, we get $\hat{\sigma} \approx \sqrt{2}\,v$, and $\hat{pd}$ must exceed

$$k = 100\left(1-\exp\left\{-\frac{\hat{\sigma}}{\sqrt{N}}t_{N-1}(\alpha)\right\}\right) \tag{5.52}$$

to be a significant decline on the prescribed α level.

Using the annual rate of decline (5.10), we can answer another question frequently asked by practitioners.

Problem 3.5.3: Suppose the concentrations are declining slowly, say only 2% annually, and we estimated $\hat{\sigma}$ = 0.6. How many years of weekly observations are needed to detect a statistically significant decline on a 5% significance level?

Solution to Problem 3.5.3: If the annual decline is 2%, then the rate of decline is $\rho = -\ln(1 - .02) \approx 0.02$. As the solution to Question 2 suggests, we need

$$P_\rho = \mu > \frac{\hat{\sigma}}{\sqrt{PW}}t_{PW-1}(\alpha),$$

where W is the number of observations collected during each year. Setting t_{PW-1} $(0.05) \approx 1.675$, we get

$$P^{3/2} > \frac{1}{\rho}\frac{\hat{\sigma}}{\sqrt{W}} \approx 83.75\frac{\hat{\sigma}}{\sqrt{W}}.$$

Consequently, we need

$$P > \left(83.75\frac{\hat{\sigma}}{\sqrt{W}}\right)^{2/3}.$$

If the standard deviation of ζ in (5.11) is $\hat{\sigma} = 0.6$ each year, then the right-hand side of the last expression is about 3.648, which has to be rounded in 4 years of paired observations, i.e., 8 years of monitoring.

Problem 4.5.3: Suppose the target of our policies is a 6% reduction over the next 10 years and the deviation of ζ_t in (5.11) is $\sigma = 0.6$. What is the shortest number of years we have to monitor to notice a significant decline in concentrations on a 5% significance level?

Solution to Problem 4.5.3: The 6% target over 10 years means that we consider an annual rate of decline $\rho = .1\ \ln(1 - 0.06)^{-1}$. Since the rate of decline is rather small and requires a long observation period, we can take $t_{PW-1}(.05) \approx 1.60$ for critical value. To answer the question, we first notice that the function

$$f(P) = P\rho - \frac{\hat{\sigma}}{\sqrt{PW}}t_{PW-1}(.05) \approx P\rho - 1.60\frac{\hat{\sigma}}{\sqrt{PW}}$$

is monotonously growing as P increases and a simple plot shows that it is crossing zero between $P = 7$ and $P = 8$ years. Consequently, $2P = 16$ years is needed to observe a significant change in the data! We can thus ask if a 6% target is not a bit too moderate when we cannot expect the change to be verifiable after the 10 years.

5.6 CHANGE ASSESSMENT IN THE PRESENCE OF AUTOCORRELATION

Field samples of atmospheric chemistry concentrations from longer periods are usually autocorrelated, and so are the corresponding data generated by (5.11). The most common models for description of stationary processes are autoregressive moving average processes. The objective of this section is to recall some features of the so-called $ARMA(p,q)$ processes and to show how they apply to the assessment of the long-term change.

5.6.1 The *ARMA*(*p*,*q*) Models

A wide class of stationary processes admits description

$$\zeta_t - \rho_1\zeta_{t-1} - \cdots - \rho_p\zeta_{t-p} = \sigma\eta_t, \tag{5.53}$$

where

$$\eta_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}, \tag{5.54}$$

$\rho_1,\ldots, \rho_p$, $\theta_1,\ldots,\theta_q$, and $\sigma > 0$ are real-valued parameters, and $\varepsilon = \{\varepsilon_t, -\infty < t < \infty\}$ are mutually independent, identically distributed random variables with zero mean and variance one. Stationary processes with $\eta_t = \varepsilon_t$, and remaining parameters zero, are known as *autoregressive* processes of the p-th order, briefly $AR(p)$, and the process η_t is generally called the *moving average* process of order q, or simply $MA(q)$. The abbreviation $ARMA(p,q)$ stands for the general stationary *autoregressive moving average* process ζ_t of orders p,q described by (5.53) and (5.54). It can be shown (Brockwell and Davis 1987, Section 7.1, Remark 3), that $ARMA(p,q)$ processes satisfy the condition (5.13) and thus obey the law of large numbers and the central limit theorem.

This section assumes that our observations are generated by the model (5.11) where ζ obeys an $ARMA(p,q)$ process. The advantage of $ARMA(p,q)$ processes is that they cover a sufficiently broad range of data, they can be identified using autocorrelation plots and methods described in Section 5.5.2, their parameters can be reasonably estimated and tested using the likelihood function, and finally, their spectral density function is a simple ratio

$$f(\lambda) = \frac{\sigma^2}{2\pi}\frac{|1+\theta_1 e^{-i\lambda} + \cdots + \theta_p e^{-iq\lambda}|^2}{|1-\rho_1 e^{-i\lambda} - \cdots - \rho_p e^{ip\lambda}|^2}. \tag{5.55}$$

We always assume that $\Phi(x) = 1 - \rho_1 x - \cdots - \rho_p x^p$ is not zero for each complex x from the unit circle $\{x{:}|x| \le 1\}$ and the polynomials in the ratio (5.55) have no common zeros. Replacing the parameters $\rho_1,\ldots, \rho_p$, $\theta_1,\ldots, \theta_q$, and σ by their maximum likelihood estimators, we have, according to (5.19),

$$v\hat{a}r(\hat{\mu}) \approx \frac{2\pi}{N}\hat{f}(0) = \frac{\hat{\sigma}^2}{N}\frac{(1+\hat{\theta}_1 + \cdots + \hat{\theta}_q)^2}{(1-\hat{\rho}_1 - \cdots - \hat{\rho}_p)^2}. \tag{5.56}$$

Example 1.6.1: The simplest example of an autoregressive process is the $AR(1)$, described by the relation

$$\zeta_t = \rho\zeta_{t-1} + \sigma\varepsilon_t, \tag{5.57}$$

where $|\rho| < 1$ and $\varepsilon = \{\varepsilon_t, -\infty < t < \infty\}$ are mutually independent random variables with zero mean and variance one. In this case,

$$R(0) = \frac{\sigma^2}{1-\rho^2} \quad \text{and} \quad f(0) = \frac{\sigma^2}{2\pi(1-\rho)^2}. \tag{5.58}$$

Consequently, according to (5.16) and (5.26), for a large sample of size N,

$$var(\hat{\mu}) \approx \frac{1}{N}\frac{\sigma^2}{(1-\rho)^2} \quad \text{and} \quad Z_\mu = \frac{\hat{\mu}-\mu}{\hat{\sigma}}\sqrt{N}(1-\hat{\rho}). \tag{5.59}$$

Due to the relations (5.23) and (5.24),

$$\nu = \sqrt{\frac{1-\rho}{1+\rho}} \quad \text{and} \quad Z'_\mu = \frac{\hat{\mu}-\mu}{\sqrt{v\hat{a}r(y)}}\sqrt{N\frac{1-\hat{\rho}}{1+\hat{\rho}}}, \tag{5.60}$$

where $\hat{\rho}$ and $\hat{\sigma}$ are either maximum likelihood or least squares estimators of ρ and σ, respectively.

Example 2.6.1: Suppose the data obey an $AR(1)$ model with $\rho = 0.500$ and we fail to consider the autocorrelation when drawing inference about the long-term percentage change, which essentially means about μ, and use the ordinary t-test statistics (5.38) instead. Then $\nu = 0.577$ and the statistics (5.60) tells us that we would have to reduce our t by nearly one half (!) to get a correct conclusion. Similarly, if $\rho = -0.500$, then $\nu = 1.732$, which means the t is nearly a half of what should be used as the correct statistics for inference. Though $\rho = \pm 0.500$ is a rather strong autocorrelation, the result is certainly alarming.

Example 3.6.1: To see how the bias and variability of the percentage decline estimator $p\hat{d}$ increase in the presence of autocorrelation, let the data follow an $AR(1)$ process with $\mu = 0$, $\sigma = 0.600$, and $\rho = 0.300$, and let each of the compared periods consist of a year of weekly data (see Example 1.5.1). Substitution of $f(0)$ from (5.58) in (5.32) and (5.33), respectively, yields

$$Ep\hat{d} \approx (1-\exp\{-\mu+\sigma^2/(1-\rho)^2/2N\}) \tag{5.61}$$

and

$$var(p\hat{d}) \approx \exp\{-2\mu+\sigma^2/(1-\rho)^2/N\}(\exp\{\sigma^2/(1-\rho)^2/N\}-1). \tag{5.62}$$

Using $N = 52$ we get $Ep\hat{d} = -0.709\%$, which compares to the true value 0%. The standard deviation of our percentage change is $\sqrt{var(p\hat{d})} = 12.013\%$. That is more than 8.364 obtained when $\rho = 0$.

Example 4.6.1: Suppose $\hat{\mu} = 0.000$, $\hat{\sigma} = 0.600$, and $\hat{\rho} = 0.300$ are obtained from a set of $N = 52$ differences (5.11) calculated from two subsequent years of monitoring. Under the $AR(1)$ model, what are the 95% confidence regions for pd? Due to (5.35) and (5.60), the $100(1 - \alpha)\%$ region for pd is (–24.888, 19.928). Consequently, given the data, an annual decline within the limits about –25 to 20% is a fairly frequent event happening 95% of the time even if no change really occurred.

5.6.2 Selection of the *ARMA(p,q)* Model

The selection of the best fitting $ARMA(p,q)$ model is based on the so-called reduced log-likelihood function and its adjustment, the Akaike's information criterion (AIC). The basics relating to our particular applications follow next.

The first step of the model selection consists of trend removal. The sample mean, or values of a more complicated curve with parameters estimated by the least squares method, for example, are subtracted from the observations of the process (5.21). The quantile and autocorrelation plots offer an idea about Normal distribution and autocorrelation of the centered data. If the quantile plot does not contradict the Normal distribution assumption, the Normal likelihood function of the $ARMA(p,q)$ model can be used for estimation. An example of the likelihood function for the $AR(p)$ process and related AIC is in Section 5.12.1. Generally, the likelihood function is calculated using the innovation algorithm (Brockwell and Davis 1987, Chapter 8). The pair p,q, for which the likelihood function is the largest, determines the proper model. Since the Normal likelihood function is not quite convenient for calculations, its minus logarithm is preferred instead. The reduced log-likelihood arises by omission of the parameter-free scaling constant from the minus log-likelihood and substitution of the variance estimator for the true parameter. It is used as the measure of fit and the smaller its value computed from the data, the better p,q. The reduced likelihood serves also for calculation of the AIC. Goodness-of-fit tests determine if the residuals of the final model do not contradict the Normality assumption.

Analysis of the autocorrelation plots, the fact that an $MA(q)$ model can be well replaced by a higher order $AR(p)$ model (Brockwell and Davis 1987, Corollary 4.4.2), and the intention to use multivariate $AR(p)$ models lead us to choose for modeling of the CASTNet air quality data $AR(p)$ models and select p using AIC. Results of the analysis are in Table 5.13, Table 5.14, and Table 5.15, respectively. Each column of the table may contain a sequence of symbols describing the slope coefficient of the linear model used for trend removal (+ or –), significance of the slope coefficient based on the common t-test criteria ignoring the autocorrelation, the resulting value of p, the letter k if the Kolmogorov–Smirnov test rejected the Normality hypothesis for residuals, and the letter c if the χ^2 test rejected it. A simple dot means a nonsignificant result.

The sometimes high values of p can be explained by the presence of trends or changes in variability not accounted for by the fitted simple linear model. A specific example provides the NSO_4 measurements exhibiting a change in trend and variability since the beginning of 1997. This observation is consistent with that made in Sickles and Shadwick (2002a). The ability of the $AR(p)$ model to adjust for this kind of inhomogeneity and still provide Normal independent residuals shows a certain degree of robustness of the procedure. Trends left in the differences (5.11) are not unusual. What is their consequence for the change assessment is discussed in Section 5.7.

5.6.3 Decline Assessment Problems Involving Autocorrelation

Next we investigate how autocorrelation of differences in (5.11) affects the solution of problems discussed in Section 5.3. We formulate the problems again to adjust for the more complicated reality.

Problem 1.6.3: Suppose we have two subsequent years of weekly concentrations following model (5.11) with a stationary process ζ. What accuracy must the concentration measurements have to admit detection of a 6% decline if $\hat{\nu} = 0.54$, the significance level $\alpha = 5\%$, and Z_μ, introduced in (5.24), is used for the assessment?

Problem 2.6.3: If the noise in (5.11) is stationary, what is the lowest percentage decline $\hat{pd}$ detectable by the statistics Z_μ as significant given a measurement accuracy?

As earlier, we denote k the critical value that must be exceeded by $\hat{pd}$ should we recognize it as significant. The inequality $\hat{pd} > k$ holds if and only if $\hat{\mu} > \ln(1 - k/100)^{-1}$, and for $\mu = 0$, $Z'_0 > u(\alpha)$ in the case $\hat{\mu} > [u(\alpha)/\hat{v}]\sqrt{v\hat{a}r(y)/N}$. Therefore, k must satisfy the equation

$$\ln\left(1-\frac{k}{100}\right)^{-1} = \frac{u(\alpha)}{\hat{v}}\sqrt{\frac{v\hat{a}r(y)}{N}}, \tag{5.63}$$

where $v\hat{a}r(y)$ is the sample variance (5.20) and $u(\alpha)$ is the α quantile of the standard Normal distribution U satisfying $Prob(U > u(\alpha)) = \alpha$. For $\alpha = 0.05$ we get $u(\alpha) = 1.643$.

Solution to Problem 1.6.3: If the critical value k = 6% then $\sqrt{v\hat{a}r(y)} = \ln(1 - 0.06)^{-1}\, 0.54\sqrt{52}/1.643 = 0.147$ (see Section 5.3). Let $\ln c_t$ and $\ln c'_t$ be Normal random variables described by (5.7) and (5.8), respectively. If $v^2 = var(\eta_t) = var(\eta'_t)$ then deviation of the differences (5.11) is $\sqrt{v\hat{a}r(y)} \approx \sqrt{2}\, v$. Hence, v should not exceed $0.147/\sqrt{2} = 0.104$. The relative measurement error (5.50) expressed in percentages must be thus under 8.33%, nearly a half of that in the case of independent data.

The solution to Problem 1.6.3 is intuitively plausible because it says that, to achieve detection of a change described by one and the same rate parameter μ, the accuracy of measurements must increase with autocorrelation of the observations.

Solution to Problem 2.6.3: The relative error (5.50) that should be available from the design of the network tells us how much $\sqrt{v\hat{a}r(y)}$ might be. The value $\hat{v}$ can be estimated from historical data. Quantities $\sqrt{v\hat{a}r(y)}$ and $\hat{v}$ determine according to (5.63) the critical value

$$k = 100\left(1-\exp\left\{-\frac{u(\alpha)}{\hat{v}}\sqrt{\frac{v\hat{a}r(y)}{N}}\right\}\right). \tag{5.64}$$

The percentage decline is significant on the prescribed α level if $\hat{pd} > k$, because the quantity k is the lower bound for declines we are still able to recognize.

Problem 3.6.3: Suppose the concentrations are declining slowly, say only 2% annually, and our observations support model (5.11). As earlier, we consider a stationary noise ζ, $\hat{v} = .54$, and sample deviation $\sqrt{v\hat{a}r(y)} = 0.6$. How many years of weekly observations are needed to detect a statistically significant decline on a 5% significance level using the statistics Z'_μ?

Solution to Problem 3.6.3: The annual decline of 2% determines the rate of decline $\rho = -\ln(1 - .02) \approx 0.02$. Due to (5.63) we need

$$P\rho = \mu > \frac{u(\alpha)}{\hat{v}}\sqrt{\frac{v\hat{a}r(y)}{PW}}, \tag{5.65}$$

where $W = 52$ and $u(\alpha) = 1.643$. Utilizing the remaining data, entering the problem we get the condition

$$P^{3/2} > \frac{u(\alpha)}{\hat{v}\rho}\sqrt{\frac{v\hat{a}r(y)}{W}} = 12.66, \tag{5.66}$$

which happens for $P > (12.66)^{2/3} = 5.43$. This makes about 5 to 6 years of paired data or 10 to 12 years of observation. As we can notice, the length of observation increased compared to the independent data case in *Problem 3.5.3*.

Problem 4.6.3: Suppose the target of our policies is a 6% reduction over the next 10 years. We expect the data collected during the monitoring process to support model (5.11) with stationary noise, have $\hat{\nu} = 0.54$ and deviation $\sqrt{v\hat{a}r(y)} = 0.6$. What is the shortest number of years we have to monitor to notice a significant decline in concentrations on a 5% significance level using the statistics Z'_μ?

Solution to Problem 4.6.3: The annual rate ρ of decline determining a 6% decline over 10 years is $\rho = .1\ \ln(1 - 0.06)^{-1} \approx 0.006$. To answer the question we notice that the function $f(P)$, defined by the relation

$$f(P) = P\rho - \frac{u(\alpha)}{\hat{\nu}}\sqrt{\frac{v\hat{a}r(y)}{PW}} \approx 0.006P - \frac{0.253}{\sqrt{P}}$$

is growing with P and crosses zero between $P = 12$ and $P = 12.2$ years. The zero point in this interval seems to be the only null point of the function $f(P)$. Consequently, at least $2P$, i.e., 24 to 26 years, is needed to observe a significant change in the data!

5.7 ASSESSMENT OF CHANGE BASED ON MODELS WITH LINEAR RATE

Until now, the observed concentrations could be described by the model (5.11). Occasionally, a linear growth or decline of the differences has to be considered. The linear change is typical for concentrations with exponentially changing long-term trend. A more detailed investigation of such a situation in the change assessment context is conducted in this section. Figure 5.8 shows an example of data with linearly changing rate of decline. The line on the left emphasizes the curvature in the data and the one on the right shows the decline. The main objective of this section is establishing of a generalization of the rate of decline for data following the more comprehensive model. The extension should admit a reasonable interpretation as well as estimation and inference about the long-term change.

5.7.1 Models with Linear Rate of Change

According to Section 5.2.1, non-negative concentrations c_t and c'_t of a chemical species collected over two equally long subsequent periods admit description (5.7) and (5.8), respectively. During this section we assume the additional property

$$\ln c_t - \ln c'_t = \alpha + \beta\,(t/W) + \zeta_t, \tag{5.67}$$

where α and β are the rate change parameters, $\zeta = \{\zeta_t, -\infty < t < \infty\}$ is a zero-mean stationary process with finite fourth moments satisfying (5.13), and W is the number of data from one year. The factor W in (5.67) acts as a weight preventing unnecessarily small values of β estimates.

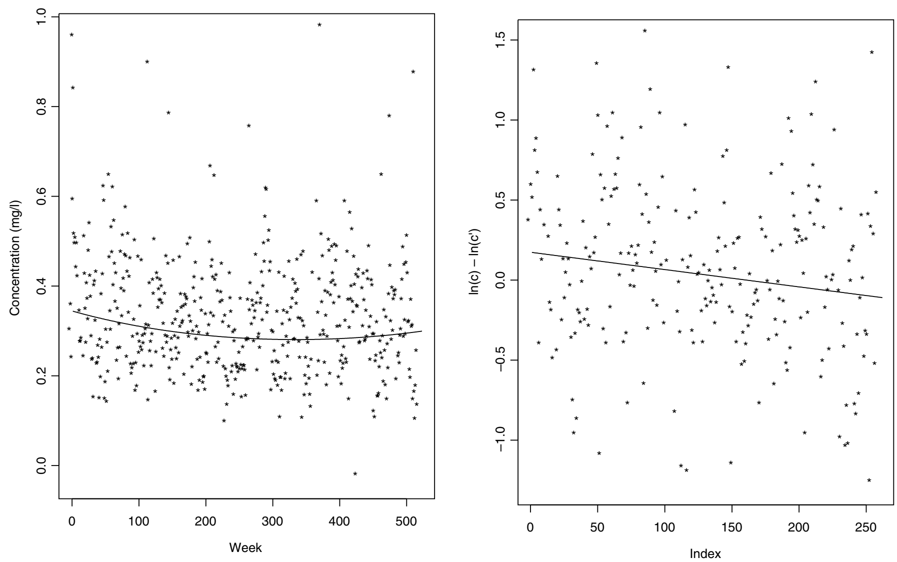


FIGURE 5.8 Observations of nylon filter HNO_3 (mg/l), collected during 1989–1998 by the CASTNet station at Glacier National Park, Montana, USA, demonstrate appropriateness of assumption (5.67). Both lines are estimated by the least squares method.

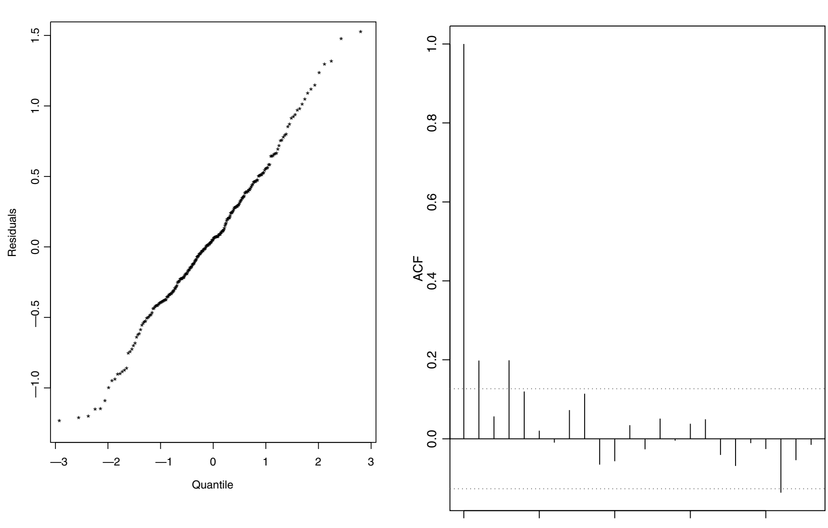


FIGURE 5.9 Diagnostic plots for normality and autocorrelation checkup on residuals from the linear model (5.67). The original data are nylon filter HNO_3 concentrations (Figure 5.8).

To interpret the parameters α and β, let the observed concentrations of a chemical species be sampled weekly, for example, over $2P$ years, in the first period and let them follow a model with

$$m_{t_n} = s + rt_n + ut_n^2 + \pi_{t_n}, \tag{5.68}$$

where $t_n = n/W$, $n = 1,\ldots, WP$, and π_t is an annual periodic component. Then it is sensible to assume that

$$m'_{t_n} = s + r(t_n + P) + u(t_n + P)^2 + \pi_{t_n} \tag{5.69}$$

and consequently,

$$\nu_{t_n} = m_{t_n} - m'_{t_n} = -rP - uP^2 - 2uPt_n \tag{5.70}$$

for each n. Comparison to (5.67) yields

$$\alpha = -(r + uP)P \quad \text{and} \quad \beta = -2uP. \tag{5.71}$$

Due to the dependence of α on the second power of P, dividing ν_{t_n} by P will not produce a quantity characterizing a one-year performance only. We may also notice that for $u > 0$ (i.e., $\beta < 0$) the concentration amounts tend to grow, whereas for $u < 0$ (i.e., $\beta > 0$) they decline starting from a point sufficiently distant in time.

5.7.2 Decline Assessment for Models with Linear Rate of Change

The percentage decline pc introduced in Section 5.4.2 requires $\beta = 0$. To remove this restriction we proceed as follows. If c_t and c'_t, $t = 1,\ldots, N = WP - 1$ are generated by (5.67), then

$$\begin{aligned}\hat{\mu} &= \frac{1}{N}\sum_{t=1}^{N}(\alpha + \beta(t/W) + \zeta_t) \\ &= \alpha + \beta\frac{P}{2} + \frac{1}{N}\sum_{t=1}^{N}\zeta_t.\end{aligned} \tag{5.72}$$

Let us denote

$$\mu = \alpha + \beta\frac{P}{2}. \tag{5.73}$$

Then

$$\hat{\mu} = \mu + \frac{1}{N}\sum_{t=1}^{N}\zeta_t. \tag{5.74}$$

Due to the stationarity, the law of large numbers is true and therefore $\sum_{t=1}^{N} \zeta_t/N \approx 0$. Hence, $\hat{\mu}$ is a consistent unbiased estimator of μ with variance (5.56). We can thus define μ by (5.73) and introduce the average percentage change and decline indicators using the old formulas (5.5) and (5.6), respectively. The corresponding estimators (5.27) and (5.28), their variance, confidence regions, etc. remain to be computed as earlier.

Example 1.7.2: Suppose we have a reason to believe that concentrations obtained from 10 years of weekly data follow model (5.67) with $\alpha = 0.03$, $\beta = -0.06$, and ζ as a stationary $AR(2)$ process with parameters $p_1 = 0.30$, $p_2 = -0.20$ and $\sigma = 0.6$. We want to compute the average percentage decline pd, the bias of its estimator, and the variance of $\hat{pd}$. In consequence of (5.56),

$$E\hat{pd} \approx 100(1 - \exp\{-(\alpha + \beta P/2) + \sigma^2/(1 - \rho_1 - \rho_2)^2/N/2\} \tag{5.75}$$

$$\begin{aligned} var(\hat{pd}) &= 10000 \exp\{-2(\alpha + \beta P/2) + \sigma^2/(1 - \rho_1 - \rho_2)^2/N\} \\ &\quad \times (\exp\{\sigma^2/(1 - \rho_1 - \rho_2)^2/N)\} - 1). \end{aligned} \tag{5.76}$$

The value of pd is −12.85, and the one calculated according to (5.75) with exact parameters is −12.75, i.e., the bias is 0.1%. The standard deviation obtained by substituting in (5.76) is 4.67%. The negative value of $\hat{pd}$ means the concentrations have grown on average over the 10 year period by 12.62%.

Inference in the presence of a linearly changing rate of decline follows somewhat different rules than in case of a constant rate. Though the null hypothesis $\mu = 0$ vs. $\mu \neq 0$ is of interest as earlier, it is worth realizing its meaning. Let $v_t = \alpha + \beta(t/W)$, $t = 0, \ldots, N$, and $N = PW$, be the trend of the process in (5.67). Then $\mu = \alpha + \beta P/2$ is the equilibrium point of the function v_t in the sense that $\mu - v_0 = v_N - \mu$, as can be shown by simple algebra. Hence, if the slope β is positive and $\mu = 0$, then the differences $\ln c_t - \ln c'_t$ started down, cross the zero level somewhere in the middle of the sample, and end above the zero level. That means the values c'_t have been declining and, by the end of the observation period, they are on average smaller than the c_t values. In other words, we have evidence of a decline in concentrations. On the contrary, if $\beta < 0$ and $\mu = 0$, the same argument provides evidence of a growth in concentrations. This clearly differs from the situation $\beta = 0$! If $\beta \neq 0$, then the failure of the test to reject $\mu = 0$ does not necessarily mean absence of a change!

5.7.3 Inference for Models with Linear Rate of Change

The parameter $\beta \neq 0$ in model (5.67) complicates the analysis because it must be estimated and tested for significance. So far, tests about μ could be based on the law of large numbers and the central limit theorem. In the presence of the linear trend, the residuals in (5.67) have to be Normal. Then it is possible to estimate β by the least squares method, test if $\beta \neq 0$ using standard regression arguments (Draper and Smith 1981, Chapter 1) and investigate the presence of autocorrelation to infer about μ.

It is also possible to fit β simultaneously with autoregressive parameters. For example, let ζ_t be a Normal $AR(1)$ process and $y_t = \ln c_t - \ln c'_t$ be the differences in (5.67). Then in consequence of (5.57) and (5.67),

$$y_t - \rho y_{t-1} = \alpha(1-\rho) + (\beta/W)(t - \rho(t-1)) + \sigma\varepsilon_t, \tag{5.77}$$

where ε_t are Normal and zero mean and variance one. This is not a linear model, however. For large W, $t/W \approx (t-1)/W$; hence, setting $\alpha' = \alpha(1-\rho)$ and $\beta' = \beta(1-\rho)$ yields a similar model

$$y_t = \alpha' + \beta'(t/W) + \rho y_{t-1} + \sigma\varepsilon_t, \tag{5.78}$$

and inference about α', β' and ρ can be carried out by standard linear methods. Differentiation of the process y_t leads to the relation

$$y'_t = \beta + \zeta'_t,$$

where $y'_t = y_t - y_{t-1}$, and $\zeta'_t = \zeta_t - \zeta_{t-1}$ satisfies the equation

$$\zeta'_t - \rho\zeta'_{t-1} = \sigma(\varepsilon_t - \varepsilon_{t-1}).$$

The process formed by variables ζ'_t is not stationary, however. To get an idea why, let us consider the solution of the difference equation (5.57) with initial condition ζ_0 representing a Normal random variable with zero mean and variance

$$var(\zeta_0) = \frac{\sigma^2}{1-\rho^2}.$$

The solution of (5.57) has the form

$$\zeta_t = \rho^t\zeta_0 + \sigma\sum_{k=1}^{t}\rho^{t-k}\varepsilon_k$$

for $t = 1, \ldots, N$, and therefore,

$$\zeta'_t = \zeta_t - \zeta_{t-1} = \rho^{t-1}(\rho-1)\zeta_0 + \sigma\sum_{k=2}^{t}\rho^{t-k}(\varepsilon_k - \varepsilon_{k-1}) + \rho^{t-1}\sigma\varepsilon_1$$

for $t = 2,\ldots, N$. Since ζ_0 is independent of each ε_t, and ρ^{t-2} $(\varepsilon_2 - \varepsilon_1)$ correlates with ρ^{t-1} ε_1,

$$
\begin{aligned}
var(\zeta'_t) &= \frac{\sigma^2}{1-\rho^2}\rho^{2(t-1)}(1-\rho)^2 + 2\sigma^2\frac{1-\rho^{2(t-1)}}{1-\rho^2} + \sigma^2\rho^{2(t-1)} - 2\rho^{t-1}\rho^{t-2}\sigma^2 \\
&= \frac{2\sigma^2}{1-\rho^2} + \frac{\sigma^2}{1-\rho^2}\rho^{2(t-1)}[(1-\rho)^2 - 2 + (1-\rho^2) - 2\rho^{-1}(1-\rho^2)] \\
&= \frac{2\sigma^2}{1-\rho^2} - \frac{2\sigma^2}{1-\rho^2}\rho^{2t-3}.
\end{aligned}
$$

The process ζ'_t is thus not stationary because its variance is time dependent. The term ρ^{2t-3} vanishes quickly with growing t; hence, the differentiation doubles variance of the noise as one would expect.

A detailed discussion of regression with autocorrelated errors and mixed autoregressive systems is available in Kendall and Stuart (1971) (Volume 3, Chapter 51, Sections 1 to 7).

5.7.4 The Absolute Percentage Change and Decline

Another possibility when assessing long-term change is to rely on the long-term trend. Let us suppose that the trend is a quadratic function plus a periodic component, $m_t = s + rt + ut^2 + \pi_t$, and we observed at nodes $t_n = n/W$, $n = 0,\ldots, N$, $N = 2WP$. If c_n is described by (5.7) and we denote $\zeta = \eta_N - \eta_0$, then the absolute percentage change from the first to the last day of observation is due to (5.71) and (5.73)

$$
\begin{aligned}
100\frac{c_N - c_0}{c_0} &= 100(\exp\{m_{t_N} - m_{t_0} + \eta_{t_N} - \eta_{t_0}\} - 1) \\
&= 100(\exp\{rt_N + ut_N^2 + \zeta\} - 1) = 100(\exp\{2rP + 4uP^2 + \zeta\} - 1) \\
&= 100(\exp\{-2\alpha - \beta P + \zeta\} - 1) = 100(\exp\{-2\mu + \zeta\} - 1).
\end{aligned}
$$

It characterizes the percentage change from the beginning to the end of the observation period assuming the model (5.7) has a quadratic trend. It is tempting to forget about ζ and introduce the *absolute percentage change* and *absolute percentage decline* as

$$apc = 100(\exp\{-2\mu\} - 1) \tag{5.79}$$

and

$$apc = 100(1 - \exp\{-2\mu\}), \tag{5.80}$$

respectively. The quantities characterize the long-term change based on the values of the trend on the first and last days of observation. Since μ has the same meaning as in the earlier sections, indicators *apc* and *apd* can be estimated and their statistical features assessed the same way as those of *pc* and *pd*, respectively. It should be realized, however, that the assumption about the quadratic shape of the long-term trend is essential and its violation may lead to serious over- or underestimation of the long-term change. Due to the problems with the estimation of β outlined in Section 5.7.3, this concept is abandoned.

5.7.5 A Model with Time-Centered Scale

A noticeable simplification can be achieved by centering the time scale, which leads to the model

$$\ln c_t - \ln c_t' = \alpha + \beta(t/W - P/2) + \zeta_t. \tag{5.81}$$

This modification of (5.67) causes the expected value of $\hat{\mu}$, described by (5.14) and computed from observations c_t, c_t' with $t = 1,\ldots, N$, $N = WP - 1$, equal directly to α, [see (5.11)]. Consequently, under the model (5.11), relations (5.72) and (5.73) lead to $\mu = \alpha$.

Model (5.81) admits a natural generalization of the annual rate of decline because if the concentration measurements follow a model with a trend described by (5.68) in the first period and by (5.69) in the second period, then for each n,

$$v_{t_n} = m_{t_n} - m'_{t_n} = -rP - 2uP(t_n + P/2).$$

Centering the time scale means to replace t_n with $t_n - P/2$, which leads to

$$v_{t_n - P/2} = -rP - 2uPt_n$$

for all n. Comparison of the last equation to (5.67) yields

$$\alpha = -rP \quad \text{and} \quad \beta = -2uP \tag{5.82}$$

and since $\mu = \alpha = -rP$, the annual rate of decline can be introduced as earlier in (5.10).

Centering the time axis simplifies the inference about the rate of decline substantially because $\hat{\mu}$ can always be interpreted as the estimate of μ obtained by means of the model (5.81). It has to be emphasized that centering of the time scale has no impact on the sign of β or its statistical significance.

There is a question whether it is sensible to characterize the change of concentrations the way we did it until now if more complicated trends occur in the data. Even if the simplest model (5.11) is true, the decline in the pollutant concentrations will stop one day, for certain amounts are in the air through very natural causes. Hence, rather than develop more elaborate models, it may be sensible to test shorter periods of data, and once the decline over short periods stops and the detected amounts are sufficiently low, the next goal will be to prevent them from rising again.

5.8 SPATIAL CHARACTERISTICS OF LONG-TERM CONCENTRATION CHANGES

Besides changes at individual stations, it is desirable to have an indicator of change over the whole monitored region. For that purpose, Section 5.4.3 presented overall percentage changes calculated from rates obtained by a flat averaging over all stations. No statistical features of these indicators have been discussed due to a concern about high correlation of data from different stations, a problem that can be justified by Figure 5.10. The aim of the next section is to fill this gap and present a method for assessment of the overall change. Section 5.4.3 has also indicated that the percentage decline of some chemical species monitored by CASTNet stations might be bigger in northeast than in southwest U.S. This section shows how to quantify this kind of differences and test if they are significant. To do so, the estimated CASTNet rates of change are described by means of a weighted linear regression model with latitude and longitude as explanatory variables. This model is subsequently used to get simultaneous 95% confidence intervals for the true rates of change over the whole region.

5.8.1 The Spatial Model for Rates of Change

Correlation between measurements collected in time and space is widely anticipated (Whittle 1962; Cressie 1993) but rarely commented on in the CASTNet data context. With regard to the effort we expended analyzing the autocorrelation of CASTNet data from individual stations, it is desirable to build a spatial model accommodating the old results and, in addition, the correlation between the stations. Logarithms of Alhambra, Bondville, Vincennes, and Salamonie Reservoir teflon filter SO_4 plotted against each other in Figure 5.10 signal that the correlation has to be addressed. Differences (5.67) used for calculation of the rate of change inherit the correlation.

To introduce the overall percentage change followed by data from several locations, we denote $c_{t,z}$ and $c'_{t,z}$ concentrations of a chemical species observed at time t and a period P years later, respectively, at a station with coordinates $z = (z_1, z_2)$, where z_1 denotes the longitude and z_2 the latitude. The concentrations are assumed to satisfy the spatio-temporal model

$$\ln c_{t,z} - \ln c'_{t,z} = \alpha + \beta(t/W - P/2) + \theta_1(z_1 - \bar{z}_1) + \theta_2(z_2 - \bar{z}_2) + \zeta_{t,z}, \qquad (5.83)$$

where $\zeta = \{\zeta_{t,z}: -\infty < t < \infty, z \in \boldsymbol{Z}\}$ is a multivariate stationary process with zero mean.

Setting

$$\mu_z = \alpha + \theta_1(z_1 - \bar{z}_1) + \theta_2(z_2 - \bar{z}_2)$$

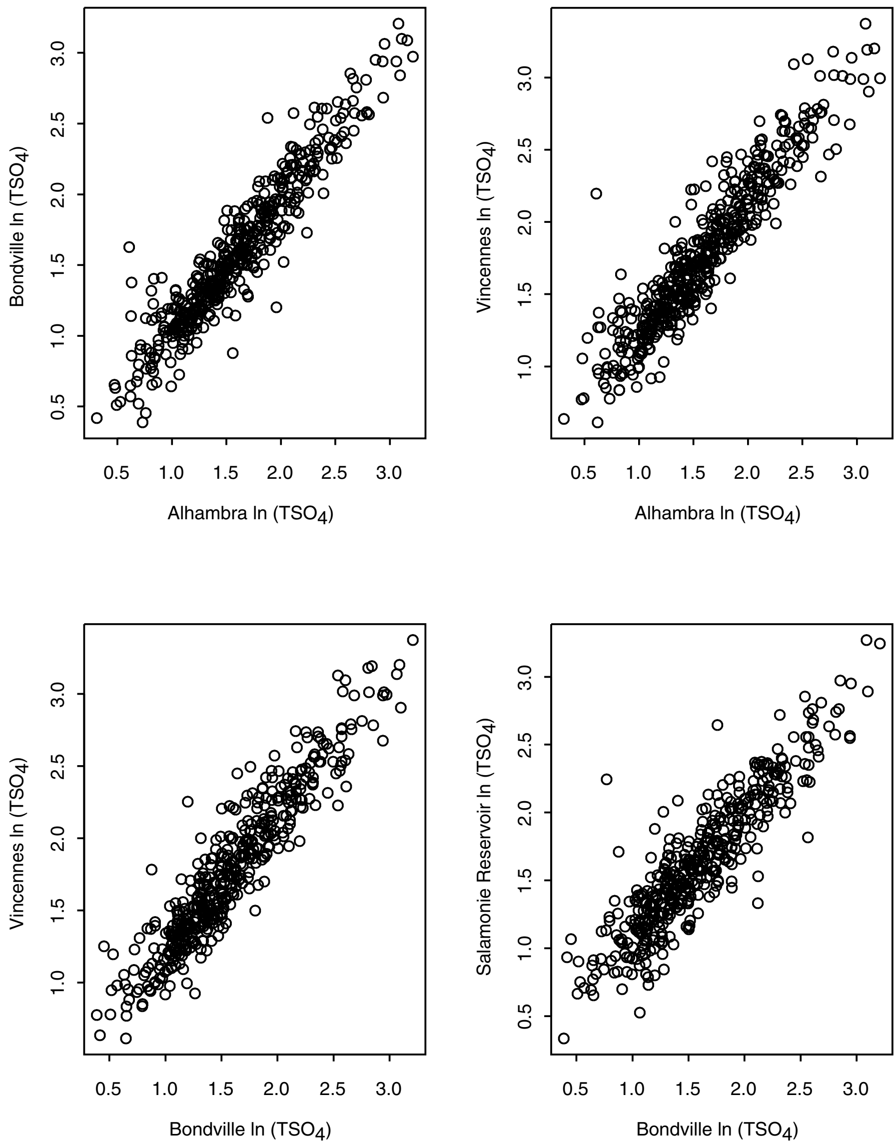


FIGURE 5.10 Correlation between some CASTNet stations is high.

allows us to rewrite (5.83) in the form

$$\ln c_{t,z} - \ln c'_{t,z} = \mu_z + \beta(t/W - P/2) + \zeta_{t,z}, \tag{5.84}$$

which is model (5.81). The spatio-temporal model (5.84) is thus in no conflict with the models used previously for description of the time series.

The estimator $\hat{\mu}_z$ of the local rate of decline μ_z calculated from data collected at the stations with coordinates $z \in \boldsymbol{Z}$ in weeks $t = 1,\ldots, N,\ N = WP - 1$, satisfies in consequence of (5.83) the relation

$$\hat{\mu}_z = \alpha + \theta_1(z_1 - \bar{z}_1) + \theta_2(z_2 - \bar{z}_2) + \zeta_{\cdot,z} \tag{5.85}$$

where $\zeta_{\cdot,z} = \sum_{t=1}^{N} \zeta_{t,z}/N$. The last equation will be referred to as the spatial model.

Before further investigating the spatial model for estimated rates of declines, let us recall that the reason for investigation of relations between the stations and introduction of the spatio-temporal model is assessment of the rate of decline over the whole region. Due to the presence of the averages $\bar{z}_1$ and $\bar{z}_2$ in (5.85), if M is the number of locations, then

$$\frac{1}{M}\sum_{z\in Z}\hat{\mu}_z = \alpha + \frac{1}{M}\sum_{z\in Z}\zeta_{t,z}.$$

Consequently, α represents the overall average discussed in Section 5.4.3. We can also say it represents the overall rate of decline over the region. To describe the percentage change or decline over the period, we thus use formulas (5.5) and (5.6) with μ replaced by α from model (5.85). Estimation of the rate α and probability distribution of the α estimator are discussed in the following sections. The distribution is necessary for construction of the confidence regions covering the true rate and regions covering the true overall long-term percentage decline.

5.8.2 Covariance Structure of the Spatial Model

Fitting of the model (5.83) requires an explicit description of the probability distribution of the process ζ, which is not easy to obtain even for one station should all diagnostic procedures be conducted. The estimation problem is therefore solved in two steps. We use procedures described in Section 5.6.2 to derive the best-fitting model (5.84) and, consequently, the estimate of the variance $\hat{var}(\hat{\mu}_z)$. As the second step, we realize that according to (5.74) and the CLT, which holds due to the stationarity and (5.13), μ is approximately Normal with mean μ_z and variance $\hat{var}(\hat{\mu}_z)$. Due to the normality of $\hat{\mu}_z$, the model will be fully determined by specifying the covariance matrix of the process $\zeta_{\cdot,z}$. Let us recall that the Normal distribution of μ does not require Normal distribution of the time series data. Next we describe the covariance function of $\zeta_{\cdot,z}$. To do so, some basic terminology from multivariate time series theory is needed. A detailed introduction to the topic is available in Brockwell and Davis (1987). Utilization of time series for spatio-temporal modeling is also described in Bennet (1979).

Let us assume that for each fixed location z, ζ is a stationary process with covariance function $R_z(h)$ that satisfies the condition (5.3) assuring existence of the spectral density $f_z(\lambda)$. In addition, we must assume that ζ is stationary in the sense that the covariance of observations from two different locations, collected at two different times t and $t + h$, depends only on the delay h between the two sampling times.

Formally, this is expressed by the *cross-covariance* function $R_{z,z'}(h)$ defined by the relation

$$cov(\zeta_{t,z,}\zeta_{t+h,z'}) = R_{z,z'}(h). \tag{5.86}$$

If the cross-covariance satisfies the condition

$$\sum_{h=-\infty}^{\infty} | R_{z,z'}(h) | < \infty, \tag{5.87}$$

then $R_{z,z'}$ determines the cross-spectral density $f_{z,z'}(\lambda)$. If we set for a moment $\zeta_t = (\zeta_{t,z}, \zeta_{t,z'})^T$, $\hat{\mu} = (\hat{\mu}_{\hat{z}}, \hat{\mu}_{z'})^T$, and $\mu = (\mu_z, \mu_{z'})^T$, then

$$\hat{\mu} = \mu + \frac{1}{N}\sum_{t=1}^{N}\zeta_t \sim \mathcal{N}\left(\mu, var\left(\frac{1}{N}\sum_{t=1}^{N}\zeta_t\right)\right), \tag{5.88}$$

where

$$var\left(\frac{1}{N}\sum_{t=1}^{N}\zeta_t\right) \approx \frac{2\pi}{N}\begin{pmatrix} f_z(0) & f_{z,z'}(0) \\ f_{z,z'}(0) & f_z(0) \end{pmatrix}. \tag{5.89}$$

The proof follows from Brockwell and Davis (1987) (Chapter 11, Proposition 11.2.2). The asymptotic covariance between the two estimators is thus described by $f_{z,z'}(0)$ and the asymptotic correlation is

$$cor(\hat{\mu}_z, \hat{\mu}_{z'}) \approx \frac{f_{z,z'}(0)}{\sqrt{f_z(0) f_{z'}(0)}}. \tag{5.90}$$

The right-hand side of (5.90) with 0 replaced by λ is also known as the *coherency* function. It is now clear that stationarity of ζ and the assumptions (5.13) and (5.87), satisfied for each pair of locations, assure that for sufficiently long periods of observations the distribution of (5.85) agrees with the asymptotic distribution of the random field $\{\hat{\mu}_z: z \in \boldsymbol{Z}\}$.

5.8.3 Multivariate *ARMA*(*p*,*q*) Models

Formally, a *multivariate ARMA*(p,q) model differs little from the univariate case described by (5.53) and (5.54). The model is composed from a multivariate autoregressive and a multivariate moving-average model denoted again $AR(p)$ and $MA(q)$, respectively. Let us assume for a moment that the stationary process $\zeta = \{\zeta_{t,z}: -\infty < t < \infty, z \in Z\}$ is $ARMA(p,q)$ and describes M stations. Then ζ_t and η_t in (5.53) and (5.54) are M dimensional column vectors; $\rho_1,\ldots, \rho_p, \theta_1\ldots, \theta_q$ are $M \times M$ rectangular matrices, perhaps not of full rank; and ε_t are mutually independent M dimensional column vectors of zero-mean random variables with the same covariance matrix $E\varepsilon_t\varepsilon_t^T = C$. In that case, the spectral density function admits the description

$$f(\lambda) = \frac{1}{2\pi}\Phi^{-1}(e^{i\lambda})\Theta(e^{i\lambda})C\Theta^T(e^{-i\lambda})\Phi^{-1T}(e^{-i\lambda}), \tag{5.91}$$

where

$$\Phi(e^{i\lambda}) = 1 - e^{-i\lambda}\rho_1 - \cdots - e^{-ip\lambda}\rho_p \tag{5.92}$$

$$\Theta(e^{i\lambda}) = 1 + e^{-i\lambda}\theta_1 - \cdots - e^{-iq\lambda}\theta_q, \tag{5.93}$$

and 1 is the $M \times M$ identity matrix. Hence, the process ζ has covariance matrix

$$var(\hat{\mu}) \approx \frac{1}{N}\Phi^{-1}(1)\Theta(1)C\Theta^T(1)\Phi^{-1T}(1). \tag{5.94}$$

Example 1.8.3: Let us consider concentrations from two stations obeying model (5.85) with $\theta_1 = \theta_2 = 0$ and ζ described by a bivariate $AR(1)$ model

$$\begin{pmatrix} \zeta_{t,1} \\ \zeta_{t,2} \end{pmatrix} = \begin{pmatrix} \rho_{1,1} & \rho_{1,2} \\ \rho_{2,1} & \rho_{2,2} \end{pmatrix}\begin{pmatrix} \zeta_{t-1,1} \\ \zeta_{t-1,2} \end{pmatrix} + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}\begin{pmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \end{pmatrix}, \tag{5.95}$$

where $\varepsilon_{t,1}$, $\varepsilon_{t,2}$ are sequences of mutually independent random variables with mean zero and variance one. We use integers 1 and 2 instead of z and z'. Then

$$\Phi(1) = \begin{pmatrix} 1-\rho_{1,1} & -\rho_{1,2} \\ -\rho_{2,1} & 1-\rho_{2,2} \end{pmatrix} \tag{5.96}$$

and

$$\Phi^{-1}(1) = \frac{1}{(1-\rho_{1,1})(1-\rho_{2,2}) - \rho_{1,2}\,\rho_{2,1}} = \begin{pmatrix} 1-\rho_{2,2} & \rho_{1,2} \\ \rho_{2,1} & 1-\rho_{1,1} \end{pmatrix}. \tag{5.97}$$

Consequently,

$$var(\hat{\mu}_1) \approx \frac{1}{N}\frac{\sigma_1^2(1-\rho_{2,2})^2 + \sigma_2^2\rho_{1,2}^2}{[(1-\rho_{1,1})(1-\rho_{2,2}) - \rho_{1,2}\rho_{2,1}]^2} \tag{5.98}$$

$$var(\hat{\mu}_1, \hat{\mu}_2) \approx \frac{1}{N}\frac{\sigma_1^2\rho_{2,1}(1-\rho_{2,2}) + \sigma_2^2\rho_{1,2}(1-\rho_{1,1})}{[(1-\rho_{1,1})(1-\rho_{2,2}) - \rho_{1,2}\rho_{2,1}]^2} \tag{5.99}$$

and

$$var(\hat{\mu}_2) \approx \frac{1}{N} \frac{\sigma_1^2 \rho_{2,1}^2 + \sigma_2^2 \rho_{1,2}(1-\rho_{1,1})^2}{[(1-\rho_{1,1})(1-\rho_{2,2}) - \rho_{1,2}\rho_{2,1}]^2}. \tag{5.100}$$

As we can see, the complexity of the covariance structure between two $AR(1)$ processes is high. Stationarity of the process requires that all eigenvalues of $\Phi(x)$ are in absolute value smaller than one, i.e., $det\ (\Phi(x)) \neq 0$ for all complex x from the unit circle $\{x: |x| \leq 1\}$.

Example 2.8.3: Let $\hat{\mu}_1$ and $\hat{\mu}_2$ be estimated from data observed at two stations and the data be generated from the model in Example 1.8.3. Suppose we estimate the overall rate of change α in the region based on the data from the two stations as the simple average $\hat{\alpha} = (\hat{\mu}_1 + \hat{\mu}_2)/2$. Then

$$var(\hat{\alpha}) = \frac{\sigma_1^2(1-\rho_{2,2}+\rho_{2,1})^2 + \sigma_2^2(1-\rho_{1,1}+\rho_{1,2})^2}{4N[(1-\rho_{1,1})(1-\rho_{2,2}) - \rho_{1,2}\,\rho_{2,1}]^2}, \tag{5.101}$$

$$Ep\hat{d} = 100\left(1-\exp\left\{-\alpha + \frac{1}{2}var(\hat{\alpha})\right\}\right), \tag{5.102}$$

and

$$var(p\hat{d}) = 10000\exp\{-2\alpha + var(\hat{\alpha})\}(\exp\{var(\hat{\alpha})\}-1). \tag{5.103}$$

5.8.4 Identification of the Spatial Model

Fitting the multivariate $ARMA(p,q)$ model determining (5.85) to data from all stations is difficult because, in theory, the presence of all parameters in the model should be justified by proper tests. Due to the asymptotic normality, we need only the cross-spectrum of pairs; hence, building simple models for pairs of stations should be sufficient. Since different pairs of data sets may lead to different models for observations from one and the same station, the desired covariance between two rates is obtained as

$$c\hat{o}v(\hat{\mu}_z, \hat{\mu}_{z'}) = \sqrt{v\hat{a}r(\hat{\mu}_z)v\hat{a}r(\hat{\mu}_{z'})}\,c\hat{o}r(\hat{\mu}_z, \hat{\mu}_{z'}), \tag{5.104}$$

where $v\hat{a}r(\hat{\mu}_z)$ and $v\hat{a}r(\hat{\mu}_{z'})$ are estimated using the best univariate models for stations with coordinates z and z', respectively, and the correlation $c\hat{o}r(\hat{\mu}_z)\ c\hat{o}r(\hat{\mu}_{z'})$ is estimated using the best bivariate model for the series from stations located at z and z'. Estimation of $v\hat{a}r(\hat{\mu}_z)$ is addressed in Section 5.6.1. Estimation of $c\hat{o}r(\hat{\mu}_z, \hat{\mu}_{z'})$ is done using the coherency function (5.90) and estimated parameters of the best model. Relation (5.104) assures that the variance of $\hat{\mu}_z$, estimated earlier, remains unchanged under the multivariate model. This approach might be called semi-parametric

(Angulo et al. 1998). As with most procedures of this kind, it does not guarantee that the resulting covariance matrix is positive definite, a condition satisfied by each covariance matrix. Identification of the bivariate model is substantially simplified if only the class $AR(p)$ is used, because the spectral density of a multivariate $ARMA(p,q)$ model with $p > 1$ and $q > 1$ does not uniquely determine all model parameters, (see Brockwell and Davis 1987, Section 11.6 for details).

5.8.5 Inference for the Spatial Data

Parameters of the model (5.85) can be estimated either by ordinary least squares (OLS) or by weighted least squares (WLS). The latter agrees with the maximum likelihood method. A classical reference on regression analysis methods is Draper and Smith (1981). Both techniques are most easily described using the matrix description of the model (5.85),

$$\boldsymbol{Y} = \boldsymbol{X}\theta + \sigma\boldsymbol{V}^{1/2}\varepsilon, \tag{5.105}$$

where $\boldsymbol{Y}$ is the M dimensional column vector of rates of change estimated from individual stations, $\boldsymbol{X}$ is an $M \times 3$ matrix of explanatory variables with rows of the form $(1, z_1 - \bar{z}_1, z_2 - \bar{z}_2)$, σ is a positive scaling parameter, and $\boldsymbol{V}^{1/2}$ is an auxiliary symmetric regular $M \times M$ matrix with the property $\boldsymbol{V}^{1/2}\,\boldsymbol{V}^{1/2} = \boldsymbol{V}$. The covariance matrix $\boldsymbol{V}$ has components

$$v_{z,z'} = c\hat{o}r(\hat{\mu}_z, \hat{\mu}_{z'}),$$

$\theta = (\alpha, \theta_1, \theta_2)^T$ is the column vector of unknown parameters and ε is an M dimensional column of mutually independent, zero-mean Normal random variables with variance one.

The OLS estimator is described by the relation

$$\hat{\theta}_{OLS} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}, \tag{5.106}$$

which has in consequence

$$\hat{\theta}_{OLS} = \theta + \sigma(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{1/2}\varepsilon. \tag{5.107}$$

The OLS estimator thus has variance

$$E(\hat{\theta}_{OLS} - \theta)(\hat{\theta}_{OLS} - \theta)^T = \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}. \tag{5.108}$$

To estimate σ^2 we calculate first the variance of the residuals:

$$\boldsymbol{r} = \boldsymbol{Y} - \boldsymbol{X}\hat{\theta}_{OLS} = \boldsymbol{Y} - \boldsymbol{X}\theta - \sigma\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{1/2}\varepsilon = \sigma\boldsymbol{M}\boldsymbol{V}^{1/2}\varepsilon, \tag{5.109}$$

where

$$M = I - X(X^T X)^{-1} X^T. \tag{5.110}$$

The variance is computed according to the formula

$$var(r) = Err^T = \sigma^2 MVM. \tag{5.111}$$

Since $r \sim \mathcal{N}(\mathbf{0}, \sigma^2 MVM)$, the quadratic form $r^T(MVM)^- r/\sigma^2$, where the minus sign denotes the generalized inverse of a matrix, has χ^2_{M-p} distribution with $M-p$ degrees of freedom. The quantity

$$\hat{\sigma}^2_{OLS} = \frac{1}{M-p}(Y - X\hat{\theta}_{OLS})^T (MVM)^- (Y - X\hat{\theta}_{OLS}) \tag{5.112}$$

is thus an unbiased estimator of σ^2.

The WLS estimator is calculated as

$$\hat{\theta}_{WLS} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y \tag{5.113}$$

and therefore,

$$\hat{\theta}_{WLS} = \theta + \sigma(X^T V^{-1} X)^{-1} X^T V^{-1/2} \varepsilon. \tag{5.114}$$

The variance of the estimator is

$$E(\hat{\theta}_{WLS} - \theta)(\hat{\theta}_{WLS} - \theta)^T = \sigma^2 (X^T V^{-1} X)^{-1}. \tag{5.115}$$

The parameter σ^2 is estimated using the residual sum of squares:

$$\hat{\sigma}^2_{WLS} = \frac{1}{M-p}(Y - X\hat{\theta}_{WLS})^T V^{-1} (Y - X\hat{\theta}_{WLS}). \tag{5.116}$$

It is an unbiased estimator and $\hat{\sigma}^2_{WLS}/\sigma^2 \sim \chi^2_{M-p}$.

5.8.6 Application of the Spatial Model to CASTNet Data

Model (5.85), fitted to the 40 CASTNet sites listed in Table 5.11 with a full 10 years of data using the WLS method, indicated a reasonable fit as to randomness of the residuals and Normal distribution of the data (see Figure 5.11). The ACF plot is not particularly informative in this situation because a plain trend removal has no effect

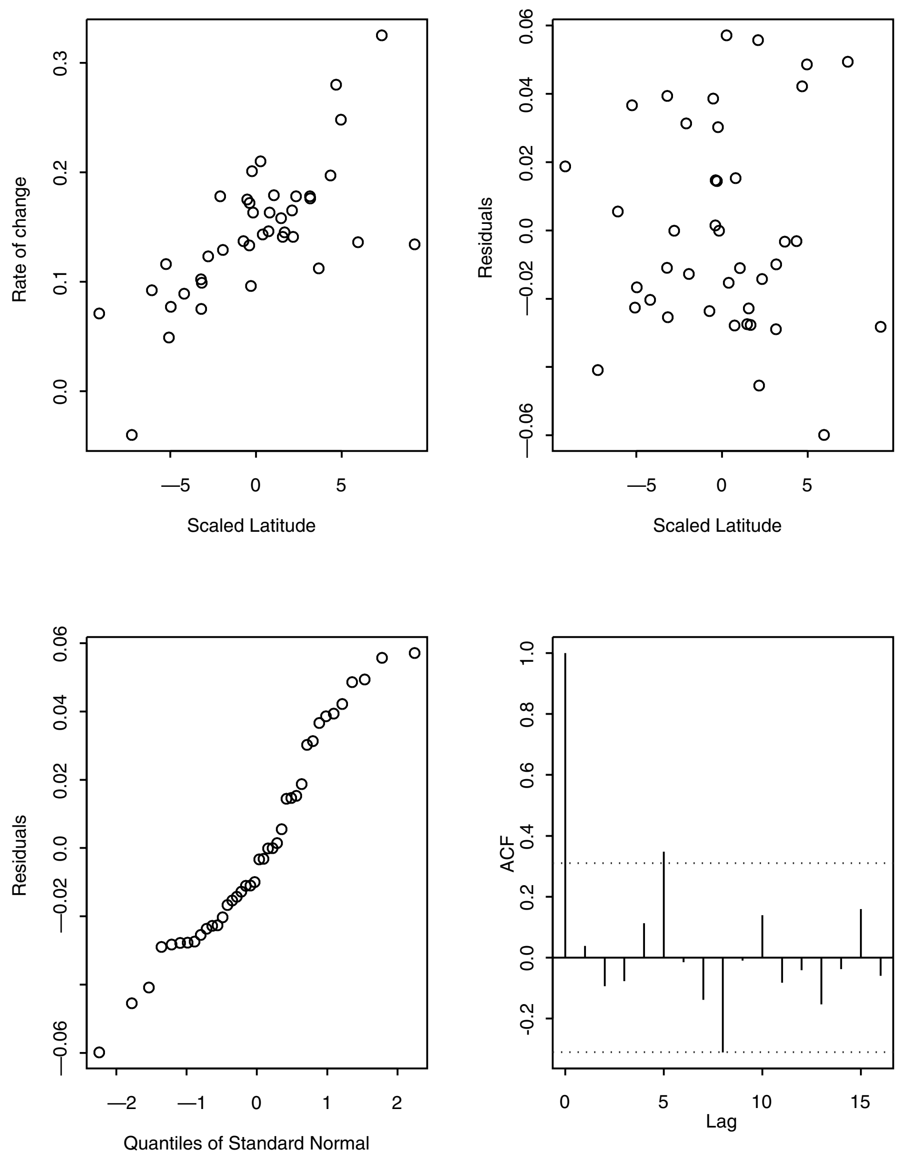


FIGURE 5.11 Residuals $Y - X\hat{\theta}_{WLS}$ of the model (5.85) fitted to 10 years of TSO_4 data sampled at 40 CASTNet locations listed in Table 5.11.

on autocorrelation of the data and the outcome depends on labeling of the data in our file. More sophisticated methods are needed to assess if the model removes correlation successfully (Cressie 1993). The significance of model coefficients, the evaluated given outcome of the *t*-test derived from (5.115) and (5.116), is summarized in Table 5.16.

TABLE 5.16
Significance of Explanatory Variables in Model (5.85) Fitted to Data Sampled at the 40 CASTNet Stations Listed in Table 5.11 Using the WLS Method

Variable	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Intercept	+	−	+	+	−	+	−
Longitude	+	−	+	+	−	−	−
Latitude	+	+	+	−	−	−	−

TABLE 5.17
Significance of Explanatory Variables in Model (5.85) Fitted to Data Sampled at the 17 CASTNet Stations Listed in Table 5.12 Using the WLS Method

Variable	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Intercept	+	−	−	+	−	+	−
Longitude	−	−	−	+	−	+	+
Latitude	−	+	+	−	−	−	−

TABLE 5.18
The Average Percentage Decline Over Ten Years for Species Monitored in the Air by 40 of the CASTNet Stations Listed in Table 5.11 Evaluated Using the Estimate of the Intercept α of the Model (5.85). Comparison to Figures 5.4 and 5.6 Indicates Overestimation for Some Species

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Percent	11	4	18	37	3	16	−7

The nonsignificant coefficients were omitted and the models were fitted again. Percentage declines obtained using α, as an overall, average rate of decline are listed in Table 5.18. The high values of TNH_4 and WNO_3 are rather annoying because Figure 5.4 and Figure 5.6 show that they grossly overestimate the sample averages. The WLS estimator is unbiased and has the smallest variance among all unbiased estimators. The suspicious estimates suggest that the samples in question could be generated by a different model. Hence, the OLS method was considered along with WLS.

Comparison of Table 5.16 and Table 5.17 with results in Table 5.20 and Table 5.21 shows that the WLS method is more sensitive to the presence of the explanatory variables than OLS. On the other hand, the OLS method seems better for calculation of the confidence regions for the overall averages. Fitting parameters of model (5.85) by the OLS produced α estimates and overall percentage decline estimates that

TABLE 5.19
The Average Percentage Decline Over 4 Years for Species Monitored in the Air by 17 of the CASTNet Stations Listed in Table 5.12 Evaluated Using the Estimate of the Intercept α of the Model (5.85). Comparison to Table 5.10 Indicates Overestimation for Some Species

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Percent	–7	2	0	33	5	–22	–10

TABLE 5.20
Significance of Explanatory Variables in Model (5.85) Fitted to the 40 Stations Listed in Table 5.11 with Full 10 Years of Observation Using the OLS Method

Variable	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Intercept	+	–	+	+	–	+	–
Longitude	–	–	+	–	–	–	–
Latitude	–	–	+	–	–	–	–

TABLE 5.21
Significance of Explanatory Variables in Model (5.85) Fitted to the 17 Stations Listed in Table 5.12 with Full 4 Years of Observation Using the OLS Method

Variable	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Intercept	–	–	–	+	–	+	–
Longitude	–	–	–	–	–	+	–
Latitude	–	–	–	–	–	–	–

agreed with those in Table 5.9. The significance of the model coefficients, assessed using a *t*-test derived from (5.108) and (5.112), is in Table 5.17. Interestingly, the growth of teflon-sampled SO_4 in the northwest direction is not considered significant. Both WLS and OLS agree as to the significance of the intercept though. The generalized inverse matrix was computed using the singular value decomposition (Press et al. 1986, Section 2.9.) Properties of generalized inverse matrices are studied in Rao (1973, Section 1b.5).

A similar procedure was applied to the 17 CASTNet Stations with a full 4 years of monitoring listed in Table 5.12. Results analogous to those derived for the 10-year period data are in Table 5.20, Table 5.21, Table 5.19, and Table 5.23, respectively.

TABLE 5.22
95% Confidence Regions for the Overall Percentage Declines in Table 5.9 Derived from the OLS Estimator

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Lower bound	5	–16	8	30	–11	3	–40
Upper bound	21	14	14	49	12	30	–2

TABLE 5.23
95% Confidence Regions for the Overall Percentage Declines in Table 5.10 Derived from the OLS Estimator

	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WNO_3
Lower bound	–14	–11	–17	16	–6	–42	–40
Upper bound	5	11	5	60	15	–22	10

While Table 5.22 shows a significant decline of TNH_4, WSO_2, and TSO_4 over the decade of monitoring, Table 5.23, representing the 4-year data shows a significant increase in WSO_2 concentrations, and since the upper bound of the confidence region for TSO_4 and TNH_4 is also close to zero, TSO_4 and TNH_4 seem to be growing as well. In fact, because the confidence regions for the 4- and 10-year TSO_4 and TNH_4 data are disjoint, their overall means differ significantly. Though the results appear contradictory because all periods end the same year, an answer follows from Table 5.14 and Table 5.15. The tables contain the sign of the slope parameter β determining the linear trend of $\ln (c_t/c_t')$.

For example, the slope of TSO_4 concentrations is negative for all but five of the eastern U.S. stations, which indicates that the concentrations were growing in the second period. TSO_4 has thus risen, though not enough to wipe out the declines over the first period. Similarly, only 4 eastern stations had a positive slope for TNH_4 and 10 out of 46 stations had a positive slope for WSO_2. Concentrations of both chemicals were thus growing at most places in the last 5 years ending 1998. Notice that we do not take in account the significance of the trend, only the direction estimated by the least squares method.

Since the CASTNet data are used here as an application of the long-term percentage decline indicator on a serious data set rather than an exhausting CASTNet data study, we abandon any further investigation of CASTNet dry deposition data. The discrepancy between 4- and 10-year periods of observations suggests the changes have a certain temporal dynamics. More could be learned by examining shorter subsequent periods of data. Considering regions as in Holland et al. (1999) would also produce more specific results.

5.9 CASE STUDY: ASSESSMENT OF DRY CHEMISTRY CHANGES AT CAPMON SITES 1988–1997

CAPMoN is the Canadian counterpart of CASTNet, though with a more limited number of sites. CAPMoN analyzes air for the same chemicals as CASTNet except WNO_3, which is replaced by WSO_4, the cellulose filter sulfate. There are a number of differences in the sampling protocols of CASTNet and CAPMoN, the latter of which is described in Ro et al. (1997), but the daily data collection by CAPMoN is the most important one. It not only provides a larger amount of data but the relatively short collection time causes the data to have much higher variability. A typical plot of CAPMoN daily sulfate data is in Figure 5.13 shows that the model (5.11), which is central to the percentage change and decline indicator definition, is appropriate in the sense that it removes the seasonal trend. The problem of the next section is to specify the noise term in (5.11) to accommodate the seasonal changes in variability, clearly present in Figure 5.13 (right), and to see how this change affects estimation and inference about the indicators. Though time series theory has models dealing with changing variance (Engle 1982; Bollerslev 1987), those models are designed for forecasting purposes. Since the attention here concentrates on the inference problem, a different approach is chosen.

5.9.1 Extension of Change Indicators to Data with Time-Dependent Variance

From the mathematical point of view, no changes to the indicator definitions are really necessary. To simplify the notation and reading of these notes, however, model (5.11) is rewritten in the form

$$\ln c_t - \ln c_t' = \mu + v_t \zeta_t, \tag{5.117}$$

where v_t is a known, deterministic function of time with strictly positive values, and ζ_t is a stationary process like the one in Section 5.2.2. The v_t values are called the *weights*. The model (5.117) is not stationary any more unless $v_t = 1$ for all t. The average $\hat{\mu}$ remains a sensible estimator of μ, but its variability is now more difficult to determine. The *weighted average*, defined for observations $y_t = \ln c_t - \ln c_t'$, $t = 1,\ldots, N$, by the relation

$$\hat{\mu} = \frac{\sum_{t=1}^{N} y_t/v_t}{\sum_{t=1}^{N} 1/v_t} \tag{5.118}$$

is thus considered instead.

In consequence of this definition,

$$\hat{\mu} = \mu + q\frac{1}{N}\sum_{t=1}^{N}\zeta_t, \tag{5.119}$$

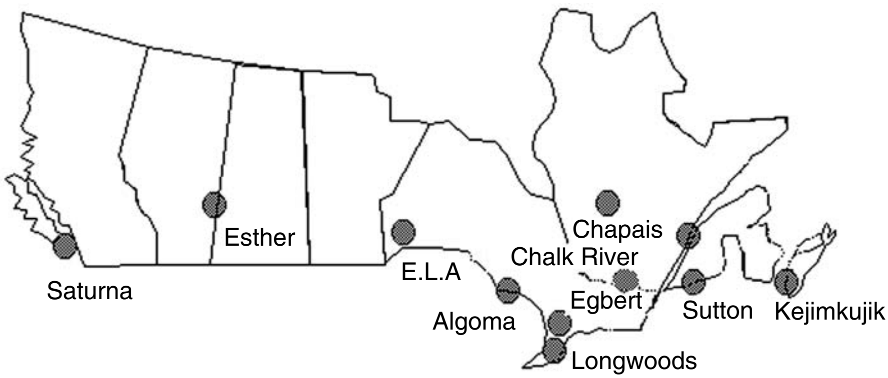


FIGURE 5.12 CAPMoN stations, Canada.

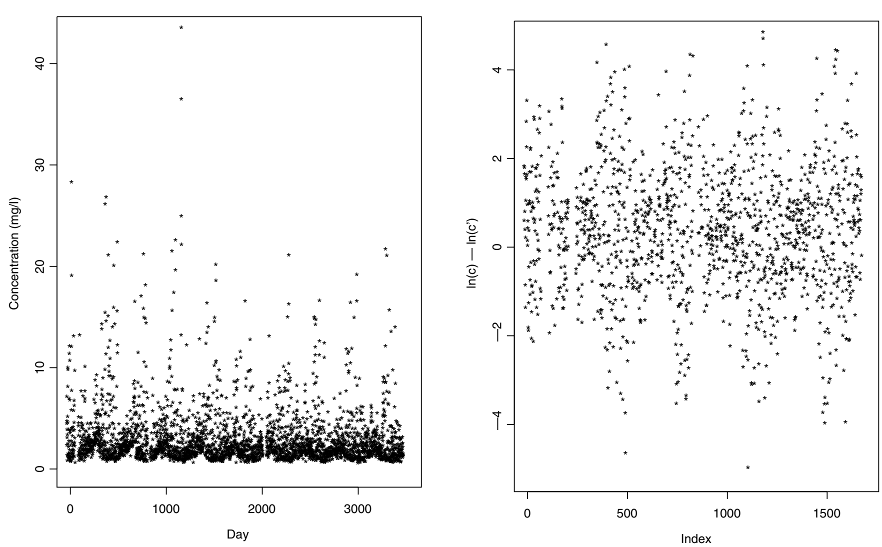


FIGURE 5.13 Observations of teflon filter SO_4 (mg/l), collected during 1988–1997 at CAPMoN station Algoma, Ontario, Canada, demonstrate how model (5.11) removes the trend but keeps seasonal changes in variability.

where

$$\frac{1}{q} = \frac{1}{N} \sum_{t=1}^{N} \frac{1}{v_t} \tag{5.120}$$

is a known constant. Variance of the average on the right of (5.119) has already been examined in Section 5.6.1. If $v_t = 1$ for all t, then the weighted average agrees with the ordinary one. It is thus acceptable to denote it $\hat{\mu}$ again because the definition of our indicators will not change and modifications to the interpretation in Section 5.4.2 are minor. Let us remember that now, for large samples,

$$var(\hat{\mu}) \approx \frac{q^2}{N} 2\pi f(0). \tag{5.121}$$

In particular, if ζ_t is an $AR(p)$ process, then

$$v\hat{a}r(\hat{\mu}) = \frac{q^2}{N} \frac{\hat{\sigma}^2}{(1 - \hat{\rho}_1 - \cdots - \hat{\rho}_p)^2} \tag{5.122}$$

is an acceptable variance estimator. The estimator μ and its variance determine the statistics Z_μ for test of significance of the decline and for computation of the confidence regions covering μ and pd, respectively (see Section 5.4.4). For example, the statistics for the test of the null hypothesis that the true rate of decline equals μ, adjusted for autocorrelation and time-dependent variance, has the form

$$Z_\mu = \frac{\sum_{t=1}^{N} (y_t/v_t)/N - \mu/q}{\sqrt{2\pi\hat{f}(0)}} \sqrt{N}. \tag{5.123}$$

5.9.2 Optimality Features of $\hat{\mu}$

Those familiar with weighted linear regression will find $\hat{\mu}$ in (5.118) annoying, because this is not the WLS estimator and, as demonstrated here, it indeed is not optimal in terms of variability and speed of convergence. It has a major advantage though. Its variance can be computed explicitly [see formula (5.122)]. The WLS estimator is also a weighted average, defined by the relation

$$\hat{\mu}_{WLS} = \frac{\sum_{t=1}^{N} y_t/v_t^2}{\sum_{t=1}^{N} 1/v_t^2} \tag{5.124}$$

and like in the case of the weighted average (5.122), its consistency is not straightforward and must be verified.

If q defined by (5.120) approaches a fixed value with growing sample size, then the consistency of μ follows from (5.121). A function $1/v_t$, for which q has such a property, is called *ergodic*. Introduction to the ergodic theory is provided in Loéve (1977, Section 34). Continuous periodic functions are ergodic. Hence, if we assume that v_t is positive, periodic, and continuous on the whole domain, which is reasonable given Figure 5.14 (left), then $1/v_t$ is also a continuous periodic function and the average in (5.120) converges to a finite positive value.

Next we show that $\hat{\mu}_{WLS}$ has smaller variance than $\hat{\mu}$. Let v_t be positive, periodic, and continuous on the whole domain. Then there is a constant K such that $0 < K \leq v_t$

for all t and therefore

$$var(\hat{\mu}_{WLS}) = \sigma^2 var\left(\frac{\sum_{t=1}^{N} y_t/v_t^2}{\sum_{t=1}^{N} 1/v_t^2}\right) \le \frac{\sigma^2}{K^2}\frac{\left(\sum_{t=1}^{N} 1/v_t\right)^2}{\left(\sum_{t=1}^{N} 1/v_t^2\right)^2} var\left(\frac{\sum_{t=1}^{N} y_t/v_t}{\sum_{t=1}^{N} 1/v_t}\right)$$
$$= \frac{\sigma^2}{K^2}\frac{\left(\sum_{t=1}^{N} 1/v_t\right)^2}{\left(\sum_{t=1}^{N} 1/v_t^2\right)^2} var(\hat{\mu}). \tag{5.125}$$

In consequence of the so-called Cauchy's inequality $\left(\sum_{t=1}^{N} 1/v_t\right)^2 \le N\sum_{t=1}^{N} 1/v_t^2$,

$$var(\hat{\mu}_{WLS}) \le \frac{\sigma^2}{K^2}\frac{N}{\sum_{t=1}^{N} 1/v_t^2} var(\hat{\mu}). \tag{5.126}$$

The last inequality means that $\hat{\mu}_{WLS}$ is a consistent estimator. The above relations also suggest that the WLS estimator is more efficient and approaches the true parameter faster than $\hat{\mu}$. For the concepts of unbiasedness, minimum variance, and efficiency, see Lehmann (1983, Chapter 2) or Rao (1973, Chapter 5, Section 5a).

5.9.3 Estimation of the Weights

The next problem is to determine the function v_t. For shorter periods of time, v_t can be composed of sine and cosine waves (Mohapl 2000b). This strategy has not worked for this particular CAPMoN data set, however. To get a more specific idea about the variance changes, the estimator $\hat{v}_t$ of v_t is calculated as the standard deviation of the subsample $y_{t-p},\ldots, y_t,\ldots, y_{t+p}$ for $2p+1 = 91$ days. Due to the lack of data at the beginning and the end of the sample, that is, for $t = 1,\ldots, p$ and $t = N - p + 1$, respectively, $\hat{v}_t$ is set equal to $\hat{v}_{p+1}$ and $\hat{v}_{N-p}$, respectively. We call $\hat{v}_t$ the moving standard deviation. The choice of a quarter of a year is strictly *ad hoc*. This period seems long enough for $\hat{v}_t$ to be nearly deterministic on the one hand and still flexible enough to yield a sensible result on the other. The moving standard deviation, estimated from the data shown in Figure 5.13 (left), is in Figure 5.14 (left). The pattern in Figure 5.14 (left) reflects the annual periodic changes in variability clearly visible in Figure 5.13 (left) very well. The annual waves are not entirely symmetric, though. Hence, instead of a deterministic function, the moving standard deviation estimate of v_t with 91-d time window was used directly in further calculations whenever variance stabilization seemed appropriate.

5.9.4 Application of the Nonstationary Model

CAPMoN stations considered in this study are in Figure 5.12. As in the CASTNet case, the daily sampling is not immune to missing observations. The percentages of pairs used for analysis are listed in Table 5.24. If sampling started in the middle of the season, for most stations in June, the corresponding part of the data in the second period was omitted. The percentages are thus not influenced by the start of the season as in the CASTNet case.

Once the pairs have been calculated, model (5.81) was fitted using the method presented in Section 5.6.2. If the analysis of residuals showed disagreement between

TABLE 5.24
Percentage of Pairs Used for Analysis from the Total Available Given the Duration of Monitoring

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WSO_4	Years
Algoma	86	37	85	71	84	75	69	10
Chalk River	83	40	81	68	81	74	69	10
Chapais	87	24	81	59	69	47	38	10
Egbert	94	88	94	85	94	91	89	10
E.L.A.	73	38	71	60	70	56	46	10
Esther	97	80	96	87	97	93	90	6
Kejimkujik	93	58	92	81	91	76	66	10
Longwoods	85	84	85	78	85	84	82	10
Saturna	95	92	93	89	93	94	93	6
Sutton	83	68	81	79	83	81	78	10

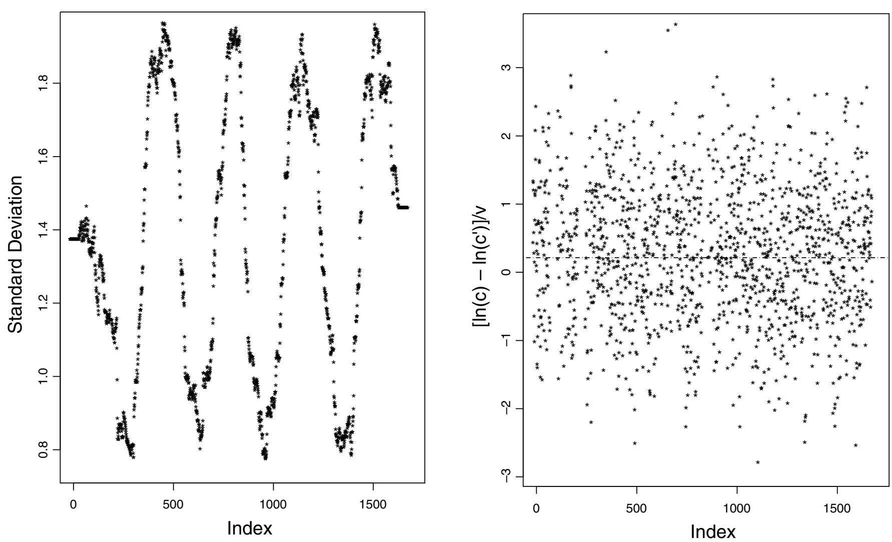


FIGURE 5.14 The moving standard deviation for data in Figure 5.13 (left) and the normalized quantities $[\ln(c_t) - \ln(c_t')]/v_t$. The dashed line on the right is the sample average.

the Normal distribution and the model residuals, the model (5.117) was evaluated again with weights estimated using the moving standard deviation. The resulting models are in Table 5.25. The model determines the value of the statistics Z_μ for testing if the percentage change of the concentrations over the period is significant. The declines and outcome of the tests are disclosed in Table 5.26.

As mentioned in Section 5.8.6, weighted averages may be over- or underestimated. Hence, to allow a fair comparison, Table 5.27 provides percentage declines computed without weighting, i.e., with $v_t = 1$ for all t and locations. The similarity between declines in WSO_2 and WSO_4 is rather remarkable.

TABLE 5.25
Description of Models Used for CAPMoN Data Analysis Is $\ln(c_t/c'_t) = \alpha + \beta t + v_t\zeta_t$, Where ζ Is an $AR(p)$ Process. Each Column Contains; Sign of β,* Is Significant, the Order p, k if Null Hypothesis Is Rejected by KS Test and c, if It Is Rejected by χ^2 Test. A Dot Means a Non-Significant Result. Symbol v Means the Weights Are Estimated Using Moving Standard Deviation, Otherwise $v_t = 1$, i.e., No Weighting Was Used

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WSO_4
Algoma	−*4..v	−*5..	−.4..	+*9..	−.4..	−.2..	+.1..v
Chalk River	−*2..v	−*10..	−.2..v	+.5k.v	−*2..	−*1..	−*1..
Chapais	−*6k.v	−*4..	−.2..v	+.1..	−.6..	−*1..	−*1..v
Egbert	−*2..v	−*6..	−*2..	+*12..	−*2..	−*2k.v	−*2..v
E.L.A.	+.2k.v	−*1..	+.2..v	+*10..	−.6kcv	−.1..	+.1..
Esther	+*2..	+*1..v	+*1k.v	−.10..	−.3..	+*1..	+*3k.v
Kejimkujik	−*2..v	−.1..	−.2..	+*11..v	−*2..	−.3..	−.4..
Longwoods	−.4..v	−*2..	−*4..	+*11k.v	+.2..	−*3..v	−*3kcv
Saturna	−.3..	−*1..	−*4k.v	−*12k.v	−*1..	−*4..	−*3..
Sutton	−*2..v	−*1..	−*2..	+7..v	−*1..	−*2..	−*2..

TABLE 5.26
Percentage Declines Estimated at CAPMoN Sites. The Asterisk Denotes a Significant Change Based on the Statistics Z_μ. Model for Z_μ Is in Table 5.25

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WSO_4	Years
Algoma	20*	−116*	2	45*	10	35*	35*	10
Chalk River	20*	−82*	6	54*	15*	34*	35*	10
Chapais	29*	−102*	9	54*	10	43*	43*	10
Egbert	22*	−8	13*	48*	16*	33*	34*	10
E.L.A.	14*	−17*	−1	44*	8	20*	18*	10
Esther	−5	−46*	−19*	11	−34*	−21*	−19*	6
Kejimkujik	22*	−69*	18*	54*	19*	42*	39*	10
Longwoods	17*	−8	2	49*	19*	26*	24*	10
Saturna	11*	−4	15*	35	17*	5	3	6
Sutton	25*	−27*	13*	53*	23*	35*	33*	10

The number of CAPMoN sites is too small for spatial analysis, but the confidence regions for the overall average percentage declines still require use of the model (5.85) with $\theta_1 = \theta_2 = 0$. The relatively high correlation, exhibited by raw CAPMoN samples plotted against each other, is not reflected by residuals of the model (5.117). The matrix $\boldsymbol{V}$ in the model (5.85) has thus zeros everywhere except variances of the $AR(p)$ models described in Table 5.25 on the main diagonal. Results obtained from model (5.85) represent the estimate of the parameter α, which is the overall mean and its 95% confidence regions. They are presented in Table 5.28.

TABLE 5.27
Percentage Declines Estimated at CAPMoN Sites Using $p\hat{d}$ in (5.27) and the Sample Average (5.14), That Is, No Weights Have Been Used

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WSO_4	Years
Algoma	19	−116	2	45	10	35	35	10
Chalk River	19	−82	5	55	15	34	35	10
Chapais	26	−102	6	54	10	43	43	10
Egbert	21	−8	13	48	16	33	34	10
E.L.A.	14	−71	−2	44	8	20	18	10
Esther	−5	−41	−19	11	−34	−21	−18	6
Kejimkujik	21	−69	18	54	19	42	39	10
Longwoods	17	−8	2	50	19	25	24	10
Saturna	11	−4	15	37	17	5	3	6
Sutton	25	−27	13	53	23	35	33	10

TABLE 5.28
Overall Percentage Declines at CAPMoN Sites with 10-Year Monitoring Period and the 95% Confidence Regions

Station	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2	WSO_4
Lower bound	17	−109	1	43	8	26	29
Percentage decline	20	−50	6	51	15	31	34
Upper bound	23	−8	10	58	22	36	38

Compared to CASTNet, the WLS method provides estimates that reasonably agree with OLS results. The use of WLS for derivation of the confidence regions is thus justified. The confidence regions derived from OLS are broader, which suggests more power of WLS for testing if, for example, a value prescribed by a policy equals the estimated value against the alternative that the policy value is different. The confidence regions in Table 5.22 and Table 5.28 intersect for all species. The overall averages from both networks could thus be just a random deviation from one and the same value particular to each species.

5.9.5 CAPMoN and CASTNet Comparison

When comparing Table 5.27 to the CASTNet tables, we should keep in mind that CAPMoN data cover a period shifted 1 year back. Because of a suspected trend reversal in the second period, CAPMoN declines might be somewhat higher. The possible growth in pollution for most species is indicated by the minus sign of β in the model description. CAPMoN percentage declines conform to our hypothesis that some species experienced a higher concentration drop in the northeast than in the southwest. The 6-year period data from Saturna and Esther must be compared to results obtained from western U.S.

TABLE 5.29
Except TNO_3, the Range of CAPMoN 10-Year Average Percentage Declines Contains the CASTNet Upper Bound for the Overall Average

Value	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2
CAPMoN Minimum	14	−116	−2	44	8	20
CASTNet Upper bound	21	14	14	49	12	30
CAPMoN Maximum	26	−8	18	55	23	43

TABLE 5.30
Data from Egbert. Lower Bound for CAPMoN Data Rate of Decline, the Rate Estimated from CASTNet Data and the Upper Bound for CAPMoN Rate of Decline

Value	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2
Lower bound	0.03735	−0.03574	0.00797	0.10392	0.01685	0.07386
Rate of decline	0.02041	0.02041	0.03094	0.47080	0.03629	−0.07421
Upper bound	0.06025	0.02363	0.04683	0.15408	0.05315	0.08894

One simple way to compare the average percentage declines at CASTNet and CAPMoN sites is to look at the range of CAPMoN values. Interestingly, except for teflon nitrate, the upper bounds for the CASTNet overall average percentage declines are in the relatively tight range of CAPMoN values. Since CAPMoN stations are further in the northeast of the CASTNet monitoring region, this result agrees with the trend observed in the CASTNet averages (see Table 5.29, Figure 5.4, and Figure 5.5).

A special investigation has to be made into results obtained from the Egbert station in Ontario. The station is used by CAPMoN and CASTNet for comparison purposes. Since CAPMoN samples daily and CASTNet weekly, the comparison studies are rather complicated (Mohapl 2000b). In addition, our Egbert samples are collected from two different periods of different lengths. The most natural procedure is to compare the annual rates of decline. As emphasized in Section 5.2.1, the percentage change resists network biases. The rates of decline estimated from the two networks at the Egbert site should be thus comparable. The estimator $\hat{\rho}$ of the annual rate of decline is described in (5.36). Since $\hat{\mu}$ is asymptotically Normal, so is $\hat{\rho}$. The use of a two-sample t-test for assessing the significance of the difference between the rate estimates might not be appropriate because of the potential correlation between the samples caused by the close distance of the sampling devices. Hence, for comparison we decided to use the confidence regions computed for CAPMoN annual rates of decline according to the formula (5.37). The main argument for choice of CAPMoN is the bulk of data available from the Egbert station compared to the rather restricted CASTNet sample. The confidence regions for the annual rate of decline for Egbert and the annual rate estimated from CASTNet data are shown in Table 5.30.

TABLE 5.31
Data from Pennsylvania State University Location. Lower Bound for the CASTNet Data Rate of Decline, the Rate Estimated from CAPMoN Data and the Upper Bound for the CASTNet Rate of Decline

Value	TSO_4	TNO_3	TNH_4	NSO_4	$NHNO_3$	WSO_2
Lower bound	0.02160	–0.04013	0.00978	0.06467	0.00073	0.00253
Rate of decline	–0.05940	–0.01140	–0.06292	–0.60874	–0.13338	–0.03017
Upper bound	0.04160	0.03093	0.03422	0.11013	0.03287	0.05147

The table shows three rate estimates outside their respective confidence regions. The low rate of teflon sulfate and cellulose filter sulfur declines can be justified by the length of the compared periods. It is the same effect observable in Table 5.22 and Table 5.23 computed from CASTNet 4- and 10-year data. Sulfate and sulfur concentrations simply increased the last 4 years ending 1998. The rate of the nylon filter sulfate is way out of line, however, even compared to the CASTNet overall average. In fact, the total decline of 61% is the highest in the network for this particular species. Since Egbert is used for network comparison and calibration, this result is certainly annoying.

Another site hosting a CAPMoN and CASTNet monitoring station is the Pennsylvania State University location. Data available from CAPMoN stretch over a relatively short period, from June 1988 to June 1990, barely enough to provide for a 2-year comparison. The data are collected daily, however, which still makes a reasonable sample size. The annual rate from the CAPMoN data and the 95% confidence regions, computed from 10 years of Pennsylvania State data of CASTNet observations, are shown in Table 5.31.

Before judging the numbers in Table 5.31 we must realize that the CASTNet data cover a period from 1989 to 1998, compared to the short CAPMoN period stretching over 1988 to 1989. We thus have results of CAPMoN operation at the early stage of the emission reduction process and CASTNet data covering 10 years after. The increases in concentrations measured over the first 2 years by CAPMoN are thus not that surprising. What is curious, however, is again a huge discrepancy in nylon sulfur annual rates. In contrast to Egbert, the long-term percentage decline at the Pennsylvania location measured by CASTNet is one of the lowest among CASTNet sites with 10 years of operation. This raises again the question of suitability of the Pennsylvania State location for comparison studies.

5.10 CASE STUDY: ASSESSMENT OF DRY CHEMISTRY CHANGES AT APIOS-D SITES DURING 1980–1993

The Acid Precipitation in Ontario Study — Daily Network (APIOS-D) is operated by the Air Resources Branch of the Ontario Ministry of Environment and Energy. The dry chemistry data discussed next were collected during 1980 to 1993. The ten locations considered here are shown in Figure 5.15. Five of them operated only 2 years.

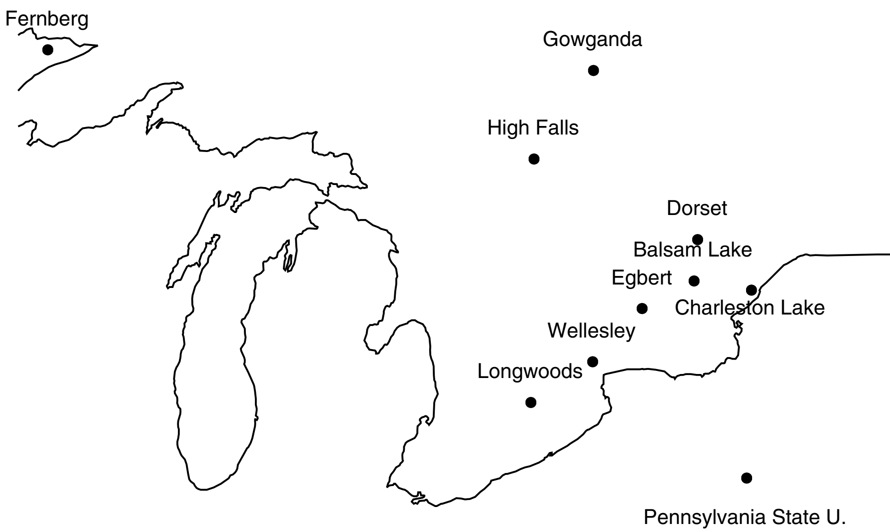


FIGURE 5.15 Locations of APIOS-D stations in Ontario, Canada. The lines denote the Canada-USA border and Great Lakes shore.

The chemical species monitored by APIOS-D and analyzed next can be interpreted as SO_4^- and NO_3^-. For details on sampling see the Environment Canada Internet pages.

5.10.1 APIOS-D Analysis

The periods of operation have been somewhat erratic, as shown in Table 5.32. Hence, the methods developed earlier were used to analyze all available data. Table 5.32 shows that even long periods of nitrate can be successfully analyzed by an autoregressive model, though the order of the model grows with the length of observations. The fit of the model to the sulfate data is somewhat worse, especially in the long run. The only significant decline is observed at the Charleston Lake location in sulfate concentrations. Otherwise, pollution was either unchanged or deteriorated with time.

Changes of SO_4, sampled by APIOS-D, agree with those reported (Husain et al. 1998, p. 968) for Mayville in Ohio over the period 1984 to 1991, and for Whiteface Mountain in New York State from 1981 to 1992. The lack of change in dry deposition data agrees with the lack of change observed in the precipitation SO_4 measurements collected at APIOS-D sites during 1980 to 1993 (Mohapl 2003a). No apparent temporal trends in precipitation sulfate and other monitored chemicals have been reported during 1985 to 1989 in the lower Ohio River Valley (Saylor et al. 1992) and during 1979 to 1990 in the Eastern Canada at CAPMoN sites (Sirois 1993). Since the APIOS-D stations were located along the prevailing flow direction of the air mass traveling from the Ohio River Valley, it seems that no apparent declines in air pollution happened in eastern North America during 1980 to 1991. This conclusion

TABLE 5.32
Analysis of the APIOS-D Data. Percentages of Pairs Used for Analysis from the Total Theoretically Available for the Period (% of Pairs). Model Used for Analysis—See Table 5.25 for Notation. Long-Term Percentage Decline Estimated Using $\hat{pd}$ in (5.27). Decline Estimated by Means of Weights as Indicated by Model Item (W% Decline) Including Significance Based on Z_μ. In the Last Column Is the Period of Monitoring

	% of Pairs		Model		% Decline		W% Decline		
Station	SO_4	NO_3	SO_4	NO_3	SO_4	NO_3	SO^-	NO_3	**Years**
Balsam Lake	83	83	+.2..v	+*1..	–19	10	–17	10	88–90
Charleston Lake	89	82	+.3..	–.9..	15	–22	15*	–22*	81–92
Dorset	96	63	+*4kcv	+*5..	4	–52	5	–52*	80–93
Egbert	94	94	+.10..	+.1..	4	2	4	2	88–90
Fernberg	56	33	–*9kcv	+.1..	–32	–52	–34*	–52*	81–92
Gowganda	98	93	+.1..v	+.8..	–2	–19	–7	–19	88–90
High Falls	96	87	–.2..	–.8k.v	–8	–11	–8	–11	86–90
Longwoods	95	93	+.2..v	+.7..	2	–47	5	–47*	81–92
Penn. State U.	90	89	+.1k.v	+*6..	–10	–27	–8	–27	88–90
Wellesley	91	91	–.4..	+.1..	–16	–53	–16*	–53*	88–90

contradicts the trend analysis results (Lynch et al. 1995, p. 1245) claiming evidence of clear declines of precipitation SO_4 in all regions of the U.S. over the period 1980 to 1992.

5.10.2 APIOS-D and CAPMoN Comparison

This section investigates how the changes reported by APIOS-D agree with information provided by CAPMoN stations. APIOS-D and CAPMoN had three stations operating jointly for short periods of time. Two of them, at Pennsylvania State University and at Egbert, Canada, have been running together for a litle less than 2 years. But since APIOS-D provides daily measurements, it is still possible to recover periods of reasonable length, summarized in Table 5.31, suitable for comparison. At Longwoods, the stations have been in joint operation about 4 years.

We use for comparison the spot change pc_ζ defined for each station and species by (5.4). Let us recall that in our notation, c represents data from Period 1, c' data from Period 2, and pc_ζ is invariant towards biases due to differences in sampling techniques (see Section 5.2.1). Our goal is to compare the rates of decline obeyed by data from the two networks. The daily spot percentage changes are not generated by Normal random variables. Hence, we transform them to Normal distribution using the function

$$\tau(pc_\zeta) = -\ln\left(1 + \frac{pc_\zeta}{100}\right). \tag{5.127}$$

TABLE 5.33
Comparison of APIOS-D and CASPMoN, Basic Information Summary. Ordinary *t*-test Results, Modified *t* Test Results and Periods of Joint Operation Used for Comparison. + Sign Denotes a Significant Result, – a Non-Significant

	Ordinary *t*		Modified *t*			
Station	SO_4^-	NO_3^-	SO_4^-	NO_3^-	**Period 1**	**Period 2**
Egbert	–	–	–	–	9/30/88–6/1/89	9/30/89–6/1/90
Wellesley	–	–	+	+	5/7/88–3/31/90	5/7/90–3/31/92
Penn. State U.	–	+	–	+	6/15/88–5/15/89	6/15/89–5/15/90

Due to the model (5.11) we get

$$\tau(pc_\zeta) = \mu + \zeta = \ln c - \ln c', \tag{5.128}$$

where μ is the rate of decline to be examined. The Normal distribution of $\tau(pc_\zeta)$ is a consequence of the Normal distribution of ζ studied in the earlier sections. Let us set $y = \ln c - \ln c'$ and denote $y_t^{(1)}$ and $y_t^{(2)}$ values of y measured by CAPMoN and APIOS-D, respectively, at a particular time t. Testing about the parameter μ, which under the zero hypothesis is the same for both series $y_t^{(1)}$ and $y_t^{(2)}$, can be done by the standard t-test. Correlation between the compared series is not a concern because the paired comparison is used. Results of a blunt t-test application are in the second and third columns of Table 5.33. The table shows that except for NO_3^- collected at the Pennsylvania location, the daily percentage declines calculated from CAPMoN and APIOS-D over the selected periods are on average the same.

The plain t-test has little credibility, however, unless the differences $y_t^{(1)} - y_t^{(2)}$ are generated by mutually independent, identically distributed Normal random variables. Differences $y_t^{(1)} - y_t^{(2)}$ calculated from SO_4^- data inherit seasonal changes in variability pointed out in Section 5.9 and must be stabilized using the moving standard deviations before the data enter the t-test. Ordinary linear regression showed that the differences $y_t^{(1)} - y_t^{(2)}$ lack trends and form after variance stabilization a series resembling data from a distribution with constant mean and variance. Numerous series of concentration data investigated earlier were autocorrelated. Autocorrelation of the second or third order was detected in most $y_t^{(1)} - y_t^{(2)}$ series, scaled or not. All data sets passed the Normal distribution goodness-of-fit test, which means that after adjusting for autocorrelation we obtain results somewhat different from those from the ordinary t-test (see the fourth and fifth column of Table 5.33). The statistics (5.123), which is accommodated to the unsteady variability and autocorrelation, seems more sensitive towards discrepancies between the data sets in this particular case. Even so, comparison of the two networks shows that the estimated rates of declines are fairly similar.

Another matter of interest might be the agreement between rates of decline of total sulfur SO_2 and nitrogen N computed from the CAPMoN data, and the APIOS-D measured SO_4^- and NO_3^- rates of decline, respectively. The same procedure as earlier, but with $y_t^{(1)}$ computed from SO_2 and N concentrations, respectively, showed no significant differences between the rates of decline. Only the Longwoods data required an adjustment for autocorrelation, otherwise the data showed no disagreement with the ordinary t-test assumptions.

5.11 CASE STUDY: ASSESSMENT OF PRECIPITATION CHEMISTRY CHANGES AT CASTNET SITES DURING 1989–1998

The percentage decline indicator and associated methods have so far been used in the context of dry air chemistry data. This section applies the indicator to wet chemistry observations gathered by CASTNet over the period 1989 to 1998. Precipitation concentration measurements come from samples logged at network sites by precipitation collectors. Like the filter packs, CASTNet's collectors are emptied weekly, tested for pH, and analyzed for concentrations of SO_4, NO_3, Cl, NH_4, Na, Ca, Mg, and K, chemicals directly determining the water acidity pH. Compared to the dry depositions, rain or snowfall are needed to obtain the "wet" data, and long-lasting droughts can dramatically reduce the number of pairs useful for pollution decline assessment. Though the dry and wet deposition collecting stations do not overlap, a comparison of Table 5.34 to Table 5.3 and Table 5.4 provides some idea about the reduction.

Modeling of the time-series formed by the concentration measurements does not require any particular innovative ideas, except pH. The value of pH is essentially the logarithm of the free hydrogen $H^{(+)}$. Compared to the other chemical concentrations, there is thus no theoretical lower or upper bound on the pH value.

Section 5.2.1 gives a list of reasons for the use of the percentage decline as a measure of change in observed air pollutant concentrations. The two most important are resistance towards biases between networks and conservation of the Lognormal distribution by the ratio of the concentrations. The possibility to attain negative values and the investigation (Mohapl 2001; Mohapl 2003b) indicate that pH observations have a distribution close to the Normal. Since a ratio of two Normal distributions is not Normal, it is preferable to assess changes of pH using the plain differences $pH_t - pH'_t$, where pH_t denotes the acidity measured on day t and pH'_t the value observed P years later. If a bias between the networks exists, it can be well described using a multiplier as well as an additive quantity. In both cases, the difference remains invariant towards network biases. Next we assume that

$$pH_t - pH'_t = \mu + \zeta_t, \tag{5.129}$$

where μ describes the change of the pH data over the two compared periods, and ζ_t is a zero-mean stationary stochastic process satisfying assumption (5.13). If the model is true, then μ can be treated as the rate of decline in the earlier sections and

TABLE 5.34
CASTNet Precipitation Collecting Stations. The Percentage of Pairs Used for Analysis Out of the Total Available Theoretically Given the Duration of Monitoring

Station	pH	SO_4	NO_3	Cl	NH_4	Na	Ca	Mg	K	Years
Abington	79	79	79	79	79	79	79	79	78	4
Alhambra	49	49	49	49	48	48	48	48	42	10
Ann Arbor	59	59	59	59	59	59	58	57	51	10
Arendtsville	73	73	73	73	73	73	73	73	68	10
Beaufort	69	69	69	69	69	69	69	69	65	4
Candor	59	59	59	59	59	59	59	58	54	8
Cedar Creek	71	71	71	71	71	71	71	71	68	10
Chiricahua NM	31	31	31	31	31	31	31	31	26	8
Cranberry	73	73	73	73	73	73	73	73	69	10
Deer Creek	64	64	64	64	64	64	64	63	61	10
Gothic	48	48	48	48	48	48	47	47	41	8
Lye Brook	63	63	63	63	63	63	63	63	62	2
Lykens	54	54	54	54	54	54	53	53	50	10
Perkinstown	60	60	60	60	60	60	59	59	54	10
Prince Edward	51	51	51	51	51	51	51	51	48	10
Scotia Range	79	79	79	79	79	79	79	79	74	6
Speedwell	63	63	63	63	63	63	63	63	58	10
Sumatra	52	52	52	52	52	52	52	52	49	10
Unionville	50	50	50	50	50	50	50	50	42	10

interpreted as the mean decline of acidity over the compared periods. If no change of pH occurred then $\mu = 0$. A negative estimate of μ means a decline of precipitation acidity (higher pH in the second period), a positive value means an acidity increase.

The observed differences (5.129) contain too many high and low values to qualify as observations from a Normal distribution. A competing model with stabilized variance was thus applied to see the difference between results of tests based on the asymptotic distribution of the raw data and on the Normal distribution of the transformed ones, respectively. The variance was stabilized by the transformation $\tau(x) = \arctan(x)$, the inverse to the tangent function and a model for the transformed data was built as described in Section 5.6.2. The transformation pulls the high and low values more towards the center of the sample and avoids unnaturally high variance in the tests about μ. Model (5.129) assumes that if the null hypothesis is true, then the differences (5.129) have a distribution symmetric about zero. The function $\arctan(x)$ is symmetric about the origin. Hence, under the null hypothesis, the distribution of the transformed data will have zero mean. The transformation seems to have no effect on the outcome of the test about the significance of μ computed from the CASTNet pH data, except for Perkinstown, where the pH value of μ is significant afterwards. Interestingly, the transformed values appear to have somewhat higher autocorrelation. Results of the tests about the changes are in Table 5.36.

TABLE 5.35
The Model for CASTNet Precipitation Data Is $\ln(c_t/c'_t) = \alpha + \beta t + \zeta_t$, pH Measurements Satisfy $pH_t - pH'_t = \alpha + \beta t + \zeta_t$, Where ζ Is an $AR(p)$ Process. Each Column Contains: Sign of β,* if β Is Significant, the Order p, k If Null Hypothesis Is Rejected by KS Test and c, If It Is Rejected by the χ^2 Test. A Dot Means a Non-Significant Result. If Necessary, Arctan($pH_t - pH'_t$) Is Used for Testing

Station	pH	SO_4	NO_3	Cl	NH_4	Na	Ca	Mg	K
Abington	–.0..	–.0..	+.0..	+.6..	+.0..	+.3..	+.0..	+.6..	–.5..
Alhambra	–.11k.	–.0..	–.0..	+.0..	–.0kc	+.1.c	–.0..	+.0..	+*6..
Ann Arbor	+.0..	+.0..	+.3..	+*0..	+.0..	+*0..	+.0..	+.0..	+*10..
Arendtsville	+.0k.	–.4..	–*0..	–.0..	–*0..	+.1..	–.2.c	+.0..	+*0..
Beaufort	–.0..	–.0..	+.0..	+.2..	+.0..	+.2..	–.0..	+.0..	+.0..
Candor	+.7..	+.0..	+.0..	–.0..	+.0kc	+.0..	–.1.c	+.0..	+.0...
Cedar Creek	+.0..	–.2..	–.0..	–.2..	–.12..	+.2..	–*0..	–.0..	–.1..
Chiricahua NM	–.1..	–.1..	–.0k.	–.1..	–.1.c	–.2..	–.1..	–.1..	–.10..
Cranberry	+.0..	–.0..	–.0..	–.1k.	+.0k.	+.0..	+*0..	+.0..	+*0..
Deer Creek	+.4..	+.0..	+.0..	+.4..	–.0..	+*0..	+.3..	+.0..	+*0..
Gothic	–.1..	–*5..	–*1..	–*5..	–*1..	+.1..	–*1..	–.1..	–*1..
Lye Brook	+*1..	–.0..	–.0..	–.4..	–.0..	–.5.c	–*0..	–*0..	–*1..
Lykens	+.0..	–.1..	+.0..	–.1..	+.8..	+.0..	–.0..	–.0..	+.3..
Perkinstown	+.1kc	–*0..	–.0..	–*0kc	–.1..	+.6..	–.1..	–.1..	+*4..
Prince Edward	+*0..	–.0..	–.0..	–.0..	–*0..	–.0..	–*0..	–.1..	+.0..
Scotia Range	+.1..	–.2..	–.0kc	+.0..	+.1..	+.6..	+0..	+.0..	–.4..
Speedwell	+.0..	–*0..	–*0..	–.0.c	–*0..	+.1..	–*1..	–.0..	+.0..
Sumatra	–.0..	+.0..	+.0..	–*0..	–*0k.	–.0..	–.0..	–*1..	+.1..
Unionville	+.0..	–*1..	–*0..	–.1..	–*0..	+.1..	–*0..	–.0..	–.2..

As in the case of the dry deposition data, it is desirable to plot the changes calculated from 10-year observations against the latitude and longitude, respectively, to find out if the percentage changes have the tendency to grow or decline in the northeast direction. Changes of SO_4, NO_3, and NH_4 are of particular interest because they are monitored among the filter pack species. First, let us notice that disregarding the large jump in acidity at Perkinstown we might say that the magnitude of pH change declines as we go to the north. This corresponds to the growing magnitude of TSO_4 changes. Interestingly, the growth of precipitation SO_4 changes is not as convincing. With a bit of imagination, the growth from south to the north might be there, but only NO_3 and NH_4 have rates that depend significantly on the latitude. This can be shown by simple linear regression. The amount of missing pairs prevents us from using the more efficient methods from Section 8. It is to be pointed out that because the data show no specific growth in the latitude or longitude direction it does not mean the directional dependence is not present. Examples of more sophisticated methods designed for the analysis of directional data are in Mardia (1972).

TABLE 5.36
Estimates of the Average Change of pH and the Long-Term Percentage Decline $\hat{pd}$ Estimated from CASTNet Precipitation Data. The Asterisk Denotes a Significant Change Based on the Statistics Z_μ. Model for Z_μ Is in Table 5.35

Station	pH	SO_4	NO_3	Cl	NH_4	Na	Ca	Mg	K	Years
Abington	0.03	8	1	8	–1	18	17	24*	37*	4
Alhambra	–0.02	29*	26*	34*	33*	42*	34*	40*	47*	10
Ann Arbor	–0.04	22*	7	6	5	16	26*	29*	–13	10
Arendtsville	0.07	18*	3	11	8	23*	7	10	12	10
Beaufort	0.16	–28*	–18	8	–59*	12	5	20	14	4
Candor	–0.02	21*	15	2	16	24*	34*	23*	22*	8
Cedar Creek	–0.08*	24*	11*	9	10	9	24*	20*	–2	10
Chirichaua NM	0.41*	29*	11	14	4	24	44*	30*	8	8
Cranberry	0.00	11	–1	1	–9	23*	23*	22*	–9	10
Deer Creek	0.01	21*	16*	21*	12	35*	27*	31*	22*	10
Gothic	0.04	15	–8	–5	–11	24	9	28*	–23	8
Lye Brook	–0.03	7	6	24	–18	29	14	26	–7	2
Lykens	–0.09	–1	–9	3	–21*	11	–35*	–48*	–24	10
Perkinstown	0.16*	33*	23*	23*	28*	43*	42*	44*	33*	10
Prince Edward	0.03	18*	3	17	4	37*	38*	34*	24*	10
Scotia Range	0.06	8	7	22*	1	50*	15	31*	51*	6
Speedwell	0.09	7	–6	0	–6	20	0	14	–6	10
Sumatra	0.06*	9	2	6	16	8	12	15	29*	10
Unionville	–0.03	15*	8	9	16	25*	13	5	15	10

Another important question is whether the percentage changes of precipitation SO_4, NO_3, and NH_4 correlate with declines of the corresponding species from the filter pack data. Simple linear regression conducted on the rates of change of species in Figure 5.19 showed no significant dependence, which agrees well with what we see in the plots. The rates of change are preferred over the percentage declines because the rates have Normal distribution, which is a common assumption for regression and correlation analysis. The invariance of pH values toward the significant changes of SO_4 and other concentrations is also interesting.

The number of significant changes detected in CASTNet precipitation data is rather remarkable because the multivariate analysis of the CAPMoN wet deposition data described (Mohapl 2001; Mohapl 2003b) show no changes over the same period. Since the filter pack data of the two networks agreed well as to the change magnitudes, the failure to detect any changes at the Canadian sites is likely a consequence of the inference method used. It is to be emphasized that CAPMoN samples daily. The use of pairs would thus reduce the sample size considerably. Development of a sensible analogy of the paired data analysis is thus highly desirable.

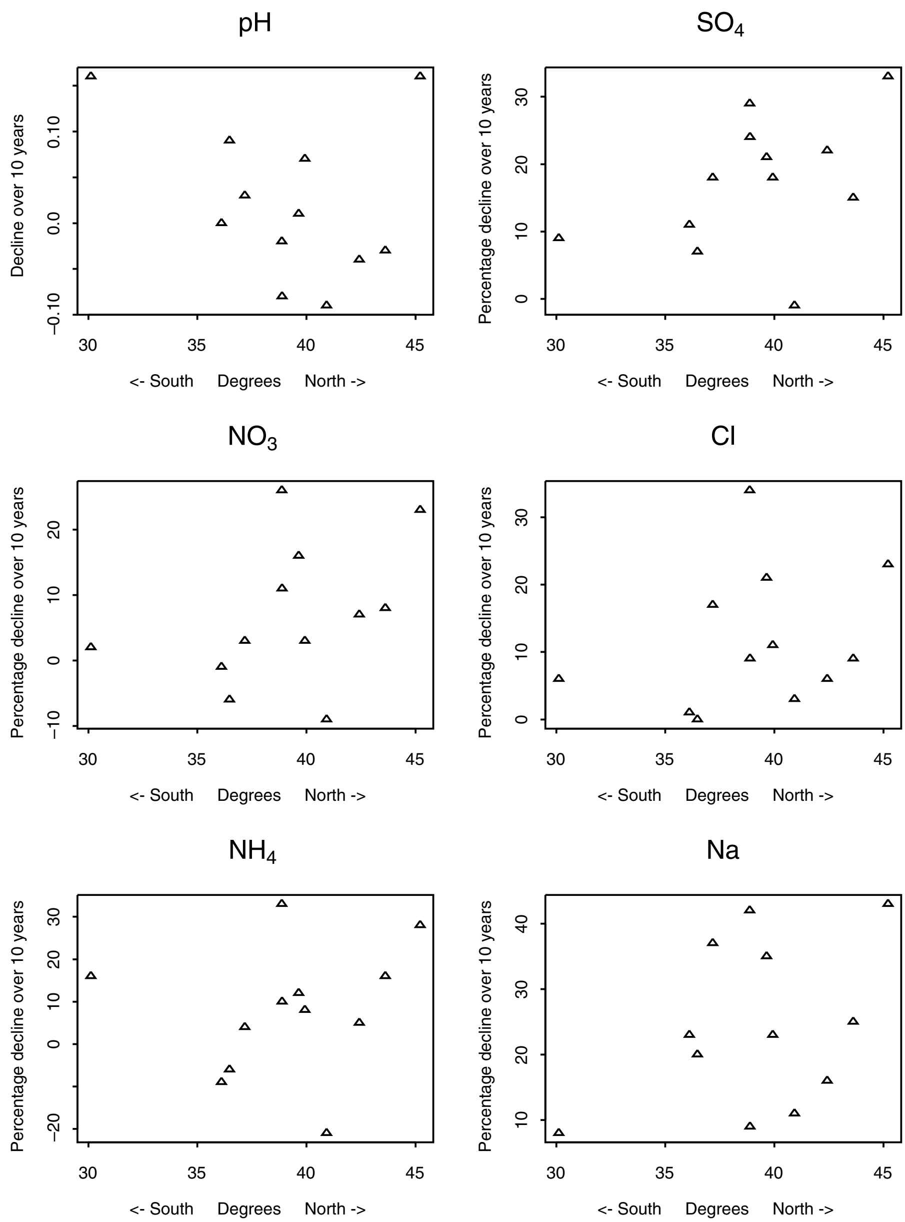


FIGURE 5.16 Change over a 10-year period observed in precipitation at the CASTNet stations in direction from south to north.

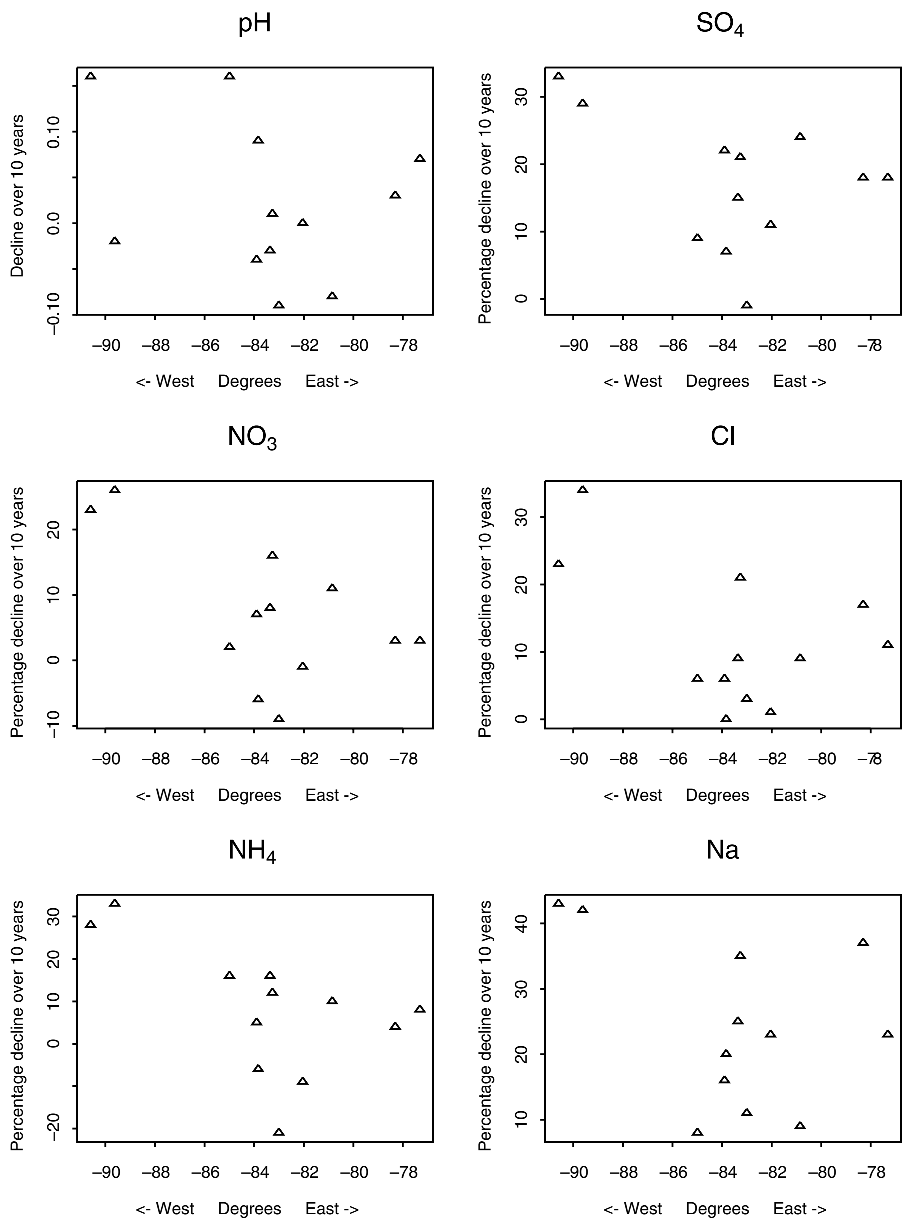


FIGURE 5.17 Change over a 10-year period observed in precipitation at the CASTNet stations in direction from east to west.

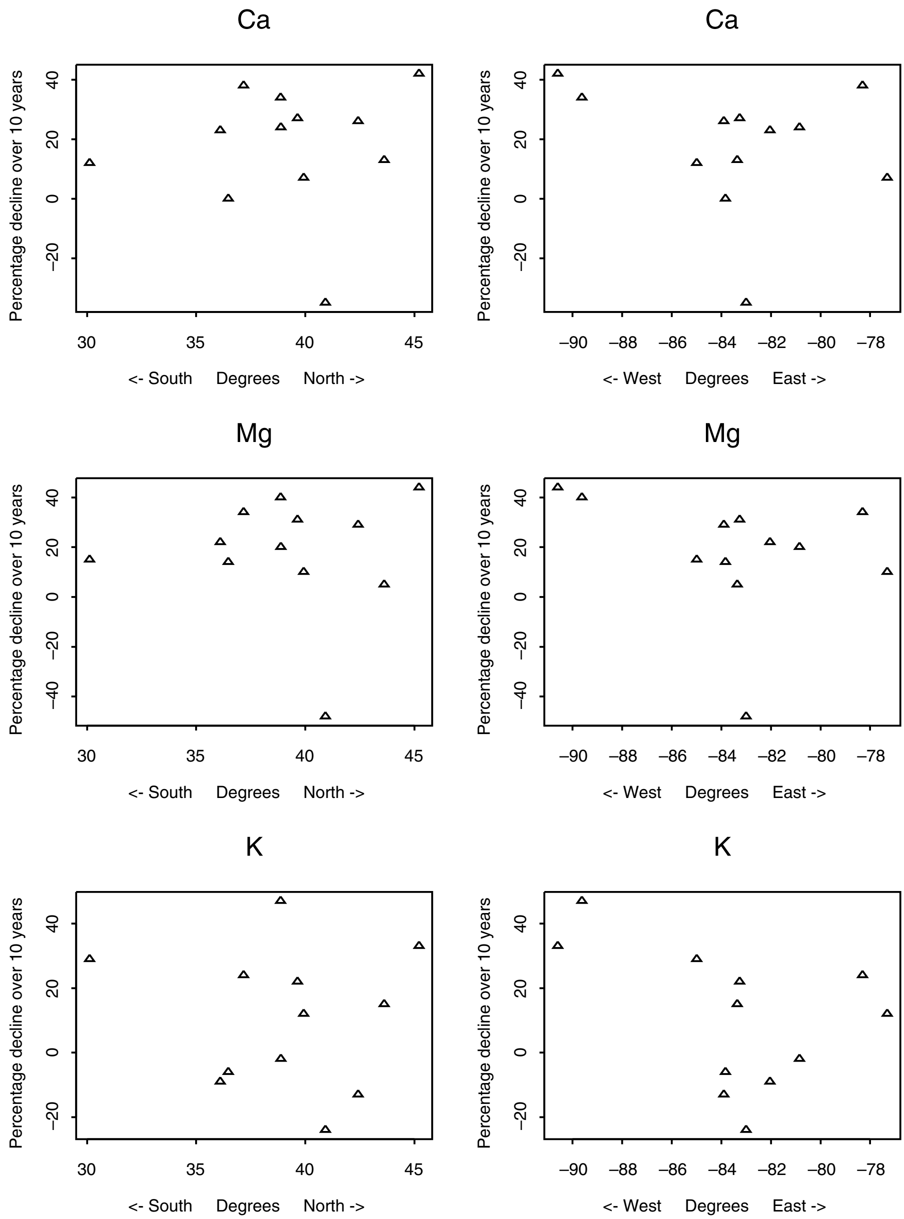


FIGURE 5.18 Change over a 10-year period observed in precipitation Ca, Mg, and K concentrations over a 10-year period at the CASTNet stations.

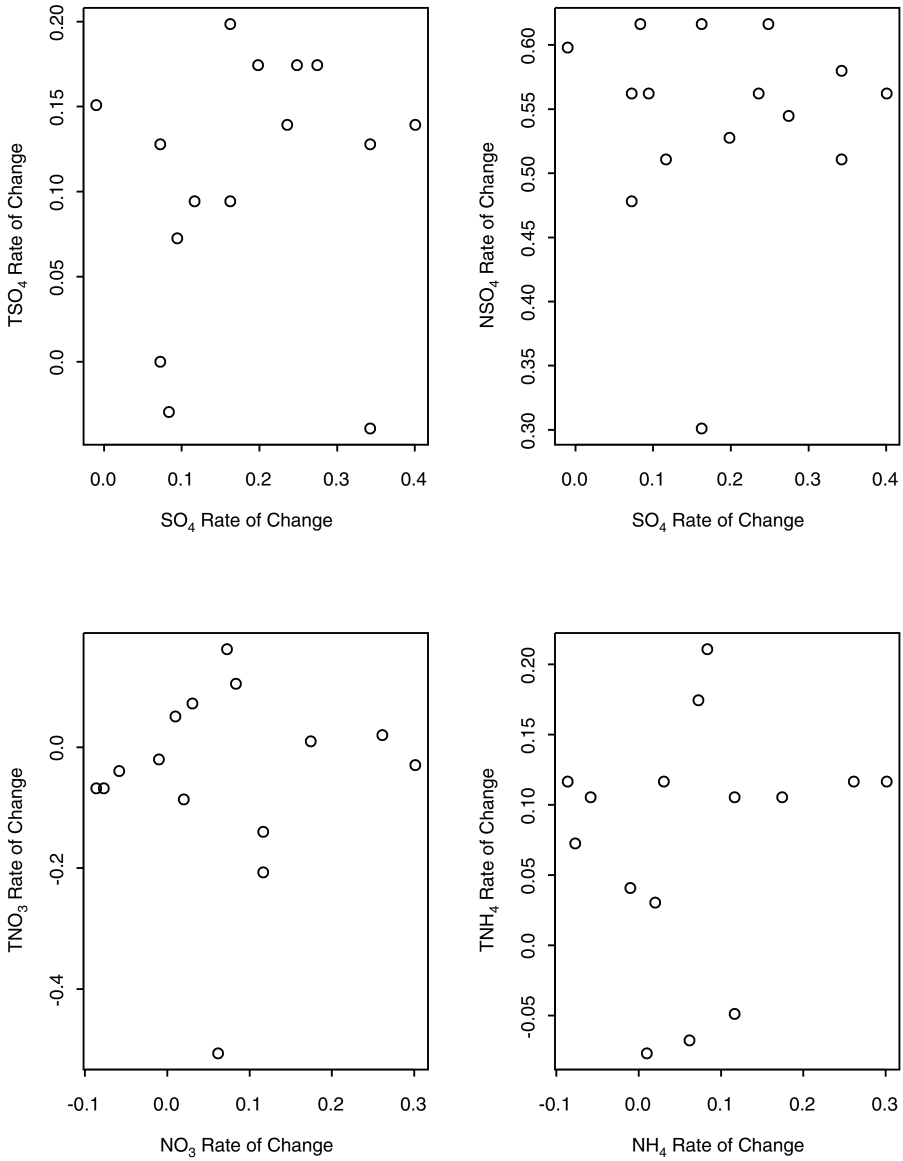


FIGURE 5.19 Rates of change determined from dry and wet CASTNet data.

5.12 PARAMETER ESTIMATION AND INFERENCE USING *AR*(*p*) MODELS

The mean μ of a stationary time series is in practice estimated either by the sample average or the maximum likelihood (ML) method. In the above theory, μ describes the rate of decline of an air pollutant under the assumption the pollutant concentrations conform the stationary process (5.11). Asymptotically, there is no difference between the distribution of the sample mean and the ML estimator. The literature on time series theory recommends thus the use of the arithmetic mean over the ML estimator (Brockwell and Davis 1987, Section 7.1) because there is little to be gained given the effort the ML calculation requires. Since the rate of decline is central to the above assessment method and inaccuracies of the rate estimates are magnified upon substitution into the percentege change estimators, it is worth discussing which of the two estimators is preferable.

This section argues that the sample mean provides a better estimator of μ than the ML method because the mean is unbiased, has smaller variance, and is not affected by incorrect specification of the model, which makes it better for plain reporting purposes. In addition, for $AR(p)$ processes, the statistics Z_μ in (5.24), based on the asymptotic properties of the average, is more powerful for the test of hypothesis $\mu = 0$ against the alternative $\mu \neq 0$ than the corresponding ML test, conducted in the presence of unknown (i.e., estimated) autoregressive parameters.

5.12.1 ML Estimation for *AR*(*p*) Processes

Next we recall the familiar conditional likelihood function for $AR(p)$ processes and the ML estimation procedure. In this section we assume that the observations $y_1,\ldots, y_N$ form a stationary time-series and satisfy the relation

$$y_t = \mu + \zeta_t, \tag{5.130}$$

where ζ_t is described by (5.53). The process (5.53) is considered with $\eta_t = \varepsilon_t$, where ε_t are mutually independent, identically distributed Normal random variables with zero mean and variance one. Under such a model, the probability density function of the data conditioned on the first $y_1,\ldots, y_p$ observations is

$$L(y_{p+1}\ldots,y_N) = \frac{1}{(2\pi)^{(N-p)/2}\sigma^{N-p}} \times \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=p+1}^{N}(y_t - \rho_t y_{t-1} - \cdots - \rho_p y_{t-p} - \mu')^2\right\}, \tag{5.131}$$

where

$$\mu' = \mu(1-\rho_1 - \cdots - \rho_p). \tag{5.132}$$

The ML method is derived from the idea that the parameters should be chosen in such a way that what was observed has the best chance to happen under the model. That is why it considers the probability density (5.131) as a function of the parameters with data fixed. As a function of parameters only, (5.131) is called the conditional *likelihood function*. The ML estimates maximize the likelihood function and, due to the monotony of the natural logarithm, minimize the log-likelihood function arising from $-2 \ln L$ by omission of the factor $(N - p) \ln (2\pi)$:

$$l(\mu', \rho_1, ..., \rho_p, \sigma^2) = (N-p)\ln\sigma^2 + \frac{1}{\sigma^2}\sum_{t=p+1}^{N}(y_t - \rho_1 y_{t-1} - \cdots - \rho_p y_{t-p} - \mu')^2. \tag{5.133}$$

Quantities obtained by solution of the equation $l(\mu', \rho_1,\ldots, \rho_p, \sigma^2) = 0$ are called *maximum likelihood estimators* of μ', $\rho_1,\ldots, \rho_p$, and σ^2, respectively. The ML estimator of σ^2 is

$$\hat{\sigma}^2(\mu', \rho_1, ..., \rho_p) = \frac{1}{N-p}\sum_{t=p+1}^{N}(y_t - \rho_1 y_{t-1} - \cdots - \rho_p y_{t-p} - \mu')^2 \tag{5.134}$$

and coincidentally, the ML estimates of μ', $\rho_1,\ldots, \rho_p$ also minimize the function $\hat{\sigma}^2 (\mu', \rho_1,...,\rho_p)$. Substitution of $\hat{\sigma}^2$ in the log-likelihood yields the *reduced log-likelihood function*

$$l(\mu', \rho_1, ..., \rho_p, \sigma^2) = (N-p)\ln\hat{\sigma}^2(\mu', \rho_1, ..., \rho_p) + (N-p), \tag{5.135}$$

which shows that as long as p is held fixed, the original likelihood is largest if $\hat{\sigma}^2(\mu', 1,\ldots, \rho_p)$ is smallest. The order of the model p can be obtained by minimizing the *Akaike's information criterion* (AIC) with respect to p. The AIC has the form

$$AIC = (N-p)\ln\hat{\sigma}^2(\hat{\mu}', \hat{\rho}_1, ..., \hat{\rho}_p) + 2p, \tag{5.136}$$

where $\hat{\mu}'$, $\hat{\rho}_1,\ldots, \hat{\rho}_p$ are the ML estimators (Brockwell and Davis 1987, Section 9.3; Shumway 1988, Section 3.3, p. 154; and Kendall and Stuart 1999, Volume 3, Section 50.11, p. 620).

Minimiziation of (5.134) combined with (5.53) leads to the equation

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \sigma\varepsilon, \tag{5.137}$$

where $\boldsymbol{Y} = (y_{p+1},\ldots, y_n)^T$, $\boldsymbol{X}$ is a $(N - p) \times (p + 1)$, matrix with the first column composed of ones and the k-th column has form $(y_{p-k+1},\ldots, y_{n-k})^T$, where T denotes the transposed to a vector (matrix). Unknown $\beta = (\mu', \rho_1,\ldots, \rho_p)^T$ denotes the solution.

The model is linear in μ' but nonlinear if considered in terms of the original parameters. The ML method thus provides the estimator

$$\hat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}. \tag{5.138}$$

Due to the equation

$$\hat{\beta} - \beta = \sigma(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\varepsilon \tag{5.139}$$

that follows from (137),

$$var(\hat{\beta}) \approx \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} \tag{5.140}$$

and for large samples, there is a positive definite matrix $\boldsymbol{V}$ such that $(\boldsymbol{X}^T\boldsymbol{X})/(N-p) \approx \boldsymbol{V}$. Since $\hat{\beta} = (\hat{\mu}', \hat{\rho}_1, ..., \hat{\rho}_p)^T$, $\hat{\mu}'$ is asymptotically Normal with mean μ' and deviation $\hat{\sigma}\sqrt{w/(N-p)}$, where w is the upper-left-corner element of the matrix $\boldsymbol{V}^{-1}$.

The ML estimator of μ' has the form

$$\hat{\mu}' = \frac{1}{N-p}\sum_{t=p+1}^{N}(y_t - \hat{\rho}_1 y_{t-1} - \cdots - \hat{\rho}_p y_{t-p}), \tag{5.141}$$

and the ML estimator of μ is

$$\hat{\mu}_{ML} = \frac{\hat{\mu}'}{1 - \hat{\rho}_1 - \cdots - \hat{\rho}_2}. \tag{5.142}$$

The test of the null hypothesis $\mu = \mu_0$ vs. the alternative $\mu \neq \mu_0$ can be based on the statistics

$$Z_{\mu'} = \frac{\hat{\mu}' - \mu'}{\hat{\sigma}\sqrt{w}}\sqrt{N-p} \tag{5.143}$$

because $\hat{\mu}' \approx \hat{\mu}(1 - \hat{\rho}_1 - \cdots \hat{\rho}_2)$ and due to (132),

$$\hat{\mu}' - \mu' \approx (\hat{\mu} - \mu)(1 - \rho_1 - \cdots - \rho_p). \tag{5.144}$$

5.12.2 Variability of the Average $\hat{\mu}$ vs. Variability of $\hat{\mu}_{ML}$

Let us introduce an estimator $\hat{\mu}_T$ that arises from $\hat{\mu}_{ML}$ in (5.142) by replacing the estimated ρ's by their true values. Intuitively, $var(\hat{\mu}_T) \leq var(\hat{\mu}_{ML})$, because using the true parameters instead of their estimators will reduce the variability. Writing M

instead of $N - p$ we get

$$var(\hat{\mu}_T) = \frac{\sigma^2 \sum_{t=p+1}^{N} var(\varepsilon_t)}{M^2(1-\rho_1 - \cdots - \rho_p)^2} = \frac{\sigma^2}{M(1-\rho_1 - \cdots - \rho_p)^2}. \tag{5.145}$$

At the same time, the mean

$$\hat{\mu} = \frac{1}{M}\sum_{t=p+1}^{N} y_t \tag{5.146}$$

has variance

$$\begin{aligned} var(\hat{\mu}) &= \frac{1}{M}\sum_{|h|<M} < M\left(1 - \frac{|h|}{M}\right)R(h) < \frac{1}{M}\sum_{h=-\infty}^{\infty} R(h) \\ &= \frac{2\pi}{M} f(0) = \frac{\sigma^2}{M(1-\rho_1 - \cdots - \rho_p)^2}, \end{aligned} \tag{5.147}$$

where $R(h)$ denotes the autocovariance function and $f(\lambda)$ the spectral density of the process ζ introduced in (5.12) and (5.55), respectively. Consequently, $var(\hat{\mu}) < var(\hat{\mu}_T) \leq var(\hat{\mu}_{ML})$. The average has thus a smaller variance than the ML estimator.

5.12.3 Power of Z_μ vs. Power of the ML Statistics $Z_{\mu'}$

In this section we rely on large sample results only; hence, N is used instead of the more accurate $N - p$. The statistics Z_μ in (5.16) is considered with

$$v\hat{a}r(\hat{\mu}) = \frac{\hat{\sigma}^2}{N(1-\hat{\rho}_1 - \cdots - \hat{\rho}_p)^2} \tag{5.148}$$

and compared to the statistics $Z_{\mu'}$ defined by (5.143). In consequence of (5.144) and (5.148),

$$\begin{aligned} Z_{\mu'} &= \frac{\hat{\mu}' - \mu'}{\hat{\sigma}\sqrt{w}}\sqrt{N} \approx \frac{\hat{\mu}' - \mu}{\hat{\sigma}\sqrt{w}}(1-\hat{\rho}_1 - \cdots - \hat{\rho}_p)\sqrt{N} \\ &= \frac{\hat{\mu} - \mu}{\sqrt{v\hat{a}r(\hat{\mu})}}\frac{1}{\sqrt{w}} = Z_\mu \frac{1}{\sqrt{w}}, \end{aligned} \tag{5.149}$$

where w is the already mentioned upper-left-corner element of a matrix $\boldsymbol{V}^{-1}$ arising from the relation $(\boldsymbol{X}^T\boldsymbol{X})/N \approx \boldsymbol{V}$ in consequence of the law of large numbers. More specifically,

$$\boldsymbol{V} = \begin{pmatrix} 1 & \mu\mathbf{1}^T \\ \mu\mathbf{1} & \boldsymbol{M} \end{pmatrix}, \tag{5.150}$$

where $\mathbf{1}$ is a p-dimensional column vector of ones and $\boldsymbol{M} = Eyy^T$ is the $p \times p$ matrix of mixed second-order moments $y = (y_1,\ldots, y_p)^T$. Consequently, for large samples, w is described by the relation

$$w \approx \frac{|\boldsymbol{M}|}{|\boldsymbol{V}|}, \tag{5.151}$$

where the absolute value denotes the determinant of the matrix. Let us recall that this is the upper-left-corner element of $\boldsymbol{V}^{-1}$. Since

$$|\boldsymbol{V}| = |\boldsymbol{M}|(1 - \mu^2\mathbf{1}^T\boldsymbol{M}^{-1}\mathbf{1}), \tag{5.152}$$

(see Rao 1973, Chapter 1) w can be further determined as

$$w = \frac{1}{1 - \mu^2\mathbf{1}^T\boldsymbol{M}^{-1}\mathbf{1}} \geq 1, \tag{5.153}$$

with equality in place if and only if $\mu = 0$. We can thus conclude that $|Z_{\mu'}| \leq |Z_\mu|$.

Consequently, if the true value of μ is not zero, then $w > 1$ and Z_μ rejects earlier the null hypothesis $\mu = 0$ against the alternative $\mu \neq 0$ than $Z_{\mu'}$. Statistics Z_μ is thus more powerful for this particular alternative than $Z_{\mu'}$.

5.12.4 A Simulation Study

To see how the accuracy of the average and Z_μ compare to the ML estimator of μ and the statistics $Z_{\mu'}$, we conducted a small simulation study. It consisted of the generation of independent series from the $AR(1)$ model discussed in detail in Example 1.6.1, with $\mu = 0.2$, $\rho = 0.3$, and $\sigma = 0.7$. The series were of length $N = 100$, which corresponds to pairs from about 4 years of weekly monitoring. The parameters were chosen to reasonably reflect the values frequently estimated using model (5.11). In addition, the choice of μ was done to create a specific situation when Z_μ should reject the null hypothesis $\mu = 0$ in favor of $\mu \neq 0$ more frequently than $Z_{\mu'}$. For larger or smaller μ values, both tests will either reject the null hypothesis or be inconclusive. The statistics Z_μ has in case of an $AR(1)$ model the form as in (5.60). The results are summarized in Table 5.37. Though on average the estimates are quite similar, Z_μ tends to grow faster that $Z_{\mu'}$.

TABLE 5.37
A Typical Simulation Study Summary Statistics From 100 Trials with an *AR*(1) Model and Parameters $\mu = 0.2$, $\sigma = .7$ and $\rho = .3$. *H*0 : $\mu = 0$ vs. *H1* : $\mu \neq 0$ Was Tested. The Average Values of $\hat{\mu}$ and $\hat{\mu}_{ML}$, the Statistics Z'_{μ} and Z_{μ}, and the Corresponding Rejection Frequencies Are in the First Line. Sample Deviations Calculated from the 100 Results Are in the Second Line. We Recall That t_{99} (0.05) = 1.984

	$\hat{\mu}$	$\hat{\mu}_{ML}$	Z'_{μ}	$Z_{\mu'}$	$Z'_{\mu}+$	$Z_{\mu'}+$
Average	0.20269	0.20326	2.15100	1.97961	53	48
Deviation	0.10282	0.10551	1.12654	0.99779		

5.13 CONCLUSIONS

Provision of data for assessment of the long-term change in air pollutant concentrations is one of the main reasons for operation of large air-quality monitoring networks. This study offers methods for the assessment of the change based on the reported amounts and applies them to several data sets collected over nearly 20 years by three North American networks known as CASTNet, CAPMoN, and APIOS-D. The procedures utilize a long-term decline indicator leading to a theory providing a number of interesting statistics for policy-related decisions. A brief summary of the presented theory, methods, and case-study results are discussed in the next two subsections.

5.13.1 Method-Related Conclusions

This chapter formalizes the concentration change as a quantity on its own, rather than a by-product of a regression analysis, a feature distinguishing it from most recent studies on concentration changes. The case studies suggest that the same simple indicator, interpreted as the average long-term percentage decline, can be used for the evaluation of the concentration changes in a wide class of data sets. The main advantage of the indicator is resistance towards network biases allowing comparison of results from different networks. Though estimation and inference about the indicator require a proper, well-justified model, the indicator structure associates with models that are more simple than those commonly used in the so-called trend analysis and their relevance is easier to justify. Compared to most studies on the topic, the relation between the model and the resulting percentage change estimates and confidence regions is very straightforward.

Estimators of the percentage decline indicator value require only the concentration data. Trend analysis based models commonly involve explanatory variables describing temperature, humidity, and other meteorological factors used to remove variability caused by fluctuating meteorology. Besides frequently missing arguments that would show that the factors do really reduce the variability of the response variable, there is little evidence regarding improved accuracy of the inference about the long-term change in the presence of the additional factors. The case studies here show that sensible results can be obtained by relying on the concentration measurements only as long as the long-term change in the data is the sole point of the inference.

Flexibility as to the sampling frequency is another positive feature of the indicator and estimators demonstrated in this study. High sampling frequency may not be affordable for each country or organization. Procedures suggested here can be used for various comparisons without data aggregation leading to biases and artificial problems with variability that can influence test outcomes.

The resistance of the percentage change indicator towards biases introduced in this chapter offers the possibility to refocus from the growing number of studies on differences between sampling procedures on features that are common and of interest to policy drafting and enforcement. As well known, systematic biases among different networks are rather a rule than an exception. It is thus important for international agreements that time-invariant biases caused by differences in protocols can be eliminated for assessment purposes. Since representativeness of samples used in network comparisons is not always addressed, the possibility to ignore the bias is a real advantage. The resistance of indicators towards biases is reasonably supported by the case studies conducted here.

The percentage decline estimator is supplied with a relatively broad class of methods for calculation of confidence regions and testing about the significance of the change. The study also demonstrates how to make elementary decisions concerning the significance of change, measurement variability, and the length of monitoring that might be of interest when percentages and reduction deadlines are set in policies.

5.13.2 Case-Study Related Conclusions

The percentage decline indicator developed in this chapter appears quite sensitive to the long-term changes, as the case studies including CASTNet, CAPMoN, and APIOS-D air chemistry data show. Since no particular selection criteria are needed to demonstrate the proper magnitude of the change, a large number of stations with variable sample sizes could be analyzed. The outcome of the analysis are the estimates of the percentage change over the available monitoring period at each station, the statistical significance of the change, regional estimates of the percentage change, and proper confidence intervals for the true percentage declines at each station and over the whole region. Autoregressive models are used to describe the variability of the rate of change in time and weighted linear regression models capture spatial changes of these rates over the whole region.

The application of the methods developed in this chapter reveals a certain homogeneity in percentage changes of dry air chemistry data produced by the examined networks. It also shows that the decline in air pollution, estimated from 10-year CASTNet and CAPMoN samples, has a geographic character and grows from southwest to northeast. Short-term samples (4 years) from those networks, however, indicate rather an increase in pollution. Data from the studied networks also demonstrate that methods suggested in this chapter offer various procedures for verification if different networks, with specific sampling protocols, report the same concentration changes. They show that the change indicators are indeed invariant towards network biases and allow us to turn attention from network comparison studies to other important issues.

The reader should be aware that the statistics for long-term assessment of air quality changes presented above are sensitive not only to the pollution declines at the site but also towards biases caused by changes and violations of the sampling protocol. For example, the observation mechanism for dry air depositions involves filter packs that are replaced daily or weekly. A permanent change in the retention characteristics of the filters is likely to cause a systematic bias which could erroneously be interpreted as a long-term change in reported concentrations. Though the filter packs should be collected at regular intervals from the field location, shipped under the same conditions in the laboratory, and always analyzed in the same manner, there are no reports that would provide a useful idea about the variability of these procedures. Changes in laboratory sample analysis techniques can also be a source of biases with undesired consequences (Mohapl 2000a). Since there are no publicly available data that would allow us to assess how far the reported amounts reflect the true concentrations, what part of variability is likely due to the measurement methods, etc., the percentage changes presented here should be interpreted with a certain degree of commonsense.

REFERENCES

Aitchison, J. and J.A.C. Brown, 1957, *The Lognormal Distribution.* Cambridge: Cambridge University Press.

Angulo, J., W. González-Manteiga, M. Febrero-Bande, and F. J. Alonso, 1998, Semi-parametric statistical approaches for space-time process prediction. *Environmental and Ecological Statistics* 5, 297–316.

Bennet, R. J., 1979, *Spatial Time Series. Analysis—Forcasting—Control*. London, Pion, Ltd.

Bollerslev, T., 1987, A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics* 69, 542–547.

Brockwell, P. J. and R. A. Davis, 1987, *Time Series: Theory and Methods.* New York: Springer-Verlag.

Civeroloa, K.L., E., Brankova, S. T. Rao, and I. B. Zurbenko: 2001, Assessing the impact of the acid deposition control program. *Atmospheric Environment* 35, 4135–4148.

Clarke, J. F., E. S. Edgerton, and B. E. Martin, 1997, Dry deposition calculations for the Clean Air Status and Trends Network. *Atmospheric Environment* 31, 3667–3678.

Cressie, N., 1993, *Statistics for Spatial Data*. New York: Wiley.

Draper, N. and H. Smith, 1981, *Applied Regression Analysis,* second edition, New York: Wiley.

Dutkiewicz, V., M. Das, and L. Husain, 2000, ‘The relationship between regional SO2 emissions and downwind aerosol sulfate concentrations in the northeastern United States, *Atmospheric Environment* 34, 1821–1832.

Engle, R. F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1008.

Feller, W., 1970, *An Introduction to Probability Theory and its Applications,* third edition, Vol. I. New York: Wiley.

Finney, D. J., 1941, On the distribution of a variate whose logarithm is normally distributed. *Suppl. J. R. Statist. Soc.* 7, 155–161.

Grenander, U. and M. Rosenblatt, 1984, *Statistical Analysis of Stationary Time Series*. New York: Chelsea, second edition.

Hess, A., H. Iyera, and M. William, 2001, Linear trend analysis: a comparison of methods. *Atmospheric Environment* 35, 5211–5222.

Holland, M. D., P. P. Principe, and J. E. Sickles, 1999, Trends in atmospheric sulfur and nitrogen species in the eastern United States for 1989–1995. *Atmospheric Environment* 33, 37–49.

Husain, L., V. A. Dutkiewicz, and M. Das, 1998, 'Evidence for decrease in atmospheric sulfur burden in the Eastern United States caused by reduction in SO_2 emissions'. *Geophysical Research Letters* 25, 967–970.

Kendall, M. and A. Stuart, 1977, *The Advanced Theory of Statistics,* fourth edition, Vol. 1–3. New York: Macmillan Publishing.

Lehmann, E. L., 1983, *Theory of Point Estimation.* New York: Wiley.

Loéve, M., 1977, *Probability Theory,* fourth edition, Vol II. New York: Springer-Verlag.

Lynch, J. A., J. W. Grimm, and V. C. Bowersox, 1995, Trends in precipitation chemistry in the United States: A national perspective, 1980–1992. *Atmospheric Environment* 29, 1231–1246.

Mardia, K.V., 1972, *Statistics of Directional Data*. London: Academic Press.

Mohapl, J., 2000a, Measurement diagnostics by analysis of last digits. *Environmental Monitoring and Assessment* 61, 407–417.

Mohapl, J., 2000b, Statistical inference on series of atmospheric chemistry data. *Environmental and Ecological Statistics* 7, 357–384.

Mohapl, J., 2001, A statistical assessment of changes in precipitation chemistry at CAPMoN sites, Canada. *Environmental Monitoring and Assessment* 72, 1–35.

Mohapl, J., 2003a, An annual wet sulfate deposition index. *Stochastic Environmental Research and Risk Assessment* 17, 1–28.

Mohapl, J., 2003b, Assessment of Wet Sulfate Concentration Changes at CAPMoN sites, Canada. *Environmental Modeling and Assessment* 8, 1–15.

Ohlert, G. W., 1993, Regional trends in sulfate wet deposition. *Journal of the American Statistical Association* 88, 390–399.

Press, W.H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery: 1986, *Numerical Recipes in FORTRAN. The Art of Scientific Computing*. Cambridge: Cambridge University Press.

Rao, C. R., 1973, *Linear Statistical Inference and Its Application,* second edition. New York: Wiley.

Ro, C., D. Vet, and A. Holloway, 1997, *National Atmospheric Chemistry Database.* Downsview, Ontario, Canada: Atmospheric Environment Service.

Saylor, R. D., K. M. Butt, and P. L. K.: 1992, Chemical characterization of precipitation from a monitoring network in the lower Ohio River Valley. *Atmospheric Environment* 26A, 1147–1156.

Shumway, R. H., 1988, *Applied Statistical Time Series Analysis*. Toronto: Prentice-Hall.

Sickles, J. E. and D. S. Shadwick, 2002a, Biases in Clean Air Status and Trends Network filter pack results associated with sampling protocol. *Atmospheric Environment* 36, 4678–4698.

Sickles, J. E. and D. S. Shadwick, 2002b, Precision of atmospheric dry deposition data from the Clean Air Status and Trends Network. *Atmospheric Environment* 36, 5671–5686.

Sirois, A., 1993, Temporal variation of sulfate and nitrate concentration in precipitation in Eastern North America. *Atmospheric Environment* 27A, 945–963.

Stensland, G. J., 1998, Acid rain. In: *Encyclopedia of Environmental Science and Engineering*. New York: Gordon and Breach Science Publishers, pp. 1–13.

Vyas, V. M. and G. Christakos: 1997, Spatiotemporal analysis and mapping of sulfate deposition data over eastern U.S.A. *Atmospheric Environment* 31, 3623–3633.

Whittle, P., 1962, Topographic correlation, power law covariance functions and diffusion. *Biometrika* 49, 305–314.

6 Atmospheric Monitoring

A. Köhler

CONTENTS

6.1 INTRODUCTION

The troposphere in space and time, containing the main part of the Earth's *turbulent* kinetic energy, is a heterogeneous "environmental compartment" whose meteorological, physical, and chemical parameters continually fluctuate. Further, as a carrier of trace gases, aerosols, and suspended particulate matter, numerous physical and chemical processes take place in its "open-air laboratory," induced and biased by incoming and outgoing radiation, meteorological phenomena such as precipitation, and through direct, (re)active contact with the adjacent compartments including soil, flora, ocean, and water bodies. Atmospheric constituents participate in chemical reactions, can be removed by fallout, washout, and rainout, and can be resuspended and transformed after contacting other compartments. These processes create pronounced vertical and horizontal concentration gradients of gaseous and particulate constituents and other gradients of physical properties (such as humidity, temperature, and wind speed) particularly close to the compartments' borders.

Vitally debated in this regard are anthropogenic activities that contribute to the atmosphere's continually changing composition from local to global dimensions. Thereby much attention is given to the potential of certain substances regarded as pollutants to influence the hygienic or radiative quality of the lower atmosphere and, together with the so-called greenhouse gases, weather and climate in different time and space scales. Trace gases and fine particles with very long residence times even reach the highest atmospheric layers. Hence, the quality of the human environment may be altered in manifold ways, more or less pronounced, and as indicated before, actually from very local and regional to continental and even global ranges. At the

1-56670-641-6/04/$0.00+$1.50

same time though, it has to be kept in mind that many of the substances regarded as "pollutants" always have been part of the various natural processes and cycles and, to a substantial degree, are not manmade.

No doubt, monitoring and measuring in the atmosphere is always a demanding — and fascinating — challenge; it is not easy to secure reproducible and dependable data of known accuracy that provide meaningful answers to individual questions. Therefore, when conceiving a new monitoring program, chances are that the best approach may be missed at first sight. T.A. Edison put it thus: "I have successfully discovered ten thousand ways that don't work," which can be taken as advice that the time spent in careful planning of a study is a most fruitful investment, no matter how simple or complicated the questions asked may appear.

In the following section, atmospheric monitoring will be considered from a more general point of view rather than describing anew common or established standard procedures that have been dealt with in detail in the existing literature.[1–5] However, some emphasis has been laid on issues and details not easily obvious or notable but which may influence results or their interpretation while escaping attention.

6.2 AIR MONITORING STUDY PREPARATION

It need not be emphasized that all questions expected to be answered in a study have to be precisely examined and formulated well before the project starts. This will reduce chances of misinterpreting results, producing insufficient or excessive data populations, considering too many eventually related parameters, or omitting some essentially needed for the correct interpretation of findings. The more detailed in space and time the expected results have to be, the higher the efforts in material, staff etc. Conscientiously delineating the objections for each parameter to be monitored can well reduce expenditures and cut start-up time. Experts from scientific and technical disciplines presumably involved from start or likely to get involved at a later stage (e.g., when interpreting first or final results) may be asked to contribute to first consultations, given their academic background, and to the benefit of both sides.

Supporting information required for interpretation and application of data to be generated can possibly be obtained from public, social, health (e.g., epidemiological data) commercial, industry, and administrative bodies. The selection of relevant information may have to be agreed upon well in advance, and the format it can be provided in may need to be adjusted for easier data handling or for it to operate decently with classified material.

Depending on the components monitored, meteorological parameters should be continuously recorded. They are needed to facilitate data evaluation, data selection, and to identify periods of extreme weather conditions, for which particular data need special treatment or have to be rejected.

When thinking of suitable instruments and procedures, it may not be sufficient to know the full range of values the parameter to be monitored can occupy. Helpful as well is information about both their span in daily and seasonal cycles, the conditions (meteorology, source strengths) under which extreme values might occur, and whether a particular instrumental setup will also work dependably under exceptional ambient situations. This further helps to decide, for example, whether continuous or discontinuous monitoring is required, and in a case of discontinuous monitoring, what exactly the most suitable and practicable on/off intervals are. Discontinuous monitoring can give equally useful results with less effort by permitting simpler instrumentation and less intense servicing (e.g., avoiding frequent change and overloading of filters and absorbing liquids of impingers, or problems with microorganisms).

Further, agreement should be found on the accuracy (truth) and precision (reproducibility) required, for both single and average values, and on the error ranges still acceptable in case resources or time available becomes limited. Experience shows that overall requirements for manpower, maintenance, cost, and time are often a power function of the degree by which the accuracy of data obtained is to be improved. If in this regard the technical specifications of instrumentation to be chosen do not provide enough details, the manufacturer (and a competitor) should be asked to submit them. Besides, another early useful investment is to examine the relevant scientific and technical literature for reports published in recent decades on somehow related efforts. While older reports may not meet present academic standards and demands, they likely could offer basic information that no longer attracts attention but is still equally valid today.

As mentioned in the previous paragraph, performance under extreme (e.g., meteorological aspects and modes of operation) data quality requirements should also be considered at an early stage. Such requirements may be known for customary standard instrumentation. For new or proposed procedures, an important issue is the accuracy obtainable under conditions of sampling and concentrations typical for the parameter monitored. The level of accuracy may not actually have been established. Some information could be gained, at least about the reproducibility of observations, by operating identical instruments side by side repeatedly under different characteristic sampling conditions in a pre-study. Further, for certain aerometric parameters, it is academically impossible to measure their correct value with sufficient accuracy since the instrument itself, its sensors, or air intake notably disturb the quantity to be measured, for example, by deviating the original air flow (representativity) and to some extent even altering the composition of the air sampled, especially when material already precipitated adheres to the air ducts. This typically applies to the measurements of many meteorological parameters. (A quotation found in a very old textbook states: "A thermometer shows the temperature of the mercury it contains; if a relation with the true ambient air temperature exists, and which has to be established"). In such cases the best or even the only method applicable or practicable is to use strictly standardized instruments and procedures like the standing, internationally agreed-upon procedures for measurement of temperature, wind speed, and the amount of wet precipitation and snowfall. Standardized, i.e., comparable, meteorological parameters acquired thus are anyway needed for the interpretation of the measurements made in a study. Besides, important information on trace gases, air pollution, or biogeochemical cycles becomes available by analyzing rain, snow, and ice samples.

Another point that should be considered in the planning stage concerns logistic requirements, such as the continued provision of routine supplies like calibrating gases, filters, and similar material. This would also include ways of treatment, transport, storage of samples, and the routine maintenance of samplers and monitors.

Monitoring the atmosphere in open terrain naturally evokes public interest. This may not be avoidable. On the other hand, it can be beneficial. When providing information to officials and the media, great care should be taken to avoid misunderstanding of objectives or in anticipating results for which the study was not truly designed.

6.3 SITING

The prerequisites and requirements of accuracy and representativity of observations are principally different at the impact and background levels. At the former, the "background noise" is less important given the relatively high concentrations, but the variability is also very high considering the abundance of air pollution in small time and space scales. The comparability of observations made at different locations and different times is limited by the frequent occurrence of small-scale local interferences and the difficulty in finding sampling sites not at all influenced by individual sources but still being able to provide important information on the frequency of high concentrations under a wide range of meteorological and emission conditions.

At the background level, natural concentrations of substances regarded as pollutants are often in the same (low) order of concentration and variability as are anthropogenically caused concentrations. To differentiate between natural and man-made levels, which requires less resolution with time, very sensitive sampling, siting, and analytical methods must be applied.

The spacing and placing of the sampling sites which are usually very distant from each other is also difficult, considering that at ideally located sampling sites it can be hard to provide the necessary infrastructure. The particular purpose(s) of a study determines the placement and number of stations required, the geographical/topographical area to be covered, and the expected duration of operation.

Impact level monitoring would deal with local air-pollution problems; compliance with quality standards; impact of certain sources and processes (industry, traffic); and research on hygienic and epidemiological subjects. Scientific experiments on the transition of airborne elements into adjacent environmental compartments (or biota), often called "integrated monitoring," has applications at both the impact and background level. A more dense net of impact level stations furnishes information with high resolution in time about both average and extreme values of the substances monitored. Taking samples more frequently instead of accumulating a longer-term sample is to be preferred since it provides more detailed information.

Longer-term and large-scale background monitoring (from the regional up to continental and global scales) mainly deal with questions of long-range transport of pollutants, establishing typical background levels of relevant substances in and around large areas or countries or for areas where changes in land use (for industry, road construction, or town planning) are anticipated, or to study the impact of radiation to life on Earth. Since the 1960s there is much concern about the potential of (anthropogenic or greenhouse) trace gases and aerosols to modify regional and global climate. Interestingly enough, the final purposes of background monitoring may not all be known at the beginning. For example, carbon dioxide was at first selected as an indicator of energy production and thereby as a general measure of atmospheric pollution and was monitored because of its very long residence times. While a relatively sparse coverage of stations is sufficient at the background level, very exacting standards are required as regards the representativity (for a larger area) of the selected sites. Since background concentrations are lower, and when continuous sampling is not feasible, longer sampling intervals are normally required to obtain a sample big enough for analysis. Generally, in background monitoring, emphasis is placed on longer-term changes.

Sampling locations perfect in every respect are very difficult to find for practical (logistics, safety) reasons and also when the topography in the area under study may provide barriers or channels to lateral diffusion. Topographical characteristics can allocate the sources to be considered in a pattern different from that over a plane area under different meteorological situations including atmospheric stability, stagnation, precipitation, and wind directions. Varying land use in the vicinity of a sampling site can also cause systematic errors. Buildings, trees, and vegetation can alter wind speed, direction, and lateral and even vertical turbulence. As a rule of thumb, a sampler should be put up at least as far away laterally from an upwind structure as the height of deflecting objects. Such precaution still cannot make up for additional inhomogeneities caused by adsorption or resuspension of deposited material at the ground surface. Further, the varying "roughness" of terrain in the vicinity of the sampling point can change the composition and content of (individual) constituents monitored. In mathematical models for turbulent diffusion, usually such qualities are also taken into account.

If at all possible, wind direction and speed, relative humidity, temperature, and precipitation should be routinely recorded on or near the site. Thereby, data obtained can be assigned to typical classes of meteorological parameters or situations. Details on meteorological peculiarities at a prospected site should be obtained from nearby meteorological stations. The national weather service certainly can provide statistics for different averaging periods. Interpretation of the general meteorological situation by a meteorologist with regard to constituents to be monitored will be advantageous.

In a case where records exist on similar aerometric observations made in the area considered, information on the representativeness of the site selected can be obtained by preparing frequency distributions of such data. When no data from previous studies exist, this should be done as soon as sufficient data have been obtained. Such distributions will reveal that the observed frequency distribution very likely embodies several single distributions, often three to five, which are superimposed and of Gaussian type. Peak fit computer programs are very useful to this end. The lowest median value (a narrower distribution) may represent the conditions for the "global" type mixing mode, the next medians of wider distributions belong to "continental" and "regional" modes, and the highest median value would pertain to a skewed distribution with a long asymptotic tail caused by very local interferences or unusual (instrumental, etc.) circumstances. In the operational and evaluation phase of a study, checking data from time to time for multiple frequency distributions facilitates decision on the rejection of suspicious values which may not have been determinable applying the routine statistical checks (1–3 standard deviations criteria). Having "cleaned" data sets by this method, applying standard outlier tests to the remaining populations becomes more certain and probably even reduces the total number of outliers.

Siting of stations is an extremely important issue. Missing the best fitting location can bias results by up to 100% and more, or even make them just worthless.

Detecting suitable and representative sampling sites is specifically difficult in local-scale ambient air monitoring, given the irregular and highly variable fields of concentration pertinent to the entity monitored. Also very difficult to estimate are the existing limits (ranges) in space (and time) which are typical at the site for the individual parameters measured. The smaller the number of sites operating, the less certain are such estimates. Whenever possible, and at least temporarily, one or more additional identical sampling devices should be operated side by side with the permanent one. Thereby, valuable information becomes available, not only regarding representativity but also about the reproducibility of measurements. (Ideally, but generally not affordable, one would put up three identical samplers side by side, very close to each other but separated enough to ensure that one instrument cannot perturb the airflow to the neighbor instrument's air intakes; that way, outliers in most cases can be identified.)

The representativity of a site quite understandably is an important, nonetheless vague, attribute ratable only qualitatively. However, for some constituents like hygroscopic gases, it can be estimated experimentally.

To that end, depending on the parameters to be monitored, less specific and less sophisticated alternate sampling gauges can be used. These just need to provide comparable data on parameters, related in a defined, fixed way with the parameter to be monitored in the prospective study. Active samplers may be replaced by passive, surface-active absorbers (e.g., small filter disks, fixed in petri dishes upside down and impregnated with hygroscopic salt-type substances, so called "surface active monitors") forming a solid compound when in contact with the gas of interest. Within the wider study area a relatively large number of such monitoring devices can be put up as satellites in different distances around the place considered for routine monitoring. Gauges and sampling conditions must, of course, be identical. A simple statistical evaluation of data obtained by analyzing the filter disks should then allow identification of that particular site which most frequently appears in significant correlations between any two sites operated and whose average values are closest to the general average of all sampling points in the area.

The comparability of observations within a period of time at a sample site or between different sampling points can also be biased by more or less frequent small-scale local interferences or from occasionally strong individual sources. On the basis of frequent and careful evaluation of data obtained so far, and considering the meteorological parameters as well, criteria will emerge to identify outliers or to decide if a sampling point must be discontinued. Comparability will be further impaired when sampling is not made simultaneously at all points in the study area.

Air quality monitoring often undertakes to investigate the dependence of atmospheric composition and concentration of constituents on the type of land use (residential, industrial, commercial) or on special selected sources like traffic routes or selected anthropogenic activities. The higher the number of independent variables considered, the more sampling points are required, and utmost standardization in any detail is required at all participating sites to ensure comparability. This applies to height above ground of air intakes and wind sensors, equal distance from line sources, absence of aerodynamic obstacles, type of ground surface (concrete, vegetation, etc.). Of course, the air intake's design must be symmetrical with respect to wind direction. Another problem will occur when in a running study, for reasons not known before or underestimated, a sampling site has to be relocated. It should be realized that moving a station just a short distance can bias comparability of data drastically, the more the distance, the more inhomogeneous are the concentration fields under study. In case a station is to be moved, another one should be established at the new site and the former site must be operated for some time to obtain information that may allow adapting measurements made so far instead of losing them completely.

Also important for site selection is knowledge about the existence of pronounced diurnal (ultimately caused by the position of the sun or by working rhythms) and weekly cycles (industrial production, traffic density, working/nonworking days), which can be different for different categories of sources, and they all possibly need to be considered at the same time. This is in fact true for constituents under study and for the meteorological parameters influencing their transport and diffusion. Evidently, exploring such cycles may also be one of the objectives of a study.

In finally selecting sampling sites, other constraints may disagree with scientific demands. The actually existing local conditions regarding safety, accessibility, and infrastructure (electrical energy, water supply, telecommunication, transport needs) require careful consideration as well in balancing advantages and disadvantages of a site envisaged.

Siting is no point of course in the case of direct source measurements, but sampling problems are similar to those mentioned in the text below.

6.4 SAMPLING AND OPERATION

In atmospheric monitoring, the process of sampling presents an intrinsic difficulty. The devices by their mere physical existence constitute a "substantial" interference for the entity to be monitored. All its surfaces — including those of a protecting shelter — may interfere with the constituents monitored, hydrodynamically, chemically, and physically. Therefore, the analytical quality of whatever sophisticated instrumentation and procedures are employed cannot be better than the quality of the air sample taken and the accuracy of the calibration processes applied. It is an indispensable requirement to design and position the air intake in a way that it remains essentially independent from wind direction and fairly independent from wind speed.

Undeniably, in aerometric monitoring, by far most of the uncertainties that inevitably build up during the successive steps from siting to evaluation are due to siting and sampling. Underrating these facts can seriously distort results or make them practically meaningless. While such impact is more or less not quite avoidable and will first of all influence the accuracy (truth) of the measurements, on the other hand much can be done to improve the precision (reproducibility) of observations at a site and further the comparability between a group of sites by strictly standardizing instruments and procedures. Similar to the problem mentioned under siting in case a site has to be moved, for a particular site the experience acquired by some time may suggest that internal modifications have to be made. An additional identical instrument should be operated at that site under both former and new conditions for some time in order to establish comparability between versions and to avoid complete loss of existing data.

Interferences occurring on surfaces near and on the air intake and along the air ducts can create other errors. Ideally, the air sample, having entered the air intake, should remain homogeneous and not change its composition with respect to the parameters monitored until it is collected (impinger, filter, flask) or has reached the analyzing system. When sampling by gravimetric deposition, it is the design of the collecting vessel, dry or wet-and-dry buckets, that largely determines the result. Many collectors in that category are far from independent of the wind vector. The only remaining effort to save some comparability, reproducibility, and reliability of data obtained is the exclusive use of strictly standardized collectors and procedures. When using open buckets, precautions should be taken to avoid formation of organic material (fungi or microorganisms) in buckets during warm and humid periods.

There is no question that the air intake is an aerodynamic obstacle disturbing the (turbulent or laminar) air flow. The size distribution of suspended particles very likely will be changed, the more particles the wider their size range. Air intakes should be so designed that neither rain and snow nor coarse particles (>10 μm aerodynamic diameter) can enter. In monitoring the ambient concentration of a constituent, particles larger than this limit, while causing immense problems, are of little scientific value given the very short distance they remain suspended. For more details on this point, see text below.

Analyzers designed to monitor one particular gas normally require flow rates of a few liters per minute only, i.e., the air enters the air intake at low speed. The protecting cover for the air intake can then act as cutoff for coarse particles when air entering the intake is directed upwards. In dimensioning the air intake, one may remember that particles of 10-μm aerodynamic diameter in still air settle with 0.3 cm/sec, 50-μm particles with 7.2 cm/sec, and 100-μm particles with 25 cm/sec. As already mentioned, particles and aerosols greater in aerodynamic diameter than 2.5 or 10 μm normally must not enter the gauges. The so-called high-volume samplers designed to monitor suspended particulate matter (SPM) (to be distinguished from total particulate matter (TPM)), operate at high flow rates (about 1 m^3/min^1) and are normally equipped with air intakes cutting off particles (not very sharply though) at one of the sizes noted. Such devices which remove certain size ranges of particles by inertial force must operate exclusively at the prescribed flow rates.

Particulate precipitating in the air ducts by inertia, electrostatic charge, or thermodiffusion provides huge additional adsorbing surfaces. It can act as a catalyst, initiating chemical transformations in the sample air, altering composition and amount (concentration) of the substance monitored. Further, most surfaces, glass in particular, are normally covered with many layers of water molecules apart from condensation by adiabatic cooling along the air ducts. Regarding the air ducts' dimensions, it should be noted that smaller cross sections of tubings reduce delay and favor turbulence, but offer a relatively large contact surface per unit of air volume moving through it.

Suitable precautions need to be taken to reduce such impact. Temperature and humidity of the running air, state of the inner surface of air ducts, and flow characteristics (laminar, turbulent) should be known and at best kept constant or standardized.

Filters, when removing coarse particulate or collecting aerosols, in variable extent, depending on the loading already accumulated, can remove gaseous and particulate matter to be monitored temporarily or finally. This is to be especially considered when constituents are monitored that coexist in gaseous and solid mode. Transition from the gas into the solid phase and vice versa may happen right on the filter material, depending mainly on temperature, humidity, and concentration of a rather reactive compound. Consequently, when a substance coexists in different modes (sulfur dioxide and sulfates), it may be more logical to regard as relevant only the total mass (in this example total sulfur) of the different chemical compounds. This may lead to more consistent results if first evaluations of different ways of sampling show results difficult to interpret. Accordingly, the sampling technique could be modified.

When determining size distributions of particulate collected on filters it must be recognized that the size of particles may have changed during sampling by coagulation.

To minimize changes in the composition of the air sampled, as already indicated, its temperature, humidity, and flow rate may need to be controlled. The air inlet and/or the whole air duct may need to be insulated and heated appropriately if the ambient relative humidity is high or condensation by adiabatic cooling cannot be excluded. Inspecting the inner surface of critical sections of the air duct should facilitate such decisions.

Vertical concentration gradients as well as those of wind speed and direction (wind shear) generally change nonlinearly near the ground. Concentrations found just above the ground, up to about 30- to 50-cm height, may no longer be representative for the ambient air. Depending on the kind of ground surface, the height of the air intake above the ground should be at least 1.5 m, but 2 m is better. However, when the experiment is to contribute to the study of hygienic impact of selected substances to humans (including children), the height of exposure (breathing) may be the dominant factor.

When an air intake is to supply air for several analyzing systems operating at the site simultaneously, it seems logical to use one common supply to enable minimum lengths of tubing for each system. A bigger tube and a higher flow rate rather than much higher speed will be favorable. Most suitable is a straight noncorrosive tube, with the inner surface polished and the blower connected perpendicularly to the tube near its lower end. The lower end should be covered with a tight-fitting, easily removable lid to allow occasional cleaning of the tube. The connections to the individual analyzers should not protrude into the cross section and point upward at the very beginning, to prevent catching of coarse material. The main air intake itself can be designed more expediently and purposefully. It may be positioned some meters above ground, thereby avoiding other problems.

The sample air usually leaves an analyzer at rather low speed and flow rate, as is common with automatic gas analyzers. Thus, the constituent under study that may have been removed in the analyzing section, or other interfering gases like the mostly irritating water vapor, can, driven by very high concentration gradients, diffuse back and may well reach from behind the sensitive parts of analyzers and produce incorrect readings. In case of humidity, a small tube containing color-reactive silica gel placed at the air outlet prevents back-diffusion (and will illustrate this phenomenon). For other gases special reagents are needed.

The analytical procedure employed by an automatic analyzing system may be based on continuous comparison with one or more running calibrating gases. The pressure drop from gas sources to the outlets, and therefore the pressures in the chambers compared, may be (slightly) different, in particular when, in order to save expensive comparison gas, the latter runs at a rather reduced flow rate. Measuring the pressure differences relative to the barometric pressure at the respective outlets can help to control such differences which might cause another systematic error.

As mentioned earlier, most surfaces, glass in particular, even under low relative humidities, are covered with many layers of water molecules which can become involved in reactions at higher surface temperature. Outburst of gas or particulate matter may be stimulated by mechanical shocks and vibrations. While this may happen only occasionally and may be detectable and correctable, a systematic cross-sensitivity to water vapor can exist, for example, with nondispersive infrared (NDIR) analyzers. In some infrared wave bands the water molecule is the main IR absorber. Freezing out avoids interference from water vapor nearly completely but requires sophisticated instrumentation and breaks for removing the ice formed. The ice surface itself can act as absorbent for the gas under study, but a rather stable equilibrium should build up in the steady state, all other conditions remaining constant provided an eventually running comparison gas is treated the same way. Running both the sample air and calibrating and reference gases through water immersions (a wet and one or two dry impingers in series to catch tiny water droplets) at temperatures very carefully kept constant a few degrees above freezing point maintains a sufficiently low and constant water vapor pressure within a very wide range of ambient humidity. Such immersion also removes particles more efficiently and at lower and constant resistance than the higher and more variable flow resistance of other devices or filters placed upstream. The procedure is less sophisticated and economical but possibly less precise than the freezing-out method. In many applications this minor inaccuracy may not markedly contribute to the total inaccuracy of a procedure.

Coarse particulate matter in common air monitoring is of little interest and relevance since it constitutes another category of atmospheric components with quite another representativity. On the other hand, "dust" can become an issue of concern locally, only an irritation in many cases but a harmful factor for certain sensitive surfaces it adheres to. To measure the amount of dust, operating powerful instrumentation able to sample the full size range of particles is probably pointless.

Well-established standard instruments and procedures exist for monitoring wet and dry deposition. Well-analyzed wet and dry fallout, as done in the early times of air pollution monitoring, can provide valuable information on substances relevant to questions like acid rain and long-range transport of air pollutants, given the processes of washout and rainout of adsorbed or absorbed material having been suspended in the troposphere over longer distances. However, relating the amount of depositions found with specific sources, to a certain extent, is feasible only by applying mathematical models estimating atmospheric transport by turbulent diffusion under parametrized meteorological conditions.

When collecting samples in open buckets, at higher ambient temperatures organic material can develop, which would interfere with the analyses to follow. This is rather unlikely for absorbing solutions in impingers but not impossible. Special agents can be put into the sampling vessels which prevent organic growth but do not actually interfere with the particular chemical procedure applied.

Such precaution is also important in view of the delays incurred by storage at the site, transport, and further storage in the laboratory. In any case these delays must be kept as short as possible. The empty bottle provided by the laboratory may need to be sterilized and kept tightly locked until used.

Whichever may be the equipment and procedures selected for starting the study and subsequent routine operation, there is one basically important permanent exercise. For all instruments operated and procedures utilized on site, general and special instructions have to be carefully documented and kept available for easy perusal at all sampling stations. The date, time, and location they were put into force or eventually replaced earlier ones must be recorded.

Still more important is to keep at any sampling site exhaustive diaries documenting details about the following:

- Day and hour sampling and monitoring instrumentation and supplementary devices (e.g., climatization) were installed, tested, serviced, recalibrated or replaced, settings corrected or changed, type of filter changed etc., why and by whom
- Special events like occurrence of unusual meteorological parameters, their noticeable impact, and damages
- Information on power and instruments' failures, bugs in electric circuits and wiring, etc.
- Change and status (pressure, volume remaining, in stp) of calibrating gases or reference gases continuously running, and reference solutions, noting their batch designations, dates of shipment, adjustment of flow rates
- Any evidence of vandalism
- Any nonroutine experiments and activities, visits
- Staff names, temporary or final changes that are most important
- Other items to be identified by the scientists responsible for the study

Such documentation may turn out to be irreplaceable information needed when evaluating data in quality assurance problems, e.g., when deciding whether dubious data can be adjusted or must be rejected. This can be particularly valuable when correlations between different ambient and monitored parameters are to be investigated.

6.5 QUALITY ASSURANCE AND CONTROL

The quality, i.e., mainly correctness, reliability, and representativeness of data emerging from a project, has to be quantified in some way since, unavoidably, data generated by any measurement, particularly in atmospheric monitoring, are subject to some errors being produced at all stages of monitoring, from the planning phase to data assessment. Some errors can be controlled; others cannot, but they can be identified and described, quantitatively or qualitatively. Data quality can be expressed as a measure delineating the types and amounts of errors occurring in the different stages of monitoring. Quality assurance objectives are plans to be established for each project in order to ensure that an appropriate level of control is exercised over the sources of error that can be controlled and to obtain, to the extent possible, information on sources of uncontrollable errors.

Quality objectives, quality assurance (QA), and quality control (QC) have been discussed for many years in numerous expert meetings organized by various scientific, technical, and administrative bodies involved in air monitoring programs of national or international scope. They unanimously emphasized the need for and the importance of applying QA procedures. Given the specific scientific objectives and technical design of any monitoring project, conclusions drawn and recommendations made had to remain more or less at a philosophical, abstract level, since much of the knowledge needed to design quality-control procedures in detail will surface during preparation and the starting phase of the individual endeavor.

In this context, the U.S. Environmental Protection Agency's definition of "quality control" deserves attention: "QC is the overall system of technical activities that measures the attributes and performance of a process, item, or service against defined standards to verify that they meet the stated requirements established by the 'customer'. QC is both corrective and proactive in establishing techniques to prevent the generation of unacceptable data, and so the policy for corrective action should be outlined. … QC activities are used to ensure that measurement uncertainty … is maintained within the acceptance criteria for the attainment of data quality objectives."[6]

With this important definition in mind, segments of previous paragraphs of this chapter can be regarded as description of quality assurance activities. Some typical QA activities, mentioned so far or not, are briefly summarized below:

- Definition of objectives, parameters, data use, geographical area and land use to be considered, accuracies required and acceptable (one does not always need data of the best quality);
- Careful selection of sampling sites, considering prevailing meteorological situations, topography, representativeness, obstructive objects, roughness of ground surface, infrastructure requirements;
- Selection of instruments of known technical specifications, accuracy, performance, and maintenance needs;
- Aerodynamic design of air intakes, its properties to influence the size ranges of suspended particles desired or acceptable, tubing, sampling probes, heating, water vapor control, and flow-rate control;
- Design of monitoring programs, sampling intervals for continuous or discontinuous sampling, scheduling routine servicing, field audits, replicate measurements, recalibration, control charts, preparation of reference materials, planning technical systems audits, and interlaboratory comparisons using reference samples, transport, storage and analysis of samples taken;

- Training and retraining of personnel;
- Preparation of proper documentation on standard operation procedures and keeping of detailed station annals;
- Assuring permanent use of strictly standardized instrumentation, and sampling and analytical procedures;
- Organizing of laboratory intercomparisons and repeated analyses of unknown test samples; similarly, pairs of identical samples are to be supplied to the laboratory to estimate analytical precision; to this end, samples big enough to allow partitioning into several equal amounts separately to be analyzed; correspondingly, filters used for gravimetric determinations should be weighed repeatedly, at least occasionally, before and after exposure; if humidities are different inside the balance and in the desiccator, dramatic weighing errors are possible since humidity is taken up rapidly;
- Whenever changes are made regarding siting, instrumentation, or procedures, and if technically feasible, both former and new variant should be run simultaneously for several successive sampling intervals.

Data reporting and data handling are other sources of errors. Recording data (additionally) on strip charts offers a very convenient and safe way to discover invalid or suspicious data by inspection of the analog signal, as compared to software controlled tests that are no replacement for a perceptive, intuitive, and experienced observer.

All data to be forwarded to the QA management or data evaluation group should always be reported only in designated units and correct dimensions (stoichiometry). At the site, generally no data selection or correction is to be made. But data obtained under slightly or notably uncommon conditions must be flagged in a computer-compatible format and as agreed with the quality assurance and data handling units. Flagging could refer to data completeness, recalibration, values below limit of detection, deviations from standard operation, repaired data (according to agreed-upon rules), interferences, replacement of calibrating gases or solutions, unusual meteorological events, and so on. One flag should be reserved for nonlisted events, which are to be explained in a separate reporting sheet. Of course, there is accompanying information to be included in data reports such as date and time, type of sampler or analyzer, mode of operation, flow rates, etc.

6.6 DATA EVALUATION AND ASSESSMENT

Measuring the amount of a constituent of a continuously varying entity like ambient air is only an estimate of the true value which, very strictly speaking, remains unknown. The accuracy of a measurement and its representativity in space and time is hard to assess. This is of course true for any measurement in any discipline, the continuously agitated atmosphere being perhaps the most extreme example. Further, as long as the magnitude of inaccuracy and precision are known and well considered, the importance and the scientific value of a study must not suffer at all. It is exactly for this reason that strict quality control and responsible data assessment are so valuable and important. Appending to numbers reported their error range (and its probability) improves the confidence a report deserves and demonstrates the authors' competence.

Here some remarks should be made regarding use and interpretation of data obtained in atmospheric monitoring.

Investigating frequency distributions was already mentioned under siting to assist discrimination among competing sites. It is an illustrative means to learn how the data are distributed, but can also be used to detect significant changes that may occur in the system. For example, when frequency distributions are established regularly, the one-peak distribution for a gaseous constituent may turn into a two-peak one quickly. If there was no change in source allocation, the reason may be that a calibrating gas was replaced. The new tanks' designated concentration may be erroneous or the certified concentration of the old tank may have drifted steadily, or the old or the new tank for some reason was not tightly connected to the system. There may well be other uses for establishing frequency distributions.

Continuously operating automatic analyzers may not need much attendance when data are immediately recorded on high-capacity electronic media or transferred online to the evaluating center. If brave enough, one may not resist the temptation to operate, at least temporarily, strip chart recorders in parallel, and not be surprised to note something peculiar in one glance what a most sophisticated software program could never detect since it is built on past experience and is not imaginative.

The concentrations of constituents monitored at times are lower than the analyzer's least detectable amount. Project management should decide for each constituent how such values are treated. Reporting zero values may not be acceptable for various reasons. Instead, a certain ratio of the lowest detectable limit could be reported; in a U.N.-supported global monitoring network, two thirds was recommended.

Care has to be taken when correlating two quantities that both show a diurnal cycle. Many parameters show pronounced diurnal cycles, and one may find a significant correlation even if there was no causal relation between them. Therefore, data for both variables should be converted to moving 24-h means. The diurnal cycle will disappear, but the number of "original" values will only be a little less (data points at beginning or end of the period considered cannot be used). The variability of running means will be smoothed, but since this applies for all variables, an actually existing correlation should remain discernible.

This procedure is also very useful when one of the two variables to be correlated shows strong cycles, which is often the case when correlating aerometric variables with epidemic type information (first day of hospital admission for a particular diagnosis, absenteeism from work or kindergarten). The pronounced weekly cycle of such data (several times more cases occur on Mondays than on weekends) must be removed by calculating running 7-d-means for all variables involved.

When interpreting statistical quantities, it should also be kept in mind that a correlation coefficient existing with even a 99 or 99.9% probability, strictly speaking, cannot necessarily be regarded as proof that there exists a causal relation between the quantities compared; neither does it tell, in case there is such correlation, which variable is the independent and which is the dependent one. This question, for example, is sometimes being raised when explaining or commenting on the relation observed between (global) temperature and the mixing ratio of atmospheric carbon dioxide.

In summing up, there is no doubt that in the need or ambition to observe, monitor, measure, interpret, and anticipate by quality and quantity the kind and amount of atmospheric components and properties and their interactions the researcher is faced with unusually sensitive and complex processes, responsibilities, and new tasks regarding project planning, monitoring strategy, instruments, material, or data handling and assessment.

As compared to the classical approach and philosophies followed in traditional scientific research, in environment-related disciplines, and especially in atmospheric monitoring, recent research is endangered by public and political interest and influence. It may be materially rewarding to put a little more emphasis on selected results that meet or seem to meet the expectations of nonscientific communities which can afford to substitute profound knowledge, experience, and modesty with eloquence, fiction, and dramatizing. The researcher, in order to remain creative and ingenious, needs recognition and self-confidence; nevertheless, he must remain decent and imperturbable and persist in being his own most skeptical critic. Never should he be completely satisfied with his results. The 1922 Nobel laureate, Niels Hendrik David Bohr, put it this way: "The true doubts of the researcher start sometimes with certainty."

REFERENCES

1. M. Katz, Ed., *Methods of Air Sampling and Analysis*, 2nd ed. American Public Health Association (APHA Intersociety Committee), 1977.
2. R.W. Boubel, D.L. Fox, D.B. Turner, and A.C. Stern, Eds., *Fundamentals of Air Pollution*, 3rd ed., Academic Press, San Diego, CA, 1994.
3. A.C. Stern, *Air Pollution*, Vol 3: *Measuring, Monitoring, and Surveillance*, 3rd ed., Academic Press, New York, 1976.
4. ASTM, Standard Practice for Planning the Sampling of the Ambient Atmosphere, ASTM Standard D1357-95 (March 1955). American Society for Testing and Materials, Philadelphia, PA.
5. ASTM, Standard Terminology Relating to Sampling and Analysis of Atmospheres, ASTM 1356-00 (March 2000). American Society for Testing and Materials, West Conshohocken, PA.
6. U.S. EPA, *EPA Quality Assurance Handbook for Air Pollution Measurement Systems*, Vol. II, Part 1, Section 10. Environmental Protection Agency, Washington, D.C., 1998.

7 Opportunities and Challenges in Surface Water Quality Monitoring

S.M. Cormier and J.J. Messer*

CONTENTS

7.1 INTRODUCTION

In a famous paper published in 1908, Kolkwitz and Marsson[1] identified 300 plant species that were indicative of the different levels of wastewater impact on receiving waters. This project was possible because even at that early date these researchers had access to a very large number of analyses from impacted water bodies whose numbers had increased in Germany since 1870 due to the rise in volume of municipal sewage and wastewater from industry and agriculture. Water monitoring in the U.S. also was becoming common enough to require some standardization. In 1889, the National Academy of Sciences published five methods for the "sanitary examination of wastewater," and in 1905 the American Public Health Association published the first edition of *Standard Methods of Water Analysis,* covering "physical, chemical,

* This chapter has been approved for publication by the U.S. Environmental Protection Agency, but the opinions are those of the authors and do not necessarily reflect the official views or policies of the agency. We thank Naseer Shafique for providing review and references on airborne and space-based sensors; Michael Paul for review of the bioassessment section; Mark Rodgers and Kristen Brenner for review of the microbiological sections; David Lattier, Ann Miracle, and Michael Blum for input on emerging molecular biology technologies; and Bhagya Subramanian for general encouragement and assistance.

1-56670-641-6/04/$0.00+$1.50

microscopic, and bacteriological methods of water examination."[2] A century later, *Standard Methods* is in its 20th (1998) edition with 1325 pages devoted to more than 400 methods. The topics include quality assurance, radioactivity, toxicity testing, statistical analyses, and electronic methods not dreamed of at the dawn of the 20th century. In this chapter, we provide a brief overview of some of the challenges and opportunities in surface water quality monitoring as we enter the 21st century.

This overview deals with monitoring intended to form the technical basis for various types of policy decisions, including identifying and analyzing environmental problems, setting goals and priorities, managing programs, and monitoring for results (see Chapter 22 this book). By monitoring, we mean the systematic collection of data for the purpose of checking on the physical, chemical, and microbiological suitability of water for human consumption, recreation, and habitat. The U.S. Environmental Protection Agency (EPA) recently has developed and updated its technical support materials on water monitoring to support policy decisions by the states. This overview relies heavily on many of these materials, which are available on the EPA's website.[3] Rather than trying to cover the entire spectrum of water monitoring, we focus on four areas that have received a great deal of recent attention:

- Monitoring design
- Bioassessment
- Microbiological monitoring
- Remote sensing

In the first three areas, older techniques are being replaced by newer ones in order to meet increasingly stringent demands for data quality and relevance. Remote sensing has not yet been widely adopted for routine monitoring of inland waters and estuaries, but research and development has been very active worldwide, and remote sensing may soon offer a cost-effective alternative to some ground-based methods. The overview will focus on surface water monitoring, because groundwater monitoring has been well covered in Chapters 8–9, this volume.

7.2 MONITORING DESIGN

One of the first steps in designing any monitoring program is to ensure that the data will be sufficient to provide a sound technical basis for the intended policy decision. Texts by Green[4] and Ward et al.[5] provide a foundation for such designs. Green's "Ten Principles" anticipated the seven-step Data Quality Objective (DQO) approach the EPA uses to ensure that data will be adequate to support management decisions[6]:

1. Identify the problem, the decision-maker, and the time and financial constraints
2. Identify the decision to be made
3. Identify the source of data to be used
4. Define the space and time boundaries that will apply to the decision
5. Develop a decision rule that provides a decision-maker with a basis for choosing among alternative decisions
6. Specify the allowable tolerance limits for decision error
7. Optimize the sampling design to meet or exceed the DQOs

Timmerman et al. stress that early (pilot) data collection, analysis, and application to the problem should feed back into the first steps in the process in an "information cycle."[7]

Critical factors in achieving the DQOs include the representativeness of the data with respect to the decision, appropriateness and quality of the sample collection and handling methods, laboratory analysis, data management, statistical analysis, interpretation of the data (including the effect of measurement error), and communication of the information to decision-makers and the public.[8,9] *Representativeness* is the degree to which monitoring data accurately and precisely represent the variations of a characteristic either at a sampling point (e.g., a stream 100 m below a wastewater discharge) or over an entire population (e.g., all streams in a watershed or region). *Sampling design* addresses the problem of representativeness and the effect of sampling and measurement error on the water quality management decision, and is the focus of this section.

Standard Methods and manuals developed by other organizations[10] cover most of the remaining DQOs, including sample collection and handling, analytical quality, and measurement error, and the EPA recently has proposed an approach to identifying indicators of the resulting data quality.[11] Martin et al. provide an interesting look at how statistical analysis affects the comparability of information used to choose among water quality management alternatives.[12]

Sampling designs fall into two main categories, *probability designs* and *judgmental designs*.[8] Probability designs apply sampling theory to ensure that any sampling unit (e.g., sampling location) has a known probability of selection. This important feature allows the characteristics of the entire site or population not directly measured to be estimated with known uncertainty. It also ensures that the results are reproducible within that uncertainty, and it enables the water quality manager to calculate the probability of decision error. Judgmental designs rely on expert knowledge or judgment to select the sampling units. They can be easier and less expensive to implement than probability sampling. (Locations chosen at random can be difficult or even impossible to access.) The representativeness of the results depends on the quality of the professional judgment, but even in the best of cases there is no way to make quantitative estimates of uncertainty.

The value of quantitative uncertainty estimates cannot be over-emphasized because all but the most simple water quality management decisions involve substantial uncertainty. For example, most water bodies in the U.S. are routinely sampled once or a few times a year, and the resulting data are used by the states to determine whether they meet the chemical and/or biological standards appropriate to their designated use (e.g., fishing, swimming).[13] What is the resulting uncertainty, based on some number of samples per year, that the water body does or does not meet the standards, and thus may be underprotected or subject to unnecessary management costs? What about the water bodies that are never sampled? Can estimates be made with known uncertainty based on the water bodies that are sampled? For instance, the U.S. EPA significantly reduced sulfur dioxide emissions from utility boilers under the Clean Air Act Amendments of 1990. Since then, it has considered whether additional reductions were needed to ensure the recovery of acidified lakes (see Chapter 22 this book). How confident can we be, based on measurements of 50 lakes made at the beginning and end of the period, about recovery of all 6500 lakes in a region? The allowable tolerance limits for any water management decision (Step 6 in the DQO process) depend on reliable uncertainty estimates relating the monitoring data to the characteristics of the water body or population of water bodies.

Substantial progress has been made in the U.S. in implementing probability designs in water quality monitoring in the last decade. The National Surface Water Survey made reliable estimates of the numbers of acid lakes and streams in the U.S. in the late 1980s.[14,15] The Environmental Monitoring and Assessment Program (EMAP) has produced statistically reliable estimates of measures of benthic community condition, sediment toxicity, and water clarity for virtually all estuaries in the U.S., and for streams in the Mid-Atlantic highlands (see Chapter 29 this book). A recent report used a design linking a judgmental sampling design to a probability sample to show that the number of acidic lakes had declined in two regions by 33% in the 1990s.[16] Some states have begun to use probability designs in conjunction with judgment samples in assessing attainment of designated uses and reporting on progress,[17] and the EPA recently has recommended that all states do so.[18] These examples all involve estimating the condition of populations of lakes, streams, or estuaries, but probability designs apply equally to sampling specific water bodies, e.g., for the National Pollution Discharge Elimination System (NPDES), permitting or assessing the status of threatened and endangered species.[8]

Probability sampling is not simply choosing sampling locations at random. EPA,[8] Green,[4] Ward et al.,[5] and Helsel and Hirsch[19] provide guidance on alternative sampling designs. Olsen et al.[20] discuss how most judgmental monitoring designs can be improved by probabilistic sampling, and Overton et al.[21] show how "found" monitoring data from judgmental samples can be used to augment a probability sample. Larsen et al.[22] show how allocating sampling effort within and among sites in a network can help to separate effects of within- and among-year variations from site-specific differences (usually the characteristic that drives water management decisions). They note that the effect on trend detection of year-to-year variability across sites caused by regional phenomena such as droughts or temperature extremes is not responsive to design choices. If the effect is large, it may be necessary to extend the time frame for decision-making based on monitoring data, again emphasizing the iterative nature of monitoring design.

Monitoring designs to determine loads of constituents leaving a watershed or entering another water body may be very different from those appropriate for water quality assessment.[23] The U.S. Geological Survey has developed substantial expertise in this area, since 1995, through the operation of the National Stream Quality Accounting Network (NASQAN) at 39 stations in four of the largest river basins of the U.S. A special issue of the journal *Hydrological Processes*[24] provides examples of the approaches used by NASQAN and their associated limitations in monitoring loads in rivers and streams. The usual approach to calculating fluxes is to develop regression models relating concentrations to discharges and seasons in order to interpolate between chemical observations.[25] The model statistics can be quite complex and may differ considerably for suspended sediments, dissolved constituents, and constituents that interact with suspended sediments.[26,27] Smith et al. provide an example of how flux data can provide important information about the sources of pollutants in watersheds and about instream processes that remove them as they travel downstream.[28]

7.3 BIOASSESSMENT

The goal of the U.S.'s Clean Water Act (CWA) is "to restore and maintain the physical, chemical, and biological integrity of the nation's water."[29] The CWA requires each state to establish designated uses for all water bodies in a state (one of which is support of aquatic life), to establish water quality standards needed to support the designated uses, and to list those that do not meet the standards. The causes of impairment for listed streams must be identified, and Total Maximum Daily Loads (TMDLs) of pollutants must be established that would not violate the standards. Implementation plans must be developed to achieve all TMDLs, and the states must assess and report biannually on attainment of designated uses for their water bodies. Bioassessment, the process of evaluating the biological condition of a water body based on surveys of the invertebrates, fish, algae, and/or aquatic plants, plays an increasingly important role in all of these requirements.

Although Kolkwitz and Marsson[1] developed a system of plant and animal bio-indicators almost a century ago, monitoring and assessment in the early days of water quality management focused heavily on physical, chemical, and microbiological characteristics at points above and below outfalls of point sources. As modern wastewater treatment brought chronic, point source pollution under control, problems involving nonpoint sources, infrequent events associated with storms and spills, and habitat alteration became relatively more important causes of nonattainment of water quality standards. The latter problems require different monitoring approaches than the former. Bioassessment, which integrates the short- and long-term effects of point and nonpoint pollution, events, and habitat alteration, is an important example. It is also a more direct indication of the biological integrity of surface waters than physical or chemical properties alone. As a result, bioassessment has become an integral part of programs in the U.K.,[30] Europe,[31] Australia,[32] Canada,[31] New Zealand,[33] and South Africa (see Chapter 28, this volume), and based on inquiries to the authors, there is a growing interest in South America and Asia. In the U.S., 57 of 65 states, tribes, and other entities that responded to a survey in 2001 reported that they used bioassessment for water resource management and interpreting aquatic life use attainment.[34]

This overview focuses on bioassessment in streams and shallow rivers. It relies heavily on materials presented at the EPA's first National Bioassessment and Biocriteria Workshop held in Coeur D'Alene, Idaho, in April 2003, and on EPA's Consolidated Assessment and Listing Methodology (CALM),[35] which should be consulted for primary references not cited in this overview. Bioassessment in streams and wadeable rivers also is described in recently updated EPA guidance,[36] and tutorials can be accessed on the EPA's Watershed Academy Website.[37] Bioassessment in lakes and reservoirs,[38] wetlands,[39] and estuaries[40] is based on the same principles but lags behind bioassessment in streams, perhaps because of the higher cost of collecting biological samples in these systems.

Bioassessment typically uses either *multimetric* or *multivariate* approaches. Metcalfe-Smith et al.[31] and Reynoldson et al.[41] provide useful reviews. Multimetric approaches involve adding up scores for dimensionless metrics (characteristics of the biota that change in some predictable way along a disturbance gradient) to create a multimetric index. Index values for populations of reference sites (usually the least impacted streams in an ecoregion) or for streams along a disturbance gradient are used to delineate different aquatic life use classes. Multivariate approaches rely on statistical methods (e.g., discriminant analysis) to predict the probability that a biological assemblage at an assessed site is different from assemblages observed at reference sites or is a member of a particular aquatic life use class.

Karr[42] pioneered the multimetric index with his Index of Biotic Integrity (IBI) for fish communities in Ohio streams. IBIs have been developed for communities of fish, benthic invertebrates, periphyton, and even amphibians. IBIs can include various measures of taxa richness (e.g., number of species of ephemeroptera), community composition (e.g., percent of ephemeroptera), trophic or habit measures (e.g., number or percent of predator taxa), presence or absence of species tolerant of various kinds of pollution (e.g., number of intolerant taxa or the Hilsenhoff Biotic Index[43]), and individual organism condition (e.g., frequency of disease or deformities).

The individual metrics, which can include numbers with various units, percentages, ratios, etc., typically are scaled to unitless values ranging from 0 to 100 and calibrated against populations of reference sites or sites along a disturbance gradient. The metrics are tested to identify those that best discriminate between reference and impaired sites and to eliminate redundant metrics, and the scaled metrics then are combined into an overall unitless index. The State of Ohio uses this approach to classify its streams into one of four levels of "aquatic life use" based on two IBIs for fish and one for benthic invertebrates.[44] For each ecoregion in Ohio, the lower limit for the highest life use class is set at the 75th percentile of IBI scores for all reference sites, and for the next highest level at the 25th percentile. The IBIs have been codified into numeric biological criteria, which serve as the standards used to evaluate streams for all water programs in the state.

The State of Maryland Biological Stream Survey uses multimetric indices for both fish[45] and benthic invertebrates,[46] and has conducted an analysis of the reliability of the indices to correctly classify stream segments. They found that one composite sample was adequate to accurately characterize a stream (relative standard error of 8%), and that misclassification of streams into four life use categories occurred in only 8 of 27 streams, never by more than one category.[47] The State of Florida recently recalibrated its multimetric index by developing an independent index of human disturbance based on landscape disturbance, habitat alteration, hydrologic modification, and the presence of pollutants for each of its subecoregions. By analyzing the statistical power of the relationships between the human disturbance index and each of the biological metrics, Florida was able to adjust the metrics in its index to provide greater discriminatory power and to provide an independent basis for its index scores.[48]

Klemm et al.[49] analyzed the discriminatory power of various multimetric macroinvertebrate indices in streams in the Mid-Atlantic highlands of the U.S. and found that the number of taxa, the number of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa, and the Hilsenhoff Biotic Index performed well across the region. Paul et al.[50] evaluated various benthic multimetric indices for estuaries between Virginia and Delaware. That paper provides a good introduction to the literature on estuarine multimetric indices and a point of comparison to freshwater benthic invertebrate indices.

Rather than starting with metrics, the multivariate RIVPACS (River Invertebrate Prediction and Classification System) approach developed by Moss et al.[51] begins with taxa lists for a sample of reference sites. Cluster analysis is used to classify the reference sites according to their biological similarity. A multivariate discriminant model is used to predict the probabilities of reference sites belonging to different biologically defined classes, and also the probability of capturing each taxon at any particular site, based on natural environmental features such as watershed area and elevation.[52] The model produces a list of expected taxa (E) for streams in each natural class. The number of taxa observed (O) in a stream to be assessed is compared to the number of expected taxa (E) to produce a ratio, O/E, and statistical tests are conducted to determine if O/E is significantly different than the reference condition, based on sampling theory and model error.[53]

The State of Maine uses a slightly different multivariate approach.[54] Independent experts were given taxa lists from 376 unidentified streams in a calibration data set. The experts placed each stream into one of four aquatic life use classes based on taxa abundance, richness, community structure, and ecological theory. The taxa lists then were used to generate hundreds of metrics, and linear discriminant models were used to assign test streams to one of the four classes based on the metrics that had the highest discriminatory power. A four-way test discriminated between the three attainment classes and the nonattainment class based on nine metrics, and a series of three two-way tests assigned the remaining streams to one of the three-tiered attainment classes based on four to seven metrics. The approach proved very good at discriminating classes and has the advantage of calculating probabilities of classification error for any assessed stream. This is a multivariate approach because even though it uses metrics, the metrics are not scaled against disturbance gradients, but are simply used by the model as discriminating variables.

Which approach works best? There are a very limited number of studies comparing multimetric and multivariate approaches. Reynoldson et al.[41] compared the strengths and deficiencies of multimetric and multivariate methods (including the procedures for establishing reference conditions) for assessing benthic assemblages. The comparison used a single data set of benthic macro-invertebrates identified with the family level collected at 37 reference sites and 6 test sites in rivers in western Canada. The precision and accuracy (ability to consistently and correctly associate a sample with a reference site) of two multivariate methods, the Australian River Assessment Scheme (AusRivAS) and the Benthic Assessment of Sediment (BEAST), were consistently higher than the multimetric methods. When streams were classified by ecoregion, stream order, and biotic group, precision ranged from 80% to 100% for multivariate models and from 40% to 80% for multimetric models. Accuracy was 100% for both multivariate models, but only 38% to 88% for the multimetric models. The authors also preferred BEAST and AusRivAS to multimetric approaches because they required no *a priori* assumptions about biological assemblages that represent reference conditions or the degree of difference required to assign a test stream to an impaired category. AusRivAS uses only the presence or absence of families and assigns a probability that a site belongs to a particular group. BEAST uses the relative numbers of families in the community, but does not calculate the probability of misassignment.

Perhaps partially as a result of the Reynoldson et al.[41] study, more states are considering multivariate approaches and are collecting data to compare the two approaches before adopting a single one for their management standards. State of Oregon scientists reported at the Coeur D'Alene Conference that they recently had compared the performance of a multimetric approach with the RIVPACS model. They concluded that the multivariate models possessed slightly better discriminatory power and exhibited less within-site variability, but that they still considered the two approaches to be complementary.

To be sure, all of these approaches have features in common. They require careful sampling protocols to insure comparability among samples from observed streams and those used to construct the indices (see previous section). Some metrics must be scaled to account for differences in watershed size, and populations of streams must be classified to separate natural factors affecting biological communities from human ones (e.g., biota differ among ecoregions, and between naturally warmwater and coldwater streams). Analysis of discriminatory power is critical in all cases to the development of indices if they are to withstand legal challenges.[55] The approaches all represent a way to rapidly, objectively, and reproducibly assign an observed water body to an aquatic life class, but they all depend at the outset on expert judgment (e.g., for selecting reference streams or assigning streams to categories based on biota). Thus, professional judgment ultimately is the basis for determining what is acceptable with respect to aquatic life uses and attainment of standards.

Although bioassessment is now widely used to determine attainment of aquatic life uses, it would be even more useful as a regulatory tool if it were able to identify the causes of impairment.[56] Some 40% of river miles in the U.S. are in nonattainment based on bioassessment and under court-ordered deadlines to develop TMDLs, but the causes of impairment (and thus the appropriate TMDLs) remain unknown for many of these streams.[57] Paul et al.[50] have calibrated benthic estuarine metrics against several important stressors (e.g., sediment toxicity and low dissolved oxygen), but most of the freshwater metrics are at best calibrated only against general indicators of stress such as the "human disturbance gradient" used by the State of Florida. Although the identification of diagnostic bioassessment indicators remains an active area of research,[58] most of the literature tends to demonstrate that the metrics used in indices are nonspecific, i.e., they respond to several stressors.[59,60] Thus bioassessment remains primarily an attainment assessment tool, but not a diagnostic one.

Sampling for most of the programs discussed above is performed using judgmental designs. The previous section on monitoring design described the value of probability designs. At the Coeur D'Alene workshop, the benefits of probability-based design were clearly recognized by the states. For example, Idaho, Florida, Vermont, Oregon, Maine, and Ohio have extensive data sets that have been collected for permits or to establish reference conditions. Some of these states intend to supplement these and other data collected, using judgmental designs for targeted questions with probability designs to assay the condition of their streams overall. Roth et al.[47] have shown how probability sampling in Maryland can be used to develop unbiased estimates with known confidence of the performance of its metrics across all streams in the state. Hopefully this trend will continue with the encouragement of the U.S. EPA for the states to include probability sampling as a component of the monitoring programs.[18]

7.4 MICROBIOLOGICAL MONITORING

A recent expert review identified waterborne disease as one of the critical challenges in the U.S. in the 21st century, and cited estimates that up to 40% of gastrointestinal illness in the U.S. and Canada may be water related.[60] Szewzyk et al.[61] reviewed the increasing importance of "emerging" pathogens from fecal sources, including *Cryptosporidium parvum*, *Campylobacter*, and rotavirus. They noted that in the absence of modern water and wastewater management, exposure to pathogens in water results primarily from contamination by human waste, but with modern sewage treatment, animal reservoirs have become an increasingly important source. They also noted that most of these pathogens are not new but may have gone previously undetected or have acquired pathogenic traits through gene transfer. Waterborne disease outbreaks (especially gastroenteritis) at bathing beaches also have been a growing concern.[62] Studies have suggested that there is a significant correlation between swimming in water with high densities of indicator bacteria and the incidence of adverse health effects.[63]

Protecting humans from pathogens from both human and animal sources, whether in drinking water supplies, in shellfishing waters, or at bathing beaches, requires monitoring methods that are sensitive and rapid. Like waters that fail to meet standards for chemicals or biological communities, waters that fail to meet microbiological standards in the U.S. also require that the fecal sources of the microorganisms be identified and reduced to a TMDL.[64] Therefore, we also need microbiological methods that can distinguish between human and animal sources.

Because pathogens are often difficult to detect in surface waters at the low levels required for infection, detection has relied primarily on microbial indicators. Such indicators should be nonpathogenic (so that they are present in fecal material at higher concentrations than pathogens), rapidly detected, easily enumerated, highly correlated with the occurrence of pathogens in wastewater, and survive similarly to pathogens in the environment. Total and fecal coliform bacteria have been long-standing indicators of choice, but that is changing because they do not necessarily correlate with the incidence of disease or with human vs. animal sources. Both microbial source tracking and sensitive, rapid detection of pathogens present significant challenges, but some progress is being made on both fronts.

The EPA recently has issued new guidelines and performance indicators for beach monitoring based on its Ambient Water Quality Criteria for Bacteria — 1986, which showed that either *E. coli* or enterococci are better indicators of fecal contamination than coliforms or fecal coliforms at freshwater beaches, while enterococci alone are best at coastal beaches.[65] Beaches at the highest risk should be monitored one or more times a week during the swimming season at one or more points on the beach at knee depth. Lower risk beaches should be sampled once a week or less depending on proximity to suspected pollution sources, beach use, historical water quality, and other risk factors. Additional sampling is recommended after a standard is exceeded following a sewage spill or rain event, or after a beach is reopened. The membrane filter tests for both groups of organisms involve one or two filtration steps and an overnight incubation of the filter on an appropriate medium, which does not allow virtual real-time monitoring.[65]

State, local, and tribal governments must provide documentation to support the validity of methods that they wish to use other than those currently recommended or approved by EPA, which has led to several recent evaluations of the indicators. Kinzelmann et al.[66] compared *E. coli* against enterococci using two available enterococci methods at five swimming beaches in southwest Wisconsin. Each beach was sampled 5 d per week. Not only did they find very little correlation among indicators, but based on EPA's single event guidelines (235 CFU/100 ml for *E. coli*, and 61 CFU/100 ml for enterococci), using enterococci as an indicator would have resulted in 56 more beach closures during the monitoring period than the *E. coli* indicator. California changed its standard for ocean recreational beaches in 1999 from total coliforms alone to a combination of total coliforms, fecal coliforms, and enterococci, and Noble et al.[67] compared the indicators in a study of over 200 sites during both a dry period and one day after a rain event. They found that in marine waters, 99% of the exceedences of standards were detected with enterococci, compared to only 56% for fecal coliforms and 40% for total coliforms. Twice as many exceedences occurred in wet weather and more than five times as many in dry weather when the enterococci standard was used instead of the former total coliform standard. Additional evaluations of this type will certainly reduce the risk of a day at the beach, but the techniques still require more than 24 h from sample collection to reporting. Research is ongoing to develop near real-time tests, but they remain a challenge.

Another challenge of microbiological monitoring is distinguishing between fecal wastes from animal and human sources. Microbial source tracking requires that indicator organisms be distinct among the various sources of fecal contamination in watersheds, including humans and various farm animals, pets, and wildlife. Two recent reviews provide a look at the state-of-science of microbial source tracking as of 2002. Scott et al.[68] provide a broad treatment of the range of then-current approaches and future directions, while Simpson et al.[69] focus more directly on the strengths and weaknesses of the various techniques for developing TMDLs in the U.S. Both reviews include tables comparing strengths and weaknesses of various approaches. This section relies heavily on these reviews, which should be consulted for details and primary sources where specific references are not otherwise indicated below.

The ratio of fecal coliforms (higher in humans) to fecal streptococci (higher in animals) was a rapid, simple test widely used beginning in the 1970s to distinguish between human and animal sources. During the last decades, evidence has mounted that differences in prevalence and survival rates among the coliforms (and *E. coli* in warm environments) make them unreliable indicators of fecal contamination in surface water. The ratio has been questioned because it is affected by the variable survival of fecal streptococci, disinfection of wastewater, different detection methods, mixed sources, pH of >9 and <4, and salinity. A number of microbial and chemical indicators have been evaluated. *Bifidobacterium* is common in humans but rarely found in animals, especially isolates capable of fermenting sorbitol. This anaerobic bacterium cannot reproduce in surface waters, which avoids false positives, but it can have very low survival rates, thus leading to false negatives. The HSP40 bacteriophage, a virus that infects *Bacteroides fragilis*, a bacterium very specific to humans, is easy to detect, and the phage cannot replicate in the environment, but it has not been found in highly polluted waters in some areas. Two groups (II and III) of F+ RNA coliphages (viruses that infect *E. coli*) have been associated with human fecal wastes and sewage, while a different group (IV) has been associated with animal wastes, and the groups can be further differentiated by immunologic methods. The method is easy to perform, but requires a concentration step, and not all human wastes contain the phage.

Human enteric viruses also are not reliably indicated by coliform counts.[68] Techniques to measure enteric viruses directly have been demonstrated in coastal waters and tap water, but concentrations are low, and methods are labor intensive. Because more than 100 viruses are known to be associated with humans, the absence of a few may not indicate that the water is free of pathogens. Cysts or oocysts of protozoa, such as *Giardia* and *Cryptosporidium*, also can come from either human or animal sources, and although detection methods for the cysts have been greatly improved in the past decade,[70] there do not appear to be methods to separate the sources. Chemical indicators such as caffeine (from beverages) and fecal sterols and stanols (products of human digestion) have been proposed as indicators, but their fates in receiving waters are not well understood. Because they are present at very low levels, analysis is expensive.

The more promising approaches to distinguishing between human and animal fecal sources appear to be phenotypic and genotypic methods. Phenotypic methods include multiple antibiotic resistance (MAR) and immunologic methods. MAR relies on the fact that antibiotics used by humans and in animal therapy and animal feed confer different patterns of antibiotic resistance on their respective gut flora. Target organisms are isolated, cultured, and plated onto media containing a suite of antibiotics at different concentrations. Their susceptibility "fingerprint" is then compared to a reference library derived from human and animal isolates. Although the technique has been shown to discriminate *E. coli* from humans and a number of different animal species, it requires large and specific libraries for each geographic area tested (reflecting geographic differences in antibiotic exposure), and it may not be effective if antibiotic resistance is carried on plasmids or in cases where microbes from different species show no distinct patterns of antibiotic resistance.[69] Immunologic methods involve testing for the presence of somatic antigens, which have been shown to differ among *E. coli* from human and various animal sources. Parveen et al.[71] correctly identified 77% of 100 human and nonhuman source isolates, and found little overlap between human and animal-derived isolates but, again, the technique requires large libraries of antisera.

Nucleic acid-based methods that employ variations on the Polymerase Chain Reaction (PCR) have the advantage of not requiring the target microorganisms to be cultured, and coupled with fluorescence-based detection methods, they are capable of near–real time monitoring.[69] Examples of applications of several of these techniques are discussed, along with their pros and cons, in the reviews by Scott et al.[68] and Simpson et al.[69] They include ribotyping, length heterogeneity PCR, terminal restriction fragment length polymorphism (T-RFLP), repetitive PCR (rep-PCR), denaturing gradient gel electrophoresis, pulsed-field gel electrophoresis (PFGE), and amplified fragment-length polymorphism (AFLP). Some of the techniques are library independent, and thus may be more universally applicable than MAR and immunologic techniques. The primary drawback to the PCR-based techniques is interference from cations and humic substances, which becomes worse if sample concentration is required. They conclude that MAR, rep-PCR, AFLP, and T-FLP seem to offer the best prospects for distinguishing between different sources of fecal bacteria, with MAR being the most practical.

Scott et al.[68] provide several criteria such as TMDLs which are important in microbial source tracking used in water quality management decisions. These generally involve confidence that the microorganism came from the presumed source, the ability to distinguish between sources, the stability of the indicator over time and matrix (does the genetic make-up or antibiotic resistance of the indicator change in the environment?), and the geographic scope (can the indicator be used outside a small watershed with well-defined sources of fecal contamination?). They point out the importance of monitoring design (often the victim of limited budgets) in ensuring that adequate numbers of fecal sources and isolates per fecal source are examined to provide the desired level of statistical certainty in the source identification. The relationships of the indicators to anticipated public health outcomes, ease of communication to the public and decision-makers, and the costs and logistics of the techniques (they note that most of the PCR-based techniques are technically demanding and many require expensive equipment) are also important. They conclude that most of these techniques need to be evaluated on larger watersheds and with more and varied sources of fecal pollution before they are ready for routine application, and that indices of correct classification of source indicators should be calculated in a similar manner across watershed studies for adequate comparison.

7.5 REMOTE SENSING

The cost of ground-based monitoring represents a significant impediment to water quality monitoring. The Association of State and Interstate Water Pollution Control Administrators has estimated that available funding for water quality monitoring in the U.S. is less than half what is needed.[72] Remote sensing has long been seen as a way to reduce monitoring costs. It also is unique in terms of offering representativeness without a probability design because the entire water body or population of water bodies can be measured. While remote sensing has made substantial contributions to monitoring the open oceans, its contributions to monitoring water quality in estuaries and inland waters has been more modest.

One of the primary problems has been the trade-off between resolution and return period for satellite-based sensors. For the past decade, the available sensors offered frequent returns and low resolution (e.g., NOAA–AVHRR—1.1 km resolution, returns every several hours) or higher resolution with longer return periods (e.g., Landsat MSS and Landsat TM — 30 m resolution, returns every 16 d, and SPOT HRV — 20 m resolution, returns every 26 d). Relative to the temporal and spatial dynamics of processes in most lakes, streams, and estuaries that can be monitored from space (e.g., algal blooms, turbidity plumes), space-based sensors have not yet proved particularly useful for routine monitoring of water quality in these systems.[73] Sensors borne on aircraft are capable of finer resolution, but the costs of operating aircraft for routine monitoring are prohibitive. Despite these limitations, research being conducted on aircraft-based sensors may soon be applicable to space-based sensors.

The U.S., Russia, France, the U.K., India, Germany, and Japan operate satellite-based sensors,[74] and most other countries have access to at least some of the data. Our brief review of the widely available scientific literature published in 2002 and early 2003 revealed more than two dozen papers on remote sensing of inland and inshore waters by investigators in more than a dozen countries, covering applications ranging from estuaries to salt ponds to wastewater lagoons. Table 7.1 provides examples of some of these studies. Much of the research and development has concentrated on measurement of chlorophyll-*a* and other pigments (indicators of phytoplankton blooms), and measurement of suspended sediment, turbidity, and colored dissolved organic matter (indicators of runoff plumes, which often are associated with other pollutants). The most commonly used space-based sensor for these studies is the Landsat Thematic Mapper (TM). Research is active on other space-based and airborne sensors, however, including hyperspectral sensors like the airborne Compact Airborne Spectrographic Imager (CASI). CASI measures several hundred narrow wavelengths, rather than 4 to 7 broad ones, thus greatly expanding the capability to differentiate between objects with similar reflectance spectra.

Most of the projects in Table 7.1 involve individual applications rather than monitoring systems, but Kloiber et al.[75] have evaluated the possibility of monitoring lake clarity on a regional scale using Landsat MSS and TM imagery. Brightness values over the period 1973 to 1998 were calibrated against historical data on Secchi disc transparency, and analysis of data from a large population of lakes was used to adjust synoptic satellite data from different dates to a common, late-summer reference date. By analyzing historical weather data, Kloiber et al.[76] had already determined that there was a very high probability of acquiring at least one clear sky image over the region on one of the three passes of the satellite during 12 of the 16 years covered by the study. The data revealed that only 10% of the assessed lakes in the region showed significant trends in estimated Secchi depth, with more than twice as many lakes tending toward increasing transparency. This approach demonstrated that the obstacles of coarse resolution and infrequent return periods (many of which were obscured by clouds) can be overcome, at least for lakes, if the measure of interest is long-term trends in a variable that is not highly dynamic in time and space. Evaluating remote sensing in the context of the entire monitoring design should be more widely practiced.

TABLE 7.1
Recent Remote Sensing Applications in Inland and Near-Shore Water Quality Monitoring

Investigator	Sensor	Location of Application	Pixel Size	Return Period	Water Quality Parameters
Space-Based Sensors					
Keiner, L. and Xiao-Hai, Y., A neural network model for estimating sea surface chlorophyll and sediments from thematic mapper imagery, *Remote Sens. Environ.*, 66, 153, 1998.	Landsat TM	Delaware Bay	30 m	28 d	CHLA and SS
Kloiber, S., Brezonik, P., and Bauer, M., Application of Landsat imagery to regional-scale assessments of lake clarity, *Water Res.*, 36, 4330, 2002.	Landsat TM	Minneapolis Lakes	30 m	28 d	CHLA and SS
Ritchie, J. and Cooper, C., Remote sensing of water quality: application to TMDL, in *Proc. of the TMDL Science Issues Conference*, Water Environment Federation, Alexandria, VA, 2001, pp. 367–375.	Landsat TM	Lake Chicot, AK	30 m 60 m (T)	28 d	CHLA, SS, and TEMP
Nellis, M.D., Remote sensing of temporal and spatial variations in pool size, suspended sediment, turbidity, and secchi depth in Tuttle Creek Reservoir, Kansas: 1993, *Geomorphology*, 21, 281, 1998.	Landsat TM	Tuttle Creek Reservoir, Kansas	30 m	28 d	SS, turbidity, and Secchi depth
Brivio, P., Giardano, C., and Zilioloi, E., Validation of satellite data for quality assurance in lake monitoring applications, *Sci. Total Environ.*, 268, 3, 2001.	Landsat TM	Lake Iseo and Lake Garda, Italy	30 m	28 d	Quality assurance of spectral data
Dekker, A., Vos, R., and Peters, S., Comparison of remote sensing data, model results, and *in situ* data for total suspended matter (TSM) in southern Frisian lakes, *Sci. Total Environ.*, 268, 197, 2001.	Landsat TM and SPOT-HRV	Frisian lakes, Netherlands	30 m, 20 m	28 d 26 d	SS
Lafon, V. et al., SPOT shallow water bathymetry of a moderately turbid tidal inlet based on field measurements, *Remote Sens. Environ.*, 81, 136, 2002.	SPOT	Lagoon of Arcachon, France	20 m	26 d	CHLA, phaeopigments, SS, and CDOM

Dor, I. and Ben-Yosef, N., Modeling effluent quality in hypertrophic wastewater reservoirs using remote sensing, *Water Sci. Technol.*, 33, 23, 1996.	SPOT Multi-spectral	Wastewater lagoons, Israel	20m 10m	26 d	CDOM and SS
Choubey, V., Monitoring turbidity with IRS-1A data, *Hydrol. Process.*, 11, 1907, 1997.	IRS-1A-LISS-I	Tawa Reservoir, India	72.5 m	22 d	Turbidity
Woodruff D. et al., Remote estimation of water clarity in optically complex estuarine waters, *Remote Sens. Environ.*, 68, 41, 1999.	AVHRR	Pamlico Sound, North Carolina	1 km	1 d	CHLA, SS, and CDOM
Froidefond, J. et al., Spectral remote sensing reflectances of coastal waters in French Guiana under the Amazon influence, *Remote Sens. Environ.*, 80, 225, 2002.	Sea WiFS	Coastal French Guiana/ Amazon River	1 km	2 d	CHLA, water color, SS, and minerals
D'Sa., E. and Miller, R., Bio-optical properties in waters influenced by the Mississippi River during low flow conditions, *Remote Sens. Environ.*, 84, 538, 2003.	Sea WiFS	Mississippi River/ Gulf of Mexico	1 km	2 d	CHLA, phytoplankton, and CDOM
Airborne Sensors					
Koponen, S. et al., Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data, *Remote Sens. Environ.*, 79, 51, 2002.	AISA	Lakes, Finland	2 m	N/A	CHLA, turbidity, and Secchi depth
Mustard, J., Staid, M., and Fripp, J., A semianalytical approach to the calibration of AVIRIS data to reflectance over water application in a temperate estuary, *Remote Sens. Environ.*, 75, 335, 2001.	H-AVIRIS	Narragansett Bay, RI	10 m	N/A	Phytoplankton, CDOM, and SS
Richardson, L., et al., The detection of algal photosynthetic accessory pigments using airborne visible-infrared imaging spectrometer (AVIRIS) spectral data, *Data Mar. Technol. Soc. J.*, 28(11), 1994.	AVIRIS	San Francisco Bay, CA	20 m	N/A	Algal pigments
Herut, B. et al., Synoptic measurements of chlorophyll-*a* and suspended particulate matter in a transitional zone from polluted to clean seawater utilizing airborne remote sensing and ground measurements, Haifa Bay (SE Mediterranean), *Mar. Pollut. Bull.*, 38, 762, 1999.	H-CASI	Haifa Bay, Israel	3 m	N/A	CHLA and SS

(*continued*)

TABLE 7.1 (Continued)
Recent Remote Sensing Applications in Inland and Near-Shore Water Quality Monitoring

Investigator	Sensor	Location of Application	Pixel Size	Return Period	Water Quality Parameters
Shafique, N. et al., The selection of narrow wavebands for optimizing water quality monitoring on the Great Miami River, Ohio, using hyperspectral remote sensor data, *J. Spatial Hydrol.*, 1, 1, 2001.	H-CASI	Great Miami River, OH	3 m	N/A	CHLA and turbidity
Thiemann, S. and Kaufmann, H., Lake water quality monitoring using hyperspectral airborne data — a semiempirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany, *Remote Sens. Environ.*, 81, 228, 2002.	CASI and HyMap	Mecklenburg Lake District, Germany	3 m 10 m	N/A	CHLA, Secchi depth
Gould, R. and Arnone, R., Coastal optical properties estimated from airborne sensors: reply to the comments by Hu and Carder, *Remote Sens. Environ.*, 79, 138, 2002.	CASI	Fort Walton Beach, FL	2 m	N/A	Optical properties of water, absorption, and scattering coefficients
Anstee, J., Jupp, D., and Byrne, G., The shallow benthic cover map and optical water quality of Port Phillip Bay, 4th International Conference on Remote Sensing for Marine and Coastal Environment, Orlando, FL, March 17–19, 1997.	CASI	Phillip Bay, Australia	5 m	N/A	Shallow benthic cover map, and optical properties of water
Althuis, J., Suspended particulate matter detection in the North Sea by hyperspectral airborne remote sensing, *Aquatic Ecol.*, 32, 93, 1998.	CASI	North Sea	1 m	N/A	CHLA, water color, SS, and minerals
Hakvoort, H. et al., Towards airborne remote sensing of water quality in The Netherlands — validation and error analysis, *ISPES J. Photogramm. Remote Sens.*, 57, 171, 2002.	H-EPS-A	Coastal zone, Netherlands	3 m	N/A	CHLA, SS, and CDOM

Hedger, R., Olsen, N., Malthus, T., and Atkinson, P., Coupling remote sensing with computational fluid dynamics modelling to estimate lake chlorophyll-*a* concentration, *Remote Sens. Environ*., 79, 116, 2002.	ATM	Lake Levin, Scotland	10 m	N/A	CHLA, phytoplankton, and CDOM

Landsat TM: Landsat Thematic Mapper
Spot: System pour l'Observation de la Terre
HRV: High Resolution Visible
IRS-1A-LISS-I: Indian Remote Sensing Satellite 1A-Linear Imaging Self Scanning
SeaWiFS: Sea-viewing Wide Field-of-view Sensor
AVHRR: Advanced Very High Resolution Radiometer
AISA: Airborne Imaging Spectrometer for Applications
CASI: Compact Airborne Spectrographic Imager
AVIRIS: Airborne Visible to Infra Red Imagine Spectrometer
EPS-A: Environment Protection System Series A
ATM: Airborne Thematic Mapper
CHLA: Chlorophyll α
CDOM: Colored Dissolved Organic Matter
SS: Suspended Solids

Notes: Landsat TM: Landsat Thematic Mapper; SPOT: System pour l'Observation de la Terre; HRV: High Resolution Visible; IRS-1A-LISS-I: Indian Remote Sensing Satellite 1A — Linear Imaging Self-Scanning; AVHRR: Advanced Very High Resolution Radiometer; AISA: Airborne Imaging Spectrometer for Applications; CASI: Compact Airborne Spectrographic Imager; AVIRIS: Airborne Visible to Infrared Imaging Spectrometer; EPS-A: Environment Protection System Series A; ATM: Airborne Thematic Mapper; CHLA: Chlorophyll-*a*; CDOM: Colored Dissolved Organic Matter; SS: Suspended Solids.

7.6 CONCLUSIONS

We have been able to touch briefly on only a few areas of water quality monitoring in this overview. Many others are rapidly evolving. The development of rapid and inexpensive gene chip technology may soon provide information about recent exposures of aquatic communities to specific pollutants (e.g., Miracle et al.[78]). Molecular methods are being evaluated that will provide information about the genetic diversity of aquatic populations, a more sensitive indicator of vulnerability than enumerating the presence of different taxa (e.g., Forbes[79]). Remote sensing from space is being supplemented by remote sensing techniques based on *in situ* sensors, including biological early warning systems that employ a variety of aquatic organisms to immediately detect pollutants or bioweapons in water entering treatment plants (e.g., Allen et al.[80]). Advances in physico-chemical methodologies including miniaturization offer the promise of advancing even these older methods for monitoring water quality, even while the movement toward citizen monitoring of water quality presents new technical challenges to developing ever more simple but accurate monitoring methods.[81]

Try to imagine the size and scope of the 25th edition of *Standard Methods* (if indeed it is still published on paper in that year). It is likely to include advances in water monitoring that Kolkwitz and Marsson could not have imagined only 100 years ago!

REFERENCES

1. Kolkwitz, R. and Marsson, M., Ecology of plant saprobia, *Berichte der Deutchen Botanischen Geselschaft*, 26a, 505, 1908, English translation in Keup, L., Ingrahm, W., and Mackenthun, K., Eds., *Biology of Water Pollution*, U.S. Department of Interior, Washington, D.C., 1967.
2. APHA, *Standard Methods for the Examination of Water and Wastewater*, 15th ed., American Public Health Association, Washington, D.C., 1980.
3. www.epa.gov/OST
4. Green, R., *Sampling Design and Statistical Analysis for Environmental Biologists*, John Wiley, New York, 1979.
5. Ward, R., Loftis, J., and McBride, G., *Design of Water Quality Monitoring Systems*, Van Nostrand Reinhold, New York, 1990.
6. EPA, Guidance for the Data Quality Objectives Process (QA/G-4), EPA/600/R-96/055, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C., 2000.
7. Timmerman, J., Ottens, J., and Ward, R., The information cycle as a framework for defining information goals for water-quality monitoring, *Environ. Manage*., 25, 229, 2000.
8. EPA, Guidance for Choosing a Sampling Design for Environmental Data Collections (QA/G-5s), EPA/240/R-02/005, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C., 2002.
9. National Water Quality Monitoring Council, http://water.usgs.gov/wicp/acwi/monitoring/, 2003 (variously updated).

10. For example, U.S. Geological Survey, National Field Manual for the Collection of Water-Quality Data: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 9, chaps. A1–A9, 2 v., variously paged, 1997 to present (updated and available online at http://pubs.water.usgs.gov/twri9A).
11. EPA, Peer Review Draft, Guidance for the Data Quality Indicators (QA/G-5i), Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C., 2001.
12. Martin, L. et al., The Role of Data Analysis Methods Selection and Documentation in Producing Comparable Information to Support Water Quality Management, Technical Report 01-01, National Water Quality Monitoring Council, http://water.usgs.gov/wicp/acwi/monitoring/, Colorado State University, Fort Collins, CO, 2001.
13. Water Quality: Identification and Remediation of Polluted Waters Impeded by Data Gaps, GAO/RCED-00-88, U.S. General Accounting Office, Washington, D.C., 2000.
14. NAPAP, *Acid Deposition: State of Science and Technology.* Vol. II. *Aquatic Processes and Effects*, National Acid Precipitation Assessment Program, Washington, D.C., 1991.
15. Baker, L. et al., Acid lakes and streams in the United States: the role of acid deposition, *Science*, 252, 1151, 1991.
16. EPA, Response of Surface Water Chemistry to the Clean Air Act Amendments of 1990, EPA/620/R-02/004, Office of Research and Development, Research Triangle Park, NC, 2003.
17. Indiana Department of Environmental Management, State of the Environment Report, www.in.gov/idem/soe2002/water/surface.html, 2002.
18. EPA, Elements of a State Monitoring and Assessment Program, EPA 841-B-03-003, Office of Wetlands, Oceans, and Watersheds, Washington, D.C., 2003.
19. Helsel, D. and Hirsch, R., *Statistical Methods in Water Resources, Techniques of Water-Resources Investigations of the United States Geological Survey*, Book 4, Chapter A3, U.S. Geological Survey, Reston, VA, 2002.
20. Olsen, T. et al., Statistical issues for monitoring ecological and natural resources in the United States, *Environ. Monit. Assess.*, 54, 1, 1999.
21. Overton, J., Young, T., and Overton, S., Using "found" data to augment a probability sample: procedure and a case study, *Environ. Monit. Assess.*, 26, 65, 1993.
22. Larsen, D. et al., Designs for evaluating local and regional scale trends, *BioScience*, 51, 1069, 2001.
23. For example, Blevins, D. and Fairchild, J., Applicability of NASQAN data for ecosystem assessments on the Missouri River, *Hydrol. Process.*, 17, 1347, 2001.
24. Hydrol. *Process.*, 17(2), 1089, 2001.
25. Hooper, R., Aulenbach, B., and Kelley, V., The National Stream Quality Accounting Network: a flux-based approach to monitoring the water quality of large rivers, *Hydrol. Process.*, 17, 1089, 2001.
26. Horowitz, A., Elrick, K., and Smith, J., Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme, *Hydrol. Process.*, 17, 1107, 2001.
27. Holtschlag, D., Optimal estimation of suspended-sediment concentrations in streams, *Hydrol. Process.*, 17, 1133, 2001.
28. Smith, R., Schwarz, G., and Alexander, R., Regional interpretation of water quality monitoring data, *Water Resour. Res.*, 33, 2781, 1997.
29. 22 USC, 1251–1387.
30. Wright, J., Furse, M., and Armitage, P., RIVPACS: A technique for evaluating the biological quality of rivers in the U.K., *Eur. Water Pollut. Control*, 3, 15, 1993.

31. Metcalfe-Smith, J. et al., Biological water-quality assessment of rivers: use of macroinvertebrate communities, in *The Rivers Handbook: Hydrological and Ecological Principles,* Vol. 2, Calow, P. and Petts, G., Eds., Blackwell Science, Cambridge, MA, 1994, pp. 144–170.
32. Marchant, R. et al., Classification and prediction of macroinvertebrate assemblages from running water in Victoria, Australia, *J. N. Am. Benthol. Soc.*, 16, 664, 1997.
33. Maxted, J., Evans, B., and Scarsbrook, M., Development of macroinvertebrate protocols for soft-bottomed streams in New Zealand, *N. Z. J. Mar. Freshwater Res.*, (in press) (also see http://limsoc.rsnz.org/publications.htm).
34. EPA, Summary of Biological Assessment Programs and Biocriteria Development for States, Tribes, Territories, and Interstate Commissions: Streams and Wadeable Rivers, EPA-822-R-02-048, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 2002.
35. EPA, Consolidated Assessment and Listing Methodology — Toward a Compendium of Best Practices, 1st ed., Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 2002 (available at http://www.epa.gov/owow/monitoring/calm.html).
36. Barbour, M. et al., Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, 2nd ed., EPA 841-B-99-002, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 1999.
37. http://www.epa.gov/watertrain/index.htm
38. Gerritsen, J. et al., Lake and Reservoir Bioassessment and Biocriteria Technical Guidance Document, EPA 841-B98-007, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 1998.
39. Danielson, T., Wetland Bioassessment Fact Sheets, EPA843-F-98-001, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 1998 (available at http://www.epa.gov/waterscience/criteria/wetlands/).
40. Bowman, M. et al., Estuarine and Coastal Marine Waters: Bioassessment and Biocriteria Technical Guidance, EPA-822-B-00-024, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 2000.
41. Reynoldson, T. et al., The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates, *J. N. Am. Benthol. Soc.*, 16, 833, 1997.
42. Karr, J., Assessment of biotic integrity using fish communities, *Fisheries*, 66, 21, 1981.
43. Hilsenhoff, W., An improved biotic index of organic stream pollution, *Great Lakes Entomol.*, 20, 31, 1987.
44. Yoder, C. and Rankin, E., Biological criteria program development and implementation in Ohio, in *Biological Assessment And Criteria: Tools For Water Resource Planning And Decision Making*, Davis W.S. and Simon T.P., Eds., Lewis Publishers, Boca Raton, FL, 1995, pp. 109–144.
45. Roth, N. et al., Maryland biological stream survey: development of a fish index of biological integrity, *Environ. Monit. Assess.*, 51, 89, 1998.
46. Klauda, R. et al., Maryland biological stream survey: a state agency program to assess the impact of anthropogenic stresses on stream habitat quality and biota, *Environ. Monit. Assess.*, 51, 299, 1998.
47. Roth, N. et al., Biological Indicator Variability and Stream Monitoring Program Integration: A Maryland Case Study, EPA/903/R-02/008, Office of Environmental Information and the Mid-Atlantic Integrated Assessment Program, U.S. Environmental Protection Agency, Washington, D.C., 2001.

48. Barbour, M. et al., A framework for biological criteria for Florida streams using benthic macroinvertebrates, *J. No. Am. Benthol. Soc.*, 15, 185, 1996.
49. Klemm, D. et al., Methods development and the use of macroinvertebrates as indicators of ecological condition for streams in the Mid-Atlantic highlands region, *Environ. Monit. Assess.*, 78, 169, 2002.
50. Paul, J. et al., Developing and applying a benthic index of estuarine condition for the Virginian biogeographic province, *Ecol. Indicators*, 1, 83, 2001.
51. Moss, D. et al., The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data, *Freshw. Biol.*, 17, 41, 1987.
52. Hawkins, C. et al., Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations, *J. North Am. Benthol. Soc.*, 19, 541, 2000.
53. Hawkins, C. et al., Development and evaluation of predictive models for measuring the biological integrity of streams, *Ecol. Appl.*, 10, 1456, 2000.
54. Davies, S. and Tsomides, L., Methods for Biological Sampling and Analysis of Maine's Rivers and Streams, DEP LW0387-B2002, Maine Dept. of Environmental Protection, Augusta, ME, 2002. (available at www.state.me.us/dep/blwq/docmonitoring/biomonitoring/index.htm).
55. Fore, L., Karr, J., and Wisseman, R., Assessing invertebrate responses to human activities: evaluating alternative approaches, *J. N. Am. Benthol. Soc.*, 15, 212, 1996.
56. Suter, G., II, Norton, S., and Cormier, S., A methodology for inferring the causes of observed impairments in aquatic ecosystems, *Environ. Toxicol. Chem.*, 21, 23, 2002.
57. Mehan, G. T., III, Assistant Administrator for Water, U.S. Environmental Protection Agency, Washington, D.C., 2003, http://www.epa.gov/water/speeches/031703tm.html
58. Norton, S. et al., Can biological assessment discriminate among types of stress? A case study from the eastern cornbelt plains ecoregion, *Environ. Toxicol. Chem.*, 19, 1113, 2000.
59. Simon, T., *Biological Response Signatures*, CRC Press, Boca Raton, FL, 2002.
60. Levin, R. et al., U.S. drinking water challenges in the twenty-first century, *Environ. Health Perspect.*, 110 (Supp. 1), 43, 2002.
61. Szewzyk, U. et al., Microbiological safety of drinking water, *Annu. Rev. Microbiol.*, 54, 81, 2000.
62. Rose, J. et al., *Microbial Pollutants in Our Nation's Water: Environmental and Public Health Issues*, American Society for Microbiology, Washington, D.C., 1999.
63. Pruess, A., Review of epidemiological studies on health effects from exposure to recreational water, *Int. J. Epidemiol.*, 27, 1, 1998.
64. EPA, Protocol for Developing Pathogen TMDLs, EPA-841-R00-002, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 2001.
65. EPA, National Beach Guidance and Performance Criteria for Recreational Waters, EPA-823-B-02-004, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., July 2002.
66. Kinzelmann, J. et al., Enterococci as indicators of Lake Michigan recreational water quality: comparison of two methodologies and their impacts on public health regulatory events, *Appl. Environ. Microbiol.*, 69, 92, 2003.
67. Noble, R. et al., Comparison of total coliform, fecal coliform, and enterococcus bacterial indicator response for ocean recreational water quality testing, *Water Res.*, 37, 1637, 2003.
68. Scott, T. et al., Microbial source tracking: current methodology and future directions, *Appl. Environ. Microbiol.*, 68, 5796, 2002.

69. Simpson, J., Santo Domingo, J., and Reasoner, D., Microbial source tracking: state of science, *Environ. Sci. Technol.*, 36, 5279, 2002.
70. Mc Cain, R. and Clancy, J., Modifications to United States Environmental Protection Agency Methods 1622 and 1623 for detection of *Cryptosporidium* oocysts and *Giardia* cysts in water, *Appl. Environ. Microbiol.*, 69, 267, 2003.
71. Parveen, S. et al., Genotypic and phenotypic characterization of human and nonhuman *Escherichia coli*, *Water Res.*, 35, 379, 2001.
72. Savage, R., Sample problem, *The Environmental Forum*, September/October 2002.
73. Cracknell, A., Remote sensing techniques in estuaries and coastal zones — an update, *Int. J. Remote Sens.*, 19, 485, 1999.
74. NASA, Remote Sensing Tutorial, Earth Observing System (EOS) Goddard Program Office, National Aeronautics and Space Administration, Greenbelt, MD, 2003 (available at http://rst.gsfc.nasa.gov/).
75. Kloiber, S., Brezonik, P., and Bauer, M., Application of landsat imagery to regional-scale assessments of lake clarity, *Water Res.*, 36, 4330, 2002.
76. Kloiber, S. et al., Trophic state assessment of lakes in the Twin Cities (Minnesota, U.S.A.) region by satellite imagery, *Arch. Hydrobiol. Spec. Issues Adv. Limnol.*, 55, 137, 2000.
77. Miracle, A. et al., The path from molecular indicators of exposure to describing dynamic biological systems in an aquatic organism: microarrays and the fathead minnow, *Toxicology*, 2003.
78. Forbes, V., Ed., *Genetics and Ecotoxicology*, Taylor and Francis, Philadelphia, PA, 1998.
79. Allen, J. et al., Monitoring watersheds: biomonitors and other measures, *J. Urban Technol.*, 9, 1, 2000.
80. http://www.epa.gov/owow/monitoring/volunteer/epasvmp.html

8 Groundwater Monitoring: Statistical Methods for Testing Special Background Conditions

C.J. Chou

CONTENTS

8.1 INTRODUCTION

Statistical methods are often required in groundwater monitoring programs to determine if a unit regulated by Resource Conservation and Recovery Act (RCRA) impacts groundwater quality beneath the site. Typical methods (referred to as inter-well comparisons) involve difference testing of contaminant indicator parameters (e.g., specific conductance, pH, and total organic carbon) or contaminant concentrations between upgradient and downgradient wells. In these cases, independent

1-56670-641-6/04/$0.00+$1.50

and identically distributed populations of upgradient and downgradient concentrations are assumed. The difference between upgradient and downgradient water quality is attributed to the site. However, in practice, this condition may not be met, and other methods are needed.

Three general situations where the above condition cannot be met are:

- When there is no upgradient well
- When there is natural variability in constituent concentrations among network wells
- When there is a preexisting background for the constituent of interest

The first section provides a stepwise approach for addressing the first two situations using control chart methods for what is referred to as intra-well testing of monitoring data. The method used is a sequential quality-control scheme called a combined Shewhart–CUSUM (cumulative sum) control chart. The method was first referenced by Westgard et al.[1] and further developed by Lucas.[2] For groundwater applications, testing procedures are discussed in U.S. Environmental Protection Agency (EPA) guidance,[3,4] Starks,[5] Gibbons,[6] and ASTM.[7] Chou et al.[8] applied the methodology in an RCRA-regulated facility. The control chart method will be insensitive to detect real changes if a preexisting trend is observed in the background data set. The second section provides an approach for illustrating how to deal with the third situation. In particular, how to detect a significant trend and remove that trend using a transformation suggested by Gibbons[6] are presented for addressing the preexisting background situation. Once the data are "de-trended," the control chart method can be applied.

8.2 NO-UPGRADIENT WELL AND/OR NATURAL VARIABILITY CASE

The no-upgradient well situation can occur as a result of the presence of a groundwater mound created by liquid discharges to the ground by the disposal facility being monitored (i.e., radial flow occurs in all directions away from the source), a flat water table such that there is no definable hydraulic gradient, or the loss of an upgradient well (or wells) due to changes in groundwater flow direction or declining water table (i.e., wells went dry and could not be replaced). In these cases, statistical tests designed for upgradient and downgradient comparisons that rely on the use of upgradient concentrations to establish background are not feasible.

In addition, there may be differences in constituent concentrations because of spatial variability among network wells. Spatial variability affects the mean concentrations of naturally occurring constituents among network wells but not typically the variance within each well, whereas contamination can affect both mean concentrations and variance. For example, chloride, a naturally occurring constituent, may be a constituent of concern for certain facilities (e.g., process brines, water softener regenerant, and surface runoff containing road salt). In such cases, the natural background of chloride in groundwater beneath the facility must be considered.

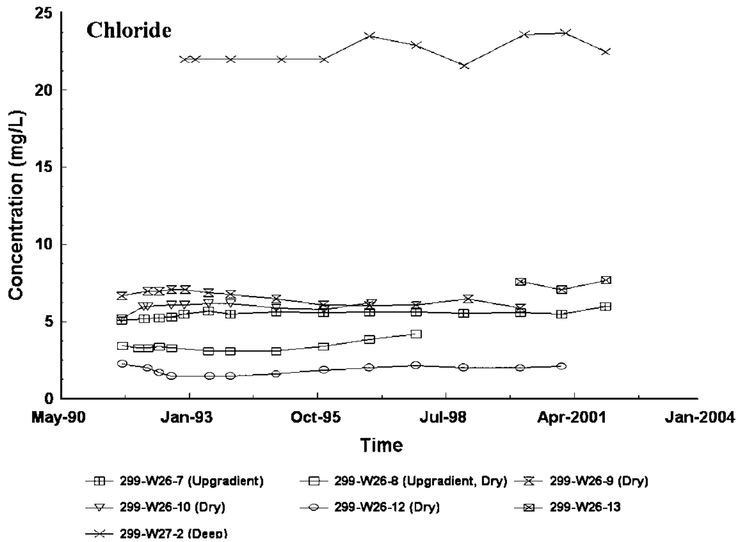


FIGURE 8.1 Chloride concentrations vs. time plot for regulated facility.

Figure 8.1 illustrates the natural variability in groundwater chloride concentrations for a monitoring facility consisting of two upgradient, four downgradient, and one deeper well. This figure shows that although the mean chloride concentrations vary from ~1.87 to 22.5 mg/L, observations over time for each well fluctuate very little from their respective mean concentrations. In this case, the spatial variability invalidates the inter-well comparison method because it assumes the only impact between the upgradient well and downgradient wells is from the facility. Intrawell testing using a control chart method to detect departures from the natural background is one alternative discussed in the following sections.

The combined Shewhart–CUSUM method can be implemented for each well in the monitoring network separately after a baseline period of eight or more independent observations is acquired. This procedure assumes that groundwater data obtained in a well during the baseline period and a future timeframe are independent and normally distributed with a fixed mean μ and variance σ^{-2}. The assumption of normality can generally be satisfied by data transformation. Even in situations where the normality assumption is violated to a slight or moderate degree, the control chart methods will still work reasonably well. The control charts will not work well if monitoring data collected from a well are correlated over time. To ensure groundwater quality data are independent, wells should not be sampled too frequently. Gibbons[6] (pp. 163, 185) recommends that groundwater not be sampled more than quarterly to reduce the likelihood of obtaining dependent data. Useful techniques to check the validity of these assumptions are provided below followed by a description of the Shewhart–CUSUM methodology that can be used once the assumptions of independence and normality are satisfied.

8.2.1 Minimum Time Interval to Obtain Independent Samples

To ensure statistical independence between sampling events, adequate time should elapse to allow the aquifer to return to an unperturbed state. Generally, the recovery time needed depends on the groundwater flow rate and the size of the disturbed zone created during a typical well purging and sampling event. The EPA[3] (pp. 3-1–3-10) suggested using the following steps for determining the minimum time needed to acquire independent samples by reference to the uppermost aquifer's effective porosity (n_e), horizontal hydraulic conductivity (K_h), and hydraulic gradient (i).

1. Calculate the horizontal component of the average linear velocity of groundwater ($\bar{v}_h$) using Darcy's equation (Freeze and Cherry[9] [p. 71])

$$\bar{v}_h = \frac{(K_h * i)}{n_e} \tag{8.1}$$

   where the dimension of $\bar{v}_h$ is the same as K_h (or distance divided by time, L/T).
2. The minimum time interval between independent sampling events is obtained by:

$$\text{Minimum Time Interval} = \frac{W_d}{\bar{v}_h} \tag{8.2}$$

   where $\bar{v}_h$ is determined in Step 1, and W_d is the diameter of the monitoring well. This approximation is most applicable to zero or low purge sampling conditions.

The Darcy equation is not valid in turbulent and nonlinear laminar flow regimes. In those cases where Darcy flow cannot be assumed (e.g., in karst or pseudo-karst aquifers, EPA[3] pp. 3–11), the groundwater velocity must be determined by more direct methods such as tracer travel time between two wells or with a flow meter.

In those cases where Darcy flow conditions apply, well purging might create a much larger effective disturbed zone diameter than just the diameter of the well screen. This can occur if large volumes of purge water are withdrawn prior to sampling. Therefore, to account for this effect, an estimated disturbed zone diameter (2 times r where r is the radius of the affected area) can be substituted for W_d in Equation (8.2) to calculate a more conservative estimate of the time interval required to obtain independent samples. An approximation for the disturbed zone radius r is as follows in Equation (8.3):

$$r = \sqrt{\frac{3V_w}{\pi * h * n_e}} \tag{8.3}$$

where V_w is the volume of water purged prior to sample collection (typically three bore volumes), h is the length of the wetted well screen, and n_e is the effective porosity. The EPA[3] (pp. 3–5) provides default values for effective porosity for use in time-of-travel analyses.

It also should be noted that when groundwater flow rate is very slow, the time required between sampling events might be impractical. For example, considering a common monitoring well diameter of 0.1016 m (4 in.) and a typical flow rate of 0.002 m/d in an aquifer with low permeability, Equation (8.2) yields a minimum time of 51 d between sampling events to ensure independence. Under this scenario, it would require 200 d to obtain four independent samples for use in the default analysis of variance method as required in the regulations. Thus, the regulatory requirement of collecting four independent samples during each semiannual period cannot be satisfied. An alternative method (such as the control chart method discussed in Section 8.2.3) that relies on collection of one sample from each well during each semiannual period should be considered in this case.

8.2.2 Check Assumption on Normally Distributed Data Set

A normal probability plot of each constituent of interest could be constructed first to examine whether a normal distribution could be used to describe the groundwater data. If a straight line can approximate these data points, a normal distribution is assumed to be a reasonable representation of the monitoring data. Also, goodness-of-fit tests such as the Shapiro and Wilk's W test and the Lilliefors test for normality of data as described in Conover[10] (pp. 357–367) are effective methods for testing the null hypothesis that these data were drawn from an underlying normal distribution. Because environmental data are highly skewed, the assumption of normality can generally be satisfied by log-transforming the data or by other techniques such as the Box–Cox transformations. Equation (8.4) gives the Box–Cox transformations:

$$
\begin{aligned}
x(\lambda) &= \frac{(x^{\lambda} - 1)}{\lambda} \; for \lambda \neq 0 \\
x(\lambda) &= \ln(x) \, for \lambda = 0
\end{aligned}
\tag{8.4}
$$

The problem now becomes how to determine the value of λ to use in the above equation. Given groundwater concentration observations $x_1, x_2, \ldots, x_n$, one way to select the λ is to use the λ that maximizes the logarithm of the likelihood function:

$$
f(x,\lambda) = -\frac{n}{2}\ln\left[\sum_{i=1}^{n}\frac{\left(x_i(\lambda) - \bar{x}(\lambda)\right)^2}{n}\right] + (\lambda - 1)\sum_{i=1}^{n}\ln x_i \tag{8.5}
$$

where $\bar{x}(\lambda) = \frac{1}{n}\Sigma_{i=1}^{n} x_i(\lambda)$ is the arithmetic mean of the power transformed data.

In practice, one can choose a value of λ from a selected range (−2, 2) or even (−1, 1) first as suggested by Draper and Smith[11] (p. 225) and extend the range later if needed. After calculating the log-likelihood function $f(x,\lambda)$ for several selected values of λ, plot the values of $f(x,\lambda)$ against λ, and draw a smooth curve through these points. The value of λ that maximizes the log-likelihood function is the maximum likelihood estimator $\hat{\lambda}$ of λ. Several transformation conditions based on λ are worth mentioning. For example, do not transform the data at all if $\lambda = 1$; a square root transformation, if $\lambda = 0.5$; an inverse transformation, if $\lambda = -1$, a log-transformation, if $\lambda = 0$. Draper and Smith[11] (pp. 226–232) provide a detailed discussion on this useful family of transformations.

8.2.3 Combined Shewhart–CUSUM Testing Procedures*

This method combines the advantages of a Shewhart control chart with that of a CUSUM control chart. It allows monitoring data from a well to be viewed graphically over time so changes over baseline conditions can be detected. The Shewhart portion of the test checks for any sudden upward shift in groundwater quality parameters based on a single observation. The CUSUM checks for a gradually increasing trend in the groundwater quality parameters. The procedure can be implemented as follows: Let x_i' be a series of independent baseline observations $i = 1, \ldots,$ n (n = 8). Let x_i be a series of future monitoring measurements $i = 1, 2, 3, \ldots$. Then, using the baseline data, the following steps are applied:

1. Using methods discussed in Section 8.2.1 and Section 8.2.2, determine if the x_i' can be assumed to follow a normal distribution with mean μ and standard deviation σ. If not, transform the x_i' using the appropriate Box–Cox transformation, and work with the transformed data.
2. Next use the baseline data to compute the estimates

$$\bar{x}_b' = \sum_{i=1}^{n} x_i'/\text{n} \text{ for } \mu \quad \text{and} \quad s_b' = \sqrt{\sum_{i=1}^{n} (x_i' - \bar{x}_b')^2/(n-1)} \text{ for } \sigma. \tag{8.6}$$

3. Determine the upper Shewhart control limit (SCL) for the procedure by calculating $SCL = \bar{x}_b' + z_s s_b'$ where z_s is a percentile of the standard normal distribution used to set the false negative and false positive values of the SCL. The value of z_s most often suggested for groundwater use is 4.5 by Starks[5] and EPA.[3] Other values may also be used, depending on the sampling scheme and whether verification sampling is used to modify the false positive and false negative error rates. If the Shewhart control scheme test were used alone, without the CUSUM portion of the test, then the false positive values of this portion of the test alone would be given by $1 - \Phi(z_s)^{\text{m}}$, which is the probability of at least one of the m comparisons (number of wells in the network times the number of water

* From Chou, C.J., R.F. O'Brien, and D.B. Barnett, Application of intrawell testing of RCRA groundwater monitoring data when no upgradient well exists, *Environ. Monit. Assess.*, 71, 91–106, pp. 97–99, 2001. Reproduced with permission of Kluwer Academic Publishers.

quality parameters) exceeding their respective SCLs, where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution.

For illustrative purposes, let us assume a waste disposal site has eight wells and three water quality parameters. Using the above-suggested value of 4.5 for z_s, this would translate to a false positive rate of 0.00008. For values of z_s being 2, 3, and 4, the respective false positive rates would be 0.42, 0.03 and 0.00076. Thus, higher values of z_s ensure that the probability of falsely declaring the site has affected groundwater quality in any one sampling period when, in fact, it has not, is small. The false negative rate for the network, for various shifts from the baseline means, needs to be computed by simulation and is computed in conjunction with the CUSUM test in Section 8.2.4.

4. Determine the upper CUSUM control limit (CCL), with $CCL = \bar{x}_b' + z_c s_b'$. The value of z_c suggested by Lucas,[2] Starks,[5] and EPA[3] is $z_c = 5$. This value can also be adjusted to reach desired false negative and false positive error rates.
5. Determine the amount of increased shift in the mean of the water quality parameter of interest to detect an upward trend. This value is referred to as k and is usually measured in σ units of the water quality parameter. Lucas,[2] Starks,[5] and EPA[3] suggest a value of k = 1 if there are fewer than 12 baseline observations and a value of k = 0.75 if there are 12 or more baseline observations.

Using the monitoring data after the baseline measurements have been established:

6. Compute the CUSUM statistic as $S_i = \max\{0, (x_i - ks_b') + S_{i-1}s_b'\}$ as each new monitoring measurement, x_i becomes available, where i = 1,2,3, … and $S_0 = 0$.
7. As each new monitoring measurement becomes available, compute the Shewhart and CUSUM tests; a verification sampling will be conducted if either $x_i \geq$ SCL or $S_i \geq$ CCL. A well is declared to be out of compliance only if the verification result also exceeds the SCL or the CCL. If both $x_i <$ SCL and $S_i <$ CCL, then continue monitoring.

If resampling is implemented during monitoring, the analytical result from the resample is substituted into the above formulae for the original value obtained, and the CUSUM statistic is updated. Note in the above combined test that the Shewhart portion of the test quickly detects extremely large deviations from the baseline period. The CUSUM portion of the combined test is sequential; thus, a small positive shift in the mean concentration over the baseline period will slowly aggregate in the CUSUM statistic and eventually cause the test to exceed the CUSUM control limit, CCL.

8.2.4 Computing Power Curves

Monte Carlo simulations should be conducted to assist in finding appropriate Shewhart and CUSUM control limits that will (1) keep the network-wide false positive rate at approximately 5% when water quality is at baseline levels, and (2) provide high power (low false negative rate) to detect real contamination when it occurs. These simulations assume that the groundwater quality data are independent and can be transformed appropriately to a normal distribution with mean 0 and standard deviation 1. Therefore, it is important to check the validity of these assumptions using methods described earlier. Chou et al.[8] used the following two-step process to generate the power curves.

1. For each incremental shift in the mean concentration, probabilities are calculated by first finding the probability of one groundwater quality parameter in one well of the network exceeding the combined Shewhart–CUSUM control limit after either the 1st, 2nd, 3rd, 4th, or 5th sampling period. The SCL and CCL values are set at various levels of interest (e.g., SCL = CCL = 3, 4, 4.5) and for shifts in the mean concentration from 0 to 5 standard deviation (sigma) units in increments of 0.2 sigma units. For each incremental shift in the mean concentrations, a large number of simulations (e.g., 10,000 runs) can be made to estimate the probability of an exceedance. In each simulation, a pseudo-random deviate is generated from a normal distribution with a fixed incremental shift in the mean and a standard deviation of 1. The probability of an exceedance in the nth sampling period (n = 1 to 5) for a specified shift in the mean is calculated as the proportion of the total pseudo-random numbers that exceeds the control limit in the nth sampling period (i.e., the number of times an exceedance occurred in the nth sampling period divided by 10,000).
2. In this step, for each incremental shift in the mean, the cumulative probability of one groundwater quality parameter in one well exceeding the control limit by at least the nth sampling period is calculated by adding the individual probabilities of an exceedance in each sampling period from n = 1 up to n = 5. When presented graphically, these probabilities are the power curves of the proposed test that should be compared with the EPA reference power curves as discussed below.

8.2.5 Power Evaluations

To assist the regulatory community in selecting alternative tests, EPA provides reference power curves EPA[4] (p. B-6) that correspond to 8, 16, 24, and 32 background samples (Figure 8.2) for comparisons with the power curves of proposed tests. Figure 8.2 shows the probability (referred to by EPA as the effective power) of detecting a shift from 0 to 5 sigma units in mean concentration for a single water quality parameter in any well in the network. The power at 0 sigma unit, called the false positive rate or size of the test, is the probability of falsely concluding that a well has increased concentrations of a water quality parameter when, in fact, it is at baseline concentrations. It should be noted that a probability of approximately 1% near 0 sigma unit is shown in Figure 8.2. This is the requirement mandated by EPA for any individual comparison; however, EPA also mandates a 5% overall network-wide false positive rate (across all wells and constituents) be kept by any

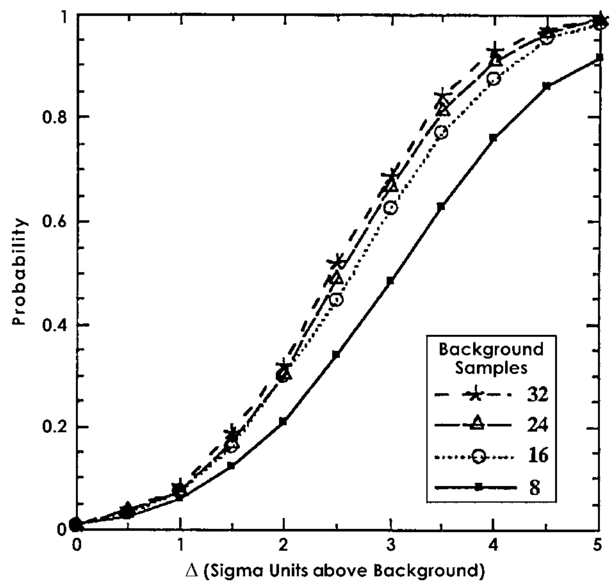


FIGURE 8.2 EPA reference power curves. [From EPA (1992), Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities — Draft Addendum to Interim Final Guidance, EPA/530-R-93-003, U.S. Environmental Protection Agency, Washington, D.C., p. B-6.]

testing scheme during each sampling event. A test is judged to be acceptable if its power is comparable to the EPA reference power curves while maintaining an approximate 5% overall false positive rate.

8.2.6 Average Run Length

Before implementing a new sequential quality control scheme, its properties need to be evaluated. The average run length (ARL) is one such criterion. The ARL is the average number of sampling events it would take for the combined Shewhart–CUSUM test to declare that the site exceeds the control limits. In general, one would like the ARL of the control chart to be long when the mean concentration is close to the baseline level and very short when the mean concentration has shifted too far from the baseline level.

The ARL for a Shewhart control chart is $1/p$, where p is the probability for a sample value to fall outside the established control limits. Therefore, for a 3 sigma control limit Shewhart chart, the probability to exceed the upper control limit is 0.00135 and to fall below the lower control limit is also 0.00135. Their sum is 0.0027 (Note: the probability values are obtained from a standard normal distribution table by setting $z = 3$ and $z = -3$, respectively.) Thus, the ARL equals $1/0.0027 = 370.37$. This means even if a site remains at baseline condition, we can expect, on the average, a false alarm every 371 sampling events. The calculations of the ARL for the combined Shewhart–CUSUM chart are quite involved, and Monte Carlo simulations are needed to obtain the ARL values for various control limits.

8.3 PREEXISTING BACKGROUND CASE

In this case, the constituent of concern (COC) exists in groundwater beneath the regulated unit, and the preexisting background concentration may show an increasing or decreasing trend due to a prior release from the site and/or an upgradient source. If the trends are not corrected, the calculated SCL and CCL will be grossly overstated and render the control chart method less effective to detect real changes. While caution must be used in applying the control chart method in the above cases, there are supplemental statistical techniques that can be used to account for preexisting trends observed in the background COC concentration. For example, Gibbons[6] (p. 165) describes a technique for removing the effects of an upward trend (or downward trend) in a single monitoring well. Once the data are detrended, the previously described control chart method can be applied.

8.3.1 Estimating Trend Using Nonparametric Method

If a concentration vs. time plot suggests the presence of a linear trend, one may estimate the true slope (change in concentration per unit of time) and the intercept (concentration at time zero) by the least squares method as described by Montgomery and Peck[12] (pp. 8–50). The underlying assumptions of the linear regression method are normally and independently distributed error terms with mean zero and constant variance. Gross violation of these assumptions may yield an unstable model. In addition, other disadvantages of linear least squares fit are poor extrapolation properties and sensitivity to outliers or gross errors. Furthermore, the presence of nondetects in the data will often invalidate the estimators (e.g., the intercept and slope) obtained by the least squares regression method. A robust method that is not greatly influenced by the presence of outliers, missing data, and nondetects is a nonparametric procedure developed by Sen.[13] Step-by-step procedures to obtain Sen's estimator of trend and to test the hypothesis of zero slope (i.e., no trend) are provided below:

1. Compute the S_{ij} sample slope estimates for each well as in Equation (8.7) below:

$$S_{ij} = \frac{y_j - y_i}{j - i} \tag{8.7}$$

   where $i < j$ and y_i and y_j are concentration values measured at time j and i, respectively. If some of the concentration data (y_i's) are nondetects, we may use one half the detect limit for these not-detected data [Gilbert[14] (p. 218)].

2. The median of these N' values of S_{ij} is Sen's estimator of trend. Let $S_{[1]} \le S_{[2]} \le \cdots \le S_{[N']}$ denote the ordered values of S_{ij}. Then Sen's estimator denoted as S is:

$$S = S_{[(N'+1)/2]} \quad \text{if } N' \text{ is odd, or} \tag{8.8}$$

$$S = \frac{1}{2}[S_{[N'/2]} + S_{[(N'+2)/2]}] \quad \text{if } N' \text{ is even} \tag{8.9}$$

   If a single measurement is collected from each well during each sampling event, then in this case $N' = \frac{n(n-1)}{2}$, where n is the number of sampling periods.
3. After obtaining the ordered values $S_{[1]} \le S_{[2]} \le \cdots \le S_{[N']}$ of the N' slopes S_{ij}, a $(1 - \alpha)\%$ distribution-free confidence interval for the true slope can be obtained as below (see Gilbert[14] [p. 218]). This procedure is based on normal theory and is valid for small sample sizes (i.e., $n \le 10$).

   (i) Set $M_1 = \dfrac{N' - C_\alpha}{2}$ and $M_2 = \dfrac{N' + C_\alpha}{2}$ (8.10)

   (ii) Compute $C_\alpha = Z_{1-\alpha/2}[VAR(S)]^{1/2}$ (8.11)

   where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)^{th}\%$ quartile of the standard normal distribution

   (iii) Compute $VAR(S) = \dfrac{1}{18}[n(n-1)(2n+5)]$, or (8.12)

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5)\right] \tag{8.13}$$

   Use Equation (8.12) if there is no tied value (i.e., nondetects or same concentration values). Otherwise, use Equation (8.13); where g is the number of tied groups and t_p is the number of occurrences of a particular tied value.

   (iv) The lower confidence limit is $\beta_L = S_{[M_1]}$, and the upper confidence limit is $\beta_U = S_{[M_2+1]}$
4. To test the hypothesis of zero slope (no trend) against the alternative hypothesis of an upward trend, ASTM[7] recommends one first compute a one-sided 99% lower confidence limit (i.e., use $Z_{0.99} = 2.326$) and reject the hypothesis of no trend if $\beta_L = S_{[M_1]}$ is greater than zero. If the alternative hypothesis is the presence of a decreasing trend, then one should reject the hypothesis of no trend in favor of the alternative hypothesis when $\beta_U = S_{[M_2+1]}$ is less than zero.

5. If a significant trend is found, one may use the methods discussed in the next section to remove the trend in the background data set.

8.3.2 Trend Removal Method

When a preexisting trend is observed in the background data set, the calculated background mean and standard deviation will be overstated and render the control chart method insensitive to detect real changes. Gibbons[6] (p. 165) suggests using the following transformation to remove a preexisting trend. Prior to proceeding with the transformation, a significant trend must be established using the method discussed in Section 8.3.1.

$$\begin{aligned} y_i^* &= \beta_0 + [y_i - (\beta_0 + \beta_1 * t)] \\ &= y_i - \beta_1 * t \end{aligned} \tag{8.14}$$

where y_i and y_i^* denote the original and the transformed data values, respectively; $i = 1, 2, \ldots, n$ denotes the sampling periods, and β_0 and β_1 are the intercept and the slope of the trend line, respectively. Note in the above transformation we only need an estimate of the true slope of the trend line, which can be obtained by using Sen's estimator. The next section provides a case study describing how a trend observed in the background data set can be removed using the transformation suggested by Gibbons.[6]

8.3.3 Case Study

Table 8.1 presents specific conductance data from a waste management site. The facility is a land-based RCRA disposal unit with seven upgradient wells and 10 downgradient wells. For illustration purposes, only one upgradient and one downgradient well are used. Downgradient specific conductance data obtained on June 3, 1998, was not used in the statistical analysis because it proved to be a gross error due to a meter calibration problem. The observed upward specific conductance trends in both wells are caused by an offsite nitrate plume upgradient of the facility (e.g., the 216-B-62 Crib). Locations of monitoring wells and specific conductance contours near the facility are presented in Figure 8.3.

Using procedures described in Section 8.3.1, we first calculated 28 and 21 slopes for the upgradient well and downgradient wells, respectively, and arranged them in increasing order. From the ordered sets of calculated slopes, we calculated their median values (S), and the lower 99% confidence limits. The results are shown in Table 8.2. For the upgradient well, the median slope is the average of the 14th and 15th slope values of the ordered set and $VAR(S) = \frac{1}{18}[n*(n-1)*(2n+5)] = \frac{1}{18}*8*7*21 = 65.33$ (μS/cm per 6-month period) because there are no tied data values. Furthermore, given $\alpha = 0.01$ ($Z_\alpha = 2.326$, one-sided case), M_1 equals 4.6. From the list of ordered sample slope values provided in Table 8.2, we obtain $S_{[4]} = 7$ and $S_{[5]} = 7.5$; hence, $S_{[4.6]} = 7.3$ (μS/cm per 6-month period) found by linear interpolation. For the

TABLE 8.1
Specific Conductance Data from a Waste Disposal Facility

Well	Sample Date	Time Period	Specific Conductance (μS/cm)
Upgradient	06/02/98	1	379
	12/03/98	2	386
	06/03/99	3	404
	12/15/99	4	410
	05/23/00	5	426
	12/18/00	6	432
	06/05/01	7	444
	12/04/01	8	449
Downgradient	06/03/98	1	289[a]
	12/03/98	2	435
	06/07/99	3	494
	02/03/00	4	507
	05/23/00	5	510
	12/15/00	6	520
	06/06/01	7	524
	12/05/01	8	555

[a] Invalid because instrument used was out of calibration on the sample date.

downgradient well case, we obtain $S_{[M_1]} = S_{[2.76]} = 5.27$ (μS/cm per 6-month period). Therefore, we concluded significant trends exist in both upgradient and downgradient wells because the lower 99% confidence limits on the slope $\beta_L = S_{[M_1]}$ are greater than zero.

Substituting Sen's slope estimators into Equation (8.14), for each well, we can remove the trend effects from the original data and obtain the following detrended data sets:

$$y_{Upgradient^*} = y_i - 7.3 * t_i \tag{8.15}$$

and

$$y_{Downgradient^*} = y_j' - 5.27 * t_j \tag{8.16}$$

where y_i and y_j' are original specific conductance data from upgradient well and downgradient well, respectively; $t_i = 1, 2, \ldots, 8$ and $t_j = 1, 2, \ldots, 7$.

For comparison purposes, we plotted the original specific conductance data and the de-trended data in Figure 8.4. It can be seen from Figure 8.4, after we remove the trend effect from each well, the de-trended data are stable for the application of the control chart method.

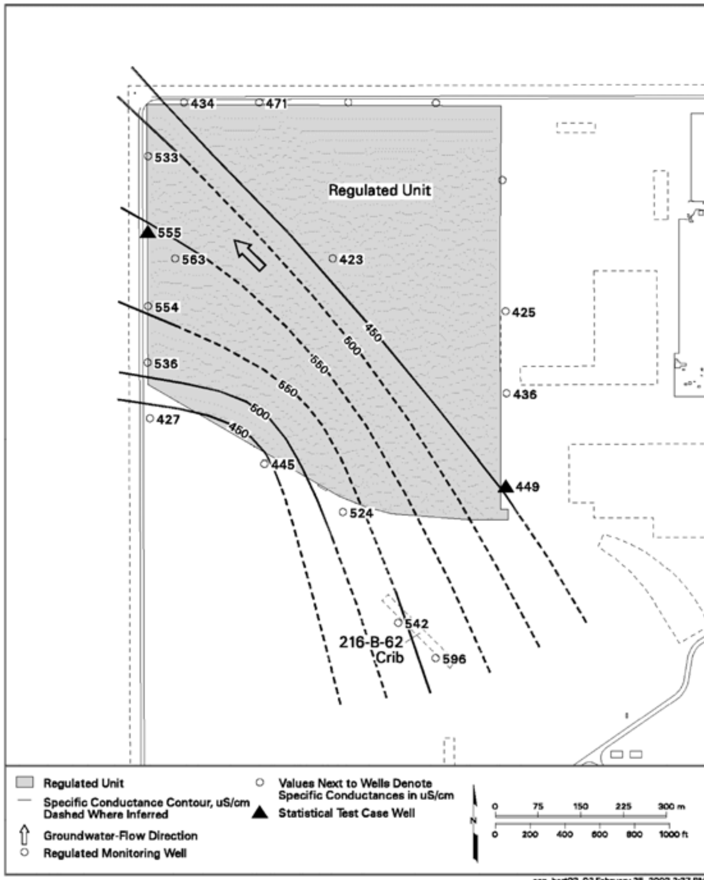


FIGURE 8.3 Map of specific conductance contours in the vicinity of regulated unit. (Note: Another contaminant plume from a different upgradient source that crosses the northeast corner of the regulated unit is not shown.)

8.4 SUMMARY

This chapter illustrates application of a powerful intra-well testing method referred to as the combined Shewhart–CUSUM control chart approach which can detect abrupt and gradual changes in groundwater parameter concentrations. This method is broadly applicable to groundwater monitoring situations where there is no clearly defined upgradient well or wells, where spatial variability exists in parameter concentrations or when groundwater flow rate is extremely slow. Procedures for determining the minimum time needed to acquire independent groundwater samples and useful transformations for obtaining normally distributed data are also provided.

TABLE 8.2
Sen's Estimators and the Lower 99% Confidence Limits on the True Slopes for the Upgradient and Downgradient Wells

Rank (i)	Upgradient Well Ordered Slopes $S_{[i]}$ (μS/cm/6-Month Period)	Downgradient Wells Ordered Slopes $S_{[i]}$ (μS/cm/6-Month Period)
1	5.5	3
2	5.5	4
3	6.25	5.67
4	7	6.5
5	7.5	7
6	8.5	7.5
7	8.75	8
8	9.05	8.67
9	9.42	10
10	9.75	12
11	10	12.2
12	10.06	13
13	10.33	15
14	10.54	17.5
15	10.6	17.8
16	10.83	20
17	11	21.25
18	11.13	25
19	11.33	31
20	11.56	36
21	11.65	59
22	11.75	
23	12	
24	12.13	
25	12.5	
26	13.42	
27	16	
28	18.25	
Sen's Slope Estimator = Median Slope (S)	$S_{[14.5]} = 10.57$	$S_{[11]} = 12.2$
$VAR(S)$	65.33	44.33
$C\alpha = (\alpha = 0.01)$	18.80	15.49
M_1	4.6	2.76
$S_{[M_1]}$	7.3	5.27

The control chart method will be insensitive to detect real changes if a preexisting trend is observed in the background data set. A method and a case study describing how a trend observed in a background data set could be removed using a transformation suggested by Gibbons[6] are presented to illustrate treatment of a preexisting trend.

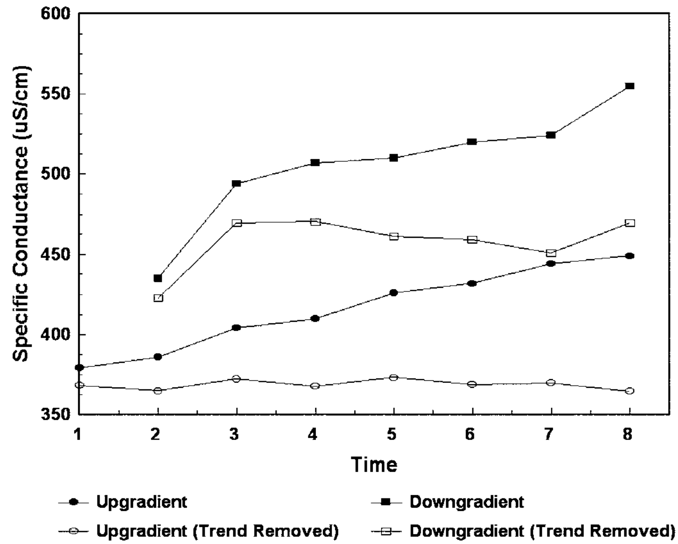


FIGURE 8.4 Original and detrended data from a waste management unit.

ACKNOWLEDGMENTS

The work reported herein was supported by the Hanford Site Groundwater Project conducted by Pacific Northwest National Laboratory for the U.S. Department of Energy, Richland Operations Office, under contract DE-AC06-76RLO 1830. Special thanks to R. F. O'Brien, V. G. Johnson, J. S. Fruchter, S. P. Luttrell, and W. J. Martin, Pacific Northwest National Laboratory, and Marv Furman, U.S. Department of Energy, for their encouragement, technical inputs, and review of the manuscript.

REFERENCES

1. Westgard, J. O., Groth, T., Aronson, T., and De Verdier, C. (1977), Combined Shewhart-CUSUM Control Chart for Improved Quality Control in Clinical Chemistry, *Clin. Chem.*, 28(10): 1881–1887.
2. Lucas, J. M. (1982), Combined Shewhart–CUSUM Quality Control Schemes, *J. Qual. Technol.*, 14(2): 51–59.
3. EPA (1989), Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities — Interim Final Guidance, EPA/530-R-93-003, U.S. Environmental Protection Agency, Washington, D.C.
4. EPA (1992), Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities — Draft Addendum to Interim Final Guidance, EPA/530-R-93-003, U.S. Environmental Protection Agency, Washington, D.C.
5. Starks, T. H. (1989), Evaluation of Control Chart Methodologies for RCRA Waste Sites, EPA/600/4-88/040, U.S. Environmental Protection Agency, Office of Research and Development, Las Vegas, NV.
6. Gibbons, R. D. (1994), *Statistical Methods for Groundwater Monitoring*, John Wiley & Sons, New York.

7. ASTM (1996), Provisional Standard Guide for Developing Appropriate Statistical Approaches for Ground-Water Detection Monitoring Programs, PS 64-96, American Society for Testing and Materials, West Conshohocken, PA.
8. Chou, C. J., O'Brien, R. F., and Barnett, D. B. (2001), Application of Intrawell Testing of RCRA Groundwater Monitoring Data When No Upgradient Well Exists, *Environ. Monit. Assess.*, 71: 91–106.
9. Freeze, R. A. and Cherry, J. A. (1979), *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ.
10. Conover, W. J. (1980), *Practical Nonparametric Statistics*, 2nd ed., John Wiley & Sons, New York.
11. Draper, N. R. and Smith, H. (1980), *Applied Regression Analysis*, 2nd ed., John Wiley & Sons, New York.
12. Montgomery, D. C. and Peck, E. A. (1982), *Introduction to Linear Regression Analysis*, John Wiley & Sons, New York.
13. Sen, P. K. (1968), Estimates of the Regression Coefficient Based on Kendall's Tau, *J. Am. Stat. Assoc.*, 63: 1379–1389.
14. Gilbert, R. O. (1987), *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold, New York.

9 Well Pattern, Setback, and Flow Rate Considerations for Groundwater Monitoring Networks at Landfills

P.F. Hudak

CONTENTS

9.1 INTRODUCTION

In a detection or source groundwater monitoring program, wells passively intercept enclaves (plumes) of contamination moving away from waste storage facilities such as landfills (Todd, 1980). Networks of source-monitoring wells should provide an opportunity for detecting contaminants released from any location within a potential contaminant source (EPA, 1994). The monitoring network should take into account the direction(s) of groundwater flow, rock and soil characteristics, and locations of neighboring properties or water supply wells (Loaiciga et al., 1992).

Liners and leachate collection systems in modern landfills reduce, but do not eliminate, the possibility of groundwater contamination. In a worst-case scenario, contaminants reaching groundwater could impact production wells supplying drinking water to households or communities. Properly positioned monitoring wells identify contaminants in groundwater and enable timely remedial action to prevent vast aquifer and supply-well contamination. Several factors influence the

1-56670-641-6/04/$0.00+$1.50

ability of a groundwater monitoring network to detect a future contaminant release. Among these considerations are well pattern, distance setback, and groundwater flow rate. This study employed computer simulations to examine the effects of pattern and distance setback on the ability of groundwater monitoring networks to detect contaminants released from hypothetical landfills in different velocity settings.

9.2 WELL PATTERN

Groundwater monitoring configurations at many landfills approximate a peripheral pattern in which wells are spaced at equal intervals (SP) along a monitoring locus at a fixed distance (DP) from the landfill's hydraulically downgradient perimeter. DP is measured perpendicular to the sides of the landfill. Figure 9.1 illustrates a five-well peripheral monitoring network for a hypothetical rectangular landfill, where DP equals 20 m and SP equals 56 m. SP was computed by dividing the total length of the monitoring locus by the number of wells, using a half-spacing at either end of the monitoring locus.

An alternative, equidistant approach to placing monitoring wells considers the local direction(s) of groundwater flow at a landfill (Hudak, 1998). In an equidistant network, wells are located an equal distance (DE) from a landfill's downgradient perimeter, but DE is measured parallel to groundwater flow. Equidistant wells are also spaced equally, but the spacing (SE) is measured perpendicular to groundwater flow. The equidistant monitoring network in Figure 9.1 also features five wells located along the centers of five flow tubes, each 33 m wide. In this network, DE equals 20 m and SE equals 33 m.

Suppose groundwater flows uniformly through an underlying aquifer at a rate of 0.01 m/d. Longitudinal and transverse dispersivity equal 1.0 m and 0.1 m, respectively, and the effective molecular diffusion coefficient equals 8.64×10^4 m^2/d. Within this setting, a computer program employing a two-dimensional analytical transport function (Domenico, 1987; Wilson et al., 1993) simulated the release of a conservative solute from a 5-m line source at each of the 9960 nodes distributed uniformly throughout the landfill's footprint. The source concentration was 1000 mg/l, and a concentration of 1 mg/l established plume boundaries.

Monitoring efficiency determinations were based on the percentage of contaminant releases detected before their associated plumes reached a buffer zone boundary located 150 m hydraulically downgradient of the landfill's downgradient corner. A leak was detected if its contaminant plume passed through one or more monitoring wells before reaching the buffer zone boundary.

The peripheral and equidistant networks detected 82.6% and 95% of simulated contaminant releases, respectively. Undetected source areas in Figure 9.1 occur as bands parallel to groundwater flow which taper in the hydraulically upgradient direction. This tapering reflects a gradual widening of plumes with increased transport distance, facilitating their detection.

Reducing SE or increasing DE can minimize undetected bandwidths. The equidistant network equalizes DE and SE among all wells and well pairs, and consequently the undetected bands in that network have similar size. In contrast, the peripheral network has a larger SE along the landfill's downgradient side most

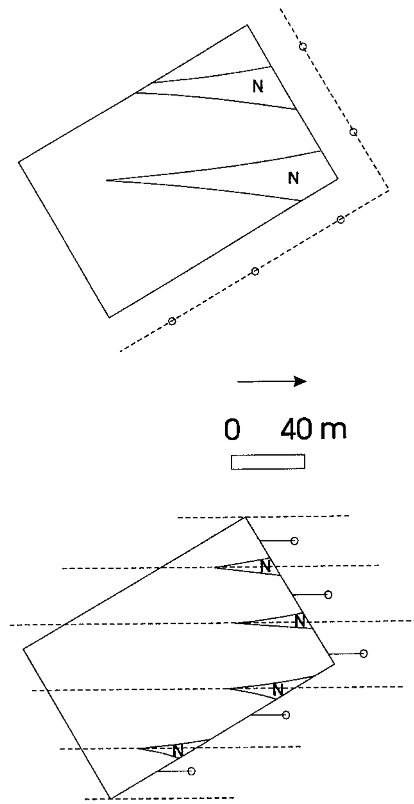


FIGURE 9.1 Map views of peripheral (top) and equidistant (bottom) monitoring networks (circles) at rectangular landfill; dashed lines delineate monitoring locus (top) and flow tubes (bottom); N denotes undetected source area; arrow denotes direction of groundwater flow. (Modified from Hudak, P.F. (1998), Configuring detection wells near landfills, *Ground Water Monit. Remediat.* 18: 93–96.)

perpendicular to groundwater flow (the shorter of the two downgradient sides in this example) and a smaller SE for the downgradient side most parallel to groundwater flow. Thus, the peripheral network performs effectively along the most parallel side but at the expense of performing poorly along the other side. Well spacing along the most parallel side is overly conservative, leaving large gaps between wells along the other side. Additionally, the peripheral network has an average DE of 35 m, compared to 20 m in the equidistant network. A smaller DE facilitates detection that is more rapid.

The preceding example shows that, for the general case of a landfill footprint oblique to the direction of groundwater flow, and for a given number of monitoring wells, equidistant monitoring networks outperform peripheral networks both with regard to the total percentage of potential contaminant plumes detected and the timing of detection. This relationship holds true regardless of the shape and orientation of a polygonal landfill, except for the special (and unlikely) case of a landfill with only one downgradient side (precisely perpendicular to groundwater flow) for which peripheral and equidistant networks are identical.

9.3 SETBACK AND FLOW RATE

Distance setback (DE in the preceding example) and groundwater velocity also affect the performance of a source monitoring network (Meyer et al., 1994; Hudak, 2002). Figure 9.2 shows an eight-well equidistant monitoring network near a hypothetical rectangular landfill (Hudak, 2002). DE equals 20 m, and the landfill is 100 m from a buffer zone boundary which establishes a distance limit of contaminant travel. Eleven networks were constructed by setting back the eight-well network from 0 to 100 m from the landfill's downgradient corner, in 10-m increments. The 0-m setback positioned wells on the landfill's downgradient perimeter, whereas the 100-m setback placed wells close to the buffer zone boundary.

The aforementioned computer program simulated conservative contaminant releases from a 1-m line source at 9643 potential source nodes distributed uniformly throughout the landfill's footprint. Concentrations were 1000 mg/l and 1 mg/l at source nodes and plume boundaries, respectively. The underlying aquifer has longitudinal and transverse dispersivity of 1.0 m and 0.1 m, respectively, and an effective molecular diffusion coefficient of 3.3×10^5 m^2/d. This set of parameters was used with each of four groundwater velocity values: 0.001, 0.01, 0.1, and 1.0 m/d.

For each of 44 scenarios (four velocity settings for each of 11 monitoring networks), the computer program calculated detection efficiency using two different solution options: (1) a buffer zone boundary with no time constraint and (2) a buffer zone boundary with a 10-year time constraint. In option (1), plumes grew until they reached a buffer zone boundary. However, in option (2), a simulation stopped if a plume traveled for 10 years but had not reached the buffer zone boundary. Plumes that did not pass through a monitoring well before reaching the buffer zone boundary or before the time limit elapsed were considered undetected.

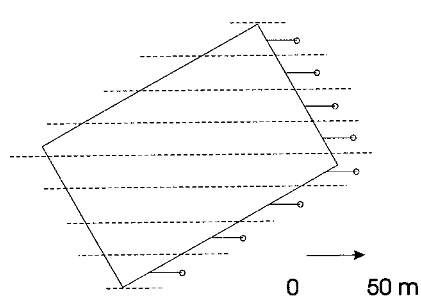


FIGURE 9.2 Map view of eight-well equidistant monitoring network (circles) at rectangular landfill; dashed lines delineate flow tubes; arrow denotes direction of groundwater flow. (Modified from Hudak, P.F. (2002), Associations between distance lags, groundwater velocities, and detection efficiencies in groundwater monitoring networks, *Environ. Monit. Assess.* 75(2): 215–221.)

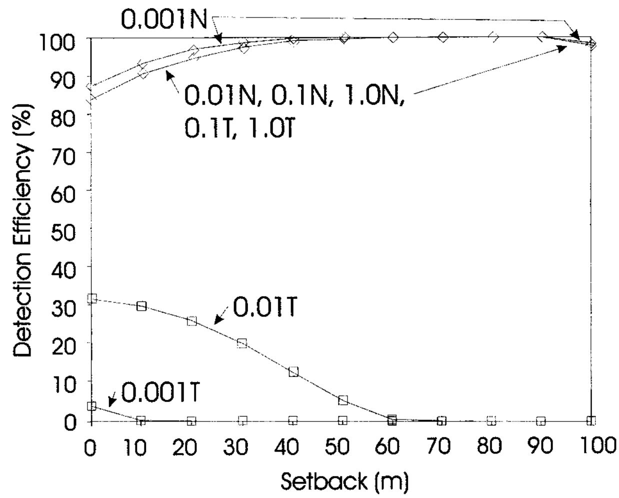


FIGURE 9.3 Detection efficiencies for groundwater velocity settings 0.001, 0.01, 0.1, and 1.0 m/d for eight-well network with 0- to 100-m setbacks. T and N denote buffer zone boundary with and without time constraint, respectively.

Figure 9.3 shows resulting detection efficiencies for distance setbacks ranging from 0 to 100 m in each velocity setting. In the lowest velocity setting with no time limit (0.001N), the detection efficiency was 87.4% for wells on the landfill's down-gradient boundary, reaching 100% at setbacks of 50 m to 90 m. The detection efficiency dropped slightly, from 100% to 99.6%, at the 100 m setback. As monitoring wells approached the buffer zone boundary, it became more difficult to detect plumes moving between wells because they could not continue to grow and widen. Thus, arbitrarily placing monitoring wells long distances from a landfill will not necessarily increase detection efficiency. Moreover, such networks will not detect contaminants until they pollute a large volume of groundwater.

Imposing a time limit substantially reduced detection efficiency in the lowest velocity setting (0.001T). The time limit rendered monitoring networks almost completely ineffective. Plumes moved slowly, most of them not reaching monitoring wells within the prescribed time limit. At the 20-m setback, detection efficiency dropped to 0%.

Similar trends were observed for the next highest velocity setting, though detection efficiencies were slightly lower for 0.01N compared to 0.001N. This was due to molecular diffusion which caused plumes to have slightly higher width/length ratios in the lowest velocity setting.

However, time-constrained detection efficiency curves differed markedly between velocity settings 0.001T and 0.01T. Time limits reduced detection efficiencies in 0.01T, but not by as much as in 0.001T. Plumes could travel farther, reaching more wells in the higher velocity setting. Detection efficiency dropped to 0% at the 70-m setback in 0.01T, compared to the 20-m setback in 0.001T.

Scenarios 0.01N, 0.1N, and 1.0N displayed virtually identical detection trends (Figure 9.3). At higher velocities, the effects of molecular diffusion become negligible, and velocity will not affect the width/length ratio of a plume that has traveled a fixed distance. In this case, there was very little reduction in detection efficiency beyond a groundwater velocity of 0.01 m/d.

Time constraints on plume travel had no effect on detection efficiencies in velocity settings 0.1T and 1T (Figure 9.3). In these highest velocity settings, plumes reached monitoring wells and the buffer zone boundary before the time limit expired. Hence, wells at all setbacks had an opportunity to detect contaminant releases.

9.4 CONCLUSION

Equidistant groundwater monitoring networks explicitly account for local directions of groundwater flow and thus outperform peripheral networks at landfills. Reducing the spacing (measured perpendicular to groundwater flow) between monitoring wells and increasing the setback (measured parallel to groundwater flow) between wells and a landfill generally increases the detection efficiency of a source monitoring program. However, setbacks made arbitrarily considerably increase the volume of contaminated groundwater and decrease detection efficiency when wells approach a buffer zone boundary.

Detection efficiency decreases as groundwater velocity increases, but there is a site-dependent velocity threshold beyond which subsequent detection efficiency decreases are minimal. In low-velocity settings, time limits severely hamper detection efficiency, reducing it to zero at farther setbacks.

REFERENCES

Domenico, P.A. (1987), An analytical model for multidimensional transport of a decaying contaminant species, *J. Hydrol.* 91: 49–58.

EPA (U.S. Environmental Protection Agency) (1994), RCRA Ground Water Monitoring, Government Institutes, Rockville, MD.

Hudak, P.F. (1998), Configuring detection wells near landfills, *Ground Water Monit. Remediat.* 18: 93–96.

Hudak, P.F. (2002), Associations between distance lags, groundwater velocities, and detection efficiencies in groundwater monitoring networks, *Environ. Monit. Assess*. 75(2): 215–221.

Loaiciga, H.A., Charbeneau, R.J., Everett, L.G., Fogg, G.E., Hobbs, B.F., and Rouhani, S. (1992), Review of ground-water quality monitoring network design, *J. Hydraul. Eng.* 118: 11–37.

Meyer, P.D., Valocchi, A.J., and Eheart, J.W. (1994), Monitoring network design to provide initial detection of groundwater contamination, *Water Resour. Res.* 30: 2647–2659.

Todd, D.K. (1980), *Groundwater Hydrology*, John Wiley & Sons, New York.

Wilson, C.R., Einberger, C.M., Jackson, R.L., and Mercer, R.B. (1993), Design of groundwater monitoring networks using the Monitoring Efficiency Model (MEMO), *Ground Water* 30: 965–970.

10 Selection of Ecological Indicators for Monitoring Terrestrial Systems

G.J. White

CONTENTS

10.1 INTRODUCTION

In recent years, the importance of assessing the condition of ecological systems including wilderness and other protected lands from atmospheric pollutants and other anthropogenic and natural factors has become widely recognized. Monitoring and assessment of natural systems are increasingly focusing on the application of indicators of ecosystem status, and substantial efforts are currently being devoted to the identification and development of suitable indicators (National Research Council, 1986; Noss, 1990; Messer et al., 1991; Bruns et al., 1991, 1997; Kurtz et al., 2001). However, accurate assessment of impacts to ecological systems has been hampered by a general lack of information in many key areas or by the failure to collect and/or consider the information that is available.

Assessment of the condition of ecological systems is further complicated by the vast diversity in structure, extent, and composition of these ecosystems, and in many cases by the harsh environments and difficult access associated with many sites. Given the diversity of ecological systems, data collected in one geographic area may not be fully applicable to others even if the two areas are located near one another. Furthermore, extensive physical, chemical, and biological monitoring programs are often impractical due to cost constraints and other factors. The challenge is to develop a program that will answer the pertinent monitoring questions in the most cost-effective manner.

1-56670-641-6/04/$0.00+$1.50

During the 1990s, federal land management agencies in the U.S., including the Forest Service, Park Service, Fish and Wildlife Service, and Bureau of Land Management, began to develop and document processes for establishing pollutant effects monitoring programs in Class I wilderness areas. The focus of these monitoring programs was to provide early detection of the effects of atmospheric pollutants on ecological systems. Toward that end, several guideline documents were published describing how pollutant effects monitoring programs should be designed (Adams et al., 1991; Schmoldt and Peterson, 1991; J. Peterson et al., 1992; D.L. Peterson et al., 1992; Peine et al., 1995). These documents relied heavily on the use of indicators of ecosystem status and served to illustrate some of the difficulties encountered in making such assessments. In most cases, these documents concluded that adequate baseline information was rarely available, greatly increasing the difficulty associated with selection of indicators of ecosystem status. Complicating this selection process is the fact that monitoring programs that utilize ecological indicators must be established on a site-by-site basis. This is important not only because each potential area of interest is unique geologically, hydrologically, and ecologically, but also because each factor conferring change on the system is at least somewhat unique. Ecological monitoring programs must be designed to address the specific stressor and protected area independently, as a program designed for one scenario will not necessarily be applicable to another.

To establish an effective assessment program based on the implementation of indicators, the following questions should be answered:

1. Which resources (or critical receptors) are potentially of concern, and where are they located?
2. Which perturbation factors are potentially responsible for impacting these receptors?
3. Which indicators will best detect the impacts of the perturbation factors on the sensitive receptors?
4. At what specific locations should the indicators be examined?
5. At what frequency should these indicators be examined?
6. What degree of change indicates cause-and-effect?

All of these questions should be addressed within the context of sound science.

Monitoring involves the continual systematic time series observation of an *appropriate* suite of predetermined chemical, physical, and/or biological parameters within the *appropriate* components of the *appropriate* ecosystem, for an *appropriate* period of time that is sufficient to determine (1) existing conditions, (2) trends, and (3) natural variations of each component measured (Segar, 1986). To accomplish this, monitoring programs must be designed properly. The most important step in the design of any monitoring program is the definition of the objectives of the program. Only when specific objectives such as these have been established can the scientific method of establishing and testing hypotheses be applied (Segar, 1986). These objectives must adequately define:

- The specific receptor to be evaluated
- The specific effect to be monitored
- The level of effect that bounds acceptable vs. unacceptable conditions

To effectively meet the objectives of an ecological monitoring program, monitoring must be designed to detect changes in indicators that are both measurable and significant. It is not realistic to design a monitoring program to assess the concentration of every potential perturbation factor in all media at all locations that might be impacted, in order to detect any change in the degree of perturbation. Similarly, ecological monitoring cannot be conducted to identify any change in the abundance, health, growth rate, reproductive rate, etc., of any species or community which is caused by any potential perturbation factor (Segar, 1986). Such goals are neither realistic nor attainable.

By definition, indicators must be indicative of some unmeasured or unknown condition (Suter, 2001). As will be discussed later in greater detail, the selection of ecological indicators must consider the roles of these indicators within the dynamics of the system to be monitored, the degree to which these roles are understood, and the certainty associated with observed levels of the indicators. Candidate indicators should therefore represent measures that, based on expert knowledge and available literature, will provide useful information concerning the condition of the ecosystem being monitored. Criteria must be established that can be used to assess the effectiveness of indicators to ensure that:

1. The resulting data will be sufficient to answer the pertinent questions regarding the status of the ecological system of interest.
2. The resulting data are of known and acceptable quality.
3. The monitoring program can be implemented in a cost-effective manner.

These criteria should then be applied to the selection of indicators of the condition of the ecological systems in question, and monitoring programs based on the measurements of these indicators may then be designed in a manner that will provide cost-effective, scientifically based assessment of ecosystem status. Without applying a consistent, scientific approach, it is difficult to predict which indicators will best reflect the potential effects due to specific perturbation factors, or to select the most effective methods for monitoring these effects.

10.2 OBJECTIVE AND APPROACH

The purpose of this chapter is to describe criteria for selecting ecological indicators for use in monitoring the status of ecological systems. Although the approach described in this document is intended to be generic in that it is applicable to virtually any situation, the output must be considered site-specific at both ends of the stressor/receptor continuum. Furthermore, although the emphasis is on terrestrial systems, the process can be applied equally to developing monitoring for aquatic systems. A series of criteria is proposed by which potential indicators of cause and effects relationships may be evaluated. By applying these criteria during the planning stage, it is anticipated that monitoring programs can be more readily developed to provide defensible, quality-assured data in the most cost-effective manner.

The general approach proposed here for developing a monitoring program based on the application of ecological indicators is as follows:

1. Gather pertinent site-specific information on (i) the ecosystems of concern; (ii) the potential critical receptors or components within the ecosystems; (iii) the factors that potentially impact the health of these ecosystems and/or components (e.g., disease, air pollution deposition, urban encroachment, invasion of exotic species, logging, etc.); and (iv) the relationship between the stressors and the receptors.
2. Develop a conceptual framework using this information to illustrate and understand the dynamics of the systems of interest.
3. Establish and rank criteria for evaluating potential indicators of ecosystem change and use these criteria to select the appropriate indicators for assessing changes in the status of the ecological systems.
4. Develop hypotheses to be tested using the indicators selected.

The general intent of this document is therefore to help with the development of scientifically defensible, cost-effective monitoring programs to assess the status of ecological systems. Much of the discussion is focused on relatively pristine ecological systems, as these are likely to prove more difficult in determining cause and effect relationships.

10.3 MONITORING TERRESTRIAL ECOSYSTEMS—DESIGN AND CONSIDERATIONS

The first step in designing a monitoring and assessment program for terrestrial ecosystems is to gather the information necessary to develop a conceptual design or model for the program. This involves compilation of information relating to the ecosystem of concern (including critical components of the ecosystem) and the factors that may potentially alter the status of the ecosystem or critical ecosystem components. It also involves determining the relationships between the potential perturbation factors and the critical receptors of components of concern. Collectively, this information is incorporated into a conceptual model for the system of interest. This model is then used to design the monitoring approach.

The first step in this approach is to identify what resources are of concern and where these resources of concern are located. This is obviously tied to the goal of the proposed monitoring program. Often this determination is one of scale. If the concern is die-off of sugar maple, then the receptor of interest is a single species, but the area of concern may be the entire range of the species, covering a couple of dozen states and much of southeastern Canada. Alternatively, the goal of the monitoring program could be to determine the status of ecological systems within Yellowstone National Park. Here, the area of concern is defined by the boundaries of the Park, but the ecosystem components of interest could include any or all species found in the Park. Spatial scales could be considerably smaller, however, such as a watershed or a single stand of trees.

Once the resources of concern have been identified, the next step is to determine what factors or agents may impact the status of those resources. These may include either natural factors such as fire, disease, weather and climate, or anthropogenic factors such as atmospheric pollutants, logging, or other land use activities. In some cases (e.g., fire) it may be difficult to separate the natural from the anthropogenic. In many instances, the perturbation factors are well known and provide a known impetus for establishment of an ecological monitoring program. The government programs to determine air pollution impacts to Class 1 airsheds, for example, were charged with determining the effects to terrestrial and aquatic ecosystems resulting from a specific cause (air pollution). Similarly, monitoring Douglas-fir forests for spruce budworm damage links a specific cause with a specific effect. It should be pointed out that not all monitoring programs are charged with determining a specific cause of a specific effect in an individual species at a specific location. At the other extreme, a monitoring program may be designed to determine the status and trends of "ecosystem health" throughout a given biome such as tropical rainforests or alpine tundra. In these instances, it is still recommended that specific perturbations and receptors be identified.

Once the system is defined in terms of location, perturbation factors, and critical receptors or components, it is often useful to develop a conceptual model of the system of concern. These conceptual models may take the form of a simple "box-and-arrow" diagram that describes the structure and function of the ecosystem or ecosystem components of concern (e.g., Figure 10.1). In these diagrams, each "box" represents some component of the ecosystem, while the arrows illustrate the transfer of nutrients, contaminants, or energy between components. Such diagrams can help to visualize the dynamics of pollutants in the environment. Thus conceptualized, mathematical models may be applied using the conceptual model to quantify the rates at which materials are expected to move through the system. Such an approach allows for periodic reevaluation of data sets based on model calculations, which

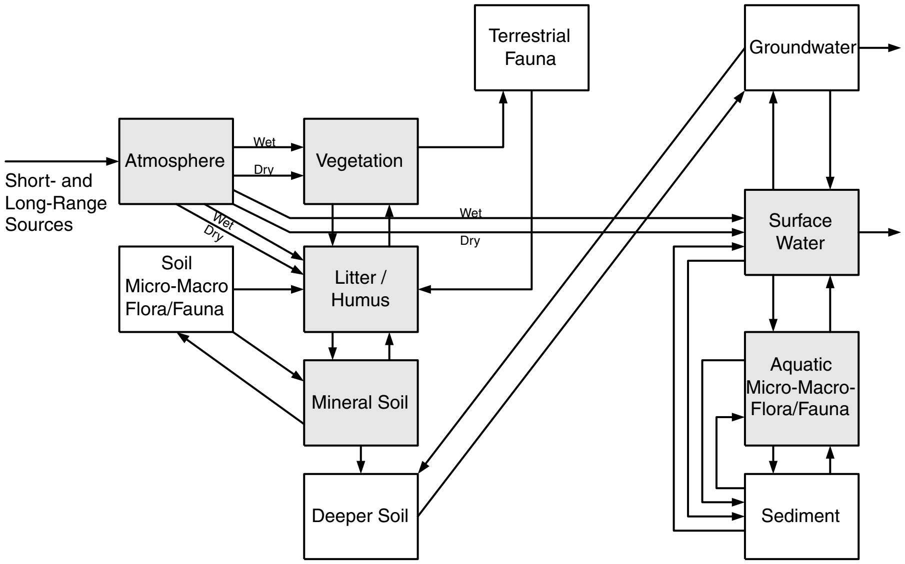


FIGURE 10.1 "Box-and-arrow" diagram used to conceptualize an ecological system during the development of a monitoring program.

ultimately may allow for the modification of the monitoring system design in such a way as to improve cost-effectiveness.

Conceptual diagrams may be considerably more complex than that shown in Figure 10.1 and may be used as heuristic tools for establishing many of the key aspects of the monitoring program design. The diagram in Figure 10.1 has been used to monitor the impacts from air pollutants on terrestrial and aquatic ecosystems in the western U.S. and elsewhere (Bruns et al., 1991). In contaminant monitoring programs, these diagrams can help determine source–receptor relationships, contaminant pathways, critical receptors, and the ultimate fate of contaminants. This is conducive to an ecosystem approach to environmental monitoring whereby interrelationships between different components of the system are considered, recognizing that alterations to one component of the system may affect other components. Conceptual models help to provide information that may be used to help determine which receptors are at risk from which stressors, and what indicators should be used to quantitatively link the stressors to critical receptors. This approach provides for the effective integration of various indicators of change that will enable the evaluation of the system as a whole. Models can also help to identify gaps in the existing data.

Once the appropriate stressors and receptors have been identified, it is important to narrow the focus of the potential relationship between source and receptor. It is not enough to determine that a stressor may cause impacts to a particular receptor. Rather, information is needed on the species or communities of plants that may be at risk, the anticipated responses of these species or communities, and the exposures necessary to elicit these responses.

- What are the effects of the identified stressors on the identified ecosystems or ecosystem components?
- At what level of biological organization do the stressors operate?
- Which stressors are responsible for these changes?
- What is the mode of action by which the effect occurs?
- What characteristics (e.g., temporal component, etc.) control the effect?
- What characteristics of the site are involved?
- To what degree can laboratory data be extrapolated to the field?

Effects of stressors on ecological systems are extremely complex and diverse. Effects from atmospheric pollutants, for example, may be classified variously as direct vs. indirect, acute vs. chronic, lethal vs. sublethal, biotic vs. abiotic, visible vs. microscopic, positive vs. negative, etc. Furthermore, it is important that effects be considered for all levels of biological organization. Not only may effects be observed at the ecosystem, community, population, or individual levels of biological organization, but at the other extreme, effects may also be observed at the cellular, biochemical, or genetic levels. Potential effects on ecological systems due to stressors must be identified even if there is no obvious evidence that this damage is occurring.

The specific stressors potentially responsible for each effect must also be determined, integrating dose/response information wherever possible. To complicate matters further, the possibility of synergistic effects brought about by a combination of stress agents must also be considered. Other potential causal or contributing factors should also be identified. These could represent additional independent stress factors (e.g., drought, pathogens, insect pests), or factors associated with the environment (e.g., soil pH, temperature, etc.) or with the organism itself (e.g., physiological, morphological, and other features of the organism that renders it susceptible). The response of organisms to stressors may vary substantially among sites, even if exposures are the same. This may be due to differences in receptor species (species composition and density, age class distribution, genetic pools) or by differences in the site (e.g., elevation, slope, aspect, solar incidence, precipitation, etc.). Soil characteristics (e.g., pH, percent organic matter, cation exchange capacity, percent base saturation, depth, sulfate adsorption capacity, fertility, buffering capacity, etc.) may be especially important.

Once the stressor/receptor relationships have been determined, the mode of action by which the effect occurs must be assessed. This requires an understanding of the mechanism of action involved with the interaction between pollutant and receptor. How is exposure duration (both instantaneous and chronic) and/or frequency involved in the manifestation of effects? Considerable information exists on the effects from short-term pollutant exposures for many plant species. However, little data are available on the effects from long-term or chronic exposures.

Organisms, not ecosystems, respond directly to stress, and higher levels of biological organization in turn integrate the responses of the various individuals through various trophic and competitive interactions before an ecosystem-level response can be observed (Sigal and Suter, 1987) without a prior organism response. Responses of organism therefore precede those of ecosystems, and in the process of monitoring the parameters of entire ecosystems, the responses of sensitive individuals and populations tend to be masked or averaged out. Observations of impacts at the organism-level biological organization are relatively easy and inexpensive to measure (Sigal and Suter, 1987). Information linking these organism-based parameters to adverse impacts on higher levels of biological organization (i.e., populations, communities, or ecosystems) are generally lacking and are confounded by natural variability, extended response times, variability of climatic conditions, influences of pathogens and insect pests, and other factors (Sigal and Suter, 1987).

Information must also be compiled on a site-specific basis. Information on the individual ecological system of interest is necessary because all ecological systems are at least to some extent unique. If vegetation is the focus, then the distribution of various species and communities are needed. Data on soil development, soil chemistry, insect and disease history, meteorological parameters, and physical parameters (e.g., slope, aspect, elevation) may also be helpful. Collection of these types of information will help in the subsequent steps in the development of an approach for monitoring the status of the system. Questions to ask include:

1. What information is available for the ecosystem or ecosystem components (i.e., receptors) of concern?
2. What information is available on factors potentially responsible for causing stress or change to these receptors?
3. What information is available from other areas sharing similar ecological, geological, and geographical properties?

The specific locations where monitoring will be most effective must also be determined. Using the information generated in the above steps, candidate locations should be identified to conduct monitoring. Criteria should be established with which to evaluate candidate sites, and then these sites should be ranked using these criteria. Monitoring locations selected may not be the same for each receptor, or for each parameter or indicator measured for a given receptor, but should be based on where the best information can be obtained in the most cost-effective manner. Once a list of candidate monitoring sites is selected, the sites must be ranked such that the "best" subset of sites is selected for monitoring the status of the resource. Although many different sites may meet the basic requirements for a monitoring location, it is desirable to select the optimum site (or sites) for each receptor to be assessed.

10.4 SELECTION OF INDICATORS OF ECOSYSTEM STATUS

Only now are we ready to change the emphasis from "what" and "where" to "how." Specifically, how can impacts to sensitive receptors best be assessed? Methods must be identified or developed with which to assess the condition of receptors of concern. For example, if aspen are identified as potentially sensitive receptors for SO_2 deposition in a particular location, methods must be applied that will allow for the assessment of the status of the aspen at that specific selected location. Indicators must be identified that will allow impacts to aspen from SO_2 to be quantified. These indicators may be chemical, physical, or biological (ecological) measurements that individually or collectively will allow for the evaluation of the aspen growing at that site.

A method is proposed here for selecting the most effective suite of indicators for assessing the status of a particular sensitive receptor or group of ecological receptors within a given geographic area. The process involves the application of a list of criteria for selecting the appropriate suite of indicators. Once the criteria list is established, the criteria may be ranked in terms of their relative importance to the success of the monitoring program. This identification and ranking of criteria is performed before actual indicators are considered; only after criteria are established are potential indicators evaluated against one another. By applying these criteria to indicator selection during the planning stage, monitoring programs can be developed to better provide defensible, quality-assured data in a cost-effective manner. Establishing selection criteria early in the overall process helps to assure that the monitoring program will adequately provide the necessary answers to questions regarding the status of the ecological systems.

The purpose of establishing criteria with which to evaluate potential indicators is to define *a priori* the characteristic properties that an indicator or indicators should possess in order to be effective. This approach is recommended to avoid some of the problems common to many existing monitoring programs whereby ecological indicators fail to provide the information necessary to evaluate the condition of the resource being monitored (D.L. Peterson et al., 1992; J. Peterson et al., 1992). The criteria developed should be used to bind potential ecological indicators in a manner that will better ensure that the data produced are of known quality and are collected in the most cost-effective manner.

As indicated earlier, ecological monitoring programs must be designed for each combination of stressor and receptor independently, as each combination is at least somewhat unique. A monitoring program designed for one scenario will not necessarily be applicable to another, and the monitoring design must therefore be established on a site-by-site basis. The criteria presented below are of varying importance, and reaching consensus opinions regarding the relative importance of each criterion may be difficult. Furthermore, the relative importance of each may vary among sites. The goal is to apply these and/or other alternative criteria to provide a consistent, generic approach to the selection of indicators. This approach can be applied in virtually any situation (i.e., any combination of source and receptor of interest), but the output must be considered site-specific.

CRITERION 1: Ecosystem Conceptual Approach — The ecosystem approach to environmental monitoring considers many features of ecosystem simultaneously rather than focusing on single, isolated features of the environment. To satisfy the ecosystem conceptual approach criterion, indicator parameters must relate in a known way to the structure or function of the ecological system to be monitored so that the information obtained provides a "piece of the overall puzzle." Individual parameters should directly or indirectly involve some physical, chemical, or biological process (or processes) associated with the atmospheric, terrestrial, and/or aquatic portions of the system.

Many different approaches can be applied to ecological monitoring, and each may be classified as either reductionist or synthesist in terms of the general strategy employed. A reductionist approach to monitoring assesses each parameter independently, whereas a synthesist strategy incorporates a more holistic approach that addresses the interrelationships between different components of the system. The reductionist approach therefore recognizes that if one component of the system is altered or stressed in some way, there will be direct and/or indirect consequences to other components as well, and that each of these, in turn, will cause further changes to occur. For most aspects of ecological monitoring programs, particularly in relatively pristine areas, it is recommended that a synthesist or "ecosystem approach" be taken to better enable overall impacts to be assessed in an integrated manner rather than as isolated, independent events.

The ecosystem conceptual approach criterion must be addressed at two levels. First, the approach should be applied to the overall monitoring program through the application of the systems conceptual models designed earlier (Figure 10.1). These models help the user visualize relationships between the receptors and stressors within the ecosystem and may therefore be used to help identify indicators of ecosystem status.

At the second level, each individual component of the monitoring program should be evaluated to see how well it fits into the ecosystem approach to monitoring. With regard to a particular indicator, the basic questions asked relating to the ecosystem conceptual approach include the following:

1. Is application of the particular indicator (or set of indicators) consistent with current concepts of ecosystem theory?
2. Does the indicator relate to some process or processes associated with the structure and/or function of the ecological system? In developing a suite

of indicators for assessing vegetation condition in Australia, a process involving 47 Australian experts recently identified 62 potential indicators of vegetation condition. These were equally representative of compositional (21), structural (20), and functional (21) attributes of biodiversity (Oliver, 2002)

3. Do the procedures to be used for measuring the indicator adequately document how that particular indicator (or set of indicators) fits within an ecosystem context?
4. Will the resulting data be useful in providing an adequate understanding of the system to be monitored?
5. If a particular indicator does not adequately satisfy the above, what alternative indicators may be recommended to meet such a requirement?

There are many good examples of indicators of ecosystem stress that meet the ecosystem conceptual approach to environmental monitoring. For example, litter decomposition and multimedia elemental analysis both provide information on the nutrient dynamics of the system. Vegetation surveys in the terrestrial system and analysis of functional feeding groups in aquatic systems can provide information on the structure of the ecosystem.

Conversely, although parameters associated with visibility may represent important measurements, these do not fit well into the ecosystem conceptual approach because visibility is primarily an aesthetic issue rather than an ecological one. Visibility is therefore more effectively treated individually.

CRITERION 2: Usability—The usability criterion relates to the level of documentation available for each indicator measurement; the relative completeness and thoroughness of the procedures for measuring indicator parameter provide the best indication of the usability of that indicator. The usability criterion is therefore satisfied for indicators for which the level of supporting documentation is complete.

Ideally, detailed standard operating procedures (SOPs) should be available (or be easily generated) for each parameter measured as part of the monitoring program, and these SOPs should represent generally accepted, standardized methods. If the methods used are not well established, then supporting documents describing earlier applications of the method should be available. Any supporting documents used to justify the choice of indicator measurements or necessary to implement the measurements should be identified and referenced within the SOPs for each parameter measured. Information on previous field testing of the SOPs and supporting documents should be available as well.

Good examples of indicator parameters that satisfy the usability criterion include the widely used methods for measuring wet deposition, water chemistry, and soil chemistry. Established procedures for monitoring wet deposition are available and have been used for over a decade as part of the National Acid Deposition Program (NADP). Procedures for analyzing the chemical properties of water and soil are also well established. These procedures have long histories of field use and generally satisfy the usability criterion. In contrast, measurements of many ecological indicators are made using variable techniques, with little or no consensus regarding the best methodology available.

CRITERION 3: Cost-Effectiveness—This criterion can be evaluated on a relative basis based on the answer to a single question: "Is the incremental cost associated with the measurement low relative to the information obtained?" In determining cost-effectiveness, consideration should be given to the time necessary for the preparation of sampling activities, collection of the samples, analysis of the samples (where applicable), and analysis and interpretation of the resulting data. The cost of field or laboratory equipment must also be considered. Where possible, measurements that may be performed using synoptic monitoring techniques are more cost-effective, although there are some cost-effective automated monitoring techniques that may be applied in some circumstances.

Aquatic chemistry parameters and litter decomposition rates are among the many examples of indicators that are relatively cost-effective. Remote sensing technologies offer promise for a variety of indicators in that they may reduce the expense associated with sending personnel to remote field sites to collect samples or to conduct measurements.

Any parameter with high equipment or analytical expenses which necessitate large time commitments in the field or laboratory may not satisfy the cost-effectiveness criterion. Atmospheric pollutant measurements, for example, are typically expensive to purchase and operate, and may not be justified based on the amount of information obtained, especially if reasonable estimations of atmospheric input can be obtained via other means (i.e., from a nearby monitoring station or by modeling). Other parameters such as relative sensitivity tables for plants and other organisms may be very useful but may be very costly to produce for a specific site, unless the information happens to be available elsewhere for the species at the site of interest.

CRITERION 4: Cause/Effect — This criterion can only be met if there is a clear understanding of the relationship between a receptor and a stress factor such that the indicator used will exhibit a clear response (effect) to a measurable increase in the level of stress (cause).

To evaluate this relationship between cause and effect, the following questions should be considered:

1. Does the indicator respond in a known, quantifiable, and unambiguous manner to a specific stressor of concern?
2. Is there dose/response information available for the indicator and the stressors of concern?
3. Are exposure thresholds or trends known for the indicator?
4. Will the indicator provide similar information for most potential sampling areas within a wide geographic region?

The primary difficulty with establishing causal effects in ecological settings that are relatively far removed from pollutant sources is that often the early symptoms of pollutant damage are indistinguishable from those caused by other stress agents. In fact, with the exception of areas where damage is severe, recognition of pollutant damage is likely to be very difficult and will take the form of general stress potentially attributable to many different factors or a combination of factors.

Odum (1985) defined stress as "a detrimental or disorganizing influence" and categorized the manifestation of stress in ecological systems as changes in (1) energetics, (2) nutrient cycling, and (3) community structure and function, as summarized in Table 10.1. For example, visible symptoms of chronic air pollution toxicity in trees and other plants are not highly specific, and in natural environments these symptoms can easily be confused with symptoms of other, unrelated stress factors including extreme climate conditions, nutrient deficiencies, and insect and disease disorders (Sigal and Suter, 1987). Climatic conditions (present and past) to which the plant is exposed, soil factors (i.e., nutrient availability), time of the year and time of the day that the plant is sampled, position within the plant and within the canopy that the plant is sampled, tissue age, genetic factors, presence of disease organisms or insect pests, etc., all present confounding variables.

Margalef (1981) stated that "stress is something that puts into action the mechanism of homeostasis." Early warning of stress will be more easily seen at the species level, although shifts here should be accompanied by changes in the rate of respiration and/or decomposition, which are more difficult to detect in large systems. When stress is detectable at the ecosystem level, there is real cause for alarm, for it may signal a breakdown in homeostasis (Odum, 1985).

Most studies of effects of air pollutants conducted to date have focused on responses of individual organisms rather than on the higher levels of biological organization. For example, visible injury to plants and reductions in biomass accumulation rates have often been cited as responses to atmospheric pollutants. However, linkages of these parameters to adverse impacts on populations and communities are lacking. Disturbance that is detrimental at one level of biological organization may actually be beneficial at another. Similarly, a disturbance may be detrimental over the short term, but beneficial over the long term. For example, Odum (1985) indicated that periodic fire in fire-adapted systems such as chaparral may cause stress to individual plants, resulting in injury or mortality, but the absence or exclusion of fire would represent the stress at the ecosystem level.

The ability to accurately quantify a response may be rendered useless if the relationship between cause and effect is ambiguous. For example, although there is a large volume of documented evidence that indicates that exposure of many species of deciduous and evergreen trees to a variety of atmospheric pollutants will result in the development of symptoms of foliar chlorosis, this represents a typical response of green plants to stress in general. True assessment of damage from atmospheric pollutants may therefore be complicated by other stress factors, including physical damage, low soil nitrogen concentrations, root fungi, bark beetles, leaf-feeding insects, drought, etc. Similarly, tree mortality has been shown to result from acute exposures of several different pollutants. In areas further removed from the pollutant sources, however, atmospheric pollutants more often represent a contributing factor to the mortality, and determining the influence of pollutants relative to other proximate stress factors is virtually impossible.

Because of difficulties in proving that an observed change is due to pollutant exposures, responses that are diagnostic of the pollutant should constitute key components of monitoring programs. Examples include accumulation of the pollutant and characteristic gross and histological injuries (Sigal and Suter, 1987).

TABLE 10.1
Trends Expected in Stressed Ecosystems

A. Energetics

1. Increased community respiration: This may represent an early-warning sign of ecosystem stress due to the acceleration of repair processes in response to damage caused by the disturbance. This requires diverting energy otherwise available for growth and production to maintenance
2. Unbalanced ratio of production to respiration: This may be either greater than or less than 1
3. Ratios of production to biomass (P/B) and respiration to biomass (R/B) tend to increase: The increased R/B occurs as organisms respond to the disorder created by disturbance
4. Auxiliary energy increases in importance
5. The fraction of primary production that is unused increases

B. Nutrient Cycling

1. Nutrient turnover rates increase
2. Horizontal transport increases and vertical cycling of nutrients decreases (cycling index decreases)
3. Nutrient loss increases (system becomes more "leaky")

C. Community Structure

1. Proportion of r-strategists (vs. K-strategists) increases
2. Size of organisms decrease
3. Life spans of organisms or parts of organisms (e.g., leaves) decrease
4. Food chains become shorter due to reduced energy flow at higher trophic levels and/or the greater sensitivity of predators to stress
5. Species diversity decreases and dominance increases; if prestress diversity is low, the reverse may occur; at the ecosystem level, redundancy of parallel processes theoretically declines

D. General System-Level Trends

1. The ecosystem becomes more open (i.e., input and output environments become more important as internal cycling is reduced)
2. Autogenic successional trends reverse (succession reverts to earlier stages)
3. Efficiency of resource use decreases
4. Parasitism and other negative interactions increase, and mutualism and other positive interactions decrease
5. Functional properties (such as community metabolism) are more robust (homeostatic-resistant to stressors) than are species composition and other structural properties

Source: From Odum, E.P., 1985, Trends expected in stressed ecosystems, *BioScience*, 35: 419–422.

CRITERION 5: Signal-to-Noise Ratio — This relates somewhat to the previous criterion, but refers more specifically to the relative ease with which changes in the indicator caused by the specific stress agent may be distinguished of changes due to natural variability. For indicators to satisfy this criterion, separation of stressor-induced changes from changes due to other factors must be relatively easy.

The signal-to-noise ratio in ecological parameters is a function of the degree of variability exhibited by the parameter in the absence of the stress factor being evaluated. To evaluate indicators in terms of this criterion, the following questions should be asked:

1. What is the natural spatial variability associated with the parameter to be measured?
2. What is the natural temporal variability associated with the parameter to be measured?
3. Are there predictable patterns in the spatial (e.g., slope, aspect, soil association, moisture) or temporal (e.g., seasonal) variability of the indicator?
4. Does the indicator possess sufficiently high signal strength, in comparison to natural variability, to allow detection of statistically significant changes within a reasonable time frame?

The successful separation of the desired "signal" from the background "noise" is generally complicated by natural variability caused by season, climate, natural succession, natural disturbance, microclimate, etc. Often, the temporal and spatial variability within the ecosystem will be substantially greater than the variability the monitoring method is designed to detect. When this is the case, assessment of the spatial and/or temporal variability necessitates enormous databases that are not available in most instances. Such monitoring methods may work well in areas of high impact, or in laboratory experiments, but may be inappropriate for wilderness systems where changes are gradual and subtle. Variability is important on a variety of spatial and temporal scales. Temporally, ecological parameters may vary on sporadic, seasonal, and/or annual basis. Many ecological parameters vary on a seasonal basis. For example, nutrient concentrations in tree foliage may change dramatically during the growing season, especially for hardwood species. Nitrogen concentrations generally increase rapidly in the spring, undergo slight declines during the growing season, and decrease rapidly at the beginning of fall senescence as the tree resorbs this element. Conversely, concentrations of boron, calcium, and some nonnutrients including aluminum and heavy metals tend to increase steadily throughout the life of the leaf. Concentrations of potassium are more difficult to predict due to factors such as foliar leaching. These types of within-year temporal patterns must be understood.

Knowledge of between-year variability is also important, as annual sampling or measurements must take into consideration the differences that occur between years. For many ecological parameters, collection and analysis data from a period of at least 5 consecutive years is necessary to minimally attempt to assess temporal variability of many ecosystem parameters. In some cases, a 5-year database may not be sufficient to assess interannual variation. Long-term data are generally not available except in isolated, existing long-term monitoring sites.

Spatial variability of ecological parameters may often exceed the range of temporal variability (Podlesakova and Nemecek, 1995). On a small scale, spatial differences may be attributable to the characteristics of the microsite, whereas factors such as slope, aspect, and elevation may be important on a larger scale.

An ideal indicator will exhibit relatively low natural variability both spatially and temporally when compared to the changes resulting from the stressors (Hinds, 1984). Unfortunately, low degrees of spatial and temporal variability are typically very difficult to attain in ecological systems. The ability to adequately define and quantify natural variability is a critical feature of the design of a monitoring program. In general, the monitoring of ecological status should be viewed as an experiment in testing the null hypothesis that the system is static.

Many past and current remote site monitoring programs have suffered from design problems that resulted in the inability to accurately determine signal-to-noise ratios in many of the parameters measured. These problems were caused by one or more of the factors listed below (Segar, 1986).

1. The species and sites used were selected according to their relative ease of sampling rather than from a scientific standpoint that would provide the most useful information.
2. Individuals (or individual samples) from a sampling site are pooled for analysis, thereby artificially reducing the spatial variability associated with the results.
3. Composite samples are used to reduce analytical costs, which also results in a reduction of spatial variability.
4. Variance estimates reported for a site are often based on analytical replicate variance only, without consideration for spatial variability. This results in the determination of statistical significance between two mean concentrations on the basis of analytical variance alone.
5. Within-year temporal variability is not considered, and/or sampling is not performed at a consistent or critical time (e.g., during spring runoff or at a critical stage of the life cycle).

Upon analysis of the data, failure to effectively consider natural spatial and temporal variability can easily lead to the wrong conclusion regarding ecological impacts. To avoid these problems, a properly designed ecological monitoring program should have the following characteristics (Segar, 1986):

1. The general objectives of the program should be clearly established (i.e., what are the resources at risk and in need of monitoring).
2. The specific objectives of the monitoring program should be clearly established (i.e., what parameters will be measured to meet the general objective of the program).
3. The limit of acceptable change (LAC) for each measured parameter should be specified and detectable.
4. Alternative null hypothesis should be established for each specific objective, stipulating the required resolution level.
5. The design of a sampling and analysis program should be established for each null hypothesis.
6. A specific null hypothesis should be selected to be tested for each specific objective. The spatial and temporal scale of the hypothesized effect that must be observed must be determined, as also the magnitude of smallest change or difference in mean value of monitored parameter that must be observed and statistically verified on the specified spatial and temporal

scales if the null hypothesis is to be disproved. When properly specified, these elements constitute the required resolution.

7. Establish and evaluate new null hypothesis if it is determined that all originally selected null hypotheses for a specific objective cannot be tested.

CRITERION 6: Quality Assurance — The criterion is satisfied if the quality of the resulting data can be reasonably assessed from a statistical and procedural standpoint. Ideally, quality assurance/quality control (QA/QC) procedures should be available for any parameter to be measured, and these procedures should be adequately referenced and outlined within the procedures used to collect the data. If no established QA/QC procedures are available, this criterion may still be satisfied if the technique lends itself to the development and application of effective QA/QC procedures.

Some parameters commonly associated with environmental monitoring programs are associated with long-established and well-accepted QA/QC procedures. For example, the wet deposition measurements collected as part of the NADP have utilized established, time-tested procedures. Similarly, many of the water chemistry procedures have good QA/QC procedures. However, many ecological parameters do not lend themselves well to effective QA/QC procedures. For example, the determination of fish age class based on the counting of scales is not generally well replicated.

CRITERION 7: Anticipatory — In many instances, an indicator applied to an ecological monitoring program should be designed to provide an early warning of widespread changes in ecological condition or processes. Measurable changes in many parameters currently being used in wilderness monitoring programs would not likely be observed until substantial damage has already occurred. For example, some programs estimate fluctuations in the populations of certain organisms. Should natural populations fluctuate measurably (i.e., to be able to distinguish from natural variability), it is likely that ecological damage has already occurred.

CRITERION 8: Historical Record — In some cases, historical data can be obtained for a parameter of interest from archived databases. Such data can be extremely valuable for establishing natural baseline conditions and the degree of natural variability associated with the parameter. For example, the U.S. Forest Service has long-term timber survey plots in many areas. In some cases, the data collected at these sites were related to timber production data only (i.e., tree species, diameter, height, crown class, etc.). In other instances, however, additional information may be available, such as the distribution of nonwoody plant species, wildlife, the presence of threatened or endangered species, etc. Similarly, many state fish and game agencies maintain substantial long-term databases on fishery status. Some lake chemistry data may also be available for many areas. Conversely, little information is generally available on parameters such as functional feeding groups in aquatic systems or other ecological parameters.

CRITERION 9: Retrospective — Some parameters allow for retrospective analysis in that new data may be generated that provide information on past conditions. For example, tree rings provide growth indices for each year of the life of the tree. Other parameters, such as metal concentrations in litter, tend to accumulate over time, such that sampling this medium provides data that are integrated over time. Most other parameters, such as ambient atmospheric monitoring, allow only for a "snapshot in time."

CRITERION 10: New Information — All parameters applied to an ecosystem monitoring program should provide new information rather than simply replicate data already collected. For example, a vegetation survey to determine the range of major plant communities in areas where the Forest Service already maintains such data would not be useful, as any observed change would invariably represent a substantial degree of impact.

CRITERION 11: Minimal Environmental Impact — Any procedure applied should result in minimal environmental impact to the area or ecosystem being monitored. Application of any indicator of damage to sensitive receptors should not in itself result in more environmental impact than the air pollution. Wherever possible, nondestructive biological surveys should be used rather than those which rely on destructive sampling techniques. Measurements that require considerable destructive sampling may not be acceptable within National Parks or wilderness areas, and should therefore be avoided. For example, tree ring chronologies and sapwood volumes are often determined from "cookies," or cross-sections of trees that are collected from a tree that must first be cut down. Such destructive sampling should not be performed unless the information generated from the sampling justifies the loss of the organisms being sampled.

Additionally, measurements that require large equipment should be avoided wherever possible. For example, most methods for measuring concentrations in ambient air involve the use of elaborate equipment housed in an instrument shelter. Although use of such equipment in some locations may be feasible, application of such techniques within most wilderness areas are not practical. Furthermore, since this equipment requires electrical power, this severely restricts the locations in which the equipment may be installed and potentially exposes the equipment to roadway pollutants.

10.5 CONCLUSIONS

This document provides an overview of a process proposed for developing ecological monitoring programs and a more detailed description of how ecological indicators should be selected for application to these programs. Whatever process is used in designing an ecological monitoring program, it must be based on sound science — i.e., monitoring activities should only be conducted if they have a sound scientific basis and if there is a reasonable probability that the resulting data will enable the status of the resources to be assessed. This dictates that hypothesis testing must be an integral part of any monitoring effort and that indicators applied to the program must meet certain predetermined criteria. Basing monitoring programs on sound scientific principles will ensure that the resulting programs are both credible and defensible.

The selection of ecological indicators according to a predetermined set of criteria provides a consistent, scientifically based process for selecting indicators, and it provides an opportunity for all stakeholders to become involved. Furthermore, the criteria list and ranking can be modified on a site-by-site basis to allow for the process to be applied to any ecosystem. Finally, as scientific knowledge progresses, the criteria may be applied to newly developed indicators or to existing indicators, for which new information exists.

As described above, the selection of ecological indicators should be based on preestablished criteria. Some of these criteria are "must" criteria because indicators that do not meet these criteria cannot adequately provide the information for which they are designed. Other, less restrictive or "want" criteria are those that are desirable but not necessarily crucial to the effectiveness of the indicator. Unlike the "must" criteria, the "want" criteria are not equally important and may therefore be ranked in terms of their relative importance.

Below is a proposed list of four "must" and seven "want" criteria for the selection of ecological indicators for most monitoring programs. Suggested ranking of the "want" criteria is also provided. It must be kept in mind that the separation of "must" from "want" criteria and the ranking of "want" criteria is at least to some degree subjective. It is further recognized that additional criteria not listed here may be important on a site-specific basis.

"Must" Criteria: The four "must" criteria are:

1. **Ecosystem Conceptual Approach —** Because our focus is on ecological indicators, any indicator selected must be related in some known way to the structure or function of the ecosystem under consideration. This criterion in not restrictive; it can be satisfied by virtually any chemical, physical, or biological parameter. However, there should be a clear understanding of the relationship between the measurement and the structure and/or function of the ecological system. For example, streamwater pH may be an effective indicator if it is understood that decreased pH alters the structure of the benthic invertebrate community.
2. **Cause/Effect —** There must be a clearly understood relationship between the stressor (cause) and changes in the parameter measured (effect). For atmospheric pollutants, this generally includes dose-response relationships.
3. **Signal-to-Noise Ratio —** Ideally, the natural variability (spatial and temporal) observed in the parameter should be relatively small in comparison to changes due to pollutant inputs. In this way, the signal-to-noise ratio is such that effects due to the pollutants of interest are readily distinguishable from natural variability.
4. **Quality Assurance —** The quality of the resulting data should be reasonably well assured from a statistical and procedural standpoint. Data generated without adequate quality assurance are not defensible scientifically.

"Want" Criteria (Ranked): The "want" criteria, ranked in order of their anticipated importance, are:

1. **Usability—**Procedures should be complete and thorough. Ideally, should use detailed and established SOPs based on standardized methods.
2. **Anticipatory—**Indicators should provide an early warning of widespread changes in ecological condition before substantial damage occurs.

3. **Result in Minimal Environmental Impact —** Non- or minimally destructive sampling techniques should be used, and measurements that require large equipment deployed over long periods of time should be avoided.
4. **Cost-Effectiveness —** The incremental cost associated with measuring a parameter should be low relative to the information obtained.
5. **Historical Record Available —** Information gained may be strengthened if a quality-assured historical database is available to provide historical time-series data.
6. **Provide Retrospective Information —** Application of some parameters will provide information on past conditions in addition to present conditions.

It should be remembered, however, that each monitoring program may rank criteria differently, depending on the objectives of the program as well as other factors.

REFERENCES

Adams, M.B., D.S. Nichols, C.A. Federer, K.F. Jensen, and H. Parrott, 1991, Screening Procedure to Evaluate Effects of Air Pollution on Eastern Region Wildernesses Cited as Class I Air Quality Areas, USDA Forest Service, Northeastern Forest Experiment Station, St. Paul, MN, General Technical Report NE-151.

Bruns, D.A., G.B. Wiersma, and G.J. White, 1997, Testing and application of ecosystem monitoring parameters, *Toxicol. Environ. Chem.*, 62: 169–196.

Bruns, D.A., G.B. Wiersma, and E.J. Rykiel, Jr., 1991, Ecosystem monitoring at global baseline sites, *Environ. Monit. Assess.*, 17: 3–31.

Hinds, W.T., 1984, Towards monitoring of long-term trends in terrestrial ecosystems, *Environ. Conserv.*, 11: 11–18.

Kurtz, J.C., L.E. Jackson, and W.S. Fisher, 2001, Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development, *Ecol. Indicat.*, 1: 49–60.

Margalef, R., 1981, Human impact on transportation and diversity in ecosystems. How far is extrapolation valid?, in *Proceedings of the First International Congress of Ecology*, Centre of Agricultural Publishing and Documentation, Wageningen, Netherlands, pp. 237–241.

Messer, J.J., R.A. Linthurst, and W.S. Overton, 1991, An EPA program for monitoring ecological status and trends, *Environ. Monit. Assess.*, 17: 67–78.

National Research Council, 1986, Indicator species and biological monitoring, *in Ecological Knowledge and Environmental Problem-Solving: Concepts and Case Studies*, National Academy Press, Washington, D.C., pp. 81–87.

Noss, R.F., 1990, Indicators for monitoring biodiversity: a hierarchical approach, *Conserv. Biol.*, 4: 355–364.

Odum, E.P., 1985, Trends expected in stressed ecosystems, *BioScience*, 35: 419–422.

Oliver, I., 2002, An expert panel-based approach to the assessment of vegetation condition within the context of biodiversity conservation. Stage 1: the identification of condition indicators, *Ecol. Indicat.*, 2: 223–237.

Peine, J.D., J.C. Randolph, and J.J. Presswood, Jr., 1995, Evaluating the effectiveness of air quality management within the Class I Area of Great Smoky Mountains National Park, *Environ. Manage.*, 19: 515–526.

Peterson, D.L., J.M. Eilers, R.W. Fisher, and R.D. Doty, 1992, Guidelines for Evaluating Air Pollution Impacts on Class I Wilderness Areas in California, USDA Forest Service Pacific Southwest Research Station, Albany, CA, General Technical Report PSW-GTR-136.

Peterson, J., D.L. Schmoldt, D.L. Peterson, J.M. Eilers, R.W. Fisher, and R. Bachman, 1992, Guidelines for Evaluating Air Pollution Impacts on Class I Wilderness Areas in the Pacific Northwest, USDA Forest Service Pacific Northwest Research Station, Portland, OR, General Technical Report PNW-299.

Podlesakova, E. and J. Nemecek, 1995, Retrospective monitoring and inventory of soil contamination in relation to systematic monitoring, *Environ. Monit. Assess*., 34: 121–125.

Schmoldt, D.L. and D.L. Peterson, 1991, Applying knowledge-based methods to design and implement an air quality workshop, *Environ. Manage*., 15: 623–634.

Segar, D.A., 1986, Design of monitoring studies to assess waste disposal effects on regional to site specific scales, in *Public Waste Management and the Ocean Choice*, K.D. Stolzenbach, J.T. Kildow, and E.T. Harding, Eds., MIT Sea Grant College Program, Cambridge, MA, pp. 189–206.

Sigal, L.L. and G.W. Suter, II, 1987, Evaluation of methods for determining adverse impacts of air pollution on terrestrial ecosystems, *Environ. Manage*., 11: 675–694.

Suter, G.W., II, 2001, Applicability of indicator monitoring to ecological risk assessment, *Ecol. Indicat.*, 1: 101–112.

11 Efficacy of Forest Health Monitoring Indicators to Evince Impacts on a Chemically Manipulated Watershed

G.B. Wiersma, J.A. Elvir, and J. Eckhoff

CONTENTS

1-56670-641-6/04/$0.00+$1.50

11.1 INTRODUCTION

The impacts of acid deposition, especially nitrogen and sulfur oxide deposition originated from anthropogenic sources, has been a focal point of research for more than three decades. Due to the long-term consequences of enhanced levels of acidifying compounds on forest ecosystems, acid deposition continues to be an area of major concern in countries around the world (Van Dobben 1999, Hallbacken and Zhang 1998, Amezaga et al. 1996, Wolterbeek et al. 1996, Meesenburg et al. 1995, Bussotti et al. 1995, Forster, 1993, Farmer et al. 1991).

In the 1980s, the U.S. EPA Science Advisory Board developed the Environmental Monitoring and Assessment Program (EMAP) to determine the current extent and condition of the nation's ecological resources (U.S. EPA 1993). The USDA Forest Service and the EPA, in cooperation with several other federal and state agencies, developed the FHM program to identify and develop indicators to assess forest health. The FHM was designed to address concerns of potential effects from air pollution, acid rain, global climate change, insects, disease, and other stressors on forest health and to assist resource managers and policy makers in managing the forest resources, evaluating policy, and allocating funds for research and development (Alexander and Palmer 1999, U.S. EPA 1994, Burkman and Hertel 1992).

Also in 1987, in response to concerns about the impacts of atmospheric acidic deposition on forest ecosystems, the EPA provided funds for the establishment of Bear Brook Watershed in Maine (BBWM) as a watershed manipulation project under the National Acid Precipitation Assessment Program (NAPAP). The BBWM was established to identify atmospheric deposition impacts on surface waters and to quantify the major processes controlling surface water acidification under increased sulfur and nitrogen atmospheric deposition (Norton et al. 1992, Uddameri et al. 1995).

The BBWM is formed by two contiguous watersheds, West Bear and East Bear. The two watersheds have similar hydrology, soils, vegetation, topography, relief, and aspect characteristics except for the experimental ammonium sulfate [$(NH_4)_2SO_4$] amendments applied to West Bear since 1989 (Uddameri et al. 1995, Fernandez et al. 1999, Norton et al. 1999a). Results from monitoring the hydrology during the first 3 years of treatment (1989 to 1992) indicated that additions of $(NH_4)_2SO_4$ produced significant changes in stream-water chemistry, including significant increases in base cations, hydrogen ions, total aluminum, sulfate, and nitrate concentrations, along with decreases in alkalinity and dissolved organic acid (DOC) concentrations (Kahl et al. 1993). Impacts of $(NH_4)_2SO_4$ on the forest vegetation at BBWM have also been examined.

Chemical analyses indicated elevated nitrogen and aluminum concentrations and lower calcium and potassium concentrations in foliage of several of the dominant tree species (White et al. 1999) and two bryophyte species, *Bazzania trilobata* (a liverwort) and *Dicranum fulvum* (a true moss) (Weber and Wiersma 1997) growing in the treated West Bear watershed.

The goal of this study was to test the efficacy of FHM indicators to evince impacts of enhanced acidic deposition on forest vegetation in the treated watershed at the BBWM. The information presented here was abstracted from a Ph.D. dissertation (Eckhoff 2000) which presents greater detail about the methodology and results including a complete description in the application of the FHM indicators at the BBWM. Study objectives were:

1. To evaluate the efficacy of five FHM indicators — forest mensuration, crown condition classification, damage and catastrophic mortality, lichen communities, and vegetation structure
2. To evaluate the efficacy of two additional indicators not part of the FHM program: canopy gap analysis and tree seed production
3. To describe the status of the vegetation at the BBWM

11.2 METHODS

11.2.1 Study Area

This study was conducted at the BBWM site which is located in eastern Maine, U.S. (44°52'15" N, 68°06'25" W) (Figure 11.1). The site lies on the upper 210 m of the southeast slope of Lead Mountain (475 m) with a mean slope of 31% (Norton et al. 1999a). The BBWM is formed by two contiguous forested watersheds, West Bear (WB) and East Bear (EB), with areas of 10.77 and 11.42 ha, respectively. Both watersheds are drained by first-order streams and have similar soils, vegetation, topography, relief, aspect, and exposure (Uddameri et al. 1995, Norton et al. 1992). Climate in the BBWM area is temperate with a temperature range of 35°C in the summer to −30°C in the winter and mean annual precipitation around 1.4 m; about 25% is in the form of snow. The soils at BBWM are predominantly haplorthods, tunbridge, rawsonville, ricker, and dixfield series soils, with well-developed spodosol mineral soils that average 0.9 m thick (Norton et al. 1999a). The BBWM forest

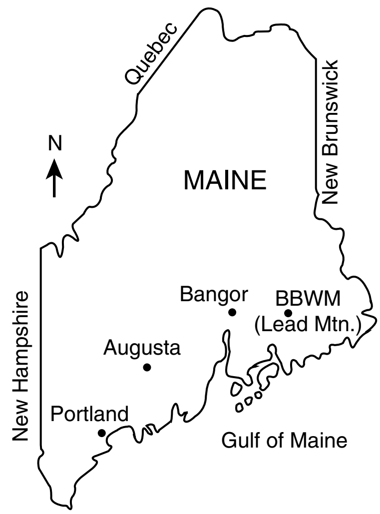


FIGURE 11.1 Location of the Bear Brook Watershed in Maine.

vegetation is dominated by five species: red spruce (*Picea rubens* Sarg.), American beech (*Fagus grandifolia* Ehrh.), red maple (*Acer rubrum* L.), sugar maple (*Acer saccharum* Marsh.), and yellow birch (*Betula alleghaniensis* Britt.), distributed in three cover types: hardwood, softwood, and mixedwood. Forests at the BBWM are mature stands with a mean age of ~50 years for hardwoods and ~90 years for softwoods (Elvir et al. 2003).

11.2.2 Treatment

During the three-year calibration period (1987 to 1989), wet plus dry deposition at BBWM was estimated to be ~600 $eq * ha^{-1} * yr^{-1}$ (~8.4 $kg * ha^{-1} * yr^{-1}$) for N and ~900 $eq^{-1} * yr^{-1}$ (~14.4 $kg * ha^{-1} * yr^{-1}$) for S (Kahl et al. 1999, Norton et al. 1999a). The manipulation of WB was initiated in November 1989, and consists of bimonthly additions of dry ammonium sulfate [$(NH_4)_2SO_4$] at the rate of 300 eq NH_4 and SO_4 $ha^{-1} *$ application^{-1}, or 1800 eq NH_4 and SO_4 $ha^{-1} * yr^{-1}$ (118.8 kg $(NH_4)_2SO_4$ $ha^{-1} * yr^{-1}$) (Norton et al. 1999a). With this treatment, atmospheric deposition in the WB is considered comparable to some areas with the highest deposition rates in the U.S. but lower than heavily polluted areas in central Europe (Lovett 1994, Lindberg and Lovett 1993, Lindberg and Owens 1993, Rustad et al. 1994, Eagar et al. 1996). Additional details on the Bear Brook site and treatment are available (see Church 1999, Norton et al. 1999a, 1999b).

11.2.3 Protocols

A brief synopsis of the protocols and methods used for sample collections and statistical analyses in the FHM indicators applied at the BBWM study by Eckhoff (2000) is given here. Complete details on the FHM program protocols are available elsewhere (Tallent-Halsell 1994).

11.2.3.1 Plot Design

The FHM plot design for the forest mensuration, crown condition, damage and catastrophic mortality assessment, and vegetation structure indicators includes a cluster of four 0.1-ha fixed-radius (17.95 m) annular plots in a triangular design (Figure 11.2). The center of annular plot 1 is the center of the overall plot. From the center of annular plot 1, the center of annular plot 2 is located 360° and 36.6 m, the center of annular plot 3 is located 120° and 36.6 m, and the center of annular plot 4 is located 240° and 36.6 m.

Within each annular plot is nested a 1/60 ha, fixed-radius (7.32 m) subplot. Within each nested subplot is a 1/750 ha fixed-radius (2.07 m) microplot. It is located 90° and 3.66 m east of the subplot center. Also within each subplot are three 1-m^2 quadrats. The 3 quadrats are each located 4.57 m from the subplot center, the first at 30°, the second at 150°, and the third at 270°. The FHM plot design for the lichen communities indicator is a 0.378 ha circular plot (36.6 m) centered in the middle of subplot 1, excluding the areas inside the four subplot boundaries (Figure 11.2).

Tree seed production and canopy gap analysis indicators are not part of the FHM program; however, the plot designs for collecting data in this study with these indicators were integrated into the existing FHM plot design. For the tree seed production indicator, one seed trap was placed 0.5 m south of each subplot center. For the canopy gap analysis indicator, six measurements are recorded at locations around the perimeter of the subplot and one at the subplot center for a total of seven.

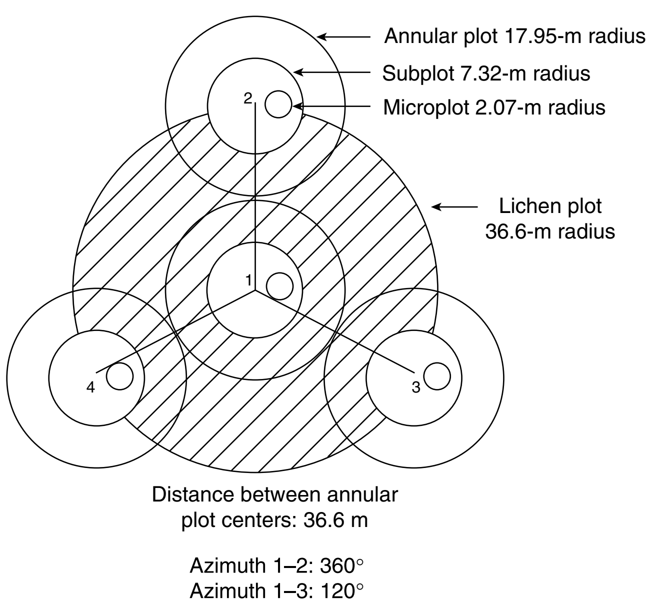


FIGURE 11.2 FHM plot design for lichen communities indicator. (Adapted from Tallent-Halsell, N.G. 1994. Forest Health Monitoring 1994 Field Methods Guide. EPA/620/R94/027. U.S. Environmental Protection Agency, Washington, D.C.)

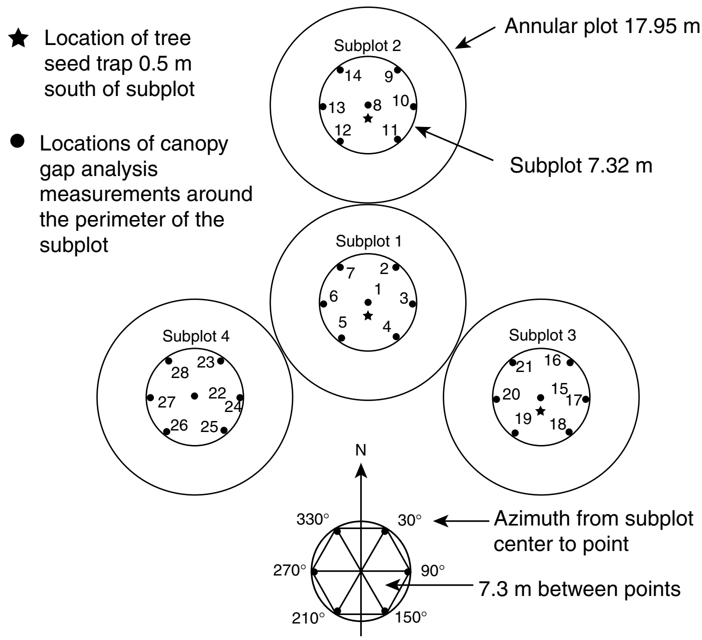


FIGURE 11.3 Plot design for tree seed production and canopy gap analysis indicators. (Adapted and modified from Tallent-Halsell, N.G. 1994. Forest Health Monitoring 1994 Field Methods Guide. EPA/620/R94/027. U.S. Environmental Protection Agency, Washington, D.C.)

The six locations around the perimeter (7.3 m from the subplot center) are at 30°, 90°, 150°, 180°, 210°, 270°, 330°, and 360° (Figure 11.3).

11.2.3.2 Plot Layout

Ten FHM design cluster-plots were established systematically in each treated WB and reference EB watersheds in 1996 (Figure 11.4). An additional five cluster-plots were established east of EB, which are referred to as the "A" plots, and an additional five cluster-plots were established west of WB, referred to as the "Y" plots. The A and Y plots were established as a strategy to increase reference areas, therefore, the treated WB plots were compared to both EB and A and Y plots.

11.2.4 Survey Methods

11.2.4.1 FHM Forest Mensuration Indicator

All standing live trees and snags (standing dead trees), ≥12.7 cm dbh, within the subplots were tallied from July to August 1997. The tree species was recorded and dbh was marked and measured at 1.37 m above the ground line on the uphill side of the tree. In the microplots, all saplings with dbh ≥2.5 cm but <12.7 cm were measured and their species was recorded. Also in the microplots, tree seedlings <2.5 cm and >30 cm in height were counted by species.

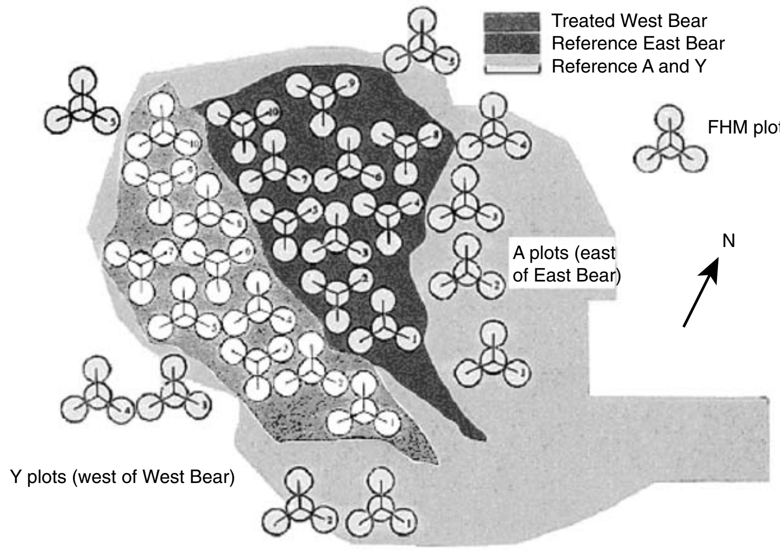


FIGURE 11.4 Location of the study plots at the Bear Brook Watershed in Maine.

11.2.4.2 FHM Damage and Catastrophic Mortality Indicator

Damage and mortality measurements were recorded for all the live trees within the subplots and the saplings within the microplots from September to October 1997. Recorded damages were prioritized first by location on the tree and then by both the type of damage sign or symptom and the severity level of the damage (Table 11.1). Severity refers to the amount of area affected for any recorded damage sign or symptom in a given location. Severity level was generally recorded in percentage classes (0 to 9%, 10 to 19%, and so on), each successive class indicates a 10% increase in the total area affected.

11.2.4.3 FHM Crown Condition Classification Indicator

Crown condition classification measurements were taken for all the live trees in the 7.3 m subplots and all the saplings and seedlings in the 1/750 ha microplots from September to October 1997. Crown classes included dominant, codominant, intermediate, and overtopped. Crown class described the extent of sunlight reaching the tree crown and the position of the tree crown in relation to its neighboring trees (Tallent-Halsell 1994). The FHM crown condition classification measurements for trees include live crown ratio, crown diameter, crown density, crown dieback, and foliage transparency. In this study, trees in the subplots were rated for only two crown condition variables: live crown ratio (LCR) and crown vigor.

LCR reflects the percentage of the total tree, sapling, or seedling height that is supporting live green foliage and was determined by:

$$\text{LCR} = \frac{\text{length of the live crown}}{\text{total height of the true, sapling, or seedling}}$$

TABLE 11.1
Location and Types of Damage Signs and Symptoms

Location of Damage Signs and Symptoms	Types of Damage and Mortality Signs and Symptoms
Roots (exposed) and stump (0.3 m) (in height from ground level)	Canker
Roots and lower bole	Conks (or other indicators of advanced decay)
Lower bole (lower half of the trunk between the stump and base of the live crown)	Open wounds
Lower and upper bole, upper bole (upper half of the trunk between stump and base of the live crown)	Resinosis or gummosis
Crownstem (main stem within the live crown area, above the base of the live crown)	Broken bole or roots less than 0.91 m from the bole
Branches (woody stems other than main stem)	Brooms on roots or bole
Buds and shoots (the most recent year's growth)	Broken or dead roots beyond 0.91 m
Foliage	Loss of apical dominance, dead terminal
	Broken or dead branches or shoots
	Excessive branching or brooms
	Damaged foliage or shoots
	Discoloration of foliage
	Other (signs and symptoms other than the types described)

Source: Adapted from Tallent-Halsell, N.G. 1994. Forest Health Monitoring 1994 Field Methods Guide. EPA/620/R94/027. U.S. Environmental Protection Agency, Washington, D.C.

The length of the live crown extends from the top (excluding dieback) of the living crown to the lowest foliage on the lowest branch at the base of the live crown.

Crown vigor assessment was based on the extent of the live crown ratio, extent of dieback in the upper half of the crown or outer exposed portion of the crown, and extent of foliage health. Each individual tree, sapling, and seedling was placed into one of three crown vigor classes: high, moderate, or low.

11.2.4.4 Tree Seed Production Indicator

For seed collection, seed traps were constructed following the design used at the Holt Research Forest, Arrowsic, ME (Witham et al. 1993). The frame or basket of the seed trap was an 80-cm tall, 14-cm radius at the top, wire basket. The bag which hung inside the basket was constructed from nylon mesh fabric. There was one seed collection basket in each subplot for a total of four per plot. Seeds were collected semiannually, one collection in midsummer and one early the following spring, for 2 years (1997 to 1999). The seeds and litter were brought to the lab where the seeds were sorted by species.

11.2.4.5 Tree Canopy Gap Analysis Indicator

The tree canopy gap analysis indicator measurements were recorded using a LI-COR LAI-2000™ plant canopy analyzer. Measurements were only recorded on clear sunny days from mid-August to mid-September in 1996 and 1997. The LAI-2000 canopy analyzer was used to determine the amount of light penetration to the forest understory. LAI-2000 measurements were expressed as percentages with values that ranged from 0% (no sky visible to the sensor) to 100% (full sun visible to the sensor) (LI-COR 1992). For the analysis in this study, the four transmittance values from the four subplots were averaged to provide one cluster-plot value.

11.2.4.6 FHM Vegetation Structure Indicator

All plant species and their abundances (percent cover) within the 1-m^2 quadrats (three quadrats per subplot) were measured from June to August 1997. The vertical distribution for measurements was subdivided into four strata above ground: 0 to 0.6 m, 0.6 to 1.8 m, 1.8 to 4.9 m, and > 4.9 m.

11.2.4.7 FHM Lichen Communities Indicator

Data were collected on the presence and abundance of lichens on woody plants in the treated WB and the reference EB only. The lichens were identified and abundances recorded in a walking reconnaissance within a 0.38-ha circular plot centered in the middle of subplot 1. The abundance codes used were (1) rare (<3 individuals), (2) uncommon (4 to 10 individuals), (3) common (>10 individuals but <1/2 of the boles and branches have that species present), and (4) abundant (>10 individuals and >1/2 of the boles and branches have that species present).

Samples of all the different lichen species found in each plot were collected and brought back to the lab for identification. Chemical methods of species identification generally followed the procedures presented by Hale (1979). Verification of the lichen samples from BBWM was done at the University of Maine in Orono by James Hinds, University of Maine Department of Biological Sciences, and Patricia Hinds of Orono, Maine.

11.2.5 Analysis

Data were analyzed by cluster plots. Data from A and Y plots were pooled for analyses. Data were analyzed among study areas using the total of plots in each area and also using plots grouped by forest types. The treated WB means were the standards to which EB and A and Y means were compared.

All statistical analyses were accomplished using the SAS® program JMP® (SAS 1998). Statistical analysis included the one-way analysis of variance (ANOVA) for continuous variables. The ANOVA assumptions were tested using Levene's test for homogeneity of variances and Shapiro–Wilk W test for normality. When transformations were necessary to meet ANOVA assumptions, log, square root, or 1/square-root transformations were used. Dunnett's test was used for multiple comparisons to assess the significance of differences among treatment means. Rank tests were performed with Standard Least Squares. Contingency table analysis was used for discrete variables. Randomization testing, also known as permutation testing, was used for determining statistical significance of species abundance. A probability of 0.05 was used to confer significance. Throughout this document, nonsignificant differences are referred to analyses with p-values >0.05.

11.3 RESULTS

11.3.1 FHM Forest Mensuration Indicator

11.3.1.1 Trees

There were 15 species recorded in the tree, sapling, and seedling categories at BBWM. No significant differences in the total number of trees/ha between the treated WB and the reference EB or A and Y were found (Table 11.2).

The major dominant tree species in all three areas were *Acer pensylvanicum* (striped maple), *Acer rubrum* (red maple), *Acer saccharum* (sugar maple), *Betula alleghaniensis* (yellow birch), *Fagus grandifolia* (American beech) and *Picea rubens* (red spruce). These species comprised 99% of all the live trees in West Bear, 98% in East Bear, and 96% in A and Y. The most abundant tree species was *Picea rubens* followed by *Fagus grandifolia*. *Acer rubrum* and *Acer saccharum* were the next most abundant tree species. WB, EB, and A and Y had similar number of trees/ha for all dominant species, except for *Acer rubrum* which was significantly lower in the treated WB (p-value 0.03). No significant differences in number of snags/ha were found between the treated WB and the reference EB or A and Y.

TABLE 11.2
Number of Live Trees and Snags per ha by Species at Bear Brook Watershed in Maine

Species Latin Name	Common Name	West Bear[a] Trees/ha	East Bear Trees/ha	A and Y Trees/ha
Picea rubens	Red spruce	285	199.5	429
Fagus grandifolia	American beech	211.5	261	148.5
Acer saccharum	Sugar maple	97.5	39	30
Betula alleghaniensis	Yellow birch	48	75	52.5
Acer rubrum	Red maple	12[b]	90	63
Abies balsamea	Balsam fir	1.5	6	19.5
Acer pensylvanicum	Striped maple	15	3	15
Betula populifolia	Gray birch	1.5	0	0
Fraxinus americana	White ash	0	1.5	0
Pinus strobus	Eastern white pine	0	0	1.5
Prunus serotina	Black cherry	1.5	0	0
Quercus rubra	Northern red oak	0	0	4.5
Sorbus americana	American mountain ash	1.5	3	3
Snags		113	120	108
Total number of trees		787.5	798	874.5

[a] Manipulated watershed.
[b] Significantly different.

TABLE 11.3
Number of Live Saplings per ha by Species at Bear Brook Watershed in Maine

Species Latin Name	Common Name	West Bear[a] Saplings/ha	East Bear Saplings/ha	A and Y Saplings/ha
Picea rubens	Red spruce	262.5	187.5	243.8
Fagus grandifolia	American beech	693.8	731.3	412.5
Acer saccharum	Sugar maple	112.5	37.5	0
Betula alleghaniensis	Yellow birch	75	75	18.8
Acer rubrum	Red maple	0	56.3	75
Abies balsamea	Balsam fir	0	168.8	18.8
Acer pensylvanicum	Striped maple	56.3	18.8	37.5
Sorbus americana	American mountain ash	0	18.8	112.5
Total number of saplings		1200	1293.8	918.8

[a] Manipulated watershed.

11.3.1.2 Saplings

Sapling species diversity was higher in the reference EB with eight species, followed by A and Y with seven species, and the treated WB watershed with five species (Table 11.3). The difference in the total number of saplings per hectare for all species combined between WB and EB and A and Y was not significant. *Fagus grandifolia* was the most abundant species in the three areas followed by *Picea rubens*. There was no significant difference between the treated WB and the reference EB or A and Y in the number of saplings/ha in any of the dominant species.

11.3.1.3 Seedlings

There were six, eleven, and twelve seedling species recorded in WB, EB, and A and Y, respectively (Table 11.4). The total number of seedlings per ha was very similar among study sites with no significant differences. No significant differences were found in number of seedlings for each dominant species except for *Acer saccharum* (p-value 0.04) being significantly higher in the treated WB.

11.3.1.4 Diameter Size

The dbh ranged from 12.7 to 71.5 cm, with a 22 cm mean for all the live trees in the three study areas (Table 11.5). The shape of the tree and sapling dbh distributions was an inverse J, with decreasing numbers of trees or saplings with increasing dbh size. There were no significant differences in dbh for any of the major dominant species between the treated WB and the reference EB or A and Y (Table 11.6).

TABLE 11.4
Number of Live Seedlings per ha by Species at Bear Brook Watershed in Maine

Species Latin Name	Common Name	West Bear[a] Seedlings/ha	East Bear Seedlings/ha	A and Y Seedlings/ha
Picea rubens	Red spruce	487.5	356.3	731.3
Fagus grandifolia	American beech	1593.8	2137.5	1800
Acer saccharum	Sugar maple	4593.8	412.5	937.5
Betula alleghaniensis	Yellow birch	1312.5	3337.5	1125
Acer rubrum	Red maple[b]	75	581.3	825
Abies balsamea	Balsam fir	0	187.5	93.8
Acer pensylvanicum	Striped maple	3581.3	1481.3	2062.5
Fraxinus americana	White ash	0	243.8	18.8
Prunus serotina	Black cherry	0	0	150
Prunus virginiana	Choke cherry	0	318.8	56.3
Quercus rubra	Northern red oak	0	0	56.3
Sorbus americana	American mountain ash	0	112.5	206.3
Tsuga canadensis	Eastern hemlock	0	56.3	0
Total number of seedlings		11643.8	9225	8062.5

[a] Manipulated watershed.
[b] Significantly different.

TABLE 11.5
Statistics for the Overall dbh of All Live Trees at Bear Brook Watershed in Maine

	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)
West Bear[a]	450	22.1	9.8	71.5
East Bear	452	21.8	8.6	67.8
A and Y	511	21.9	8.4	67.6

[a] Manipulated watershed.

11.4 FHM DAMAGE AND CATASTROPHIC MORTALITY ASSESSMENT INDICATOR

Visual signs of foliar chlorosis or necrosis associated with acidic deposition were not observed on any of the trees or saplings. Damage recorded indicated that for the tree species about 45%, 40%, and 60% in WB, EB, and A and Y, respectively, and for samplings about 22%, 26%, and 39% in WB, EB, and A and Y, respectively, had no damage signs or symptoms present. The damage category with the highest

TABLE 11.6
Tree dbh Statistics by Individual Species at Bear Brook Watershed in Maine

	Acer pensylvanicum				*Acer rubrum*			
	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)
West Bear[a]	10	15.3	2.1	18.3	8	25.6	8.8	39.6
East Bear	2	17.1	3.3	19.4	60	19.7	6.4	44.7
A and Y	10	15.3	1.7	17.4	42	22.4	8.3	44.3

	Acer saccharum				*Betula alleghaniensis*			
	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)
West Bear[a]	65	22.7	14.3	71.5	32	24.8	14.1	67.7
East Bear	26	20.4	12.6	67.8	50	25.0	11.9	55.5
A and Y	20	26.4	15.3	63.3	35	24.2	13.8	67.6

	Fagus grandifolia				*Picea rubens*			
	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)	# of Trees	Mean dbh (cm)	Stdv (cm)	Maximum dbh (cm)
West Bear[a]	141	18.8	6.6	48.3	190	24.2	8.6	66.8
East Bear	174	19.5	6.0	46	133	25.1	8.7	57.0
A and Y	99	19.4	6.4	44	286	22.5	7.3	58.6

[a] Manipulated watershed.

occurrence in all three study areas was cankers, followed by conks, and other indicators of advanced decay, open wounds, and loss of terminal leader.

Between 50 and 70% of the *Acer rubrum, Acer saccharum*, and *Betula alleghaniensis* trees had some sign or symptom of damage, especially conks and fruiting bodies, in all study areas. For *Fagus grandifolia* almost 100% of the trees in the three treatment areas had signs or symptoms of cankers. The highest occurrence of cankers was related to the beech bark disease affecting *F. grandifolia*. *P. rubens* showed the lowest percent of trees with damages (<20%), which included cankers, conks, and advanced decay. There was no significant difference between the treated WB and the reference EB or A and Y in the distribution of damage signs for trees or saplings in the conifer forest type. However, in the deciduous forest type, the treated WB showed a significantly larger number of trees with open wounds (p-value 0.01). About 33% and 30% of the open wounds in WB occurred on *Acer saccharum* and *Picea rubens* trees, respectively. In all three areas, the most common location of damage signs and symptoms on the tree or sapling was the roots and lower bole.

11.4.1 FHM Crown Condition Classification Indicator

11.4.1.1 Crown Class

Between 71 and 76% of the live trees in all three study areas were in the codominant crown class, 14 to 20% in the intermediate class, and 9 to 12% in the overtopped class. No dominant crown class was found, and therefore this class was omitted in this crown class analysis.

For the overall trees, there was a significant difference (p-value 0.04) between the treated WB and the references EB and A and Y, the number of trees in the intermediate crown class being significantly higher in WB; no differences were found in the other two crown classes. Distributions of the number of trees, saplings, and seedlings per crown classes between the treated WB and the reference EB or A and Y for each dominant species were similar.

11.4.1.2 Live Crown Ratio

Overall, the LCR ranged from 5% up to 100% with a mean of approximately 47% for all the trees in the three study areas. The LCR distributions of the trees in the three study areas approximate normal distributions, with decreasing numbers of trees in both the high and low percentages. Overall there were not significant differences in LCR between the treated WB and the reference EB or A and Y for all tree species combined or for any major tree species. Also the differences in either sapling or seedling LCR between the treated WB and the reference EB or A and Y were not significant.

11.4.1.3 Crown Vigor

Within the three study areas, WB, EB, and A and Y, ~80% of the trees/ha had a high crown vigor, ~18% had a medium crown vigor, and only 1 to 2% had low crown vigor. For all the tree species combined, there were not significant differences in the crown vigor class distributions between the treated WB and the reference EB or A and Y.

For *A. rubrum, A. saccharum, B. alleghaniensis*, and *P. rubens* between 77 and 100% of the trees had high crown vigor, regardless if they were in the codominant, intermediate, or overtopped crown class. Compared to those species, *F. grandifolia* had a lower number of trees with high crown vigor (~62%) which might be attributed to the effects of bark disease on this species. Differences in any crown vigor class for major dominant tree species between treated and reference areas were not significant, except for *P. rubens* which had a significantly higher number of trees in the medium vigor class (p-value 0.01) in the reference A and Y. Also, differences in either sapling or seedling crown vigor classes between the treated WB and the reference EB or A and Y were not significant.

11.4.2 Tree Seed Production Indicator

In the 2 years of seed collection (1997 to 1999) the annual mean of seeds production for all tree species combined or individual dominant tree species varied among study areas (Table 11.7). Although WB had an overall lower annual mean seed production, the difference was not significant compared to EB and A and Y. Among species,

TABLE 11.7
Annual Mean of the Total Number of Seeds per ha (1997–1999), of the Major Dominant Species at Bear Brook Watershed in Maine

Species Latin Name	Common Name	West Bear[a] # Seeds/ha	East Bear # Seeds/ha	A and Y # Seeds/ha
Betula alleghaniensis	Yellow birch	9,321,137	13,544,713	11,172,402
Picea rubens	Red spruce	3,339,430	2,871,951	3,976,093
Acer saccharum	Sugar maple	274,390	144,309	154,472
Fagus grandifolia	American beech	128,049	286,585	93,496
Acer rubrum	Red maple	123,984	227,642	380,751
Acer pensylvanicum	Striped maple	191,057	138,211	217,480
Other species		87,398	193,089	212,052
Total		13,465,445	17,406,500	16,206,746

[a] Manipulated watershed.

Betula alleghaniensis had the highest seed production per hectare, followed by *Picea rubens*. Differences in seed production per ha for any of the major dominant species between the treated WB and the reference EB or A and Y were not significant.

11.4.3 Tree Canopy Gap Fraction Indicator

The canopy gap fraction indicator was a method of assessing the abiotic environment in terms of light resources available to the understory vegetation. In 1996, LAI-2000 measurements were recorded for the treated WB and the reference EB only. Overall, the average amount of light under the mature canopies at BBWM during the growing season ranged between 1 and 6% of full sunlight.

In 1997, data were collected in WB, EB, and A and Y. The proportion of visible sky data in 1997 was very consistent with data collected in 1996 and no significant differences between years in the treated WB or the reference EB. Differences in the proportion of visible sky between the treated WB and reference EB or A and Y were not significant in any forest type in any year.

Results in this study were similar to other studies reporting that deciduous or conifer closed canopies screened up to 90% of the visible light during the growing season, with only 0.5 to 5% of solar radiation reaching the forest floor beneath closed canopies (Reifsnyder and Lull 1965, Chazdon and Pearcy 1991, Constabel and Lieffers 1996, Chazdon 1986).

11.4.4 FHM Vegetation Structure Indicator

11.4.4.1 Forb, Graminoid, Moss, and Lichen Species

Twenty-two forb species were found at the BBWM (Eckhoff 2000). Forbs were found only in the lower stratum (0 to 0.6 m). In the deciduous forest types, 15 forb species were recorded in both the treated WB and the reference EB, and 12 forb species were recorded in the reference A and Y. The average abundance of forbs was 9%, 11%, and 16% of the total area in WB, EB, and A and Y, respectively. Six forb species, *Aralia nudicaulis* (wild sarsaparilla), *Aster* spp. (aster), *Maianthemum canadense* (false lily of the valley), *Medeola virginiana* (Indian cucumber-root), *Trientalis borealis* (starflower), and *Uvularia sessilifolia* (wild oats) comprised up to 92% of the total percent cover for all the forbs. No significant differences were found between the treated WB and the reference EB or A and Y in the overall abundance of forbs in any forest type.

Graminoid species were found only in the lower stratum (0 to 0.6 m). In the deciduous forest types, the average abundance of *Carex* spp. (sedge) and grasses combined were very similar among study areas, covering approximately 1% of the overall area.

Moss and lichen species were found and recorded only in the lower stratum (0 to 0.6 m). Differences between the treated WB and the reference EB or A and Y in lichen or moss abundance in any forest type were not significant.

11.4.4.2 Fern Species

Six fern species were found at the BBWM. Ferns were found predominantly in the lower stratum (0 to 0.6 m) with only two species (*Dryopteris campyloptera* and *Osmunda claytoniana*) reaching the second stratum (0.6 to 1.8 m). In the deciduous forest types, six fern species were recorded in both the treated WB and the reference EB, and four species were recorded in A and Y. The most common fern in all plots was *Dryopteris campyloptera* (mountain wood-fern) (> 72% of all the fern species present) followed in a decreasing order by *Gymnocarpium dryopteris* (oak fern), *Thelypteris noveboracensis* (New York fern), *Osmunda claytoniana* (interrupted fern), *Polystichum acrostichoides* (Christmas fern), and *Dennstaedtia punctilobula* (hay-scented fern). In the coniferous forest types, only two fern species were observed, *Dryopteris campyloptera* and *Dennstaedtia punctilobula*. Differences between the treated WB and the reference EB or A and Y in fern abundance in any forest type were not significant.

11.4.4.3 Shrub and Tree Species

Eight shrub species were found in the lower stratum (0 to 0.6 m), with four of these species reaching the second stratum (0.6 to 1.8 m). The most common species were *Viburnum alnifolium* (hobble-bush or moosewood), *Rubus* spp. (blackberry and raspberry), *Ribes glandulosum* (skunk current), and *Lonicera canadensis* (fly honeysuckle). These species comprised up to 88% of the shrubs at the BBWM site. Other less abundant shrub species recorded were *Cornus alternifolia* (dogwood), *Diervilla ionicera* (bush honeysuckle), *Vaccinium* spp. (blueberry), and *Viburnum acerifolium* (moosewood). All shrub species were recorded in the deciduous forest types with no shrub species found in the coniferous forest type. Overall, shrub abundance was higher in the treated WB than in the reference EB or A and Y (p-value <0.01).

Tree species and abundance including trees, saplings, and seedlings in all the strata were the same as those reported in FHM forest mensuration indicator. Differences in the overall tree species abundance between the treated WB and the reference EB or A and Y in any forest types were not significant.

11.4.5 FHM Lichen Communities Indicator

A total of 65 different lichen species were identified at the BBWM (Eckhoff 2000) with 51 species recorded in the treated WB and 57 species recorded in the reference EB. The mean number of lichen species per plot was 26 in the treated WB and 28 in the reference EB. *Parmelia fertilis*, which occurred in both WB and EB, had never been recorded in the U.S. prior to this study. Three other lichen species also found at BBWM, *Everniastrum catawbiense*, *Melanelia exasperatula*, and *Parmotrema arnoldii*, have been sited only in a few other locations (Hinds et al. 1998). Differences in the number of lichen species per plot or per forest type between the treated WB and the reference EB were not significant.

11.5 DISCUSSION

11.5.1 Result Comparison of the FHM Indicators at the BBWM and the Northeastern Region

Comparing FHM results from BBWM (this study) and Northeastern region for 1999 (U.S. Forest Service 2002), the overall findings were similar, regardless of the manipulation at BBWM. In both reports, forests were found in good health and with good tree crown condition. The damage and catastrophic mortality indicator showed that up to 60% of the trees at the BBWM and up to 73% of the trees in the Northeastern region indicated no signs of damage. The most prominent damage in the Northeastern region was decay while at the BBWM the most prominent damage was cankers. This difference was likely due to the high *Fagus grandifolia* population at BBWM which has been affected by the bark disease.

Lichen species diversity was higher at the BBWM with 27 species per plot while for the Northeastern region 15 species per plot were reported. Sixty-five lichen species were found at the BBWM while 91 lichen species were reported in the entire Northeastern region. All species reported from the BBWM were included in the 91 species reported in the Northeastern region but one species, *Parmelia fertilis*, had never been reported in the U.S. prior to this finding at BBWM.

11.5.2 Discrepancies in the Application of the FHM Indicators at the BBWM

The EPA defines monitoring as "the repeated recording of pertinent data over time for comparison with an identified baseline or a reference system" (Eagar et al. 1992). The detection monitoring component of the FHM program assesses temporal changes in forest ecosystems using repeated indicator measurements in permanent plots, comparing data over time. The application of the FHM indicators at BBWM did not assess changes temporally but rather spatially between the treated WB, specifically designed to study experimentally enhanced nitrogen and sulfur impacts, and two reference areas, EB and A and Y. Inherent in paired watershed studies of this nature is the issue of variability due to landscape heterogeneity. For example, the number of *Acer rubrum* trees in the reference EB was significantly higher than in the treated WB. However, $(NH_4)_2SO_4$ additions to WB began in 1989, therefore these results were not related to the treatment but to premanipulation differences between watersheds. Foliar chlorosis or necrosis and any other visual foliage damage, which is often associated with enhanced acidic deposition, were not observed at the treated WB.

Forest growth decline associated with acidic deposition is reported in high elevations or along coastal areas where forests are generally under acidic rainfall or prolonged incidences of cloud or fog cover. Following wet deposition via natural rainfall, acidity deposited directly on leaf surfaces accumulates as a consequence of water evaporation, resulting in increased solution concentrations on the leaves (Cox et al. 1996, Jacobson et al. 1990, and Heller et al. 1995), especially along leaf margins and tips (Cox et al. 1996). The resulting steep gradients that are created promote inward diffusion of ionic species resulting in injury (Heller et al. 1995) and increased nutrient losses due to foliar leaching (Jagels et al. 1989). Additionally, cloud immersion has been found to enhance foliar leaching levels (Vong et al. 1991, Jiang and Jagels 1999). The dry $(NH_4)_2SO_4$ treatment in the WB is applied directly to the soil, and therefore it does not simulate acidic rainfall or fog conditions.

11.5.3 Overall Efficacy of the FHM Indicators Tested

Previous studies at BBWM have shown clear evidence of effects of the $(NH_4)_2SO_4$ on the ecosystem at the BBWM. Kahl et al. (1999) indicated that ecosystem accumulative retention of added nitrogen in WB was 80% after 8 years of $(NH_4)_2SO_4$ treatment, and they suggested that the watershed was entering N saturation conditions. Norton et al. (1999b) reported that base cations as well as NO_3 exports in stream water at the BBWM have significantly increased in the treated WB since after the first year of treatment. In soil chemistry studies, Fernandez et al. (1999) indicated that the treatment has induced depletion of base cations (Ca, Mg, and K) from soil exchangeable pools in the treated WB watershed, while Al and Fe concentrations in soil solution have increased. Weber and Wiersma (1997) demonstrated significant differences in foliar chemistry of mosses between the treated WB and the reference EB, with enhanced levels of nitrogen and depressed levels of calcium and potassium in foliar tissue, along with enhanced levels of aluminum. White et al. (1999) also reported higher N concentrations and lower base cation concentrations in foliage of dominant tree species growing in the treated WB after the first 4 years of treatment. These studies suggested that the higher availability of N and depletion of base cations induced by the treatment might affect forest growth.

The FHM indicators, however, indicated no significant difference between the treated WB and the reference EB and A and Y in radial growth using tree dbh measurements. Similar results have been reported in other studies on red spruce (Johnson et al. 1984), white pine (Johnson et al. 1984), and mixed deciduous forest trees (Brooks 1994). In a literature review covering several decades of atmospheric deposition research, Morrison (1984) concluded that conventional mensuration methods, particularly radial growth measurements using dbh measurements, did not provide evidence that acidic deposition was influencing growth rates either negatively or positively. Lack of evidence of treatment effects on tree radial growth using ring-core analyses was also reported for red spruce at BBWM (White 1996).

However, in a recent study, Elvir et al. (2003) indicated that following 10 years of treatment, basal area increment was significantly higher for sugar maple trees growing in the treated WB watershed, with yearly increases relative to the reference watershed ranging from 13 to 104%, while red spruce showed no basal area growth response to the treatment. In a similar study, Magill et al. (1997) reported a 50% increase in wood production of a 50-year-old hardwood forest at the Harvard Forest in central Massachusetts after 6 years of simulated chronic N additions (113 $kg * ha^{-1} * yr^{-1}$). In contrast, Magill et al. (1997) also reported that a 70-year-old red pine at the Harvard Forest with the same treatment showed a decline in wood production. These studies suggest that tree-ring analyses can be a better indicator detecting effects of nitrogen and sulfur deposition on forest radial growth.

Some studies have speculated that shifts in species composition may occur as a result of acidic deposition, especially in high elevation spruce stands (McNulty et al. 1996). Although differences were not statistically significant, this study showed that there was a reduced number of saplings, tree seedlings, and lichen species in the treated WB. However, the combined number of lichens per hectare was very similar among areas. The lower number of lichen species in the treated WB might indicate possible effect of the treatment on lichen populations. Studies have reported that as ecosystem conditions are altered by acid deposition, the most sensitive lichens tend to disappear, while the abundance of less sensitive lichens might be unaffected or tend to increase (Campbell and Liegel 1996). Lichens are considered an environmental component at high risk from enhanced sulfur and nitrogen deposition. Lichens' sensitivity to enhanced acidic deposition has been attributed to their poikilohydric characteristics — failure to maintain constant internal moisture content (Campbell and Liegel 1996). Therefore, lichens undergo wetting and drying cycles which concentrate pollutants within plant tissues.

The FHM indicators did not provide evidence of treatment effects on the forest vegetation at the BBWM. The ineffectiveness of the FHM indicators to detect effects of the $(NH_4)_2SO_4$ treatment on vegetation might be attributed, as discussed earlier, to the temporal aspect of actual FHM monitoring compared with the spatial application at BBWM. The FHM indicators might be more effective to detect possible responses of vegetation to the treatment with repeated measurements within the treated WB watershed through time. However, the extensive nature of FHM indicators to detect forest responses to disturbances might also be attributed to their ineffectiveness to detect specific forest responses to enhanced acidic deposition.

The FHM program includes a large number of complex objectives related to the effects of acidic deposition, insects and disease, abiotic stressors including fire, storms, and flooding as well as land use including land clearing and domestic animals (Stolte 1997). To accomplish these objectives, the FHM indicators are designed to monitor changes in wide regions all across the U.S. Effects of acidic deposition on forest ecosystems, however, are often reported in localized areas under enhanced deposition rates (e.g., high elevations of the Appalachian Mountains and along Maine coastal areas). It has been proposed in the FHM program that effects of enhanced nitrogen and sulfur may affect species diversity, including favoring of nitrophilic species or loss of some vascular plant and lichen species due to phytotoxicity (Stolte and Lund 1996, Stolte 1997). However, much of the past research with lichens linked composition and foliar chemical analysis with pollutants from localized sources (Gunn et al. 1995, Bates et al. 1996), rather than with long-distance pollutants affecting extensive areas. In this study, the FHM lichen communities indicator did show some responses in community composition that may be related to the ammonium sulfate additions. However, the results were inconclusive, and a better-defined conceptual model and more information about specific plant species or groups of species that are at risk are needed to relate this indicator to acidic deposition impacts.

The most widely used forest health indicators are visual estimators (Alexander and Palmer 1999). Assessment of visible impacts has been useful for many forms of physical disturbances (fire, flooding, etc.). The FHM indicators were also based on physical and visual measurements, including growth measurements as dbh, crown conditions, and specific damages to the entire tree. It was assumed that since the tree crown directly interacts with the atmosphere, monitoring crown variables would detect atmospheric deposition effects on vegetation growth and mortality (Brooks et al. 1992). However, acidic deposition is a chemical perturbation and a chemical parameter, foliar analysis, was not implemented. Foliar chemistry shows the impacts of enhanced nitrogen and sulfur at BBWM even when other visual signs are not apparent.

Comparison of this study's results with previous results from foliar chemical analysis with trees (White et al. 1999) and mosses (Weber and Wiersma 1997) demonstrates that the chemical boundary for the foliage was responding earlier than the physical or visual indicators used in this study. By tracking changes over time, the FHM program attempts to identify early signs of changes in forest ecosystems due to acidic deposition (Stolte and Lund 1996). Therefore, since chemical analysis can demonstrate changes before they are visible, it needs to be implemented into the FHM program. Foliar analysis is already one of the indicators in use into the Acid Rain National Early Warning System (ARNEWS) program in Canada (Forestry Canada 1991, Hall 1995a and 1995b).

REFERENCES

Alexander, S.A. and C.J. Palmer. 1999. Forest health monitoring in the United States: first four years. *Environ. Monit. Assess.*, 55: 276–277.

Amezaga, I., A. Gonzalez Arias, M. Domingo, A. Echeandia, and M. Onaindia. 1996. Atmospheric deposition and canopy interactions for conifer and deciduous forests in Northern Spain. *Water Air Soil Pollut.*, 97: 303–313.

Bates, J.W., P.J. McNee, and A.R. McLeod. 1996. Effects of sulfur dioxide and ozone on lichen colonization of conifers in the Liphook Forest Fumigation Project. *New Phytol.*, 132: 653–660.

Brooks, R.T. 1994. A regional-scale survey and analysis of forest growth and mortality as affected by site and stand factors and acidic deposition. *Forest Sci.*, 40: 543–557.

Brooks, R.T., D.R. Dickson, W.G. Burkman, I. Millers, M. Miller-Weeks, E. Cooler, and L. Smith. 1992. Forest Health Monitoring in New England: 1990 Annual Report. USDA Forest Service, Northeastern Forest Experiment Station, Resource Bulletin NE-125, 59 pp.

Burkman, W.G. and G.D. Hertel. 1992. Forest health monitoring: a national program to detect, evaluate, and understand change. *J. Forest.*, 90: 26–27.

Bussotti, F., M. Ferretti, A. Cozzi, P. Grossoni, A. Bottacci, and C. Tani. 1995. Crown status of Holm oak (*Quercus ilex* L.) trees as related to phenology and environmental stress. *Water Air Soil Pollut*., 85: 1269–1274.

Campbell, S. and L. Liegel. 1996. Disturbance and Forest Health in Oregon and Washington. General Technical Report. PNW-GTR-381. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 105 pp.

Chazdon, R.L. and R.W. Pearcy. 1991. The importance of sunflecks for forest understory plants. *BioScience*, 41: 760–766.

Chazdon, R.L. 1986. Light variation and carbon gain in rain forest understory plants. *J. Ecol*., 74: 995–1012.

Church, M.R. 1999. The Bear Brook Watershed manipulation project: watershed science in a policy perspective. *Environ. Monit. Assess*., 55: 1–5.

Constabel, A.J. and V.J. Lieffers. 1996. Seasonal patterns of light transmission through boreal mixedwood canopies. *Can. J. For. Resour*., 26: 1008–1014.

Cox, P.M., G. Lemieux, and M. Lodin. 1996. The assessment and condition of Fundy white birches in relation to ambient exposure to acid marine fogs. *Can. J. For. Resour*., 26: 682–688.

Eagar, C., M. Miller-Weeks, A.J.R. Gillespie, and W. Burkman. 1992. Forest Health Monitoring in the Northeast 1991. USDA Forest Service, NE/NA-INF-115–92.

Eagar, C., H.V. Miegroet, S.B. McLaughlin, and N.S. Nicholas. 1996. Evaluation of Effects of Acid Deposition to Terrestrial Ecosystems in Class I Areas of the Southern Appalachians. Southern Appalachian Mountains Initiative (SAMI) Terrestrial Ecosystems Report. Available online at: http://www.saminet.org/reports/aciddepeffects.pdf

Eckhoff, J. 2000. Efficacy of forest health monitoring indicators to evince impacts on a chemically manipulated watershed. Ph.D. thesis, University of Maine, Orono, ME.

Elvir, J.A., G.B. Wiersma, A. White, and I. Fernandez. 2003. Effects of chronic ammonium sulfate treatment on basal area increment in red spruce and sugar maple at the Bear Brook Watershed in Maine. *Can. J. For. Res*., 33: 862–869.

Farmer, A.M., J.W. Bates, and J.N.B. Bell. 1991. Comparisons of three woodland sites in NW Britain differing in richness of the epiphytic *Lebanon pulmonariae* community and levels of wet acidic deposition. *Holarctic Ecol*., 14: 85–91.

Fernandez, I., L. Rustad, M. David, K. Nadelhoffer, and M. Mitchell. 1999. Mineral soil and solution responses to experimental N and S enrichment at the Bear Brook Watershed in Maine (BBWM). *Environ. Monit. Assess*., 55: 165–185.

Forestry Canada. 1991. ARNEWS Annual Report 1990. Information Report ST-X-1. Science and Sustainable Development Directorate, Forestry Canada, Ottawa, 16 pp.

Forster, B.A. 1993. *The Acid Rain Debate: Science and Special Interests in Policy Formation*. Iowa State University Press, Ames, IA, pp. 12–29, 48–56.

Gunn, J., W. Keller, J. Negusanti, P. Potvin, P. Beckett, and K. Winterhalder. 1995. Ecosystem recovery after emission reductions: Sudbury, Canada. *Water Air Soil Pollut*., 85: 1783–1788.

Hale, M.E. 1979. *How to Know the Lichens*. Wm. C. Brown Company Publishers, Dubuque, IA, 246 pp.

Hall, J.P. 1995a. ARNEWS assesses the health of Canada's forests. *Forest. Chron*., 71: 607–613.

Hall, J.P. 1995b. Forest health monitoring in Canada: how healthy is the boreal forest? *Water Air Soil Pollut*., 82: 77–85.

Hallbacken, L. and L.Q. Zhang. 1998. Effects of experimental acidification, nitrogen addition and liming on ground vegetation in a mature stand of Norway spruce [*Picea abies* (L.) Karst.] in SE Sweden. *Forest Ecol. Manage*., 108: 201–213.

Heller, L.I., A.J. Shaw, and J.S. Jacobson. 1995. Exposure of red spruce seedlings to acid mist: importance of droplet composition just prior to drying periods. *New Phytol.*, 129: 55–61.

Hinds, J.W., P.L Hinds, L.G. Biazrov, and J.D. Eckhoff. 1998. First United States record of the lichen *Parmelia fertilis*. *Northeast. Nat.*, 5: 21–23.

Jacobson, J.S., L.I. Heller, K.E.Yamada, J.F. Osmeloski, T. Bethard, and J.P. Lassoie. 1990. Foliar injury and growth response of red spruce to sulfate and nitrate acidic mist. *Can. J. For. Res.*, 20: 58–65.

Jagels, R., J. Carlisle, R. Cunningham, S. Serreze, and P. Tsai. 1989. Impact of acid fog and ozone on coastal red spruce. *Water Air Soil Pollut.*, 48: 193–208.

Jiang, M. and R. Jagels. 1999. Detection and quantification of changes in membrane-associated calcium in red spruce saplings exposed to acid fog. *Tree Physiol.*, 19: 909–916.

Johnson, A.H., T.G. Siccama, R.S. Turner, and D.G. Lord. 1984. Assessing the Possibility of a Link Between Acid Precipitation and Decreased Growth Rates of Trees in Northeastern United States. In: *Direct and Indirect Effects of Acidic Deposition on Vegetation*, Linthurst, R.A. (Ed.), Acid Precipitation Series Vol. 5, Butterworth Publishers, Boston, MA, pp. 81–95.

Kahl, J.S., S.A. Norton, I.J. Fernandez, K.J. Nadelhoffer, C.T. Driscoll, and J.D. Aber. 1993. Experimental inducement of nitrogen saturation at the watershed scale. *Environ. Sci. Technol.*, 27: 565–568.

Kahl, J., S. Norton, I. Fernandez, L. Rustad, and M. Handley. 1999. Nitrogen and sulfur input-output budgets in the experimental and reference watershed, Bear Brook Watershed in Maine (BBWM). *Environ. Monit. Assess.*, 55: 113–131.

LI-COR. 1992. LAI-2000 plant canopy analyzer. Instruction manual, LI-COR, Inc., Lincoln, NE.

Lindberg, S.E. and G.M. Lovett. 1993. Deposition and forest canopy interactions of airborne sulfur: results from the Integrated Forest Study. *Atmos. Environ.*, 26: 1477–1492.

Lindberg, S.E. and J.G. Owens. 1993. Throughfall studies of deposition to forest edges and gaps in montane ecosystems. *Biogeochemistry*, 19: 173–194.

Lovett, G.M. 1994. Atmospheric deposition of nutrients and pollutants in North America: an ecological perspective. *Ecol. Appl.*, 4: 629–650.

Magill, A., J. Aber, J. Hendricks, R. Bowden, J. Melillo, and P. Steudler. 1997. Biogeochemical responses of forest ecosystems to simulated chronic nitrogen deposition. *Ecol. Appl.*, 7(2): 402–415.

McNulty, S.G., J.D. Aber, and S.D. Newman. 1996. Nitrogen saturation in a high elevation New England spruce–fir stand. *Forest Ecol. Manage.*, 84: 109–121.

Meesenburg, H., K.J. Meiwes, and P. Rademacher. 1995. Long term trends in atmospheric deposition and seepage output in Northwest German forest ecosystems. *Water Air Soil Pollut.*, 85: 611–616.

Morrison, I.K. 1984. Acid rain: a review of literature on acid deposition effects in forest ecosystems. *Forest. Abstr.*, 45: 483–506.

Norton, S., J. Kahl, I. Fernandez, T. Haines, L. Rustad, S. Nodvin, J. Scofield, T. Strickland, H. Erickson, P. Wigington, Jr., and J. Lee. 1999a. The Bear Brook Watershed, Maine (BBWM), USA. *Environ. Monit. Assess.*, 55: 7–51.

Norton, S., J. Kahl, and I. Fernandez. 1999b. Altered soil-soil water interactions inferred from stream water chemistry at an artificially acidified watershed at the Bear Brook Watershed, Maine. *Environ. Monit. Assess.*, 55: 97–111.

Norton, S.A., R.F. Wright, J.S. Kahl, and J.P. Scofield. 1992. The MAGIC simulation of surface water acidification at, and first year results from, the Bear Brook Watershed Manipulation, Maine. *Environ. Pollut.*, 77: 279–286.

Reifsnyder, W.E. and H.W. Lull. 1965. Radiant energy in relation to forests. USDA Forest Service, Technical Bulletin 1344, Washington, D.C.

Rustad, L.E., J.S. Kahl, S.A. Norton, and I.J. Fernandez. 1994. Underestimation of dry deposition by throughfall in mixed northern hardwood forest. *J. Hydrol.*, 162: 319–336.

SAS Institute. 1998. SAS User's Guide. SAS, Inc., Cary, NC.

Stolte, K.W. 1997. 1996 National Technical Report on Forest Health. General Technical FS-605. U.S. Department of Agriculture, Forest Service, Washington, D.C., 47 pp.

Stolte, K.W. and H.G. Lund. 1996. Forest Health Monitoring Program in the United States. In Bravo, C.A. (Ed.), North American Workshop on Monitoring for Ecological Assessment of Terrestrial and Aquatic Ecosystems, Mexico City, September 18–22, 1995, USDA Forest Service General Technical Report RM-GTR-284, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, pp. 55–67.

Tallent-Halsell, N.G. 1994. Forest Health Monitoring 1994 Field Methods Guide. EPA/620/R94/027. U.S. Environmental Protection Agency, Washington, D.C.

Uddameri, V., S.A. Norton, J.S. Kahl, and J.P. Scofield. 1995. Randomized intervention analysis of the response of the West Bear Brook Watershed, Maine to chemical manipulation. *Air Water Soil Pollut.*, 79: 131–146.

U.S. Environmental Protection Agency. 1994. Forest Health Monitoring: Field Methods Guide. Environmental Monitoring and Assessment Program EPA/620/R-94/027, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency. 1993. Regional Environmental Monitoring and Assessment Program. EPA/625/R-93/012. U.S. Environmental Protection Agency, Washington, D.C.

U.S. Forest Service. 2002. Forest Health Highlights: Northeastern States. U.S. Department of Agriculture, Forest Service. Available online at: http://www.na.fs.fed.us/spfo/fhm/

Van Dobben, H.F., D.J.F. ter Braak, and G.M. Dirkse. 1999. Undergrowth as a biomonitor for deposition of nitrogen and acidity in pine forest. *Forest Ecol. Manage.*, 114: 83–95e.

Vong, R.J., J.T. Sigmon, and S.F. Mueller. 1991. Cold water deposition to Appalachian forests. *Environ. Sci. Technol.*, 25: 1014–1021.

Weber, K.A. and G.B. Wiersma. 1997. Trace element concentration of mosses collected from a treated experimental forest watershed. *Toxicol. Environ. Chem.*, 65: 17–29.

White, G., I. Fernandez, and G. Wiersma. 1999. Impacts of ammonium sulfate treatment on the foliar chemistry of forest trees at the Bear Brook Watershed in Maine. *Environ. Monit. Assess.*, 55: 235–250.

White, G.J. 1996. Effects of Chronic Ammonium Sulfate Treatments on Forest Trees at the Bear Brook Watershed in Maine. Ph.D. thesis, University of Maine, Orono, 167 pp.

Witham, J.W., E.H. Moore, M.L Hunter, Jr., A.J. Kimball, and A.S. White. 1993. A Long-Term Study of an Oak Pine Forest Ecosytem: Techniques Manual for the Holt Research Forest. Technical Bulletin 153, Maine Agricultural Experiment Station, University of Maine, pp. 74–76.

Wolterbeek, H.T.H., P. Kuik, T.G. Verburg, G.W.W. Wamelink, and H. van Dobben. 1996. Relations between sulphate, ammonia, nitrate, acidity, and trace element concentrations in tree bark in the Netherlands. *Environ. Monit. Assess.*, 40: 185–201.

12 Landscape Monitoring

D. Bailey and F. Herzog

CONTENTS

12.1 INTRODUCTION

Today's landscapes are—almost all over the world and increasingly—influenced by human actions. Even in remote areas, atmospheric transport of pollutants and modifications to our climate act on the landscape. Landscape change in turn alters the living conditions of plants, animals, and the human population. The perception, monitoring, and understanding/assessment of landscape change is, therefore, a prerequisite for predicting the future conditions of life and eventually for steering landscape change in desirable directions.

Landscape monitoring is a complex issue and could be the subject of an entire book in itself, covering various landscape-related issues from biodiversity and vegetation monitoring over the follow-up of abiotic landscape components to anthropogenic

1-56670-641-6/04/$0.00+$1.50

and cultural aspects such as scenery and landscape aesthetics. Some of these issues are addressed in other chapters of this book. For our part, we will concentrate on landscape monitoring in a more narrow sense, i.e., mainly land use and land cover. We will summarize the fundamental concepts of the monitoring of land use/land cover as they are presented in the literature and, of course, as we perceive them from our personal experience. There will therefore be a certain bias towards European landscapes, towards agricultural landscapes, and towards the interactions between landscape change and habitat availability.

12.1.1 Monitoring: Perception of Change

A principal aim of monitoring is to perceive the changes to our ecosphere as influenced by the activities of humans or through natural disturbances. Monitoring systems in general observe the factors that affect our natural environment; they provide an assessment of the actual situation and also a prognosis of the potential future developments (Vahrson, 1998). In terms of landscape monitoring the main purpose is to assess whether the structure and the function of the landscape changes over time. Of particular interest for landscape monitoring systems are the changes to the ecological situation of our landscapes, often with particular reference to anthropogenic activities. For instance, we might want to assess if and to what extent an intensification of agricultural production affects the extent and quality of seminatural habitats and thus potentially the species diversity.

Landscape monitoring systems have a historical perspective. That is, they examine the past situation in order to assess the current condition of the landscape, and this information is used to make projections of the likelihood of future change. In the example mentioned above, scenarios can be established which predict the future extent of seminatural habitats depending on a given increase in agricultural intensification. Above all, this will depend upon the changes in the composition, spatial configuration, and function of the landscape (Vahrson, 1998; Bastian et al., 2002). In an ecological context, critical change factors will include the level of habitat fragmentation within the landscape, the extent of habitat loss for particular species and species groups, and the effects on flows of energy, water, and matter (e.g., nutrient leaching). The monitoring of landscape in contrast to many other programs measures in the spatial dimension and therefore analyses change on a medium to large spatial scale. Normally, the programs monitoring landscape are not confined to a single subject matter or small-scale process but observe integrative changes of structure and function.

12.1.2 Landscape Monitoring as a Basis for the Monitoring of Many Environmental Indicators

In the context of environmental indicator systems as they are being developed by the UN (UN, 2001), OECD (OECD, 2002), the EU (Smeets and Weterings, 1999), or national governments (e.g., Orians et al., 2000; Statistics Canada, 2000; OFS–OFEFP–ARE, 2002), knowing the extent and status of the ecosystems is needed for the calculation of many of the state indicators which relate to a certain portion of the land or to some specific landscape type. Also, being informed on the distribution of different land use/land cover types may facilitate the sampling of the data for many indicators. It allows the implementation of stratified sampling strategies which can reduce the number of samples in “common” land-use/land-cover types (avoiding oversampling and thus reducing cost) while making sure that minor but nevertheless important land use/land cover types, which in fully random sampling would not or only rarely be sampled, are sufficiently covered (e.g., Tolle et al., 1999).

12.1.3 Fundamental Notions

Landscape: As our planet is subdivided spatially in numerous ways (examples include cultural, religious, social, economical, political, or natural divisions), our landscapes have been defined from many viewpoints in literally hundreds of ways (Naveh and Lieberman, 1994; Schreiber, 1990). Principally, the definitions and subdivisions vary according to the area of research interest, planned management, or monitoring program.

Our planet is subdivided into a hierarchy of scales. For example, the globe can be subdivided into continents, which are further subdivided into regions, landscapes, and local ecosystems. Within this hierarchy, each level represents a different level of detail, different degrees of variability, different sizes, shapes, as well as many other attributes. At the continental scale, each continent expresses a great deal of heterogeneity, exhibiting a diversity of climates, soils, topography, vegetation, land uses, and politics. The regions, on the other hand, mostly consist of a broad geographical area with a similar macroclimate and common cultural, economic, and political background but an extremely diverse ecological context.

By way of contrast, “landscapes represent a mosaic where a cluster of local interacting ecosystems or land uses are repeated over a kilometres-wide area” (Forman, 1995, p. 39). According to this definition, a landscape will have similar and repeated clusters of spatial elements such as the soils, vegetation, fauna, and land use throughout its area. We intend to adopt this definition for the purpose of this chapter. However, it is important to note that the definition considers landscape from the anthropogenic perspective. From a human point of view we tend to perceive landscape on a broad scale of tens of kilometers and are able to take account of temporal changes (Forman, 1995; Turner et al., 2001). Most of us indeed have an instinctive sense of landscape and are able to tell the difference between certain landscapes (alpine, lowland) as well as to define elements within given landscapes (fields, hedgerows, solitary trees, small woodlands, barns). However, landscapes are not always large scale. From the perspective of other organisms, landscapes may be represented on a scale of tens of meters or less. In fact, the size of the landscape and the nature by which it is defined will depend on the organisms or processes under consideration (Wiens, 1976; Wiens and Milne, 1989; Farina, 1995, 2000; Mazerolle and Villard, 1999). To avoid reference to absolute scale while addressing the importance of spatial configuration for ecological processes, Turner et al. (2001) have deliberately defined landscape in broader terms as an “area that is spatially heterogeneous in at least one factor of interest.”

It is important to remember that landscapes do not stand alone but are nested within other landscapes and therefore also have a regional context. Depending on our object of interest and hence our definition, our landscape can either act as open system where energy, materials, and organisms will move in and out of them freely or as a closed system. For example, a landscape which is open to a bird species may act as a closed system for other organisms. When establishing a landscape monitoring system, it is therefore essential that the landscape definition is suitable for the phenomenon and processes under consideration and that the regional context and openness of the landscape are taken into account.

Landscape mosaic: According to Forman and Godron (1986), landscapes have three main characteristics, namely, their structure, their function, and their tendency to change. Structure depends upon the spatial configuration of the elements, which define the landscape of interest. The function is related to the interaction and flow of energy, materials, and species between these spatial elements. Both characteristics are subject to change. In terms of landscape, monitoring the changes to the structure of the landscape and how they will affect the landscapes' function is often a principal concern in the program.

The structure of the landscape is depicted by the *mosaic*, the cluster of interacting ecosystems or land uses repeated at the relevant scale of interest of the focal object (Forman and Godron, 1986; Forman, 1995). The mosaic can actually be broken down into a pattern of *patches*, *corridors*, and *matrix*, spatial elements comprised of similar and clustered objects. At the landscape scale we refer, in this chapter, to these spatial elements as *landscape elements.* Patches consist of relatively distinct, homogeneous, nonlinear areas (e.g., woodland, grassland, moor), and corridors are linear strips of a particular type (hedgerow, road verge, canal margin), both of which differ from the adjacent landscape elements (Forman, 1995). The matrix is in effect the background ecosystem or land-use type of the mosaic. It is characterized by its extensive cover, high connectivity and/or major control over dynamics. For example, in an agricultural landscape the matrix will be the arable land, in a suburban landscape the built-up areas will form the matrix, and in an upland landscape the moors may represent the matrix.

The mechanisms which create the pattern of the land mosaic include the geomorphology of the landscape and the associated substrate heterogeneity, the natural processes and natural disturbance factors, and the level of human activity (e.g., Krönert, 1999). Natural phenomena (floods, hurricanes, earthquakes) can result in dramatic changes to the mosaic. However, the current level of human activity, through land use, is likely to instigate both more rapid and long-term changes to the structure and function of the landscape.

Land use was described by Barsch et al. (2002) as the fundamental human process of acquiring space, and in the wider sense defines the nature of the human activity within the landscape. Of course, this will depend upon the natural resources which are available and the ease by which they can be obtained. Through monitoring changes in land use the demands that society directly or indirectly makes on the landscape can be observed (Bastian et al., 2002). Land use can modify material cycles and exchange processes in the biosphere by changing landscape structure (Haefner, 1999). Furthermore, at the landscape scale the nature of our land use has a strong impact on the adaptability, regeneration, and regulation capability of ecosystems (Volk and Steinhardt, 2001).

Land cover is a (bio)physical description of the earth's surface and represents what covers the ground at the time point at which it was observed. Often, but not always, the cover corresponds to the habitat or vegetation present. The terms "land cover" and "land use" are closely related. However, an important difference is that land cover can have more than one land use associated with it. The moors of upland, for example, may be used among other things for farming, recreation, nature conservation, and to provide habitat for game birds. A woodland cover may be used for timber production, to provide wood for specific woodland crafts, to counteract climate change, and may act as a wind belt and also serve as a recreation area.

Scale, grain, and resolution: Scale is the spatial or temporal dimension of an object or process and is characterized by both its extent and grain size (Turner et al., 2001). The extent is the size of the study area or the temporal duration of the phenomenon under observation. The grain is the finest level of spatial resolution possible within a given dataset. Together they define the upper (extent) and lower (grain) limits of the spatial resolution in the study. The spatial resolution is a measurement of precision as it dictates the smallest possible feature that can be detected in the study. It can vary enormously. For example, in terms of remote sensing and the use of satellite imagery, the grain size can vary from a few meters or less (a fine/high resolution) to several kilometers (a coarse/low resolution). Generally speaking, the finer the resolution of the image the more objects that can be detected but the less total ground area that can be covered.

Scale is either cartographic (ratio of distance on the map to the Earth's surface), absolute (actual distance, direction, shape, geometry), or relative (transformation of absolute scale to a relative scale). In cartographic terms a large scale refers to a fine spatial resolution (1:5,000) and a small scale to a coarse spatial resolution (1:200,000). This can be confusing, and when biologists speak of large and small scale, they can often mean quite the opposite. In this chapter we adopt the use of fine (small area, high resolution) and broad scale (large area, low resolution) as also suggested by Turner et al. (2001). In a landscape context, a fine-grained mosaic is made up of small spatial elements and a coarse-grained mosaic by larger clusters of objects.

Spatial or temporal scale of an object or a process is intricately linked with the ecological hierarchy theory of Allen and Starr (1982). This theory refers to how a system of discrete functional elements is linked to at least two or more scales. In a hierarchy, each level will be composed of subsystems on the lower level, constrained by the layers above, and linked to the other elements on its own level (e.g., tree, stand, forest; cereal crop, arable field, agricultural landscape). The individual elements will function as a unit; they will be subject to constraints and will exhibit an individual degree of stability or variability. Organisms and processes operate at different scales in time and space and are separated from one another by magnitudes of scale. To understand the stability of a particular element it is necessary to have knowledge of a minimum of three hierarchical levels. It is therefore essential that the study be directed at the level of scale at which the object of interest is known to operate (Farina, 2000; Turner et al., 2001). Without careful consideration of the grain size and extent of the study area, erroneous conclusions may be drawn. If the level of scale of the object of interest is unknown, it is common practice to select a finer scale than thought necessary as this can be dissolved later to a coarser scale.

The choice of scale in the design of a landscape-monitoring program is thus an extremely important issue, and it should largely depend upon the focal point of the study. The scale will influence the patterns, which can be detected in the landscape mosaic and therefore the conclusions that can be drawn from the study. However, from a human viewpoint, landscapes are usually contemplated at a broad scale, partly as environmental issues have manifested themselves over much larger areas (Forman, 1995; Turner et al., 2001). The scale used also dictates whether the results can be extrapolated to other spatial or temporal dimensions (Forman and Godron, 1986; Turner et al., 2001). A pattern or process is said to be scale dependent when the response changes with the grain or extent of the measurement. When assessing landscape patterns, it is not possible to extrapolate below the resolution of the grain or beyond the extent of the landscape (Wiens, 1989).

There are several problems associated with scale. Often the scale and extent of the study are dictated by the scale of the imagery available (e.g., spatial resolution of the aerial photographs and the coverage available) and also the computer technology which is available. In addition, much of the data needed for a landscape-monitoring project will not be at the same scale. This means that it is often necessary to extrapolate the scale or extent from one scale to another or information from one data set to another at the same scale. This process is often not very straightforward and can be extremely problematic. Indeed, a modern-day issue, for scientists and landscape managers alike, is the practical and theoretical problems associated with extrapolating results across scales, e.g., from fine to broad scales (Bierkens et al., 2000). Interpretation of the data is often difficult across scales and therefore the prediction of an ecological attribute or environmental process can be complex when extrapolating fine scale information to broad scale issues. It is, for example, still very unclear how landscape patterns change with grain and extent (Wu et al., 2002). There are, of course, also problems associated with sampling design, as the landscapes are often very large and variable in character.

12.2 MAPPING THE LAND

The design of the landscape-monitoring program will depend upon the key questions of the study. Examples include: What is the rate of change which needs to be detected, i.e., do we need updated information every 5 years or every 20 years? Over what spatial dimension are changes taking place? Do we want averaged information about the entire region only or do we need data for individual subregions? About which ecological effects of change do we need information? Are the changes acceptable to society? Box 1 illustrates these questions with the example of designing a landscape-monitoring program for Switzerland.

Depending on the answers to these questions, suitable methods for the collection, analysis, and evaluation of data can be chosen. Three major decisions have to be made (Figure 12.1), which are interdependent and are the key for defining the quantity and quality of the information, that can be derived. They also determine the cost of the monitoring program.

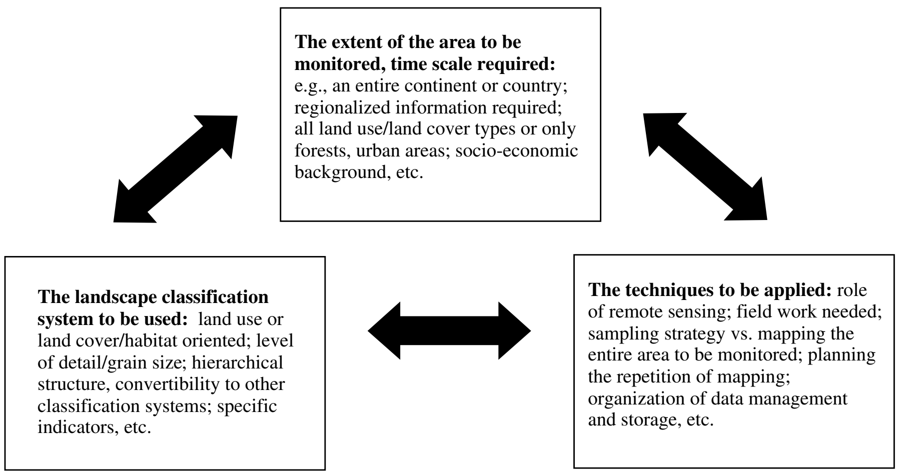
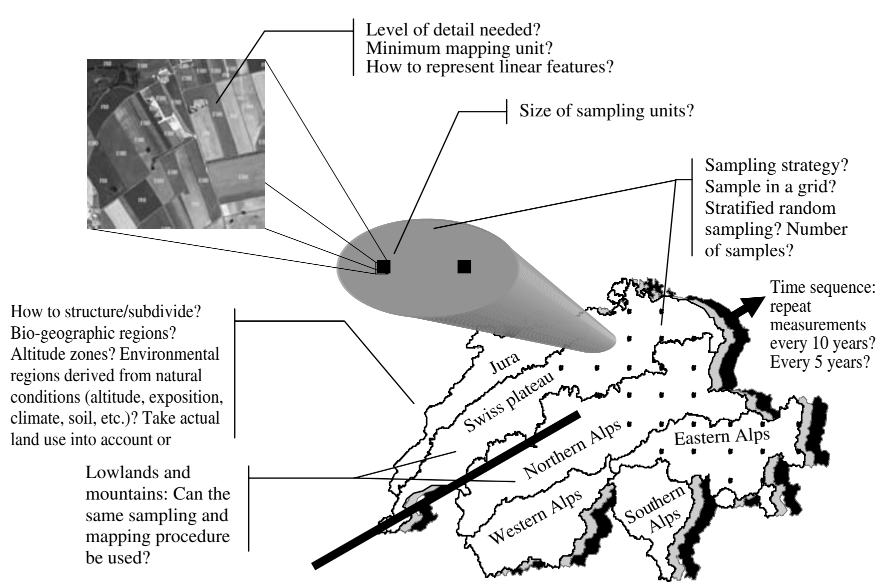


FIGURE 12.1 Major factors determining the features as well as the cost of a landscape monitoring program.

Box 1: Designing a landscape monitoring program for Switzerland.



Switzerland is a heterogeneous country with densely populated lowlands (plateau), lower mountains (Jura) and high mountain areas of different characteristics (Alps). The map represents the Swiss bio-geographic regions (Gonsethet al., 2001. With permission.)

Landscape-scale dynamics are extremely complex and difficult to explain because, in addition to the temporal component that is inherent to all monitoring programs, they have a spatial dimension. Also, to observe temporal change, the data will need to have a historical and current content. From a social and economic viewpoint it is important to have a working knowledge of the current and historical changes that have or had an impact on the ecological and landscape functions in the region of interest (Bastian, 1999). Knowledge of the driving forces behind change will help to identify the potential trends for the future development of the landscape and to suggest guidelines which aim to stop or slow down unacceptable changes. In the next section we will examine the techniques for mapping and classifying the landscapes. Principally, the data for monitoring landscapes are generated by remote sensing combined with fieldwork ("ground truthing"). In addition, statistics and censuses can be used for the interpretation of change.

12.2.1 Techniques

12.2.1.1 Remote Sensing

Remote sensing is a set of data acquisition techniques that enable the observation and monitoring of earth surface and processes at a distance. All techniques that gather data at a distance without contact with the object of interest may be called remote sensing and the technology includes aerial photography, satellite imagery, radar, and thermal imagery.

Aerial photo interpretation is a common, highly effective, and accurate method for the classification of temporal change in landscapes especially when combined with other data sources and technologies (Lindgren, 1985; Jensen, 1986; Singh, 1991; Naveh and Lieberman, 1994; Bolaños et al., 2001; Falkner and Morgan, 2002). The Food and Agriculture Organization (FAO) have been using aerial photography since the 1950s for their field projects and surveys. New possibilities due to developments in computer technology (high-resolution scanners, digital photogrammetry, digital image processing, and GIS) are now also available. It is possible, for example, to analyze vegetation dynamics at a combination of spatial resolutions, spatial extents, and temporal scales using aerial photographs within the GIS environment (Carmel et al., 1999).

Ideally, aerial photo interpretation is done in a stereoscopic way which, through overlapping images taken from different angles, offers a three-dimensional view on the landscape. The classic instrument for this purpose is the stereoscope. More recently, this can be done electronically on the PC with the help of 3-D glasses. This has the advantage that the photographs can directly be overlaid with additional GIS layers and/or, for monitoring purposes, with preceding land-use/land-cover interpretations.

Satellite technology: Since the 1970s, satellites have regularly scanned the entire planet. The continuity of this digital information makes it extremely useful for the monitoring of landscapes (Naveh and Lieberman, 1994). Solar radiation received by the earth's surface is either absorbed or reflected in specific wavelengths, most of which are not visible to the human eye. This reflectance is recorded in different wavelengths by satellite-mounted sensors and these build up the characteristic spectral signature of a particular object. The spectral signature will vary depending on the backscattering properties of the material (e.g., soils, vegetation type, water). The spectral areas that are commonly used for monitoring are those of visible, near-infrared, infrared, and the thermal part of the electromagnetic spectrum. These are not recorded in a continuous spectrum but rather in spectral bands. The data are collected by the sensors in groups of spectral bands, the number of bands depending on the particular satellite. In addition to the spectral capacities of the satellite, the spatial resolution is of interest. It determines the pixel size (grid cell) of the image and can vary tremendously (e.g., from 1 m $\times$ 1 m to 5 km $\times$ 5 km). The temporal resolution, or the frequency of scanning the same area, is another important feature. This can vary between several pictures per day (as for weather satellites) to a couple of takes per year.

Spectral signatures can be processed into photograph-like images (raster format with a given pixel size) of the different surface objects. Prior to dispatch, the raw data are normally processed through standard algorithms, cleansed (corrected for atmospheric, sensor errors, and geometric distortion) and possibly geo-referenced. The user will then classify the image using either a supervised or unsupervised classification method (Buiten and Clevers, 1993). A supervised classification method uses predetermined training pixels for which the class is known (through field verification) to order the image. Unsupervised classification assigns the pixels into clusters according to certain rules and algorithms. These clusters are then identified through field verification (Lillesand and Kiefer, 2000).

Satellite systems that are currently used to monitor the landscape include LANDSAT, SPOT, IKONOS, satellites in the Indian Remote Sensing program (IRS), and the Earth Remote Sensing (ERS) program. The commonly used NASA satellite—LANDSAT-7—has a resolution of 15 m $\times$ 15 m in the newly added panchromatic channel. The SPOT satellites record radiation in visible and near-infrared range with a spatial resolution of 20 m. They also have an additional panchromatic channel which has a resolution of 10 m. The panchromatic sensor of the Indian satellite IRS-1C has since 1995 a resolution of 5.8 m. IKONOS panchromatic data has a 1 m resolution since 1999. Table 12.1 summarizes the major properties of aerial photographs and satellite images. In Box 2 some links to metadata systems on remote sensing data availability are given.

12.2.2 Field Methods/Ground Truthing

The use of remote sensing techniques without some degree of fieldwork is normally not accurate enough for most landscape monitoring purposes. Fieldwork is usually needed at least for ground truthing. That is, selected training sites are mapped in the field, ideally at about the same time (year and season) as when the remote sensing image was taken. This information is needed to standardize the classification algorithms.

The intensity of fieldwork depends on the degree of information, which cannot be acquired by remote sensing. For example, particularly if the habitat types need also to be identified as well as the land use/land cover, this can only be carried out

TABLE 12.1
Characteristics of Aerial Photographs and Satellite Images

	Aerial Photographs	Satellite Images
Availability	Since 1930s in black and white, later in true and infrared color. Many countries are regularly covered for the production and revision of maps	Since 1970s, on average 1–4 recordings per year free of clouds (Kühbauch et al., 1990)
Type of data	Often analogous	Digital
Preprocessing	Orthorectification is needed if photos are to be overlaid with maps and additional GIS information	Geo-referencing/orthorectification
Scale/spatial resolution	On demand, often 1:2,000–1:30,000, resolution down to 0.3 m	1:50,000–1:250,000, resolution down to 1 m; the most frequently used are Landsat (30 m) and SPOT (20 m)
Interpretation	Visual or electronic; stereoscopic interpretation possible	On PC, unsupervised or supervised classification with specific software packages
User friendliness	Valid information can be gained from visual interpretation	Sophisticated hard- and software needed

Box 2: Metadata systems

Remote sensing and computer technology has resulted in the generation of vast amounts of spatial data, which are available in digital format. This can make the task of data retrieval problematic. To help solve this problem meta information systems may be used or set up by users and suppliers to aid the collection of appropriate data. Meta information systems aim to provide thematic, spatial and temporal references of what is available for particular specified area. A number which currently exist in the environmental domain include:

GEIN: German Environmental Information System (Germany, http://www.gein.de/index_en.html)
GISU: Geographic Information System Environment (German Federal Environment Agency, http://193.174.169.36/GISU/gisu.htm)
UDK: Environmental Data Catalogue (German, Austria, http://www.umweltdatenkatalog.de/wwwudk/V-UDKServlet)
CDS: Environmental Catalogue of Data Sources (European Environmental Agency, http://www.mu.niedersachsen.de/cds/)
NGSC: National Geospatial Clearinghouse (USA, http://nsdi.usgs.gov/)

Information packages are also available that are designed for viewing and distributing information on topics like land cover and nature in a user-friendly way. NATLAN (NATure/LANd cover) is one such information package designed and developed by the European Environmental Agency. The purpose of NATLAN is to give public access to the large amount of geo-referenced data on nature sites and areas, which are collected on behalf of the Agency.

in the field (see also Section 12.2.3). Sending people to the field is usually the most expensive part of any monitoring exercise and therefore needs to be planned carefully. Fieldwork will usually start once the information that is available from aerial or satellite images is extracted and, usually, digitized in GIS. The fieldwork should be undertaken in accordance with clearly predefined rules with respect to sampling (shall all patches in a monitoring site be visited or only selected ones, what is the procedure of selection) and with respect to the data to be recorded. The latter is not straightforward. Ideally, in order to record objective and reproducible data, species lists of plants should be elaborated and the habitat types then be deduced based on these lists. This, however, requires skilled botanists and is rather time consuming. Alternatively, habitats can be identified with vegetation keys and the help of predefined indicator species. This is more rapid but has the drawback that subtle changes between subsequent recordings cannot be detected.

A meaningful repetition of surveys is only possible if the same location is visited each time the fieldwork is repeated. If vegetation relevés are to be repeated, shifting even only a few meters can lead to important differences in the vegetation recorded. In order to prevent this, the exact location of sampling sites need to be recorded by possibly a combination of GPS measurements, written descriptions and sketches, land survey techniques from fixed triangulation, points, or magnets which are placed deep enough in order to prevent dislocation by frost or ploughing and which can later be located by detectors. The relocation of the survey site is facilitated if relevés are undertaken on circles instead of (conventionally used) squares because then the precise location of only the centre point has to be recorded.

12.2.3 Land-Use/Land-Cover Categories

Numerous systems to define land-use/land-cover categories have been proposed to classify landscape at a variety of scales (e.g., Murray and Cooper, 1992; Di Gregorio and Jansen, 1996; Frietsch, 1997; Delarze et al., 1999; Dramstad et al., 2001; EEA, 2002). It may seem simple at the offset, for example, to distinguish forest from agricultural land. But reality is more complicated than that. If we stick to this example and give it a second thought, we realize that it can be quite difficult to define, for example, the edge of forests in alpine regions where, towards the upper limit of tree growth, pastures, forest patches, and single trees are intermixed. Blaschke (in press) showed that different persons will interpret the tree line differently (Figure 12.2). To overcome this, clear definitions of land-use/land-cover categories are needed with rules and quantitative specifications.

Land-use/land-cover classifications systems generally have a hierarchical structure which allows the user to analyze landscapes at different levels of detail and to aggregate classes to, ideally, a common level. For example, “forest” can be subdivided into “perennial forest” and “deciduous forest,” and the latter again into different plant sociological associations if needed.

For monitoring, it is important that the definitions of the different classes remain constant, at least at the level of detail, which is relevant for the phenomenon of interest. Ideally, different classifications can be converted into each other, probably

FIGURE 12.2 Different tree-line interpretations by three individual interpreters. (From Blaschke, T., in *A Message from the Tatra: Geographical Information Systems and Remote Sensing in Mountain Environmental Research*, Widacki, W., et al., Eds., unpublished data (in press).)

at a more general level. This approach is currently developed in the EUNIS habitat classification (http://mrw.wallonie.be/dgrne/sibw/EUNIS/home.html). The system can be linked to the EU Habitat Directive, to the Bern Convention Resolution EMERALD network, to the Palaearctic Habitat Classification, to the CORINE land-cover classification, to some regional and national classifications, and also to systems such as the European Vegetation Survey (Moss and Davies, 2002a, 2002b, 2002c).

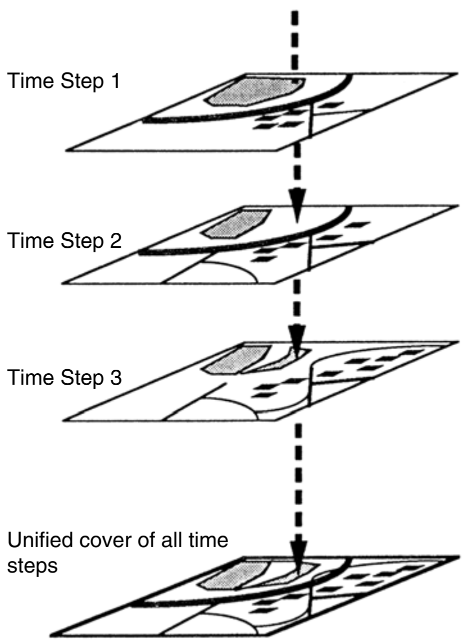
12.2.4 Repeating the Mapping

Monitoring consists of repeated measurements of a certain feature. In order to detect change we will want to reduce all other factors which act on the feature's indicator value. In landscape monitoring, these disturbance factors are, namely, the exact geographical location of measuring points or of recorded geometries and the actually assigned land-use/land-cover type.

The satellite images and aerial photographs used in the monitoring study will need to be geo-referenced. This process assigns spatial data with a coordinate system and enables the images to be associated with a specific location on the earth. The coordinate system should be projected. Many different projections exist and it is essential that the correct projection be selected for the study location. Map projections transform the spherical surface of the Earth on to a flat surface such as a map. An advantage of geo-referencing is that it allows spatial data to be overlaid and analyzed with other forms of geographic data.

If remote sensing images are translated in GIS layers and the changes are detected through the analysis of subsequent GIS maps, subsequent maps need to be based on each other. That is, land use/land cover should not be digitized from scratch each time but only the changes should be recorded. Otherwise, minor shifts in the position of individual patches or lines will occur and this will result in a false interpretation of the modifications, which are actually not real. Kienast (1993) developed a procedure to generate a spatial–temporal data model based on a subsequent set of topographical maps. By overlaying and intersecting the digitized maps, land-cover change can be analyzed for each location (Figure 12.3; see also Box 3).

As for the analysis of land-use/land-cover types from satellite or aerial photographs (Box 3), observing seasonality is a key issue. Photographs must be taken in the same season of subsequent years when vegetation development is at a similar stage. Broadleaved forest, for example, reflects differently when the trees have leaves than when there are no leaves. In agricultural landscapes the crop rotation introduces an additional difficulty because the spatial pattern changes from one year to another, sometimes even more rapidly. If individual crops are to be identified, multitemporal techniques where two or more images of the same year are combined can largely



Object table

Object no.	Time steps 1	2	3
1	forest	open land	forest
2	resident	road	road

FIGURE 12.3 Conceptual model for the analysis of subsequent time steps based on topographic maps. (From Kienast, F., *Landscape Ecol.*, 8(2), 101, 1993. With permission.)

improve the accuracy of classification. For example, in satellite images, grassland and germinating cereal crops have a similar spectral signature. Later in the year, when the crop is harvested and the grassland is still there, they will differ clearly. By combining images from both seasons and backing up by field checking and information on the agricultural management, crops can then be identified.

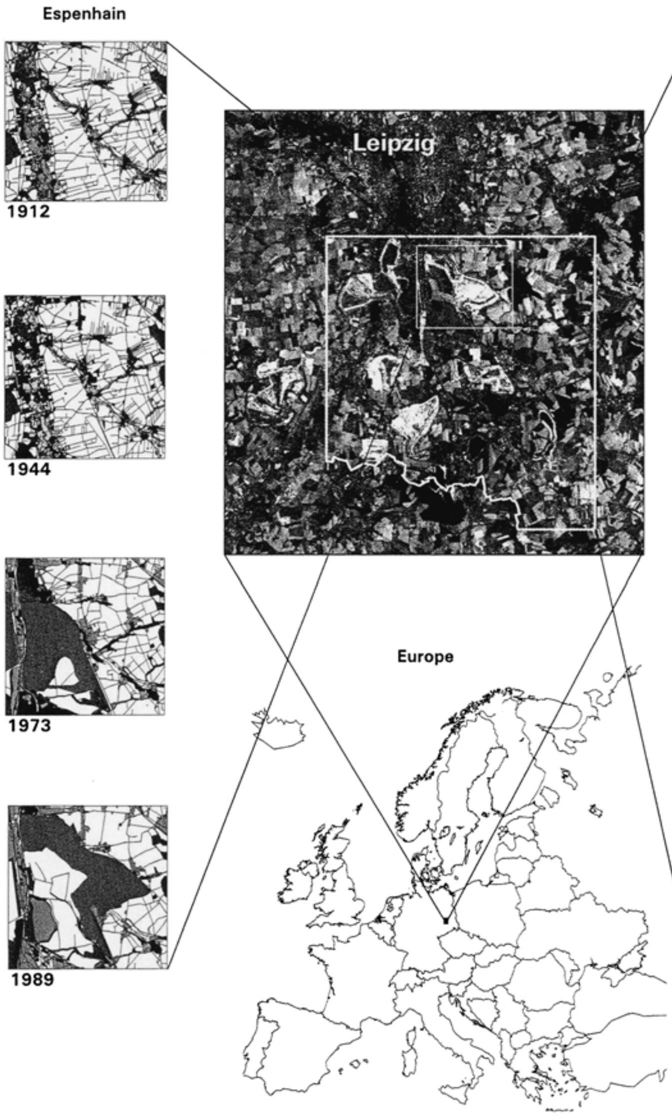
Box 3: Methodological lessons on landscape monitoring from surface mining regions

Surface mining landscapes change rapidly and radically over comparatively short time scales. This makes them extremely suitable to test landscape-monitoring techniques. A 7000-km^2 region, south of Leipzig in central Germany, has been subject to open-cast lignite mining throughout the 20th century: about one third of the land has been mined so far and this will be reclaimed by 2020. Two monitoring approaches were tested (Herzog and Lausch, 2001):

- Based on historical maps and photographs a sequence of four maps (1912–1989) was digitized manually using GIS for a sample of the "Espenhain" region.
- Based on satellite images (Spot-XS) and reclamation plans, a sequence of four maps (1990–2020) was established for the entire "Leipzig South" region.

Major lessons with respect to landscape monitoring techniques were:

- For the establishment of a series of GIS maps, start with one map and, for the subsequent maps, only digitize the changes. This approach avoids numerous small polygons which falsify changes and which have to be removed manually.
- Accounting for linear elements is a prerequisite when the changes in size and shape of polygons are of interest. For example, the increase of the size of agricultural fields—which is due to the intensification of agricultural production—can only be measured if the streets and paths between the fields are an integral part of the polygon layer.
- The "data model"—polygon based GIS maps or raster-based satellite images—has a major influence on the values of the landscape metrics which are calculated (see Section 12.3.1). If landscape metrics are to be regularly used in landscape monitoring, this would require rigorous standardization of data acquisition and processing. Most importantly, this would concern the number and definition of land-use/land-cover classes and the definition of patches (minimum mapable units; in the case of raster maps one has to define whether diagonal adjacent pixels of the same class are considered one patch or two).
- Future landscape monitoring will be based more and more on satellite images. Their spatial resolution must be sufficient to capture linear landscape features not only visually but also for their classification. This will facilitate the interpretation of landscape change with respect to the



function of landscapes; for example, as habitats. Depending on the species under consideration, some linear elements act as barriers (e.g., traffic infrastructure, rivers) and others as corridors or specific habitats (e.g., hedgerows, rivers).

Source: Herzog, F. and Lausch, A., *Environ. Monit. Assess.* 72, 37, 2001; Lausch, A. and Herzog, F., *Ecolo. Indic.*, 2, 3, 2002. With permission.

A particular difficulty in land-use/land-cover or habitat mapping is—as illustrated in the example of defining forest edge (Figure 12.2)—that the transition from one class into another is often gradual. To solve this problem it is necessary to establish rules and definitions, which enable the designation of landscape patches to specific classes. When the mapping is repeated, the boundaries between neighboring patches of the classes, which have these transitional qualities, should only be changed if the real situation in the landscape has actually changed. It is important therefore that the people who do the mapping of the subsequent time steps also take the original rules and definitions into account.

12.3 ANALYZING LAND-USE/LAND-COVER CHANGE

12.3.1 Techniques

There are basically three ways to analyze the change of land use/land cover. The most common and widely used analysis compares the increase or decrease of the extent (or share) of land-use/land-cover types over time (land-use statistics). Examples are the continuous increase of urban areas in industrialized countries or the decrease of forest in some tropical countries. At a second level of analysis, one is interested to know which land-use/land-cover types have converted into each other. For example, is the increase of urban areas mostly at the expense of agricultural land and the deforestation in the tropics due to an increase of agriculture? This analysis is done by means of a transition matrix and of related Markov transition models (Turner and Gardner, 1991; Muller and Middleton, 1994; Aaviksoo, 1995). Markov transition models enable the effects of disturbance on heterogeneous landscapes to be examined using a spatially aggregated approach. These models are stochastic as an explicit value is given to the probability of the transition from one state to another. The transition probability itself is usually related to a suitable time period. Markov transition models are useful in that they are relatively simple, can be applied to many scales, and using hierarchical data the appropriate level of detail can be selected.

At the landscape level, the combination of remote sensing data, geographic information systems, and Markov modeling have been used to predict the loss of prime agricultural land (Hathout, 2002; Wenig, 2002) or desert and other habitats to urbanization (Jenerette and Wu, 2001; Lopez et al., 2001); to explain forest-cover change as a result of socio-economic change (Brown et al., 2000); to model land degradation (Stoorvogel and Fresco, 1996); or to simulate forest succession (Jarvis, 1993). At the habitat level, metapopulation extinction through permanent loss of patch habitat, erosion of existing habitats, and random environmental catastrophes (Casagrandi and Gatto, 2002); transitions in vegetation structure (Dale et al., 2002; Li, 2002; Callaway and Davis, 1993); and habitat availability to mammals (Cuaron, 2000) have been modeled. Markov models have also been used to project backward missing land-cover data in a long-term time series (Petit and Lambin, 2002) or to generate short-term land-cover change projections in a region characterized by exceptionally rapid rates of change (Petit et al., 2001).

A more recent tool in landscape monitoring is landscape metrics. They allow us to indicate, for example, whether an increase in urban areas is focused and concentrated around existing settlements or whether built-up areas are scattered throughout the landscape. As for the loss of forest, it is important to know if it leads to fragmentation, i.e., if forest islands are created which are isolated from each other and which can become too small to sustain the original biocenoses. Correlations between landscape metrics and various landscape functions are sought (Figure 12.4). In the particular field of landscape monitoring, the application of landscape metrics has been tested in a number of studies representing a wide range of test areas and methods of data acquisition and treatment (Table 12.2). There are considerable variations in the size of the test areas, the spatial and temporal resolutions, the number of different land-use/land-cover types, and the kind of raw data used. The most frequently applied landscape indices belong to the broad category of edge and shape metrics. They quantify the occurrence of ecotones and are often related to patch area, the fractal dimension, or the discrepancy between actual and isodiametric shapes. Diversity measures are usually derived from information theory and often involve the use of Shannon's diversity index. The number and size of patches (patch area) are also often measured, whereas metrics for landscape configuration (contagion indices) were seldom applied.

There is a wealth of different landscape metrics proposed (Gustafson, 1998), starting from for example, Forman and Godron, 1986; O'Neill et al., 1988; Turner and Gardner, 1991; Baker and Cai, 1992 to more recent publications, Baskent and Jordan, 1995; McGarigal and Marks, 1995; Riitters et al., 1995; Cain et al., 1997; and Jaeger, 2000. Several programs compute landscape metrics automatically from GIS maps (Baker and Cai, 1992; McGarigal and Marks, 1995; Menz, 1997; Gardner, 1999; APACK—http://landscape.forest.wisc.edu/Projects/APACK/apack.html).

When applying landscape metrics, one is confronted with the question of selecting indicators relevant for the area and the problem under investigation. Because of the danger of "getting lost" in a wealth of data and figures, which are difficult to interpret,

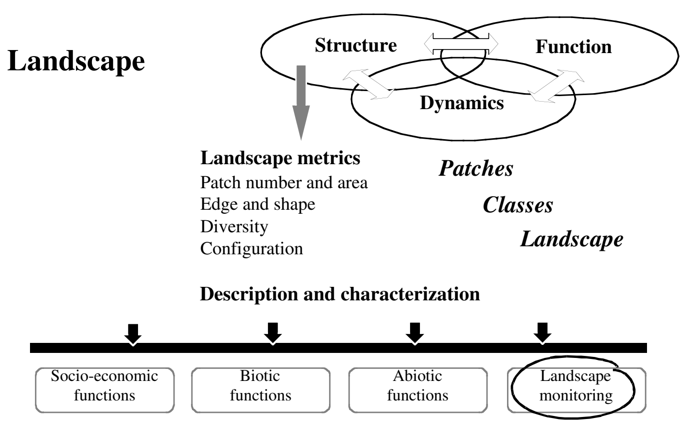


FIGURE 12.4 The application of landscape metrics. (From Lausch, A. and Herzog, F., *Ecol. Indic.*, 2, 3, 2002. With permission.)

TABLE 12.2
Selected Landscape-Monitoring Studies Combining Statistical Information on the Spatial Extent of Land-Use/Land-Cover Types (LT) with an Analysis of Landscape Pattern

Study Area	Size	Scale of Investigation	Spatial Resolution	Temporal Resolution	Data Source	Reference
Finland	3.2 km^2	1:5,000	2 m raster	ca. 16 years 1944–1991	Aerial photography	Ruuska and Helenius (1996)
Japan	4.4 km^2	1:7,500–1:20,000	Vector data	ca. 15 years 1948–1994	Aerial photography	Maekawa and Nakagoshi (1997)
Switzerland	~10 km^2	1:25,000	Vector data	ca. 8 years 1880–1982	Topographical maps	Kienast (1993)
France	32 km^2	1:25,000	Vector data	25 years 1964, 1989	Aerial photography	Poudevigne and Alard (1997)
U.S.A.	36 km^2	1:24,000	Vector data	ca. 30 years 1926–1988	Aerial photography	Thibault and Zipperer (1994)
Germany	75 km^2	1:12,000–1:25,000	Vector data	ca. 25 years 1912–1989	Topographical maps, aerial photography	Herzog et al. (2001)
Netherlands	100 km^2	1:50,000	200 m raster	ca. 20 years 1845–1982	Topographical maps	Hulshoff (1995)
U.S.A.	208 km^2	1:13,000–1:24,000	10 m raster	ca. 25 years 1935–1984	Aerial photography	Medley et al. (1995)
U.S.A.	242 km^2	1:20,000–1:40,000	Vector data	ca. 16 years 1940–1988	Aerial photography	Simpson et al. (1994)
Estonia	879 km^2	1:42,000–1:50,000	Vector data	ca. 30 years 1900–1989	Topographical maps	Palang et al. (1998)
U.S.A.	2230 km^2	1:100,000–1:250,000	80 m raster	16 years 1972, 1988	Satellite images	Luque et al. (1994)
Canada	4200 km^2	1:100,000– 1:250,000	25/50 m raster	17 years 1975, 1992	Satellite images	Sachs et al. (1998)
China/North Korea	9678 km^2	ca. 1:50,000	30 m raster	16 years 1972, 1988	Satellite images	Zheng et al. (1997)

Source: Modified from Herzog, F. et al. [Authors are Lausch, Müller, Thulke, Steinhardt, and Lehmann.], Landscape metrics for the assessment of landscape destruction and rehabilitation, *Environ. Manage.*, 27(1), 91, 2001. With permission.

Turner et al. (2001) have included a "read this first" section in their guidelines to landscape pattern analysis. Basically, there are two ways to approach the problem. Based on expertise and experience, meaningful indices, which address the problem to be investigated, can be selected (e.g., Herzog and Lausch, 2001). For example, if landscape fragmentation is to be examined, one will choose metrics, which relate to patch size, nearest neighborhood, core area, etc. The advantage is that one knows what one is doing; the disadvantage is that, due to the wealth of available indices, one is never totally sure whether one has really selected the most relevant and appropriate index.

Alternatively, statistical methods can be used. The approach then is to calculate a high number of indices and, through the application of statistical techniques such as data mining, correlation, and factor analysis, identify those indices which react to landscape change more strongly than others (e.g., Riiters et al., 1995; Cain et al., 1997; Herzog et al., 2001; Lausch and Herzog, 2002). The problem there is that the interpretation of the resulting indices may not be straightforward. The advantage, however, is that this is an unbiased method which can potentially lead to new and unexpected insights.

The first, targeted approach is more useful if metrics are to be related to specific landscape functions (e.g., the support of biodiversity) whereas for landscape monitoring we feel that at the present state of knowledge, the second, statistical approach is the better one.

There is no core set of indicators yet which have proven robust towards different data models. All of them depend on grain, scale, and whether they are computed from raster or from vector maps. That is, for the very same landscape the values of most landscape metrics differ depending on the data source, the spatial resolution, and the way of data interpretation (e.g., in the case of satellite image interpretation). This is the major drawback, which hampers the application of landscape metrics for statistical purposes; their calculation would require drastic standardization in data acquisition and analysis if time series and cross-country comparisons should make sense. Yet, some indices are more sensitive to these factors than others. Patch-size indices, indices relating patch area to edge length such as fractal indices for patch shape, and landscape composition indices are among the ones which have proven more robust than others (Lausch and Herzog, 2002).

At the same time, landscape metrics should be sensitive towards landscape change and allow an early perception of (un-)desirable developments (O'Neill et al., 1988). There is, however, an inherent contradiction between these two requirements—robustness and sensitivity—both of which can hardly be fulfilled by the same index. An index which is robust or tolerant towards different data sources and models will only detect more apparent and obvious changes whereas sensitive indices will only indicate real phenomena (instead of artifacts) if the data quality is comparable and consistent throughout time.

12.3.2 Interpretation

Once the analysis of landscape change has been successful, we will aim to interpret it, understand the driving forces behind the observed changes, make projections into the future, and evaluate whether the observed changes are desirable for society or not.

Censuses and other published statistics, used separately or in combination with remote sensing data and map data, are extremely useful in this context. Examples include the use of agricultural censuses to monitor the distribution of agricultural land and crops (Ilberry and Evans, 1989; Cardille et al., 2002; Frolking et al., 2002; Xiao et al., 2002), the use of population census data to study land degradation (Liu et al., 2003), the use of the 1990 U.S. Decennial Census data to monitor landscape change and to suggest landscape-level forest management strategies (Radeloff, 2001), the use of land-use/land-cover statistics to monitor change in metropolitan areas (Lo and Yang, 2002), and the use of socio-economic census data to monitor tropical deforestation (Geoghegan et al., 2001).

12.4 EXAMPLES OF ONGOING LANDSCAPE MONITORING PROGRAMS

At a global level there are a number of interdisciplinary projects that observe and try to understand the dynamics of land-use and land-cover change. In particular, reference is made to the potential consequences to global environmental change and sustainable development. Programs that monitor landscape change include the Land Use and Land Cover Change (LUCC) programme, prompted by the International Geosphere and Biosphere Programme and the International Human Dimensions of Global Environmental Change Programme (IHDP), NASA's Land Cover Land Use Change Program (LCLUC), and the Global Terrestrial Observing System (GTOS) cosponsored by FAO, UNEP, UNESCO, WMO, and ICSU.

The U.S. Geological Survey (USGS), the University of Nebraska — Lincoln (UNL), and the European Commission's Joint Research Centre (JRC) have developed a 1-km resolution global land cover characteristics database (Loveland et al., 2000). Released initially in 1997, there are two versions currently available for use in a wide range of environmental research and modeling applications. The data set is derived from 1-km Advanced Very High Resolution Radiometer (AVHRR) satellite data spanning the period from April 1992 to March 1993. Based on seasonal land-cover regions, the dataset provides a framework for presenting the temporal and spatial patterns of vegetation. The land-cover regions are composed of relatively homogeneous vegetation associations (similar floristic and physiognomic characteristics), which exhibit a distinctive phenology (onset, peak, and seasonal duration of greenness) and common level of primary production. Developed on a continent-by-continent basis, each continental database contains unique elements based on the geographic aspects of the specific continent. The continental databases are combined to make seven global datasets, each representing a different landscape based on a particular classification legend. Further specifications as well as the data itself are available at http://edcdaac.usgs.gov/glcc/glcc.html.

At a continental level, for Europe several land-cover databases exist. CORINE (Co-ordination of Information on the Environment) was initiated in the European Union in the 1980s. Using a 44-class nomenclature, the aim of the database is to provide consistent localized geographical information which can be used at the community level to determine and implement environmental policy. Furthermore, the database can be combined with other data (climate, inclines, soil, etc.) to make complex environmental assessments. The land-cover database provides a benchmark of the actual land cover, and through updating enables monitoring of both mid- and long-term change. The mapping scale of the database is 1:100,000 and the minimum mapping unit is 25 ha. The project is ongoing. Consequently, certain parts of Europe are still missing from the database and there are large differences in the satellite acquisition dates for the various countries which are covered. Based on CORINE, topographical data sets, and AVHRR sensor data, PELCOM (Pan-European Land Cover Monitoring) was established by the European Union in response to the need for up-to-date and reliable information for their current policy frameworks (e.g., European Environmental Outlook, Economic Assessment of Priorities for a European Environmental Policy Plan). See: http://www.eea.eu.int/ for CORINE, http://cgi.girs.wageningen-ur.nl/cgi/projects/eu/pelcom/index.htm for PELCOM.

A further land-use and land-cover database, which is currently under development in Europe, is LUCAS (Land Use/Land Cover Area Frame Statistical Survey [Gallego, 2002]). The objective is to provide annual and harmonized data which can be used for agricultural statistics and other environmental applications (e.g., to generate landscape and regional indicators). In contrast to CORINE and PELCOM, which cover the entire territory of the EU, LUCAS is to be based on a sampling strategy. The observation points are arranged in a regular grid. In addition to mere land-use/land-cover information, LUCAS is to yield information on soil erosion, natural hazards, landscape features (e.g., hedgerows, isolated trees), and farming practice. See http://www.eea.eu.int/.

Large areas (continents, countries) are often very heterogeneous with important differences in natural conditions. Subdividing them into eco-regions of similar climate, topography, and geology helps landscape monitoring because (1) these ecoregions can be used for stratification if sampling techniques need to be applied, (2) they help the definition of the land-use/land-cover classes, and (3) they can facilitate the understanding of the observed land use/land cover and its changes. Examples for ecoregions comprise the Digital Map of European Ecological Regions (Bohn, 1994), Canada's ecozones (Marshall and Schut, 1999) and, more recently, the European Environmental Regions. By defining relatively homogeneous ecological conditions, it should be possible to make meaningful comparisons and assessments on biodiversity (Painho et al., 1996).

In addition, there are sectoral monitoring programs which focus on specific land-use/land-cover types; for example, forest monitoring, or the monitoring of specific habitats which are relevant for nature conservation, such a moors, or the monitoring of the development of deserts. The most prominent example is the U.K. countryside survey which concentrates on rural Great Britain. It is the first landscape-monitoring scheme which has implemented a stratified random sampling technique (Haines-Young et al., 1994; Bunce et al., 1996; Haines-Young et al., 2000; Sheail and Bunce, unpublished). Detailed field observations are made in randomly selected 1 km grid squares. Many of the sample sites were first visited in 1978 and subsequently in 1984, 1990, and 2000, providing a time series of changes in the countryside. Collection of data such as habitat types, hedgerows, plant species, and freshwater invertebrates complements powerful satellite imagery and enables a deeper level of ecological understanding (http://www.cs2000.org.uk/).

12.5 CONCLUSIONS, RECOMMENDATIONS

We all live in landscapes. They form the basis of life not only for humankind but also for all living organisms on earth. It is therefore imperative that we are careful about what we do to our landscapes. We have to keep them "healthy" and give them a regular health check. This is now widely recognized, and landscape-monitoring programs are, as a result, mostly policy driven (e.g., O'Neill et al., 1996; Groom and Reed, 2001). While it is good that policy makers increasingly recognize the importance of landscape monitoring, policy needs can also act as a source of conflict and misunderstanding. Landscape monitoring programs are long-term undertakings. Monitoring time intervals of about a decade are appropriate and trends can only be observed after repeated measurements, hence after several decades. Policy makers and stakeholders, however, have to act and react in shorter periods, for example, election periods of four years. Thus there may be a call for more frequent measurements over a shorter time scale. Naturally, this is possible but it will increase the cost of the monitoring program exponentially due to the finer changes which will need to be detected and thus the higher measurement precision.

This brings us to the financial aspects of landscape monitoring. The real challenge, actually, is not the design of a monitoring program as such but the optimization of expected information and the cost of the program. Landscape monitoring is usually conducted with public money, which, of course, has to be spent carefully. Therefore, monitoring has to be limited to the essential information and additional, "nice-to-have" information can, in most cases, not be collected. The cost depends on the spatial and temporal accuracy and on the desired level of classification of data, which is needed. This need determines the spatial–temporal density of sampling. Below a certain density of samples—and, consequently, below a certain amount of money which is spent on the program—observed changes will no longer be statistically significant. The problem is that, when designing the program, this limit is not known. Therefore the lower acceptable limit is usually negotiated between the policy maker and the scientist. It is important that this negotiation is not driven too strongly by actual policy needs because these requirements change and it would be counterproductive to design a long-term monitoring program only to meet short-term requirements. The information gathered later on would then be useless.

Landscape monitoring programs have to be representative for the landscapes they investigate. If the entire region of interest cannot be covered, sampling strategies have to be employed. The point here is that the sampling has to be random, either based on a grid or on predefined strata (Bunce et al., 1996; Brandt et al., 2002). Biased sampling, e.g., sampling only nature protection zones or sampling only regions which can easily be accessed, will only yield information on these particular locations but not reflect the development of the entire landscape.

Another consequence of the long-term character of landscape monitoring—and this probably applies to all monitoring programs—is that when the program starts there can only be guesses on the future needs of information. This has two consequences:

1. The program has to concentrate on broad, fundamental (long-term) features, which are independent of political (short-term) characteristics.
2. Wherever possible, raw, noninterpreted data must be collected and stored which later can be reinterpreted to allow the uptake and integration of future requirements.

For example, forest-monitoring programs, which started decades ago, focused mainly on wood production. Later the interest in additional functions of forests (resource protection, recreation, etc.) increased. Ideally, these assessments can be made retrospectively on the original data sets, which were recorded decades ago. Unfortunately, this is often not the case because budget restrictions originally restrained the program to the sole purpose of assessing productivity (e.g., Bättig et al., 2002). Another example concerns the assessment of the effects of agri-environmental measures on landscape and biodiversity. An evaluation which only focuses on specially managed areas (according to the political guidelines) and assesses their effects only according to general, predefined criteria cannot be used for monitoring because these guidelines are likely to change over time and thus prevent a comparison (Hofer et al., 2002).

Long-term monitoring programs, therefore, have to be more general in nature. They have to provide fundamental information on status and trends of particular indicators. In addition, they can form the basis for other, shorter-term evaluation programs, which assess the effectiveness of specific environmental measures and policies. These evaluation programs can be specifically targeted to support decision making and meet the societal needs for sustainable landscapes.

ACKNOWLEDGMENTS

We thank André Desaules, Gabriela Hofer, Angela Lausch, Isabel Augenstein, Achilleas Psomas for their comments on preliminary manuscript versions and especially Bob Bunce, who taught us the principles of landscape monitoring. The Swiss Federal Office funded part of the work on this text for Education and Science under the credit BBW 00.0080-1.

REFERENCES

Aaviksoo, K., Simulating vegetation dynamics and land use in a mire landscape using a Markov model, *Landscape Urban Plann.*, 31, 129, 1995.

Allen, T.F.H. and Starr, T.B., *Hierarchy: Perspectives for Ecological Complexity*, University of Chicago Press, Chicago, 1982.

Baker, W.L. and Cai, Y., The r. le program for multi-scale analysis of landscape structure using GRASS geographic information system, *Landscape Ecol.*, 7(4), 291, 1992.

Barsch, H., Bastian, O., Beierkuhlein, C., Bosshard, A., Breuste, F., Kloetzli, F., Ott, K., Tress, B., Tress, G., and Weiland, U., Application of landscape ecology, in *Development and Perspectives of Landscape Ecology*, Bastian, O. and Steinhardt, U., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, p. 307.

Baskent, E.Z. and Jordan, G.A., Characterizing spatial structure of forest landscape, *Can. J. For. Res.*, 27, 1675, 1995.

Bastian, O., Description and analysis of the natural resource base, in *Land-Use Changes and Their Environmental Impact in Rural Areas in Europe*, Krönert, R., Baudry, J., Bowler, I.R., and Reenberg, A., Eds., Man and the Biosphere Series 24, UNESCO, Paris; Parthenon, New York, 1999, p. 43.

Bastian, O., Beierkuhlein, C., and Syrbe, R.U., Landscape change and landscape monitoring, in *Development and Perspectives of Landscape Ecology*, Bastian, O. and Steinhardt, U., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002.

Bättig, C., Bächtiger, C., Bernasconi, A., Brändli, U.-B., and Brassel, P., *Wirkungsanalyse zu LFI*1 *und* 2 *und Bedarfsanalyse für das LFI*3, Landesforstinventar, BUWAL, Bern, 2002, Umweltmaterialen 143, Wald.

Bierkens, M.F.P., Finke, P.A., and de Willigen, P., *Upscaling and Downscaling Methods for Environmental Research*, Kluwer Academic, Dordrecht, Netherlands, 2000.

Blaschke, T., Accounting for sustainability: An integrated GIS/remote sensing methodology applied to mountain environments, in *A Message from the Tatra: Geographical Information Systems and Remote Sensing in Mountain Environmental Research*, Widacki, W., Bytnerowicz, A., and Riebau, A., Eds., unpublished data (in press).

Bohn, U., International project for the construction of a map of the natural vegetation of Europe at a scale of 1:2.5 million—its concept, problems of harmonisation and application for nature protection, Working text, Bundesamt für Naturschutz (BfN), 1994.

Bolaños, F., Garcia del Barrio, J.M., Regato, P., and Elena-Rosello, R., Spanish forested landscapes: classification and dynamics, in *Development of European Landscapes*, Proceedings of the IALE European Conference, Mander, Ü., Printsmann, A., and Palang, H., Eds., Publicationes Instituti Geographici Universitatis Tartuensis, Tartu, Estonia, 2001, p. 7.

Brandt, J.J.E., Bunce, R.G.H., Howard, D.C. and Petit, S., General principles of monitoring land cover change based on two case studies in Britain and Denmark, *Landscape Urban Plann.*, 62(1), 37, 2002.

Brown, D.G., Pijanowski, B.C., and Duh, J.D., Modelling the relationships between land use and land cover on private lands in the Upper Midwest, U.S.A., *J. Environ. Manage.*, 59(4), 247, 2000.

Buiten, H. and Clevers, J., *Land Observation by Remote Sensing: Theory and Applications*, Vol. 3, Gordon and Breach, Warsaw, 1993.

Bunce, R.G.H. Barr, C.J., Clarke, R.T., Howard, D.C., and Lane, A.M.J., Land classification for strategic ecological survey, *J. Environ. Manage.*, 47, 37, 1996.

Cain, D.H., Riiters, K., and Orvis, K., A multi-scale analysis of landscape statistics, *Landscape Ecol.*, 12, 199, 1997.

Callaway, R.M. and Davis, F.W., Vegetation dynamics, fire, and the physical environment in coastal central California, *Ecology*, 74(5), 1567, 1993.

Cardille, J.A., Foley, J.A., and Costa, M.H., Characterising patterns of agricultural land use in Amazonia by merging satellite classifications and census data, *Glob. Biochem. Cycles*, 16(3), art. no. 1045, 2002.

Carmel, Y., Kadmon, R., and Lahov-Ginot, S., Studying long-term vegetation dynamics using image processing of historical aerial photographs, *Proceeding of the 5th World Congress*, International Association for Landscape Ecology, IALE, Snowmass Village, CO, 1999.

Casagrandi, R. and Gatto, M., Habitat destruction, environmental catastrophes, and metapopulation extinction, *Theor. Popul. Biol.*, 61(2), 127, 2002.

Cuaron, A.D., Effects of land cover changes on mammals in a neotropical region: a modelling approach, *Conserv. Biol.*, 14(6), 1676, 2000.

Dale, M., Dale, P., and Edgoose, T., Using Markov models to incorporate serial dependence in studies of vegetation change, *Acta Oecol.—Int. J. Ecol.*, 23(4), 261, 2002.

Delarze, R., Gonseth, Y., and Galland P., *Lebensräume der Schweiz*, Ott Verlag, Thun, Germany, 1999.

Di Gregorio, A. and Jansen, L.J.M., FAO land cover classification: a dichotomous, modular-hierarchical approach, Paper presented at the U.S. Federal Geographic Data Committee (FGDC) Vegetation Subcommittee and Earth Cover Working Group meeting in Washington, D.C., October 15–17, 1996; FAO, Rome, 1996, http://www.fao.org/sd/EIdirect/EIre0019.htm.

Dramstad, W.E., Fjellstad, W.J., Strand, G.-H., Mathiesen, H.F., Engan, G., and Stokland, J.N., Development and implementation of the Norwegian monitoring programme for agricultural landscapes, *J. Environ. Manage.*, 64, 49, 2001.

EEA, EUNIS Habitat Classification, European Environment Agency, 2002, http://mrw.wallonie.be/dgrne/sibw/EUNIS/home.html.

Falkner, E. and Morgan, D., *Aerial Mapping: Methods and Applications*, 2nd ed., Lewis Publishers, Boca Raton, FL, 2002.

Farina, A., Cultural landscapes and fauna, in *Cultural Landscapes of Universal Value*, van Droste, B., Plachter, H., and Rossler, M., Eds., Gustav Fischer, Jena, Germany, 1995, p. 60.

Farina, A., *Landscape Ecology in Action*, Kluwer Academic, Dordrecht, Netherlands, 2000.

Forman, R.T.T., *Land Mosaics*: *The Ecology of Landscapes*, Cambridge University Press, New York, 1995.

Forman, R.T.T. and Godron, M., *Landscape Ecology*, John Wiley & Sons, New York, 1986.

Frietsch, G., *Ergebnisse der CIR-Biotoptypen- und Landnutzungskartierung und ihre Anwendungsmöglichkeiten in der Naturschutzpraxis—Einführungsvortrag*, Sächsische Akademie für Natur und Umwelt, 1997, 3, 7, http://www.umwelt.sachsen.de/lfug/.

Frolking, S., Qiu, J.J., Boles, S., Xiao, X.M., Liu, J.Y., Zhuang, Y.H., Li, C.S., and Qin. X.G., Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China, *Glob. Biogeochem. Cycles*, 16(4), art. no. 1091, 2002.

Gallego, J., *Building Agro Environmental Indicators Focussing on the European Area Frame* Survey *LUCAS*, European Commission Joint Research Centre, Ispra, Italy, 2002, 176. http://agrienv.jrc.it/publications/ECpubs/agri-ind/

Gardner, R.H., RULE: A program for the generation of random maps and the analysis of spatial patterns, in *Landscape Ecological Analysis: Issues and Applications*, Klopatek, J.M. and Gardner, R.H., Eds., Springer-Verlag, New York, 1999, p. 280.

Geoghegan, J., Villar, S.C., Klepeis, P., Mendoza, P.M., Ogneva-Himmelberger, Y., Chowdhury, R.R., Turner, B.L., and Vance, C., Modelling tropical deforestation in the southern Yucatan peninsular region: comparing survey and satellite data, *Agric. Ecosyst. Environ.* 85(1–3), 25, 2001.

Gonseth, Y., Wohlgemuth, T., Sansonnens, B., and Buttler, A., *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard*, Umwelt Materialien Nr. 137, Bundesamt für Umwelt, Wald und Landschaft, Bern, 2001.

Groom, G. and Reed, T., Eds., Strategic Landscape Monitoring for the Nordic Countries, Nordic Council of Ministers, TemaNord, Copenhagen, 523, 2001.

Gustafson, E.J., Quantifying landscape spatial pattern: What is the state of the art?, *Ecosystems*, 1, 143, 1998.

Haefner, H., Fernerkundung als Instrument der Landschaftsoekologie, in *Angewandte Landschaftsoekologie*, *Grundlagen und Methoden*, Schneider-Silwa, R., Schaub, D., and Gerold, G., Eds., Springer-Verlag, Berlin, 1999, p. 201.

Haines-Young, R. H., Barr, C.J., Black, H.I.J., Briggs, D.J., Bunce, R.G.G., Clarke, R.T., Cooper, A., Dawson, F.H., Firbank, L.G., Fuller, R.M., Furse, M.T., Gillespie, M.K., Hill, R., Hornun, G.M., Howard, D.C., McCann, T., Morecroft, M.D., Petit, S., Sier, A., Smart, S.M., Smith, G.M., Stott, A.B., Stuart, R.C., and Watkins. J.W., *Accounting for Nature: Assessing Habitats in the U.K. Countryside*, Her Majesty's Stationery Office, London, 2000.

Haines-Young, R.H., Bunce, R.G.H., and Parr, T.W., Countryside information systems, in *Geographic Information and Sourcebook for GIS*, Green, E.R., Rix, D., and Cadeaux-Hudson, J., Eds., Taylor and Francis, London, 1994, p. 97.

Hathout, S., The use of GIS for monitoring and predicting urban growth in East and West St Paul, Winnipeg, Manitoba, Canada, *J. Environ. Manage.*, 66(3), 229, 2002.

Herzog, F. and Lausch, A., Supplementing land-use statistics with landscape metrics: some methodological considerations, *Environ. Monit. Assess.*, 72, 37, 2001.

Herzog, F., Lausch, A., Müller, E., Thulke, H.-H., Steinhardt, U., and Lehmann, S., Landscape metrics for the assessment of landscape destruction and rehabilitation, *Environ. Manage.*, 27(1), 91, 2001.

Hofer, G., Herzog, F., Spiess, M., and Birrer, S., Vegetation und Brutvögel als Ökoindikatoren im Mittelland, *Agrarforschung*, 9(4), 152, 2002.

Hulshoff, R.M., Landscape indices describing a Dutch landscape, *Landscape Ecol.*, 10(2), 101, 1995.

Ilberry, B.W. and Evans, N.J., Estimating land loss on the urban fringe: a comparison of the agricultural census and aerial photograph/map evidence, *Geography*, 74, 214, 1989.

Jaeger, A.G., Landscape division, splitting index, and effective mesh size: new measures of landscape fragmentation, *Landscape Ecol.*, 15, 115, 2000.

Jarvis, C., Modelling forest ecosystem dynamics using multi-temporal multi-spectral scanner (MSS) data, Remote Sensing of Earths Surface and Atmosphere, *Adv. Space Res.*, 14(3), 277, 1993.

Jenerette, G.D. and Wu, J.G., Analysis and simulation of land use change in the central Arizona–Phoenix region, U.S.A., *Landscape Ecol.*, 17(7), 611, 2001.

Jensen, J.R., *Introductory Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

Kienast, F., Analysis of historic landscape patterns with a geographical information system —a methodological outline, *Landscape Ecol.*, 8(2), 101, 1993.

Krönert, R., Introduction, in *Land-use Changes and Their Environmental Impact in Rural Areas in Europe*, Krönert, R., Baudry, J., Bowler, I.R., and Reenberg A., Eds., UNESCO, Paris, Parthenon, New York, 1999, p. 1.

Kühbach, W., Kupfer, G., Schellberg, J., Müller, U., Dockter, K., and Tempelmann U., Fernerkundung in der Landwirtschaft, *Luft- und Raumfahrt*, 11, 36, 1990.

Lausch, A. and Herzog, F., Applicability of landscape metrics for the monitoring of landscape change: issues of scale, resolution, and interpretability, *Ecol. Indic.*, 2, 3, 2002.

Li, B.L., A theoretical framework of ecological phase transitions for characterising tree-grass dynamics, *Acta Biotheor.*, 50(3), 141, 2002.

Lillesand, T.M. and Kiefer, R.W., *Remote Sensing and Image Interpretation*, 4th ed., John Wiley & Sons, New York, 2000, 724 pp.

Lindgren, D.T., *Land Use Planning and Remote Sensing*, M Nijhoff Publishers, Dordrecht, Netherlands, 1985.

Liu, Y.S., Gao, J., and Yang, Y.F., A holistic approach towards assessment of severity of land degradation along the Great Wall in Northern Shaanxi Province, China, *Environ. Monit. Assess.*, 82(2), 187, 2003.

Lo, C.P. and Yang, X.J., Drivers of land use/land cover changes and dynamic modelling for the Atlanta, Georgia Metropolitan Area, *Photogramm. Eng. Rem. Sen.*, 68(10), 1073, 2002.

Lopez, E., Bocco, G., Mendoza, M., and Duhau, E., Predicting land cover and land use change in the urban fringe—a case in Morelia city, Mexico, *Landscape Urban Plann.*, 55(4), 271, 2001.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., and Merchant, J.W., Development of a global land cover characteristics database and IGBP DISCover from 1km AVHRR data, *Int. J. Remote Sens.*, 21, 1303, 2000.

Luque, S.S., Lathrop, R.G., and Bognar, J.A., Temporal and spatial changes in an area of the New Jersey Pine Barrens landscape, *Landscape Ecol.*, 9(4), 287, 1994.

Maekawa, M. and Nakagoshi, N., Riperian landscape changes over a period of 46 years, on the Azusa River in central Japan, *Landscape Urban Plann.*, 37, 37, 1997.

Marshall, I.B. and Schut, P.H., *A National Ecological Framework for Canada*, Environment Canada and Agriculture and Food Canada, 1999, http://sis.agr.gc.ca/cansis/nsdb/ecostrat/intro.html.

Mazerolle, M.J. and Villard, M.-A., Patch characteristics and landscape context as predictors of species presence and abundance: a review, *Ecoscience*, 6(1), 117, 1999.

McGarigal, K. and Marks, B.J., FRAGSTATS, Spatial Analysis Program for Quantifying Landscape Structure, USDA Forest Service General Technical Report PNW-GTR-351, 1995.

Medley, K.E., Okey, B.W., Barrett, G.W., Lucas, M.F., and Renwick, W.H., Landscape change with agricultural intensification in a rural watershed, southwestern Ohio, U.S.A., *Landscape Ecol.*, 10(3), 161, 1995.

Menz, G., Landschaftsmaβe und Fernerkundung—neue Instrumente für die Umweltforschung, *Geogr. Rundsch.*, 49, 1, 1997.

Moss, D. and Davies, C.E., Cross-References between the EUNIS Habitat Classification and the Nomenclature of CORINE Land Cover, Centre for Ecology & Hydrology, NERC, Swindon, U.K., 2002a.

Moss, D. and Davies, C.E., Cross-References between the EUNIS Habitat Classification and Habitats Included on Annex I of the EC Habitats Directive (92/43/EEC), Centre for Ecology & Hydrology, NERC, Swindon, U.K., 2002b.

Moss, D. and Davies, C.E., Cross-References between the EUNIS Habitat classification and the Palaearctic Habitat Classification, Centre for Ecology & Hydrology, NERC, Swindon, U.K., 2002c.

Muller, M.R. and Middleton, J., A Markov model of land-use change dynamics in the Niagara region, Ontario, Canada, *Landscape Ecol.*, 9(2), 151, 1994.

Murray, R.T.M. and Cooper, A., A Land Classification and Landscape Ecological Study of Northern Ireland, Report to Environment Service, Department of the Environment for Northern Ireland, University of Ulster, Coleraine, 1992.

Naveh, Z. and Lieberman, A.S., *Landscape Ecology Theory and Application*, 2nd ed., Springer-Verlag, New York, 1994.

OECD, *Environmental Indicators — Towards Sustainable Development* 2001, OECD, Paris, 2002.

OFS–OFEFP–ARE, *Mesurer le développement durable. Un aperçu de MONET—le système suisse de monitoring*, Office fédéral de la statistique, Neuchâtel, 2002, http://www.statistik.admin.ch/stat_ch/ber21/.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H., and Graham. R.L., Indices of landscape pattern, *Landscape Ecol.*, 1(3), 153, 1988.

O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Jackson, B.L., Jones, K.B., Riitters, K.H., and Wickham, J.D., Scale problems in reporting landscape pattern at the regional scale, *Landscape Ecol.*, 11(3), 169, 1996.

Orians, G.H. et al., *Ecological Indicators for the Nation*, National Academies Press, Washington, D.C., 2000.

Painho et al., Digital Map of European Ecological Regions (DMEER): its concept and elaboration, Second Joint European Conference (JEC) and Exhibition on Geographical Information, Barcelona, 1996.

Palang, H., Mander, Ü., and Luud, A., Landscape diversity changes in Estonia, *Landscape Urban Plann.*, 41, 163, 1998.

Petit, C., Scudder, T., and Lambin, E., Quantifying processes of land cover change by remote sensing: resettlement and rapid land cover changes in south-eastern Zambia, *Int. J. Remote Sens.*, 22(17), 3435, 2001.

Petit, C.C. and Lambin, E.F., Long term land cover changes in the Belgium Ardennes (1775–1929), model-based reconstruction vs. historical maps, *Glob. Change Biol.*, 8(7), 616, 2002.

Poudevigne, I. and Alard, D., Landscape and agricultural patterns in rural areas: a case study in the Brionne Basin, Normandy, France, *J. Environ. Manage.*, 50, 335, 1997.

Radeloff, V.C., Hammer. R.B., Voss, P.R., Hagen, A.E., Field, D.R., and Mladenoff, A.J., Human demographic trends and landscape level forest management in the northwest Wisconsin, Pine Barrens, *Forest Sci.*, 47(2), 229, 2001.

Riiters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timons, S.P., Jones, K.B., and Jackson, B.L., A factor analysis of landscape pattern and structure metrics, *Landscape Ecol.*, 10, 23, 1995.

Ruuska, R. and Helenius, J., GIS analysis of change in an agricultural landscape in Central Finland, *Agric. Food Sci. Finl.*, 5, 567, 1996.

Sachs, D.L., Sollins, P., and Cohen, W.B., Detecting landscape changes in the interior of British Columbia from 1975 to 1992 using satellite imagery, *Can. J. For. Res.*, 28, 23, 1998.

Schreiber, K.F., The history of landscape ecology in Europe, in *Changing Landscapes: An Ecological Perspective*, Zonneveld, I.S. and Forman, R.T.T., Eds., Springer-Verlag, New York, 1990, p. 21.

Sheail, J. and Bunce, R.G.H., The development and scientific principles of an environmental classification for strategic ecological survey in Great Britain, *J. Environ. Conserv.*, 2003.

Simpson, J.W., Boerner, R.E.J., DeMers, M.N., and Berns, L.A., Forty-eight years of landscape change on two contiguous Ohio landscapes, *Landscape Ecol.*, 9(4), 261, 1994.

Smeets, E. and Weterings, R., *Environmental Indicators: Typology and Overview*, European Environmental Agency, Copenhagen, 1999, Technical Report No. 25, http://reports.eea.eu.int/TEC25/en/tab_content_RLR

Singh, R.B., *Environmental Monitoring — Application of Remote Sensing and GIS*, Geocarto International Centre, Hong Kong, 1991.

Statistics Canada, *Human Activity and the Environment*, Ontario, 2000.

Stoorvogel, J.J. and Fresco, L.O., Quantification of land use dynamics: an illustration from Costa Rica, *Land Degrad. Dev.*, 7(2), 121, 1996.

Thibault, P.A. and Zipperer, W.C., Temporal changes of wetlands within an urbanizing agricultural landscape, *Landscape Urban Plann.*, 28, 245, 1994.

Tolle, T., Powell, D.S., Breckenridge, R., Cone, L., Keller, R., Kershner, J., Smith, K.S., White, G.J., and Williams, G.L., Managing the monitoring and evaluation process, in *Ecological Stewardship*, Sexteon, W.T., Malk, A.J., Szaro, R.C., and Johnson, N.C., Eds., Vol. III, 1999, p. 585.

Turner, M.G., Gardner, R.H., and O'Neill, R.V., *Landscape Ecology in Theory and Practice: Pattern and Process*, Springer-Verlag, New York, 2001.

Turner, M.G. and Gardner, R.H., Quantitative methods in landscape ecology: The analysis and interpretation of landscape heterogeneity, *Ecological Studies*, Vol. 82, Springer-Verlag, New York, 1991.

UN, *Indicators of Sustainable Development: Framework and Methodologies*, United Nations, Department of Economic and Social Affairs, Commission on Sustainable Development, New York, 2001, Background paper No. 3.

Vahrson, W.G., Landschaftsmonitoring—einige Grundgedanken und Konzeptionen, *Ebers. Wiss. Schr.*, 2, 9, 1998.

Volk, M. and Steinhardt, U., Landscape balance, in *Landscape Balance and Landscape Assessment*, Kroenert, R., Steinhardt, U., and Volk, M., Eds., Springer-Verlag, Berlin, 2001, p. 163.

Wenig, Q.H., Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling, *J. Environ. Manage.*, 64(3), 273, 2002.

Wiens, J.A., Population responses to patchy environments, *Annu. Rev. Ecol. Syst.*, 7, 81, 1976.

Wiens, J.A., Spatial scaling in ecology, *Funct. Ecol.*, 3, 385, 1989.

Wiens, J.A. and Milne, B.T., Scaling of 'landscapes' in landscape ecology, or, landscape ecology from a beetle's perspective, *Landscape Ecol.*, 3, 87, 1989.

Wu, J., Shen, W., Shun, W., and Tueller, P.T., Empirical patterns of the effects of changing scale on landscape metrics, *Landscape Ecol.*, 17, 761, 2002.

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., He, L., and Zhao, R., Landscape-scale characterisation of cropland in China using vegetation and landsat TM images, *Int. J. Remote Sens.*, 23(18), 3579, 2002.

Zheng, D., Wallin, D.O., and Hao, Z., Rates and patterns of landscape change between 1972 and 1988 in the Changbai Mountain area of China and North Korea, *Landscape Ecol.*, 12, 241, 1997.

13 Nonsampling Errors in Ocular Assessments— Swedish Experiences of Observer Influences on Forest Damage Assessments

S. Wulff

CONTENTS

13.1 INTRODUCTION

Collecting observations for monitoring purposes requires suitable sampling design and methods for estimation, measuring, and assessment. A wise choice of methods simplifies the compiling of the data. In the textbook by Thompson,[1] a good coverage of basic and standard sampling designs and estimation methods are given. In this book Thompson concludes that "getting good estimates with observations means picking out the relevant aspects of the data, deciding whether to use auxiliary information in estimation, and choosing the form of the estimator."[1] The applicability

1-56670-641-6/04/$0.00+$1.50

of measuring and assessment methods depends on many different factors. One crucial factor, which has to be considered, is that the selected method should give results that are consistent over time. Otherwise, very large changes must occur before one can safely conclude that a change really has taken place.[2]

In theory, basic sampling assumes that the true values of a variable are measured. Errors in the estimate only occur as sampling errors due to the fact that only a part of the population is included in the sample. In real survey situations, however, an estimate may also be afflicted by nonsampling errors. Errors could arise in the data collection as missing data and recording errors. It is recommended to make every effort to minimize such errors. Still, substantial nonsampling errors as measuring errors may occur. Accurate interpretation of results demands estimations of the accuracy of the basic measurements.

A measurement error is defined as the difference between a “true” value and a recorded value. In many cases in environmental monitoring the variable of interest is a visual observation of amounts or symptoms. The measurement errors are here affected by factors such as weather conditions during the observations, the visibility of objects, and the status of the observed object.[3] The accuracy of the assessments also largely depends on the experience and personal style of the observer.[4] The personal style involves different aspects, e.g., an observer’s ability to detect variation and influence of prior expectations. This could be summarized in visual perception. In Gordon,[5] an overview is given of the many aspects of this topic. One important understanding is that perceptions are not simply inputs.

In this chapter, nonsampling errors and their possible impact on data quality will be discussed. To describe observer influence on assessments as measurement errors, an example is given from the inventory of forest health in Sweden. Important issues such as accuracy of visual observation and consistency over time are distinguished.

13.2 FOREST HEALTH INVENTORY IN SWEDEN

In Sweden, forest health has, since 1984, been assessed in the National Forest Inventory (NFI), and since 1994, in the National Forest Damage Inventory (NFDI).[6] The inventories are part of an International Cooperative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests). In the forest health inventory, as in many other kinds of environmental monitoring, several key variables are of the ocular assessments type. One variable of interest is defoliation, widely used over Europe as the most important indicator of forest health, although the symptoms are ambiguous as no single cause can explain the observed symptoms.[7]

Slightly different methods have been applied in different countries in forest health assessments throughout the years.[8] The differences relate to what reference (crown photographs, local reference tree, or imaginary reference tree) has been used. In Sweden, an imaginary reference tree is used, and each tree is compared to its presumable appearance in full foliage. The assessments are made regardless of the causes of damages. This method, which is based on extensive experience among the observers of different types of trees and forest, demands many calibration exercises to support a common view on how the assessment should be made. Still, there may be a risk for changes of reference over time.

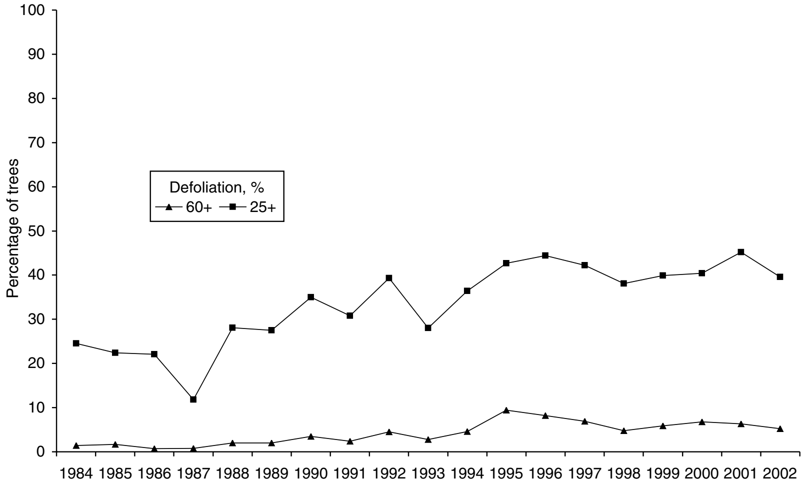


FIGURE 13.1 Proportion of trees with a defoliation >25% and >60%. Norway spruce (*Picea abies*) in older forests. Sweden, 1984 to 2002.

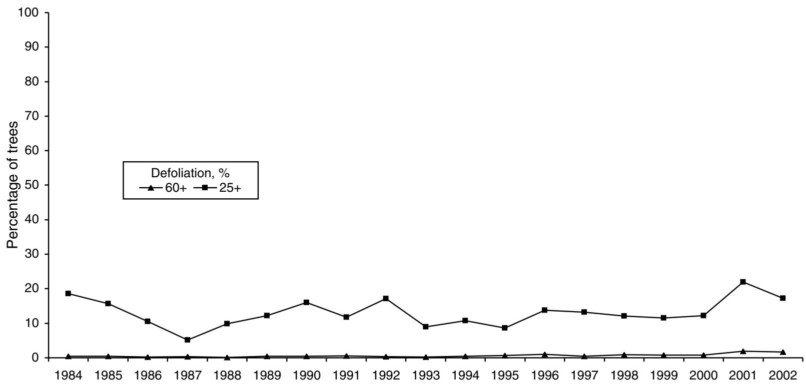


FIGURE 13.2 Proportion of trees with a defoliation >25% and >60%. Scots pine (*Pinus silvestris*) in older forests. Sweden, 1984 to 2002.

In Figure 13.1 and Figure 13.2 the development of defoliation is shown on the main tree species in Sweden: Norway spruce (*Picea abies*) and Scots pine (*Pinus silvestris*). The results from 19 years of inventory indicate over the long-term development of damage, deterioration in tree condition in Norway spruce up to the mid-1990s. On Scots pine no clear trend is revealed. However, in both figures a large variation in the degree of defoliation between single years is seen. This variation is probably due to occasional stress factors such as extreme weather conditions and/or temporary outbreak of insect attacks or fungi. A large outbreak of *Gremmeniella abietina* in 2001 is clearly noticed as an increased defoliation on Scots pine. The high level of defoliation on both tree species in the early 1980s can be connected to very dry summers. However, to some extent, the uncertainty of the observation methods also affects the results. Although the long-term development of forest damage is the most important information from the inventory, the measurement errors and their possible impact on the analysis need to be addressed.

13.3 HOW TO EXPRESS THE ACCURACY OF ASSESSMENTS?

The reliability of measurements largely depends on the selection of variables included in an inventory. Objective measurements are desirable; however, the significance and simplicity of the variables of interest are important arguments in their selection. A simple measurement that is easy to collect will lower the costs and thereby make an increased number of samples possible. A careful selection of variables is advisable as a poor choice will lead to difficulties in the analysis. Although a variable of interest is highly significant, it can be of low value due to less efficient estimations with a weak accuracy of measurements. A good knowledge of the measurement errors should therefore be of great concern.

13.3.1 Inferences Based on Continuous Data

The magnitude of measurement errors depends of both the precision and the mean deviation from a true value. Dealing with continuous data and assuming a normal distribution of the errors, analysis in terms of bias and variance components is most suitable. However, estimation of the bias component cannot be made unless an unbiased reference is determined. In many studies it is not possible to assess the bias due to difficulties in establishing an unbiased reference. On the other hand, assuming a constant bias for the measurements by a reference standard eliminates the bias when taking the difference between the mean measurements by observers and a reference standard. By assuming the same precision in the measurements of all the observers, including the reference standard, the following test statistic can be applied

$$t_0 = \frac{\overline{\text{diff}}}{\text{SD}_{\text{diff}}\sqrt{n}} \tag{13.1}$$

where $\overline{\text{diff}}$ is the mean of the differences, SD_{diff} is the standard deviation of the differences and n the number of observers. The hypothesis H_0: $\overline{\text{diff}} = 0$ was tested and rejected if $|t_0| > t_{\alpha/2n,n-1}$.

13.3.2 Measures of Agreement

In many cases, data are collected in a discrete form. To study the reliability of measurements made on a discrete scale, agreement is a useful estimate. The agreement of the observer's assessments can be estimated by calculating the percentage of the registrations that shared the same classification (actual agreement). However, for skewed distribution with many observations in a few highly agreeable classes, the actual agreement may be misleading by not taking the marginal totals in account. A better estimate of agreement is the kappa statistic (K).[9,10] The kappa statistic compares the actual agreement with the agreement expected from pure chance classification (chance agreement). The kappa statistic is given as

$$K = \frac{P_{actual} - P_{chance}}{1 - P_{chance}} \tag{13.2}$$

where P_{actual} is the proportion of actual agreement and P_{chance} is the proportion agreement expected from pure chance agreement. A kappa value of 1 points to a full agreement between the observer teams and 0, an agreement that is the same as would have been the gain from a pure chance agreement.

13.3.3 Evaluation of the Accuracy of Assessments in the Swedish Forest Health Survey

Forest damage assessments were evaluated in different calibration courses and in control surveys.[10] During annual training courses assessments were compared when several possible causes of error could be kept constant as, for instance, weather conditions. The assessments of the observers were evaluated with a national reference standard. To study the accuracy and the consistency of the observers, the test statistic described above was applied. The same test statistic was also applied on the data from the control survey. As the skill of the control teams and the ordinary field team was about the same, the two independent observations may be considered to be of equal precision. The control survey was also mostly conducted as pair-wise remeasurements by ordinary field teams.

Significant differences were found between observers in a study between 1995 and 1999, but the overall observer average assessments of defoliation did not differ significantly from the reference standard.[10] Larger differences are found when adding data from recent years to this study (Table 13.1). The data in the table refer to field tests done at the very beginning of the field season. Corresponding field tests carried out in the middle of the field season point to no significant difference between the overall observer average and the reference standard. The results from the control survey during 1995 to 2002 show a mean difference of 0.3 to 2.2% (Figure 13.3) and a standard deviation of differences (SD_{Diff}) in assessments of 6.6 to 17.8% units. By dividing SD_{Diff} with $\sqrt{2}$ retrieves, the standard deviation of single observations (SD_{Obs}) covers an interval of 4.7 to 12.6% units. Figure 13.4 indicates that low and high defoliation scores are assessed with a higher accurateness between observers.

Measurement of agreement is illustrated by assessment carried out on the grade of infection (amount of dead foliage) of *Gremmeniella abietina* in 2001. The following damage classes were used: 0 to 10%, 11 to 25%, 26 to 60%, 61 to 100%. The reliability of the assessments was evaluated by degree of agreement and estimates of kappa statistic. The results reveal the ability to detect the damages caused by the fungi. The control survey revealed an actual agreement of the assessments on 0.86 (n = 474). However, with a skewed distribution (about 75% of the assessments

TABLE 13.1
***t*-Test of the Differences between the National Standard Reference and Mean of All Observers' Assessment of Defoliation at National Calibration Courses from 1995 to 2002**

	Norway Spruce (*Picea abies*)					Scots Pine (*Pinus silvestris*)				
Year	No. of Trees	Mean Defoliation	Diff	df	$\lvert t_0 \rvert$	No. of Trees	Mean Defoliation	Diff	df	$\lvert t_0 \rvert$
1995[a]	10	32.8	–0.8	14	0.65	10	16.3	0.0	14	0.02
1996[a]	10	35.3	–2.1	19	1.80	10	23.9	0.0	18	0.01
1997[b]	20	28.7	0.2	8	0.20	18	18.4	0.2	8	0.31
1998[b]	20	26.3	0.7	7	1.04	20	16.8	–0.9	7	1.23
1999[a]	10	33.7	–0.5	11	0.25	10	18.3	–0.7	12	0.90
2000[b]	25	19.5	0.1	6	0.10	20	15.7	0.3	6	0.23
2001[b]	26	22.9	–2.5	6	2.25	21	29.4	–5.8	6	3.11
2002[b]	21	23.4	–3.3	5	2.29	20	11.8	2.5	5	3.20

[a] Single observer assessment;
[b] Observer team assessment.

Source: Modified from Wulff, S., The accuracy of forest damage assessments — experiences from Sweden, *Environ. Monit. Assess.*, 74, 295, 2002.

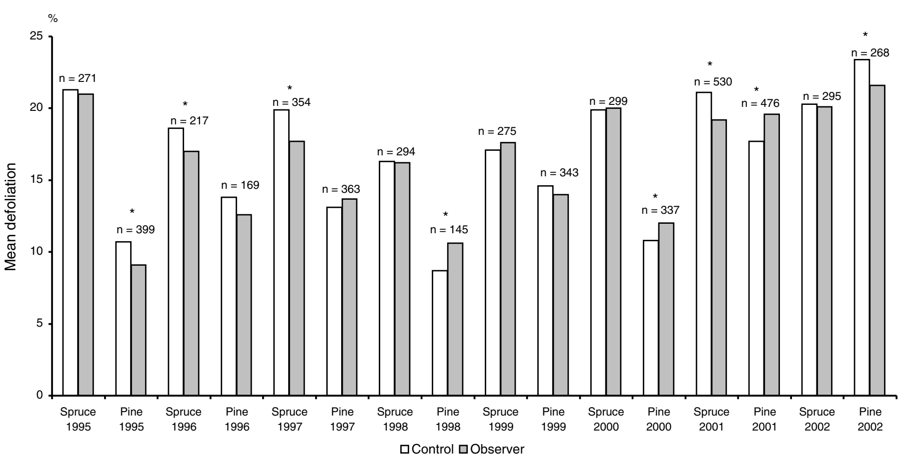


FIGURE 13.3 Mean defoliation assessment for Norway spruce (*Picea abies*) and Scots pine (*Pinus silvestris*) by observer team and controlling team, respectively. Control survey from 1995 to 2002. (*Note:* *Significant difference at 5% level.)

in class 0 to 10), the kappa statistic will give a more appropriate estimate of the agreement. The estimate of kappa statistic was 0.69. How can the estimations of kappa statistic be interpreted? Landis and Koch[11] propose a description of relative

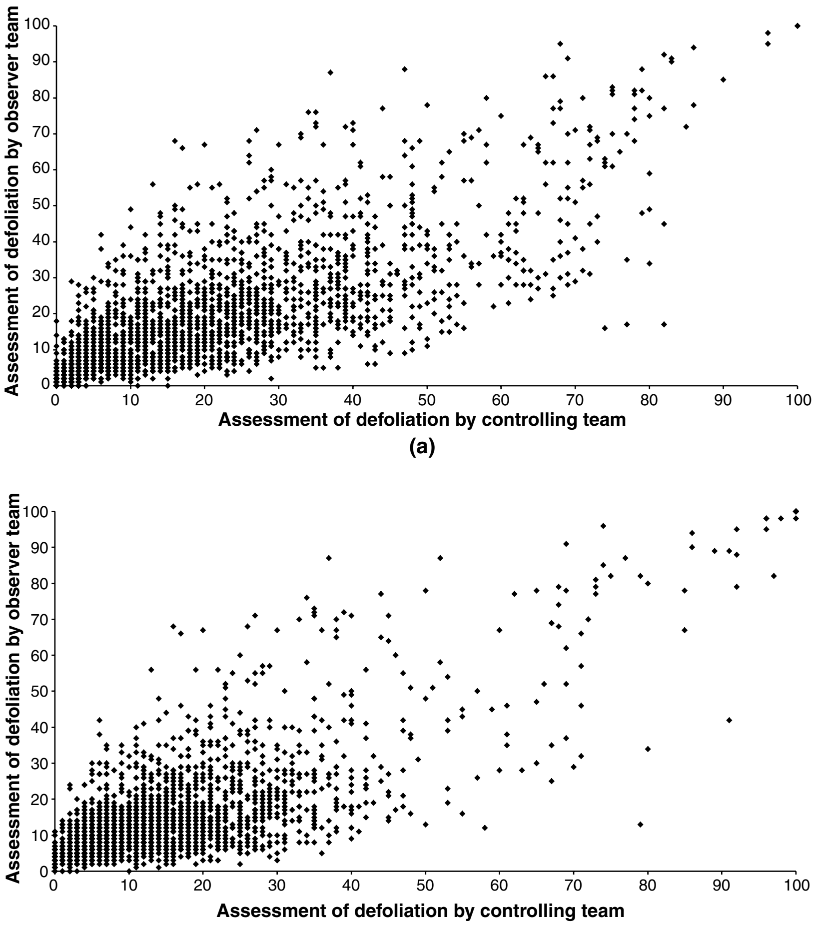


FIGURE 13.4 Assessment of defoliation on (a) Norway spruce (*Picea abies*) and (b) bellow Scots pine (*Pinus silvestris*) by observer team and controlling team in the control survey, 1995 to 2002. (Modified from Wulff, S., The accuracy of forest damage assessments — experiences from Sweden, *Environ. Monit. Assess*., 74, 295, 2002.)

strength of agreement that can be useful. The estimated kappa statistic in this case will be in the class of 0.61 to 0.80 that corresponds to a substantial agreement.

13.4 PROBLEMS WITH MAINTAINING THE CONSISTENCY OF ASSESSMENTS

To begin, we have to describe what we mean by consistency. In statistics, consistency relates to an estimator converging closer and closer around its target as sample size increases. A most important objective in many surveys, as in environmental monitoring, is to follow the development of a parameter and to be able to detect changes that may occur. To fulfill the demands made on estimation, a good estimator also needs to be consistent over time, e.g., measurement of objects in the same status will be the same over time.

Field checks on forest health assessments, as calibration courses, have been carried out on both a national and international scale for some years now.[12] As the objectives of the international courses from the beginning have been focused on harmonizing, analyzing, and discussing the differences in crown condition assessment on individually selected trees, the findings from the courses can only be used as indications. Since 2000, focus has been set on quantifying the differences between countries, and in the last few years, a permanent set of calibration plots with objectively selected trees has been established.[13] Thereby, there is a possibility to use the international results, together with the national calibration courses as further indications of the consistency of the assessments over time.

Photos have become a useful assistance in the forest health survey with the use of semiautomatic image analysis systems for crown condition assessments.[14] These estimates can be the unbiased reference previously missing and strengthen the follow up of consistency of assessment over time.

In the analysis of forest health assessments in Sweden the intrapersonal variance is exposed. Significant differences between observers indicate that observer bias also is present in this study. Observer bias interacts with both sampling error and natural variation and will influence the interpretation of the development of damages.[15] A permanent allocation of surveyor teams can reduce the effect of observer bias but may imply difficulties in interpretation of geographical patterns. However, since the study also reveals a substantial interpersonal variance among observers, why an allocation of surveyor teams should be made to the same area is not obvious.

13.4.1 Visual Perception

Differences in assessments found among observers largely depend on the experience and personal style of the observers. For instance, in a study by Ringvall et al.[2] on surveyor consistency in presence/absence sampling for monitoring vegetation, the influence of experience appears as more accurate and consistent assessments. Significant differences are more often found between observers in a less experienced group than in an experienced group. The impact of an observer's personal style involves several aspects but it can appear as shown in Table 13.2, where five different observers' assessments are compared with the assessment done by a national reference standard. The assessments were carried out in the same stand during different calibration courses. Three of the observers' performed a very consistent assessment whereas two always scored lower compared to the reference. Observers 1 and 2 should be recognized as experienced with several years as surveyors in the forest health inventory prior to 1995, while for observers 3 to 5, 1995 was only their second year in the inventory. Experience will improve their ability to detect differences/changes but also other underlying conditions which affect visual perception ability. Prior expectation also will influence the assessments as well as expectation of what is seen and what it is seen *as*. A distinction made by Fodor and Pylyshyn[16] between "seeing" and "seeing as" describes the difference: "What you see when

TABLE 13.2
Five Single Observers' Deviation from the National Reference Standard Assessment on Norway Spruce (*Picea abies*)[a]

	Assessment Carried Out by Observer									
	1		2		3		4		5	
Year	Diff	SD_{diff}	Diff	SD_{diff}	Diff	SD_{diff}	Diff	SD_{diff}	Diff	SD_{diff}
1995	–8.7	12.6	6.4	10.9	–4.4	12.5	3.4	10.6	0.7	7.7
1996	–7.6	6.2	–5.8	7.4	–4.8	9.3	–11.7	7.7	0.7	8.3
1999	–10.1	7.3	3.8	7.7	–3.9	7.6	–6.7	6.7	1.8	4.2

Note: n = 10.

[a] Data from national calibration courses in 1995, 1996, and 1999.

you see a thing depends upon what the thing you see *is*. But what you see the thing *as* depends upon what you know about what you are seeing."

13.5 CONCLUSIONS

In sample surveys, as in many kinds of environmental monitoring, nonsampling errors indeed interact with sampling errors and natural variation and will influence interpretation of the results. Not to address nonsampling errors and their impact on the analysis can be a serious misjudgement. However, it can be concluded that the long-term development of a parameter of interest, rather than short-term fluctuations, is the most important information from these kinds of inventories.

REFERENCES

1. Thompson S.K., *Sampling*, John Wiley & Sons, New York, 1992, p. 1.
2. Ringvall, A. et al., Surveyor consistency in presence/absence sampling for monitoring vegetation — case studies in boreal forest, manuscript submitted, 2003.
3. Köhl, M., Waldschadensinventuren: Mögliche ursachen der variation der nadel-/blatt-verlustschätzung zwischen beobachtern und folgerungen für kontrollaufnahmen, *Allg. Forst. Jagdztg*., 162, 210, 1991.
4. Solberg, S., Crown density assessment, control surveys and reproducibility, *Environ. Monit. Assess*., 56, 75, 1999.
5. Gordon, I.E., *Theories of Visual Perception*, 2nd ed., John Wiley & Sons, Chichester, U.K., 1996.
6. Wulff, S., Forest Health Conditions in Sweden 1998, in *Crown Condition Assessment in the Nordic Countries, Proceedings from the 5th International ECE/EU Intercalibration Course for Northern Europe on Crown Condition Assessment and SNS;* Ad Hoc Working Group Meeting on Monitoring of Forest Damage, June 16–18, 1999, Estonia, Õunap, H., Ed., Estonian Centre of Forest Protection and Silviculture, Tartu, 2000, 20.

7. Innes, J.L., *Forest Health: Its Assessment and Status*, CAB International, Wallingford, U.K.,1993.
8. UN/ECE, *Manual on Methods and Criteria for Harmonised Sampling, Assessment, Monitoring and Analysis of the Effects of Air Pollution on Forests*, 4th ed., Programme Coordinating Centre in Federal Research Centre for Forestry Products (BFH), Hamburg, Germany, 1998.
9. Cohen, J., A coefficient of agreement for nominal scales, *Educ. Psych. Measure*., 20, 37, 1960.
10. Wulff, S., The accuracy of forest damage assessments — experiences from Sweden, *Environ. Monit. Assess*., 74, 295, 2002.
11. Landis, J.R. and Koch, G.G., The measurement of observer agreement for categorical data, *Biometrics*, 33, 159, 1977.
12. Dobbertin, M. et al., Quality of crown condition data, in *Ten Years of Monitoring Forest Condition in Europe — Studies on Temporal Development, Spatial Distribution and Impacts of Natural and Anthropogenic Stress Factors*, Müller-Edzards, C., De Vries, W., and Erisman, J.W., Eds., UC-UN/ECE, Brussels and Geneva, 1997, p. 7.
13. Ferretti, M. et al., New design of international cross-calibration courses of ICP forests and the EU scheme, Draft guidelines, Unpublished manuscript prepared for the ICP Forests Expert Panel on Crown Condition Assessment, Federal Research Center for Forestry and Forest Products (BFH) and LINNÆA Ambiente Srl, Hamburg, Germany, 2002.
14. Mizoue, N., CROCO: Semi-automatic image analysis system for crown condition assessment in forest health monitoring, *J. For. Plann*., 8, 17, 2002.
15. Strand, G.H., Detection of observer bias in ongoing forest health monitoring programmes, *Can. J. Forest Res*., 26, 1692, 1996.
16. Fodor, J.A. and Pylyshyn, Z.W., How direct is visual perception? Some reflections on Gibson's "Ecological approach," *Cognition*, 9, 139, 1981.

14 Tree-Ring Analysis for Environmental Monitoring and Assessment of Anthropogenic Changes

R. Juknys

CONTENTS

14.1 INTRODUCTION

Forests cover almost one third of the earth and their environmental and economic value is extremely high. Forest ecosystems are one of the most important sources of renewable resources, playing an exceptional role in the global turnover of materials and energy. Forests provide very important habitats and their role in biodiversity protection is very important. The environmental role of forests has especially increased along with the increase in environmental pollution. Forests can be considered as one of the main carbon dioxide sinks and help to absorb some air pollutants. The exceptional environmental and economical value of forests and their sensitivity to the environmental pollution requires us to consider forests as one of the most appropriate objects of environmental monitoring and assessment.

1-56670-641-6/04/$0.00+$1.50

The growth rate of trees is one of the best indicators of the general condition and sustainability of forests. Many investigators have pointed out that the anatomic structure of tree stems enables us to estimate annual radial increment for a long retrospective period, and that the width of annual tree rings is one of the most acceptable indicators for environmental monitoring needs (Kairiukstis, 1987; Eckstein, 1989; Schweingruber, 1996; McLaughlin et al., 2002; Juknys, Stravinskiene, and Vensloviene, 2002).

Tree-ring analysis was applied for local forest damage assessment more than 100 years ago (Donaubauer, 1980). However, this technique became very common only in the beginning of the 1980s, when forest damage was recognized as a deeper and wider problem than previously supposed (Eckstein, 1985; Cook, 1987; Schweingruber, 1985). Environmental acidification effects caused by long-range transport of polluted air masses were considered as new and very powerful external factors, considerably affecting the condition and productivity of forests across vast areas of West and Central Europe (Knabe, 1981; Bauer, 1982; Hosker and Lindber, 1982; Braker and Gaggen, 1987; Pollanschutz, 1987; Jones, 1989; Mehne-Jakobs, 1990). Increased concentrations of ground-level ozone are also often considered as one of the main reasons of forest decline on a regional scale (Klap et al., 1997; McLaughlin et al., 2002).

However, it is necessary to note that at present there is no consensus on any primary reason for regional forest decline. Unfavorable climatic conditions, invasions of forest pests, diseases, errors of forest management (plantations with monocultures, intensive felling, fertilization, etc.) are often cited along with environmental pollution (Boneau, 1986; Nys, 1989; De Vries et al., 1997; Klap et al., 1997). In the opinion of most scientists, forest decline is prompted by a complex of natural and anthropogenic stressors.

Some investigators are denying any regional forest decline phenomenon at all and have presented contrary data indicating that the growth rate of trees and stands in some parts of Central Europe is currently higher than in the past (Kander and Innes, 1995). Global warming and increased nitrogen deposition along with improved forest management usually are mentioned as possible causes of accelerated tree growth. Temporal acceleration of tree growth has been determined as well where local environmental pollution exists (Juknys et al., 2002).

Such conflicting data and opinions are evidently a result of the very complicated nature of the integrated impact of numerous natural and anthropogenic factors and their different combinations in forests. Tree rings can be considered as natural monitors and are able to record information on the impact of natural and anthropogenic stressors; however, deciphering this information is a rather complicated task.

Essential reduction of air emissions as a result of very important international efforts started in Western Europe in the mid-1980s. The considerable reduction of air emissions along with transitional economy decline took place in eastern European countries after the collapse of the Soviet block in the early 1990s. Lower emissions caused the improvement of air quality and reduction of acid deposition (Sopauskiene and Jasineviciene, 1997; Juknys et al., 2002). Air pollution in the surroundings of most industrial enterprises decreased even more than on a regional scale (Armolaitis, 1998). The condition and growth of damaged forests started to improve both locally and on a regional scale as a result of these positive environmental changes (Klap et al., 1997; DeVries et al., 1997; Ozolincius and Stakenas, 1999; Juknys et al., 2002).

The assessment of anthropogenic changes of monitored indicators that appear in a matrix of natural quasi-periodical fluctuations is one of the main problems of environmental science. Particular difficulties with interpretation of collected data arise in the case of low-level environmental pollution on a regional scale. Possibilities of detection of anthropogenic signals depend mainly on the following features of indicators:

- Sensitivity of the indicator
- Spatial and temporal variability of the indicator
- Duration of observations

Retrospective tree-ring analysis provides very useful long-term information and can serve as an appropriate tool for investigation and assessment of general consequences of anthropogenic environmental changes. Different modifications of tree-ring analysis, including analysis of wood density and chemical analysis of tree rings, have been elaborated on during the last few decades (Vaganov, 1990; Watmough, 1998; McLaughlin et al., 2002). However, tree-ring width directly reflecting changes in the productivity of forests is considered as the main object of dendroecological investigations (Eckstein, 1985; Cook, 1987; Juknys et al., 2002).

Determination of "norm" (normal radial increment in the case of tree-ring analysis) is the key issue for the quantitative assessment of anthropogenic changes of monitored indicators. Depending on the approach to evaluation of "normal growth," different methods for the assessment of anthropogenic changes of radial increment (tree-ring width) can be classified into three main methods of control of:

- Populations (stands of trees)
- Individuals (trees)
- Periods (of growth as shown in tree rings)

Methods of the first group are based on the comparison of tree increment or tree increment indexes of damaged stands with relatively healthy ones which are growing remotely from the source of pollution in similar site conditions and have approximately the same dendrometric characteristics (species composition, age, average height and diameter, density, etc.). An increment of these control stands is considered as normal in comparison with those damaged (Stravinskiene, 2002). The accuracy of assessment in this case is rather subjective and depends essentially on the choice of control stands. Methods in this group can be used in cases of local pollution with an apparent gradient of pollution decrease, i.e., increasing the distance from the pollution source. In the case of regional forest decline, the choice of control stands is problematic, and methods of control for individuals or for periods are usually used.

Methods of control for individuals are based on two presumptions:

1. Intraspecific variability of tree sensitivity to the impact of environmental pollution is essential.
2. Impact of crown condition (level of defoliation) to tree growth (radial increment) is statistically significant.

The increment of relatively healthy individuals (trees) without obvious signs of crown defoliation is considered as a norm in this method group and is compared with the increment of defoliated trees inside the same stand (Kramer, 1986; Soderberg, 1991).

Methods of control for periods are based on a quantitative analysis of the tree-ring series. The annual increment of trees during the period before the essential environmental impact (pollution) started is considered as normal and the prediction of normal growth for the period of injury is made on the basis of autoregressive (Kairiukstis and Dubinskaite, 1987; Shiyatov, 1987; Stravinskiene, 2002) or climate response (Cook, 1987; Becker, 1989; Juknys and Jancys, 1998) models.

Data on long-term investigations were collected in the area surrounding one of the largest pollution sources in Lithuania — Jonava Mineral Fertilizers Plant Achema — from 1980 to 2000. They are used to present different methods for the assessment of anthropogenic changes of tree growth in the polluted environment.

14.2 MATERIALS AND METHODS

The Achema Plant is located in the central part of Lithuania (55˚05 latitude, 24˚20 longitude) at the confluence of the rivers Neris and Sventoji. Production of fertilizers started in 1965 and gradually expanded until 1978, when an even more toxic component, the Nitrophoska Division, was constructed. Total annual emissions reached almost 40,000 tons. Nitrogen fertilizers produced by fixing nitrogen from air are the basic Achema Plant product. A large amount of energy is needed for this endothermic process, and emissions of sulfur, nitrogen, and carbon oxides comprise the main part of the air pollutants. Rather large quantities of ammonia and dust from the mineral fertilizers are emitted into the air while producing these fertilizers (Figure 14.1). Apatites from Cola peninsula have been used as raw material for the Nitrophoska Division's production since 1978, and small quantities of heavy metals (Zn, Cu, Mn, Cr, Ni, Cd, etc.) have been detected in air emissions as well.

Scots pine (*Pinus sylvestris* L.) forests prevail in this region. The first signs of forest damage surrounding Achema were noticed in 1972, but this problem became extremely acute in 1979 when, after a very hard winter, obvious signs of forest damage in the direction of prevailing winds were recorded up to 10 to 12 km from the plant, and in the coniferous forests which had completely died up to a distance of 2 to 3 km from the pollution source. The zone of damaged forests in the direction of prevailing winds expanded up to 20 to 25 km by the end of the 1980s, despite the essential reduction of emissions (Figure 14.1), when different pollution mitigation measures were implemented at Achema.

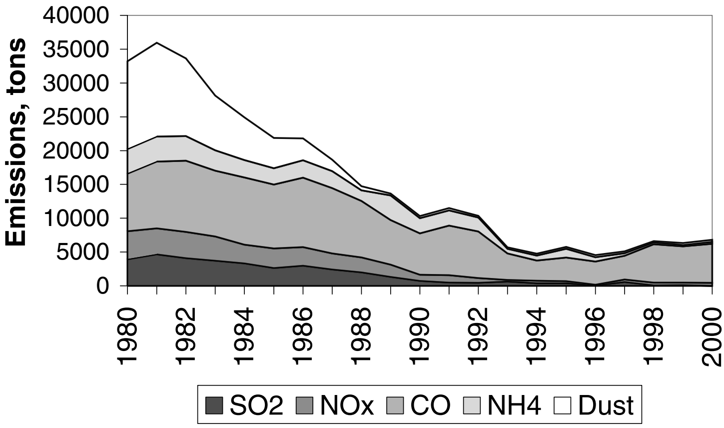


FIGURE 14.1 Annual emissions of Achema fertilizer plant.

An additional input to the reduction of emissions was a serious accident in the most polluting Nitrophoska Division in 1989, and this division was closed. As seen from Figure 14.1, total annual emissions of Achema were reduced almost eight times from the beginning of the 1980s up to the middle of the 1990s. Emissions of Achema have increased slightly during the last several years but do not exceed 6000 to 7000 tons annually.

The composition of emissions has also changed during the investigation period. Carbon monoxide comprised about 25%; sulfur dioxide, 12%; nitrogen oxides, 10%; ammonia, 10%; and mineral dust, almost 40% at the very beginning of the 1980s. The input of carbon monoxide was much bigger and made up 70 to 80% in the emissions of recent years. The input of other pollutants has essentially decreased to the point where sulfur dioxide comprised only 2 to 3%; nitrogen oxides, about 6%; ammonia, over 10%; and mineral dust, about 5%.

The total (wet and dry) annual deposition of sulfur during the last decade was reduced almost ten times near the Achema Plant and approximately three times at a distance of 10 to 20 km in the direction of the prevailing winds. The decrease of total nitrogen deposition was less essential. A four- to fivefold decrease was recorded near the source of pollution and an approximately twofold reduction at a distance of 10 to 20 km (Armolaitis, 1998).

Permanent investigations of surrounding forests were started in 1982 and have been continuing for the last 20 years. A three-stage sampling pattern was used for investigation of damaged forests: sampling of research stands; sampling of circular plots within sampled research stands; sampling of trees in circular plots for more detailed measurements and tree-ring analysis.

Estimation of tree condition (crown defoliation) was made for each sampled tree according to European forest monitoring methodology, and five defoliation classes were distinguished: (1) conditionally healthy trees with defoliation up to 10%, (2) slightly damaged trees with defoliation from 11 to 25%, (3) moderately damaged trees with defoliation from 26 to 60%, (4) severely damaged trees with defoliation from 61 to 99%, and (5) dead trees (UNECE, 1994). Instead of a rather rough classification, defoliation has been evaluated within 5% of accuracy since 1990.

Three trees closest to the center of the sampling plot were sampled during the third sampling stage, and the main stem and crown parameters (diameter, height) were measured. Wood samples for sampled trees of dominant and codominant classes without obvious injury were taken by a special borer for measuring annual tree rings, and a mean tree-ring series for each stand was established on the basis of a tree-ring series of 30 to 36 sampled trees.

Data of eight 80- to 90-year-old damaged pine stands, situated at different distances from the pollution source at (2.7 to 21.6 km), and data of three relatively healthy pine stands of similar age were sampled and measured according to the same sampling design at a distance of 50 to 55 km against the prevailing winds under similar site conditions.

14.3 IMPACT OF CROWN DEFOLIATION TO TREE INCREMENT

The methods of control for individuals are based on two main presumptions mentioned in Section 14.1. Regarding the first presumption concerning the variability of sensitivity, usually there are no essential contradictions, and it is rather obvious that trees with different levels of damage (levels of defoliation) can be found at the same site. However, rather different results were obtained during investigations of the relation between tree crown defoliation and radial increment (tree-ring width). Several authors (Kohler and Stratman, 1986) did not find any statistically significant relationship between crown defoliation and tree-ring increment. According to the others, a more obvious decrease of tree increment usually starts only when defoliation exceeds 20 to 30% (Schweingruber, 1985; Braker and Gaggen, 1987; Soderberg, 1991) or even 40 to 50% (Frantz et al., 1986; Petras, 1993).

Opposite results were presented by Philips et al. (1977) and Kontic et al. (1987). They found that decrease of tree increment in the polluted environment usually starts before obvious signs of crown defoliation can be determined. Sophisticated threshold patterns in the relations of crown defoliation and radial increment were discovered by Ozolincius (1996). Our earlier investigations have shown that the correlation between crown defoliation and radial increment is rather weak ($r = 0.3 – 0.5$), but usually statistically significant (Juknys and Jancys, 1998).

The investigations carried out by Kramer (1986) have indicated that in analyzing relations of crown defoliation and radial tree increment, quantitative crown parameters should be taken into account as well. Crown surface area and crown volume are usually considered as the main quantitative crown indicators in such a type of investigation. Analysis of our experimental data has shown that stem diameter correlates with the crown volume very closely ($r = 0.712 – 0.945$). Taking into account the very strong correlation between stem diameter and crown volume, the dependence of radial increment on crown defoliation and stem diameter was analyzed further. In this case, crown defoliation was considered as a qualitative crown indicator and stem diameter as the indicator reflecting the quantitative parameters of trees.

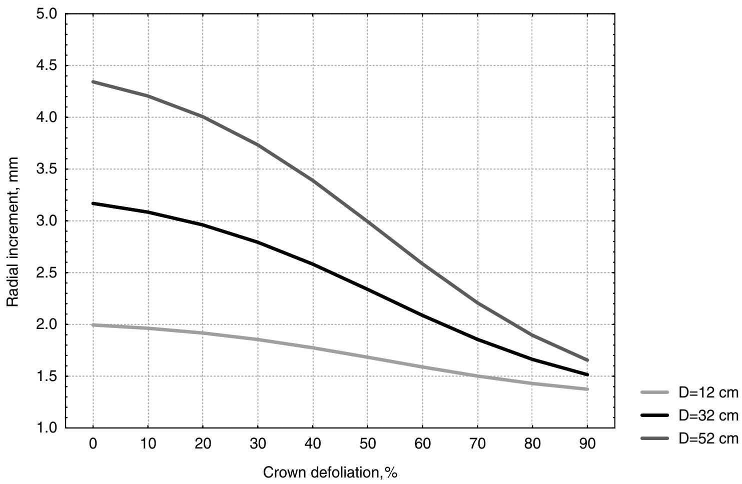


FIGURE 14.2 Dependence of pine radial increment on crown defoliation at different stem diameters (D).

The dependence of radial increment on crown defoliation (f) and stem diameter (d) was approximated according to the elaborated multidimensional regression model (Formula 14.1):

$$Zr = 1.29 + 0.0636*d - 0.066*d/(1 + 16*\exp(-0.048*f)) \quad (R^2 = 0.329) \qquad (14.1)$$

The dependence of radial increment on crown defoliation is expressed by a logistic curve (Figure 14.2). At first, radial increment decreases relatively slowly, while defoliation increases up to 20 to 25%. Further increase of defoliation leads to the essentially higher increment losses. However, having achieved 70 to 80% of defoliation, the increase of defoliation does not result in a fast decrease of radial increment.

Impact of crown defoliation on tree growth markedly depends on the quantitative parameters of the trees (in this case, the diameter). The impact of defoliation on radial increment of thin, suppressed trees is relatively much weaker than for the thick ones (Figure 14.2). While defoliation of large trees increases from 5 to 90%, their radial increment decreases almost three times, whereas the radial increment of thin trees in the same defoliation interval decreases about 35%. The annual radial increment of trees with zero defoliation can be considered as the norm and assessment of tree increment losses for differently damaged (defoliated) trees can be made on the basis of the elaborated regression model (Formula 14.1) presented above.

The main weakness of this method is that visual estimations of the main predictor—crown defoliation—is subjective and accuracy of data is rather low. Analysis of crown defoliation data in Europe has shown that there are sharp changes of defoliation levels at country borders due to methodological differences among countries (DeVries et al., 1997). In addition the relation of tree increment with crown defoliation is rather weak.

14.4 QUANTITATIVE ANALYSIS AND PREDICTION OF TREE-RING SERIES

The methods of control periods enable use of the tree-ring analysis data in the most effective way. As noted earlier, duration of observations and measurements is very important for the detection and assessment of anthropogenic changes in monitored indicators. Quite erroneous conclusions can be made with accurate but short-term data if natural, quasi-periodical fluctuations of monitored indicators are not taken into account. Methods and indicators should have a high ability to allow the use of retrospective information for a sufficiently long period.

Normal tree growth (annual increment) for the period of monitored anthropogenic impact is predicted on the basis of models created using the data from the period before the impact started. The effect of environmental impact (pollution) is assessed in accordance with the difference between the actual values of annual increment and those estimated by the model. Two main types of models are mainly used to predict a normal tree growth for the period of injury:

1. Autoregressive models, based on the quantitative analysis of internal regularities of quasi-periodical fluctuations of tree-ring series (Kairiukstis and Dubinskaite, 1987; Shiyatov, 1987; Stravinskiene, 2002);
2. Climate response models, based on multidimensional regression analysis of tree-ring width dependence on different climatic factors (Cook, 1987; Becker, 1989; Juknys and Jancys, 1998).

Taking into account that the prediction of normal annual increment as a basis for quantitative assessment of anthropogenic changes should be made for a rather long period (over 30 years), the predictive capacity of both types of the above-mentioned prediction models was first compared on the basis of the tree-ring series of relatively healthy Scots pine stands, sampled as controls at a distance of 50 to 55 km from the pollution source (Achema) and upwind of prevailing winds (Section 14.2).

Tree-ring series $\{Y_t\}$ for elaboration of autoregressive models was presented as a compound of three constituents:

$$Y_t = M_t\,(F_t + \varepsilon_t) \tag{14.2}$$

where M_t is the age trend of annual increment; F_t, periodic constituents; and ε_t, random process.

Since with the aging of trees the increment tends to diminish, the age trend has been approximated according to the exponential function (Fritts, 1976):

$$M_t = \mathrm{a}e^{ct} + \mathrm{b} \tag{14.3}$$

Harmonic analysis of periodic constituents T_j was performed according to the following method:

$$F_t = A_0 + \sum_{j=1}^{n} A_j \cos(2\pi t/T_j + \varphi_j) \tag{14.4}$$

The parameters a, b, and c of the age trend Equation (14.3) and parameters of periodic components A_j, φ_j Equation (14.4) were established by the least squares method.

Actual and model estimated tree-ring series for the control stands are presented in Figure 14.3. As illustrated by the data, 2 to 3 statistically significant (p <0.05) periodic components were singled out during harmonic analysis of the presented tree-ring series. Results of our earlier investigations have shown that for periodic components, medium terms of 9 to 15 and 20 to 25 years are the most powerful and usually take about two thirds of the general tree-ring series variance (Juknys and Vencloviene, 1998).

The end of the investigated time series was cut and a retrospective prediction was made on the basis of developed autoregressive models in order to assess the predictive capacity of these models. Three exercises were made by cutting 10, 20, and 30 years of a tree-ring series, and standard prediction error was calculated for different periods of prediction according to the following formula:

$$\mathrm{Sp} = 100^* \sqrt{(1/n)\sum_{t=k+1}^{t=k+n}((Y_t - W_t)/W_t)^2} \tag{14.5}$$

where Y_t is the actual value of the increment and W_t the estimated value of the increment.

Standard prediction error for a period of 10 years was estimated from 14.2 to 22.7%; for a 20-year prediction period, 16.9 to 26.4%; and for a period of 30 years, 19.6 to 29.4%. As seen from Figure 14.3, prediction capacity of autoregressive models based on the analysis of a comparatively short time series is rather low and, in the case of a prediction of a 30-year period, anthropogenic changes in the annual increment of not less than 20 to 30% can be reliably assessed statistically. However, autoregressive models are very powerful tools for the analysis and prediction of tree growth when relatively long-term (100 years and longer) tree-ring series are used (Kairiukstis and Dubinskaite, 1987).

Multiple climate response models are used most often as tools for the assessment of anthropogenic changes in annual tree increment (Eckstein, 1985; Cook, 1987; Schweingruber, 1996; Juknys et al., 2002). The methodology of tree-ring analysis for environmental monitoring needs is rather different from that for traditional dendroclimatology. Taking into account that annual radial increment (tree-ring width) depends on a whole complex of external and internal factors, large information noise is characteristic for both dendroecologic and dendroclimatolologic investigations. However, as was noticed by McLaughlin et al. (2002), what should be considered as a signal or as a noise depends on the goal of the investigations.

From the dendroclimatologic point of view, fluctuations of annual radial increment (tree-ring width) and its derivations from long-term average (age trend) are most informative regarding former climatic conditions. From the point of environmental monitoring, fluctuations of tree-ring width create a large information noise and major difficulties in the detection of anthropogenic signals (environmental pollution in this case). Climate response models are very helpful in selecting the main climatic factors which are responsible for tree-ring width fluctuations and enable reduction of information noise.

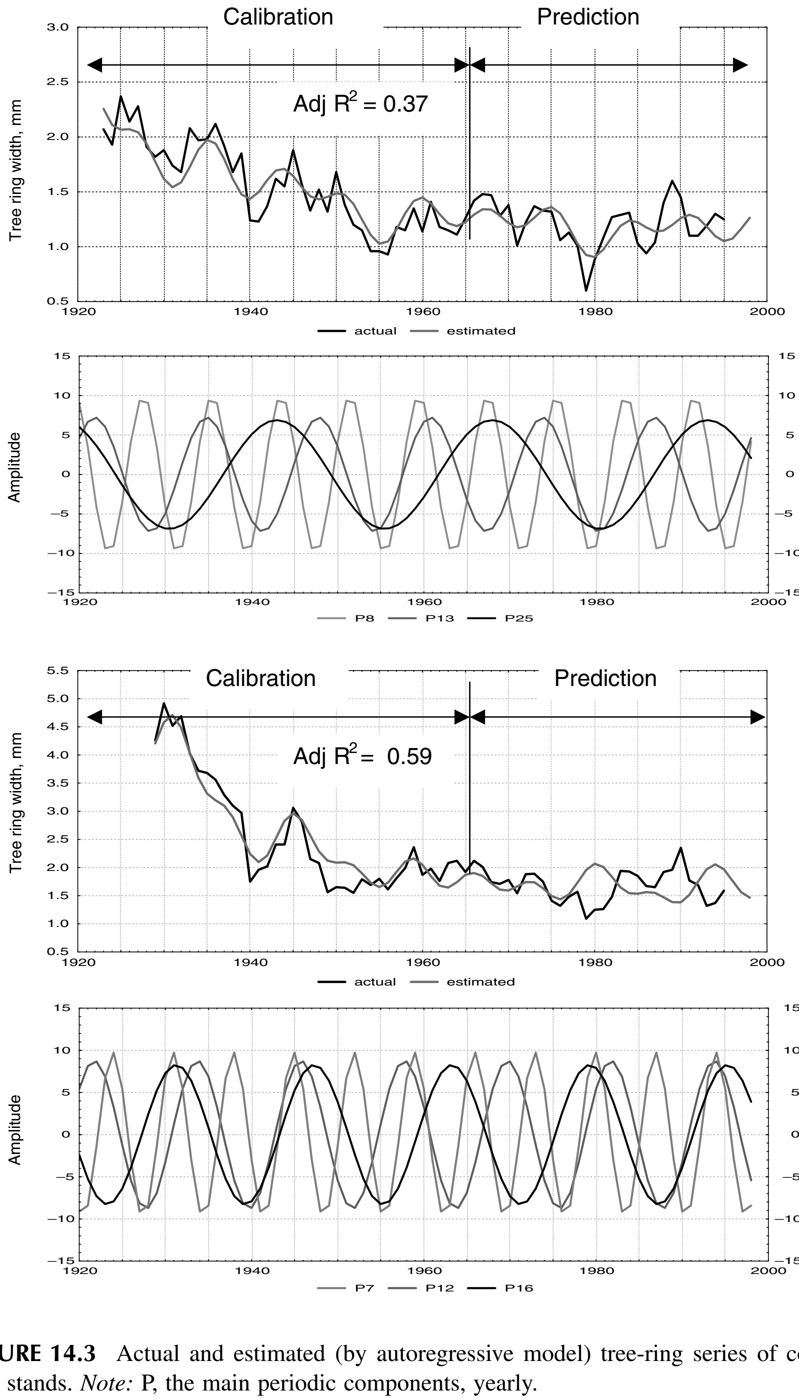


FIGURE 14.3 Actual and estimated (by autoregressive model) tree-ring series of control pine stands. *Note:* P, the main periodic components, yearly.

Another essential difference from dendroclimatologic investigations is the length of the analyzed tree-ring series. Relatively long tree-ring series (not less than 100 years) usually are used for dendroclimatologic investigations but for environmental monitoring purposes, it is necessary to work with the most common tree species, and length of tree-ring series usually do not exceed 60 to 80 years.

Different climate predictors can be included into climate response models; however, monthly temperatures and precipitation usually are most common choice (Fritts, 1976; Cook, 1987; Juknys and Jancys, 1998). Our earlier investigations (Juknys, 1994; Juknys and Vencloviene, 1998) have indicated that pine growth correlates most closely with the temperatures of late winter and early spring (February, March, and April) and temperatures of late summer (August) in our latitudes. In some cases, the temperatures of the previous September and October have had an essential impact on the radial increment of pine. The correlation of annual pine increment with precipitation is less significant in our latitudes. The closest relationships were found between tree-ring width and precipitation of late winter (February) and summer (June, July). Data on correlation of annual tree increment of relatively healthy (control) stands with monthly temperatures and precipitation are presented in Figure 14.4.

Solar activity was considered the main cause of midterm tree-ring width fluctuations from the very beginning of dendrochronologic investigations (Douglass, 1928). However, the impact of solar activity on tree growth is rather different in different regions (Shiyatov, 1987). A significant impact of solar activity on tree growth usually is not determined in the middle latitudes, and sometimes even correlation signs for different periods of the same tree-ring series are changed. According to our investigations, the correlation between solar activity and tree-ring width usually does not exceed 0.2 and it sometimes has negative values (Juknys and Vencloviene, 1998).

Predictive capacity of climate response models was evaluated in the same way as it was made for autoregressive models, and prediction error was calculated according to Equation (14.5) for 10-, 20-, and 30-year prediction periods.

Actual and estimated data of tree ring-series for the control Scots pine stands are presented in Figure 14.5 (Juknys et al., 2002). As the presented data illustrate, 4 to 6 climatic factors prove to be statistically significant ($p < 0.05$) and 37 to 67% of tree-ring series variance can be explained by climatic factors included in multiple climate response models. The list of included climatic factors is rather typical for climatic conditions of Lithuania. The temperatures of late winter and early spring, precipitations of winter, as well as the temperatures of past autumn prevail in this list.

Before the construction of the multiple climate response model, the time series of annual radial increment should be standardized, i.e., they should be transformed so that the mean of the series would be constant. Exponential function was used for excluding the age trend, and after dividing the measured values of the increment by the values of age trend, increment indexes were calculated. The correlation between increment indexes and climatic factors (average monthly temperatures and monthly precipitation) for the calibration period was estimated for deriving the regression equation. The climatic factors were included in the regression model by a stepwise

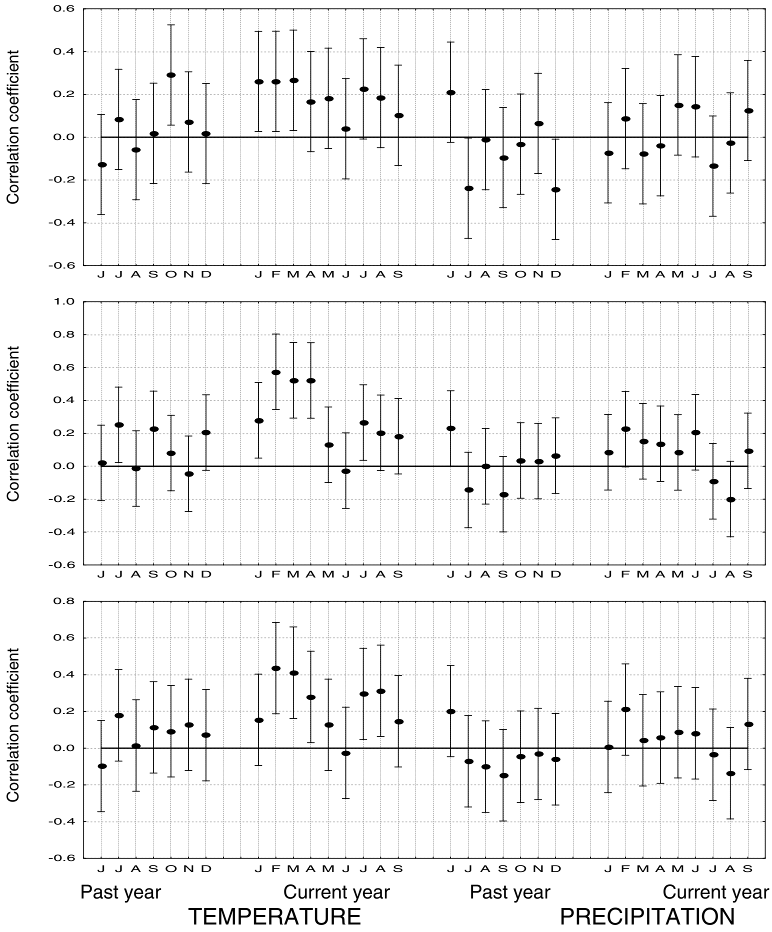


FIGURE 14.4 Correlation of the tree increment indices with climatic indicators.

method (Draper and Smith, 1986) as long as including a new variable significantly increases R^2. For a more vivid presentation of data (Figure 14.5 and Figure 14.6), the values of increment indexes were transformed again into actual values of the increment (Juknys et al., 2002).

Actual and estimated values according to multiple climate response model tree-ring series of control (relatively healthy) stands are presented in Figure 14.5. The following standard errors for different prediction periods were determined: The standard errors were set at 7.6 to 11.3% for a 10-year prediction; 10.7 to 14.6% for a 20-year prediction period; and 12.5 to 16.1% for a period of 30 years. The general conclusion that the resolving capacity of climate response models is much higher than that for autoregressive models can be drawn from the obtained results. Our earlier investigations have shown that closeness of climate/annual tree-ring width relations depends on the length of time series. If the length of time series exceeds

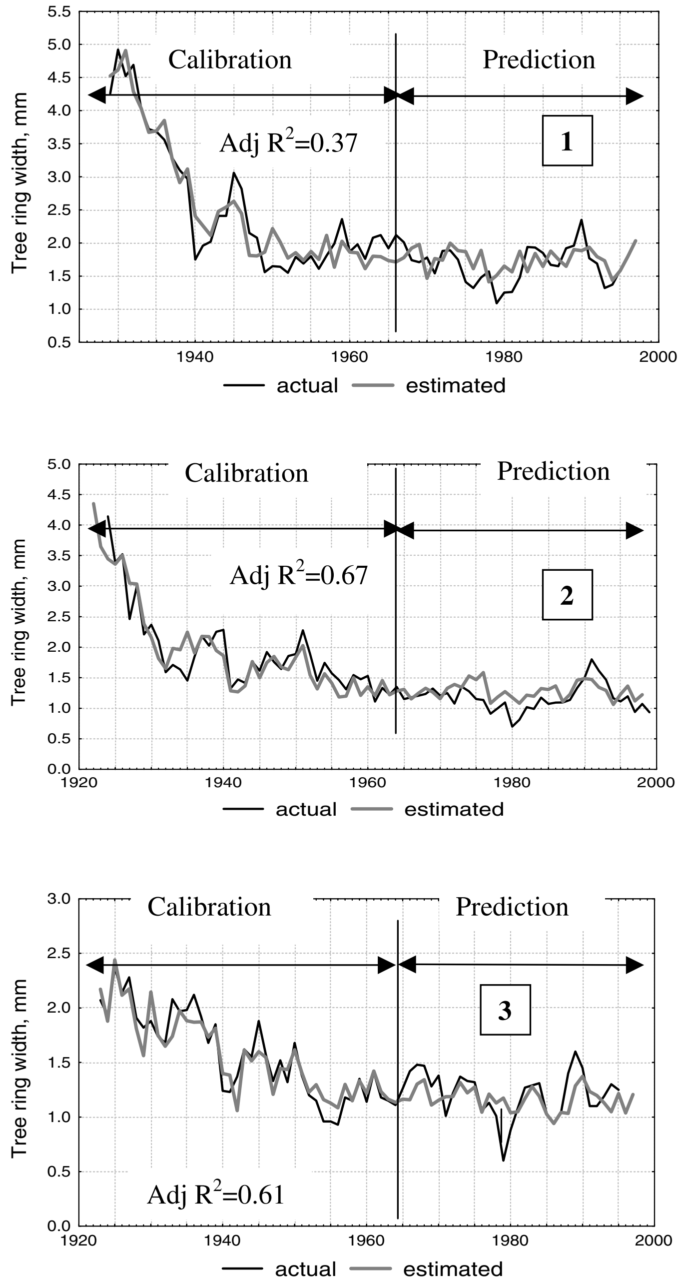


FIGURE 14.5 Actual and estimated tree-ring series of the control pine stands. (Modified from Juknys, R., Stravinskiene, V., and Vencloviene, J., Tree-ring analysis for the assessment of anthropogenic changes and trends, *Environ. Monit. Assess.*, 77, 81–97, 2002.)

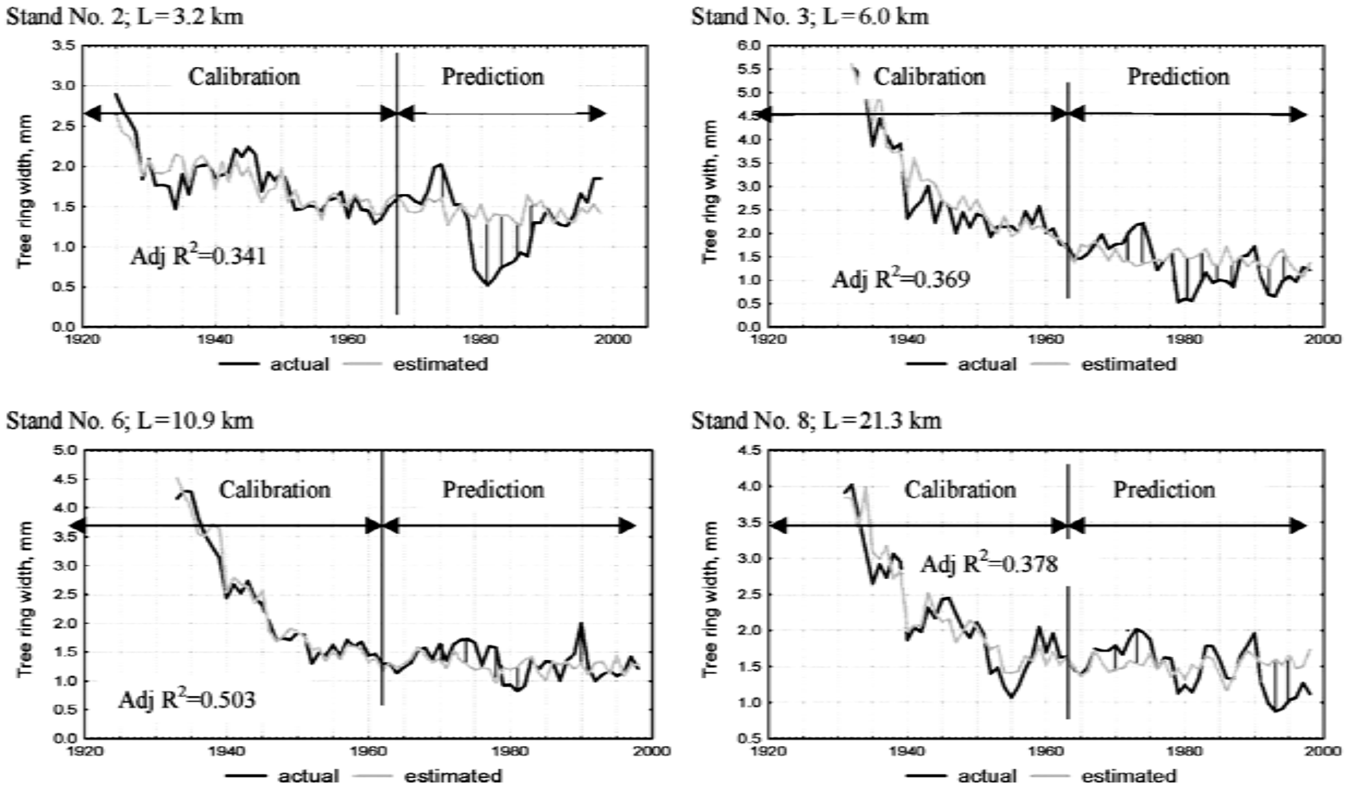


FIGURE 14.6 Actual and estimated tree-ring series of damaged pine stands. (From Juknys, R., Stravinskiene, V., and Vencloviene, J., Tree-ring analysis for the assessment of anthropogenic changes and trends, *Environ. Monit. Assess.*, 77, 81–97, 2002. With permission from Kluwer Academic Publishers.)

50 years, the closeness of these relations usually tends to decrease. As mentioned by Schweingruber (1987), the reaction of trees to external factors is changing in the process of the aging of trees and possibly it is the main reason for this phenomenon. The conclusion was that 40- to 60-year-long tree-ring series are most suitable for the calibration of climate response models used to predict normal tree growth for the forest decline studies. Usually included in multiple climate response models are three to six climatic factors, and 35 to 65% of tree-ring series variance is attributed to the impact of included climatic factors (Juknys and Jancys, 1998).

Examples of actual and estimated data of tree ring-series for damaged pine stands, situated at different distances from pollution sources, are presented in Figure 14.6 (Juknys et al., 2002). As seen, since 1965 when the Achema Plant started production, emitted nitrogen compounds had a positive impact and annual increment increased, as compared with the normal (estimated) growth. After several years the general impact of increased emissions became negative and growth depression started. An especially fast decline of annual increment has been noticed since 1979, when the most polluting Nitrophoska Division opened. The exceptionally cold winter of 1978–1979 was an additionally strong stressor for weak trees.

The recovery of damaged stands started in the mid-1980s as a consequence of essentially reduced environmental pollution (Figure 14.1). An especially apparent recovery process began in 1989, when an extremely serious accident took place in the Nitrophoska Division, and production in this toxic facility was completely stopped.

As seen in Figure 14.6, the duration of fertilization and depression periods and their severity is rather different for pine stands situated at different distances from the pollution source. The cases of statistically significant ($p < 0.05$) anthropogenic changes of tree-ring width are marked by vertical dashes in this figure.

The period of fertilization ranged from 5 years for the most distant Scots pine stands up to 8 years for the ones closest to the plant. The duration of depression period differs from 10 years for the closest and the most damaged stands to 4 years for stands situated at a distance of 20 to 21 km from Achema.

The acceleration of growth constituted 20 to 30%, compared with predicted normal growth during the period of fertilization, and the increase was rather similar at different distances from the polluting source. Presumption can be made that not only acceleration but partial depression of tree growth took place from the very beginning because of higher concentration of pollutants closer to the polluting source.

Decrease of tree-ring width was rather different at various distances from Achema during the depression period (Figure 14.6). Average losses of radial increment comprised 40 to 45% for the closest pine stands, and 20 to 25% for the most distant ones. The intensity of recovery is rather different for differently damaged pine stands. Stands with less depression of tree growth reached normal levels of annual increment already by the mid-1980s. In the closest to pollution source and the most damaged pine stands, recovery of annual increment to the normal level continued up to the very end of the 1980s; however, at the end of the 1990s, the most obvious increase of radial increment and even excess of the predicted normal tree-ring width is observed, particularly in the pine stands damaged most considerably (Figure 14.6).

Some additional depression of growth can be seen for the most distant pine stands during the last few years. Redistribution of air pollution after the reconstruction of Achema could be a cause of this phenomenon. According to the air monitoring data, maximal concentrations of sulfur dioxide have been registered at a distance of 15 km from the pollution source during the last several years.

14.5 ANTHROPOGENIC CHANGES IN TREE-RING SERIES VARIANCE

The investigated tree-ring series were split into two parts and analyzed separately for the investigation of anthropogenic changes in tree-ring series variance in the polluted environment (Juknys et al., 2002). The first part of the tree-ring series was considered as conditionally natural and included the beginning of a time series until 1965, when Achema started and considerable air pollution began. The second part of the time series (from 1965 to the end) was considered transformed by environment pollution impact. Two approximately equal (about 40 years long) parts of the time series were obtained, and the statistical significance of differences between variance of these two parts of the tree-ring index series was assessed according to F criteria.

Tree-ring series $\{z_t\}$ usually are rather strongly autocorrelated. Analysis of partial autocorrelation functions of the standardized time series (Juknys et al., 2002) has shown that it is possible to consider them as autoregression processes (ARIMA [1,0,0]) of the first rank:

$$z_t - m = p(z_{t-1} - m) = e_t \tag{14.6}$$

where m is the mean of index series, and $\{e_t\}$ the series of uncorrelated random values with zero mean and identical variance.

The investigation of different authors has indicated some decreases of tree-ring series variances with the increase of tree age (Fritts, 1976; Briffa et al., 1996; Schweingruber, 1996; Lovelius, 1997). In order to evaluate impact of age, tree-ring series of control stands were analyzed in the same manner as damaged ones (Table 14.1).

As seen from the presented results (Table 14.1), the variance of the standardized tree-ring series increased considerably in the polluted environment, and only one of most-distant-from-pollution-source stands (No. 7) makes an exception. Increase in variance of the tree-ring series in this stand is statistically insignificant ($p < 0.05$). Differences of variance for the same periods in the control stands are relatively slight and statistically insignificant (Table 14.1).

The conclusion was made that the variance of a tree-ring series increases considerably in the polluted environment, and more intensive fluctuations of destabilized systems continue despite very essential reduction of anthropogenic impact (Juknys et al., 2002).

TABLE 14.1
Comparison of Tree-Ring Series Variance Prior to and after Construction of Fertilizer Plant

State of Stand	No. of Stand	L, km	Tree-Ring Series Variance: Before 1965	Tree-Ring Series Variance: After 1965	F-Test for Homogeneity of Variances	p of F-Test
Damaged stands	1	2.6	259.0	1335.5	5.16	0.000001
	2	3.2	101.6	248.4	2.45	0.008315
	3	6.0	190.8	475.7	2.49	0.010805
	4	6.5	116.0	381.6	3.29	0.001308
	5	11.4	147.5	713.1	4.83	0.000024
	6	10.9	77.5	435.7	5.63	0.000004
	7	20.6	171.2	244.2	1.43	0.315450
	8	21.3	248.8	502.9	2.02	0.045988
Control stands	1		167.8	170.4	1.22	0.541651
	2		154.2	188.9	1.02	0.955723
	3		267.5	167.4	1.60	0.173078

Source: From Juknys, R., Stravinskiene, V., and Vencloviene, J., Tree-ring analysis for the assessment of anthropogenic changes and trends, *Environ. Monit. Assess.*, 77, 81–97, 2002. With permission from Kluwer Academic Publishers.

14.6 ANTHROPOGENIC CHANGES IN TREE GROWTH RELATIONS WITH CLIMATIC FACTORS

Our investigation has shown that not only the intensity of growth (tree-ring width) and the intensity of temporal fluctuations of tree-ring series are changed, but relations of tree growth with different climatic factors are transformed under the polluted environment as well (Juknys et al., 2002). Data on increased sensitivity of trees to the impact of unfavorable climatic conditions in the polluted environment were reported already at the very beginning of the 1990s (LeBlank and Raynal, 1990; Cook and Zedaker, 1992).

Analysis of tree-ring width relations with climate factors for the period before Achema started production and for the period of essential environmental pollution was made in a way similar to the investigation of the anthropogenic changes of tree-ring series variance. Tree-ring series were split into two parts (before 1965 and after 1965) and their relations with different climate factors (average monthly temperatures and monthly precipitation) were analyzed.

Taking into account that dependence of tree-ring width from different climatic factors are rather weak and the coefficient of correlation usually does not exceed 0.3 to 0.4 (Bitvinskas, 1989; Juknys and Vencloviene, 1998), interpretation of their anthropogenic changes is rather complicated.

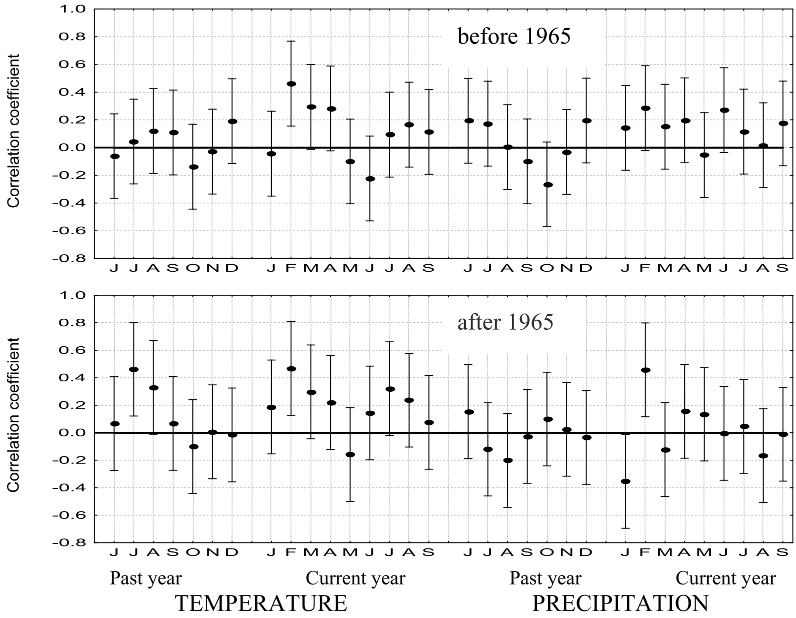


FIGURE 14.7 Impact of environmental pollution to tree-ring relations with climatic factors.

Our investigation (Juknys et al., 2002) has shown that the increase of tree-ring dependence on climate under a polluted environment is rather typical for all investigated tree-ring series of damaged pine stands. A most essential and in some cases statistically significant increase was found for the tree-ring correlation with temperatures of the summer months (July, August). Rather evident is the increase of tree-ring width correlation with past year's climatic conditions — summer and winter temperatures (Figure 14.7).

The conclusion was made that reaction of trees to the impact of climatic factors (temperature, precipitation) has changed under polluted environments and the sensitivity of trees has increased (Juknys et al., 2002).

14.7 CONCLUSIONS

Tree rings can be considered as natural monitors and are able to record information on the impact of natural and anthropogenic stressors. Nevertheless, deciphering this information is a rather complicated task. Retrospective tree-ring analysis provides very useful long-term information and can serve as an appropriate tool for investigation and assessment of the general consequences of anthropogenic environmental changes.

According to the approach taken in evaluating normal growth as a basis for the assessment of anthropogenic changes in annual radial increment, different methods for the assessment of anthropogenic changes in radial increment (tree-ring width) can be classified into the three main groups of methods of control for populations (stands), for individuals (trees), and for periods.

Methods of control for populations are based on the comparison of tree increment or tree increment indexes of damaged stands with relatively healthy ones, with those growing in areas remote from the pollution source in similar site conditions, and with trees having approximately the same dendrometric characteristics. An increment of these control stands is considered as a norm to compare with damaged stands. The accuracy of assessment in this case is rather subjective and depends essentially on the choice of control stands. Methods of this group can be used only in the case of local pollution with an apparent pollution gradient.

Methods of control for individuals are mainly based on the presumption that impact of crown defoliation on tree increment is statistically significant. Increment of relatively healthy individuals (trees) without obvious signs of crown defoliation is considered as normal in this case, and is compared with the increment of defoliation to a different extent on trees inside the same stand. The main weaknesses of this method are that the correlation of crown defoliation with tree increment is rather weak and usually does not exceed 0.3 to 0.5. Also, visual estimations of the main predictor, crown defoliation, are rather subjective and accuracy of data is rather low.

Methods of control for periods are based on quantitative analysis of tree-ring series. Annual increment of trees during the period before the essential environmental impact (pollution) has started is considered as a norm and prediction of normal growth for the period of injury is made on the basis of autoregressive or climate response models. Autoregressive models are based on the quantitative analysis of internal regularities of quasi-periodical fluctuations of tree-ring series. Climate response models are based on multidimensional regression analysis of tree-ring width depending on different climatic factors.

Autoregressive models are very powerful tools for the analysis and prediction of tree growth when relatively long-term (100 years and longer) tree-ring series are used. However, in the case of comparatively short (60 to 70 years) time series, the prediction capacity of these models is rather low, and in the case of a prediction of a 30-year period, prediction error consists of 20 to 30%.

Multiple climate response models are one of the most appropriate tools for the assessment of anthropogenic changes in tree-ring width. Resolving the capacity of climate response models is essentially higher than that for autoregressive models. In the case of 10-year predictions, standard error usually does not exceed 10%, and in the case of a longer (20 to 30 years) period, the prediction error consists of 12 to 15%.

In the case of investigated anthropogenic changes in pine increment in the surroundings of Achema, three different periods of anthropogenic transformations of tree growth were singled out: the fertilization period, the depression period, and the recovery period. When Achema started production of mineral fertilizers, emitted nitrogen compounds had a positive impact on tree growth; annual tree increment increased as compared with normal growth, predicted by climate response models. However, approximately 10 years later, the general impact of increased air pollution became negative and a growth depression period started. Recovery of damaged stands started in the mid-1980s in response to an essential reduction of environmental pollution.

The most intensive recovery of survived trees — even excess of the predicted normal tree-ring width — was recorded in the pine stands damaged most considerably.

The variance of tree-ring series increases considerably in the polluted environment, and more intensive fluctuations of destabilized systems continue despite a very essential reduction of anthropogenic impact.

In line with changes in intensity of growth (tree-ring width) and the intensity of temporal fluctuations of the tree-ring series, the relationship of tree growth with climatic factors is transformed in the polluted environment as the sensitivity of trees increases under the impact of unfavorable natural factors in the polluted environment.

14.8 ACKNOWLEDGMENTS

The author is grateful to the Lithuanian State Science and Studies Foundation, which has funded this research.

REFERENCES

Armolaitis, K., Nitrogen pollution on the local scale in Lithuania: vitality of forest ecosystems, *Environ. Pollut.*, 102, 55–60, 1998.

Bauer, F., Kommt es Forstlich zur Katastrophe? *Allgemeine Forstzeitschrift*, 37, 865–867, 1982.

Becker, M., The role of climate on present and past vitality of Silver fir forests in N.E. France, *Can. J. For. Res.*, 19, 1110–1117, 1989.

Bitvinskas, T., Prognosis of tree growth by cycles of solar activity, in *Methods of Dendrochronology. Applications in the Environmental Sciences*, Cook, E. and Kairiükštis, L., Eds., Kluweru Academic, Dordrecht, Netherlands, 1989, pp. 332–337.

Braker, O.U. and Gaggen, S.Z. Tree-ring analysis in the Swiss Forest Decline Study of 1984, in *Forest Decline and Reproduction: Regional and Global Consequences*, Kairiükštis, L., Nilsson, S., and Straszak, A., Eds., International Institute for Applied Systems Analysis, Laxenburg, Austria, 1987, pp. 124–129.

Briffa, K.R. et al., Tree-ring variables as proxy-climate indicators: problems with low frequency signals, in *Proceedings of the NATO Advanced Research Workshop: Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, 1996, pp. 3–7.

Cook, E.R. The use of climatic response models of tree rings in the analysis and prediction of forest decline, in *Proceedings of the Task Force Meeting on Methodology of Dendrochronology, Methods of Dendrochronology-1*, Kairiükštis, L., Bednarz, Z., and Feliksic, E., Eds., Krakow, Poland, 1987, pp. 269–276.

Cook, E.R. and Zedaker, S.M., The dendroecology of red spruce decline, in *Ecology and Decline of Red Spruce in the Eastern United States*, Eagar, C. and Adams, M.B., Eds., Springer-Verlag, New York, 1992, pp. 192–231.

De Vries, W., Reinds, G.J., Vel, E.M., and Deelstra, H.D., Intensive monitoring of forest ecosystems in Europe, Technical report UN/ECE, EC, 1997.

Donaubauer, E., Historical perspectives and international concerns about air pollution effects on forests, in *Proceedings of the Symposium on Effects of Air Pollutants on Mediterranean and Temperate Forest Ecosystems*, Riverside, CA, June 1980, pp. 10–12.

Douglass, A.E., *Climatic cycles and tree growth. A Study of the Annual Rings in Trees in Relation to Climate and Solar Activity*, Carnegie Institute, Washington, D.C., 1928, Vol. 1 and Vol. 2.

Draper, N.R. and Smith, H., *Applied Regression Analysis*, John Wiley & Sons, New York, 1986.

Eckstein, D., On the Application of Dendrochronology for the Evaluation of Forest Damage, in *Materials of IUFRO Conference: Inventorying and Monitoring Endangered Forests*, Schmid-Haas, P., Ed., Zurich, Switzerland, 1985, pp. 287–290.

Eckstein, D., Qualitative assessment of past environmental changes, in *Methods of Dendrochronology. Applications in the Environmental Sciences*, Cook, E. and Kairikükštis, L., Eds., Kluwer Academic, Dordrecht, Netherlands, 1989, pp. 220–223.

Frantz, F., Preuchler, T., and Rohle, H., Vitalitatsmerkmale und Zuwachsreactionen erkankter Bergwaldbestande in Bayerischen alpenraum, *Allg. Forstzft.*, 41, 962–964, 1986.

Fritts, H.C., *Tree Rings and Climate*, Academic Press, London, 1976.

Hosker, R.P. and Lindber, S.E., Review: atmospheric deposition and plant assimilation of gases and particles, *Atmos. Environ.*, 16, 889–910, 1982.

Innes, L.J. and Cook, E.R., Tree-ring analysis as an aid to evaluating the effects of pollution on tree growth, *Can. J. For. Res.*, 19, 1174–1189, 1989.

Jones, P.D., Possible future environmental change, in *Methods of Dendrochronology: Applications in the Environmental Sciences*, Cook, E. and Kairikükštis, L., Eds., Kluwer Academic, Dordrecht, Netherlands, 1989, pp. 337–340.

Juknys, R., Dendrochronological data applications at the forest monitoring system, in *Proceedings of the International Conference on Climate and Atmospheric Deposition Studies in Forests*, Solon, J., Roo-Zielinska, E., and Bytnerowicz, A., Eds., Warsaw, Poland, 1994, pp. 245–254.

Juknys, R., Trends of Lithuanian environment during transitional period, *Environ. Res. Eng. Manage.*, 1, 15–24, 1995.

Juknys, R. and Jancys, E., Dendrochronology for environmental impact assessment, in *Proceedings of the International Conference on Dendrochronology and Environmental Trends*, Stravinskiene V. and Juknys, R., Eds., Kaunas, Lithuania, 1998, pp. 169–177.

Juknys, R. and Venclovienе, J., Quantitative analysis of tree-ring series, *Proceedings of the International Conference on Dendrochronology and Environmental Trends*, Stravinskiene, V. and Juknys, R., Eds., Kaunas, Lithuania, 1998, pp. 237–249.

Juknys, R., Stravinskiene, V., and Vencloviene, J., Tree-ring analysis for the assessment of athropogenic changes and trends, *Environ. Monit. Assess*., 77, 81–97, 2002.

Kairiukstis, L. and Dubinskaite, J., Modeling fluctuations of dedrochronological indices to predict eco-climatic background variability, in *Proceedings of the Task Force Meeting on Methodology of Dendrochronology, Methods of Dendrochronology-1*, Kairikükštis, L., Bednarz, Z., and Feliksic, E., Eds., Krakow, Poland, 1987, pp. 143–162.

Kairiukstis, L., Grigaliunas, J., Skuodiene, L., and Stravinskiene, V., Physiological and dendrochronological indications of forest decline and their application for monitoring, in *Forest Decline and Reproduction: Regional and Global Consequences*, Kairikükštis, L., Nilsson, S., and Straszak, A., Eds., International Institute for Applied Systems Analysis, Laxenburg, Austria, 1987, pp. 151–169.

Kairiukstis, L. and Stravinskiene, V., Application of dendrochronology in regional monitoring of forest decline, *Tree Rings and Environment*, Lund University Department of Quaternary Geology, Lund, Sweden, 34, 1992, pp. 159–161.

Klap, J. et al., Relationships between crown condition and stress factors, in *Ten Years of Monitoring Forest Condition in Europe. Studies on Temporal Development, Spatial Distribution and Impacts of Natural and Antropogenic Stress Factors*, ICP, Brusssels, Geneva, 1997, pp. 277–307.

Knabe, W., Das Waldsterben aus Immisionsokologischer Sicht, *Geogr. Rundsch.*, 35, 249–256, 1981.

Kohler, H. and Stratmann, H., Wachstumund Benadelung von Fichten in Westharz, *Forst. Holzwirt*. 41, 152–157, 1986.

Kontic, R. and Winkler-Seifert, A., Comparative studies on the annual ring patterns and crown conditions of conifers, in *Forest Decline and Reproduction: Regional and Global Consequences*, Kairiukstis, L., Nilsson, S., and Straszak, A., Eds., International Institute for Applied Systems Analysis, Laxenburg, Austria, 1987, pp. 143–152.

Kramer, H., Relation between crown parameters and volume increment of Picea abies stands demaged by enviromental pollution, *Scand. J. For. Res.* 1, 251–263, 1986.

LeBlanc D.C. and Raynald, J., Red spruce decline on Whiteface Mountain, *Can. J. For. Res.*, 20, 1415–1421, 1990.

Lovelius, N.V., *Dendroindication of Natural Processes and Anthropogenic Influences*, St. Petersburg, Pensoft, Moscow, 1997.

McLaughlin, B., Shorlte, W.C., and Smith, K.T., Dendroecological applications in air pollution and environmental chemistry: research needs, *Dendrochronologia*, 20, 133–157, 2002.

Mehne-Jakobs, B.M., Untersuchungen zur Überprufung der Epidemiehzpothese als Erklärungsansatz zu den 'neuartiger' Waldschäden, *Allg. Forst Jagdg.*, 161, 231–239, 1990.

Ozolincius, R., *Conifers: Morphogenesis and Monitoring*. Aesti Publishing, Kaunas, Lithuania, 1996 (in Russian; summary in English).

Ozolincius, R. and Stakenas, V., *Monitoring of Forest Ecosystems in Lithuania*, Ozolincius, R., Ed., Lutut Publishing, Kaunas, Lithuania, 1999, pp. 82–106.

Petras, R., Nociar, V., and Pajtik, J., Changes in increment of spruce damaged by air pollution, *Lesnictvi*, 39, 116–122, 1993.

Philips, S.O., Skelly, J.M., and Burkhart, H.E., White pine growth retardations by fluctuating air pollution levels: interaction of rainfall, age and symptom expression, *Phytopathology*, 67, 721–725, 1977.

Pollanschutz, J., Effects of air pollutants on forest growth, in *Forest Decline and Reproduction: Regional and Global Consequences*, Kairiukstis, L., Nilsson, S., and Straszak, A., Eds., International Institute for Applied Systems Analysis, Laxenburg, Austria, 1987, pp. 125–129.

Schweingruber, F.H., Abrupt changes in growth reflected in tree ring sequences as an expression of biotic and abiotic influences, in *Materials of IUFRO Conference Inventorying and Monitoring Endangered Forests*, Schmid-Haas, P., Ed., International Union of Forestry Research Organizations, Zurich, Switzerland, 1985, pp. 291–295.

Schweingruber, F.H., Potentials and limitations of dendrochronology in pollution research, in *Proceedings of the International Symposium on Ecological Aspects of Tree-Ring Analysis*, G.C. Jacoby and J.W. Hornbeck, Compilers, National Technical Information Service, U.S. Dept. of Commerce, Springfield, VA, 1987, pp. 721–725.

Schweingruber, F.H., *Tree Rings and Environment Dendroecology*, Paul Haupt Publishers, Berne, Switzerland, 1996.

Shiyatov, S.G., A dendrochronological approach to forecasting, in *Proceedings of the Task Force Meeting on Methodology of Dendrochronology: Methods of Dendrochronology-1*, Kairikükštis, L., Bednarz, Z., and Feliksic, E., Eds., Krakow, Poland, 1987, pp. 137–142.

Soderberg, U., The relation between increment and defoliation for Scots pine and Norway spruce in Sweden, in *Proceedings of IUFRO Workshop on Monitoring Air Pollution Impact on Permanent Sample Plots, Data Processing and Results Interpretation*, Prahatice, CSFR, 1991, pp. 119–127.

Sopauskiene, D. and Jasineviciene, D., Time trends in concentration of acidic species in precipitation in Lithuania 1981–1995, *Atmospher. Physics,* 1, 35–40, 1997.

Stravinskiene, V., *Dendrochronological Indication of Climatic Factors and Anthropogenic Environmental Trends*, Lutute Publishing, Kaunas, Lithuania, 2002 (in Lithuanian; summary in English).

UNECE, Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests, Programme Coordinating Centers, Hamburg and Prague, 1994.

Watmough, S.A., An evaluation of the use of dendro-chemical analyses in environmental monitoring, *Environ. Rev.,* 3, 181–201, 1998.

15 Uranium, Thorium, and Potassium in Soils along the Shore of Lake Issyk-Kyol in the Kyrghyz Republic*

D.M. Hamby and A.K. Tynybekov

CONTENTS

15.1 INTRODUCTION

Lake Issyk-Kyol is situated in the northeast region of Kyrghyzstan, one of the independent republics of the former Soviet Union and bordered by China to the south and east, Kazakstan to the north, and Uzbekistan and Tajikistan to the west (Figure 15.1). The lake is one of the largest in Central Asia, having a surface area of 6240 km^2 and a depth of 668 m. It lies in the valley between the Terskei mountains to the north and the Kungei mountains to the south, at a surface elevation of 1550 m (CAGC, 1987). The briny lake, fed by mountain runoff which flows through about 80 small rivers and creeks, has no discharge streams. Lake Issyk-Kyol is used for swimming, boating, and fishing, but because of its salt content, it is not a direct source of drinking water. During the Soviet era, hotels along the central northern shore of Lake Issyk-Kyol were well-known vacation spots for the Soviet elite, but

* From Hamby, D. 2002. *Environmental Monitoring and* Assessment, 73(2): 101–108. Reprinted with permission by Kluwer Publ.

FIGURE 15.1 Location in Central Asia of the Kyrghyz Republic and Lake Issyk-Kyol. (From Hamby, D.M. and Tynbekov, A.K. 2002. *Environ Monitor. Assess.*, 73: 101–108. With permission.)

otherwise it remained in virtual isolation from the outside world until the early 1990s. Today, the lake attracts tourists from all over Central Asia.

A government commission in Kyrghyzstan was established from 1994 to 1996 to assess the radiation situation at Lake Issyk-Kyol and determine whether radioactivity in areas with elevated radiation levels was natural or man-made. An investigation of radiation levels along the shoreline of Lake Issyk-Kyol was conducted previously by Kyrghyz scientists (SCSC, 1990) and more recently by Hamby and Tynybekov (1999). The SCSC (1990) survey included about 400 measurements and our survey consisted of over 2200 measurements taken along the perimeter of the lake. The most recent measurements indicate that sampling locations near Genish, Kadji-Sai, Bokonbaevo, and Cholpon-Ata have radiation exposure rates in excess of ten times ambient levels (Hamby and Tynybekov, 1999). To corroborate earlier data and to determine the source of the increased radiation fields, a radiological assessment of the shoreline of Lake Issyk-Kyol was executed, including analyses of nearly 300 soil samples.

We have measured concentrations of thorium, uranium, and potassium in the shoreline soils. Each of these naturally occurring elements has isotopes that are radioactive and may increase the amount of exposure received by the populations living in the vicinity of the lake. These exposures can result in individuals receiving radiation dose in the form of external gamma radiation or internal alpha and beta emissions. Additionally, radon is a decay progeny of thorium and uranium and may result in increased radiation dose via the inhalation exposure pathway. The following study reports on the results of our soil analyses at Lake Issyk-Kyol.

15.2 METHODS

15.2.1 Sample Collection

In early 1999, several hundred soil samples were obtained from 99 locations around the shoreline of Lake Issyk-Kyol. The selection of these sampling locations was driven by results from previous assessments of external exposure rates in the region

(Hamby and Tynybekov, 1999), so as to include representative areas of both high and low gamma exposure. Samples were collected at locations near the mouths of streams emptying into Lake Issyk-Kyol, along the shoreline of the lake, and at specific locations with elevated radiation levels. Precise positional data were recorded using a portable GPS receiver.

Soil samples were collected by first recording the location and relative exposure rate at 1 m from the undisturbed surface directly over the area to be sampled. An area of 30 × 30 cm was marked and cleared of debris. Three 30 to 40 g samples (wet weight) of surface soil to a depth of 1 to 2 cm were collected at random within the marked 900 cm^2 area. The three samples were then combined into one, sifted, mixed thoroughly, and dried for 4 h in a 100°C oven. Water fractions averaged 8.9%, ranging from less than 1% to as much as 37%. Dry weights of combined samples were consequently 83.6 ± 12.1 g. Dried samples were sealed in 250 ml polyurethane bottles and set aside for a minimum of 30 d to allow the in-growth of uranium and thorium decay products (Myrick et al., 1983; Murith et al., 1986).

15.2.2 Sample Counting

Prepared soil samples in radiological equilibrium were counted in their sealed bottles on a high-purity germanium (HPGe) detector with 70% efficiency, relative to a 3 × 3" NaI. Following a 30-min counting time, count rates were recorded for five gamma energies: 0.239 MeV (^{212}Pb, with a 44.6% gamma yield); 0.352 MeV (^{214}Pb, 37.1%); 0.609 MeV (^{214}Bi, 46.1%); 0.911 MeV (^{228}Ac, 27.7%); and 1.461 MeV (^{40}K, 10.7%). Concentrations of ^{232}Th were determined from the average concentrations of ^{212}Pb and ^{228}Ac in the samples, and ^{238}U was determined from the average of the ^{214}Pb and ^{214}Bi concentrations. Radiological concentrations of ^{232}Th, ^{238}U, and ^{40}K were then converted to total elemental concentrations of thorium, uranium, and potassium in surface soils, as described in the following text. Total thorium and uranium concentrations are reported in units of ppm, while concentrations of potassium are reported in units of percent.

15.2.3 Calculation of Elemental Concentrations

Radiological concentrations in soils collected from the Issyk-Kyol shoreline are determined from measurements of the gamma rays emitted by specific radionuclides in the decay of uranium, thorium, and potassium. These concentrations, while specific only to particular radioisotopes, are used to estimate elemental concentrations in the soil samples. Since the decay progeny of ^{232}Th and ^{238}U are measured, we must rely on the establishment of secular equilibrium in the samples in order to provide an accurate measurement of total thorium and uranium, hence the 30-d in-growth time. A true measure of potassium is taking place since we are measuring ^{40}K directly.

Elemental concentrations are calculated from measured radiological concentrations in the soil samples. First, the radiological concentration of nuclide i, $C_{S,i,}$ in units of Bq per gram of soil, is calculated using

$$C_{S,i} = \frac{\dot{C}}{Y_i \cdot \varepsilon_i \cdot M_x}$$

where $\dot{C}_i$ is the measured count rate (cts/sec), Y_i is the yield of gamma rays per disintegration, ε_i is the efficiency (cts/gamma) of the detector at the energy of the nuclide i gamma ray, and M_x is the dry mass of the soil sample being analyzed. The fraction of the element in the soil sample, F_E, in units of percent or ppm, is then calculated by

$$F_{E,j} = \frac{C_{S,i} \cdot M_{A,j}}{\lambda_i \cdot N_A \cdot f_{A,i}} \cdot K$$

where $M_{A,j}$ is the atomic mass (g/mol) of element j; λ_i is the decay constant (s^{-1}) of the radioisotope being counted; N_A is Avogadro's number (6.022×10^{23} atom/mol); $f_{A,i}$ is the fractional atomic abundance of ^{232}Th, ^{238}U, or ^{40}K; and the constant, K (with a value of 100 or 1,000,000), converts the ratio of the element's mass to soil mass into a percentage or ppm.

15.3 RESULTS AND CONCLUSIONS

Concentrations of total thorium, uranium, and potassium are plotted in Figure 15.2 for our 99 sampling locations around the perimeter of Lake Issyk-Kyol. Measured concentrations over all sampling sites are 53 ± 110 ppm, 21 ± 64 ppm, and 5.7 ± 1.3% for thorium, uranium, and potassium, respectively. For comparison, Myrick et al. (1983)

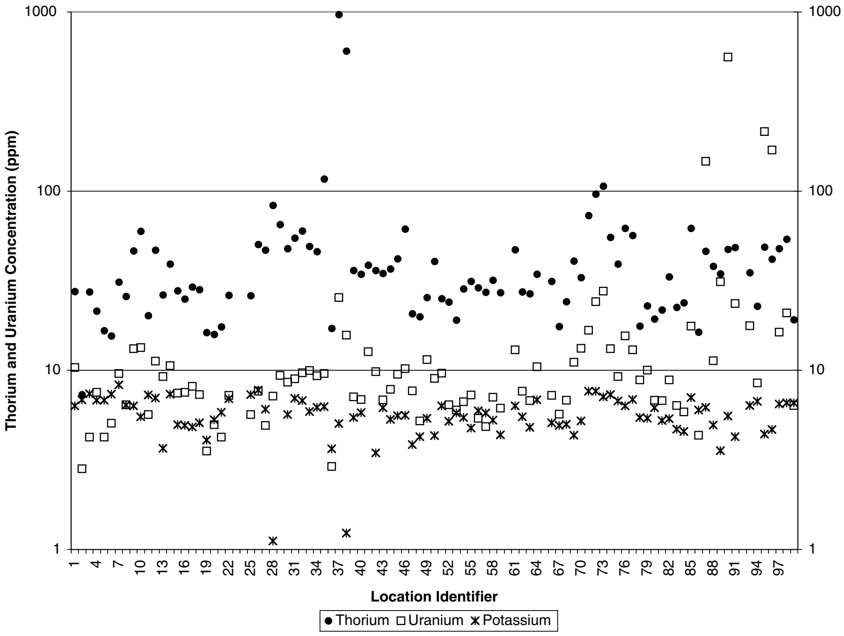


FIGURE 15.2 Thorium, uranium, and potassium concentrations in soils on the shore of Lake Issyk-Kyol. (From Hamby, D.M., and Tynbekov, A.K. 2002. *Environ Monitor. Assess.*, 73: 101–108. With permission.)

have determined arithmetic mean concentrations and standard deviations of thorium and uranium in surface soils in more than 300 samples obtained from locations around the U.S. to be 8.9 ± 4.2 ppm and 3.0 ± 2.5 ppm, respectively. Also, Chang et al. (1974) report the concentrations of thorium, uranium, and potassium in earthen building materials of Taiwan to range from 14 to 16 ppm, 1.2 to 4.3 ppm, and 0.15 to 12.8%, respectively. Potassium concentrations in a wide variety of rock types are estimated to range from approximately 0.1 to 3.5% (Kohman and Saito, 1954).

For thorium at Lake Issyk-Kyol, if the two high concentrations at locations 37 and 38 (Figure 15.2) are removed from the analysis, the concentration over the remaining soil samples is 37 ± 20 ppm, about a factor of two-to-four greater than the averages of Myrick et al. (1983) and Chang et al. (1974). Likewise, removing the four high concentrations at locations 87, 90, 95, and 96, the concentration of uranium is 10 ± 5 ppm, a factor of about three greater.

An analysis of concentrations of potassium in Lake Issyk-Kyol shoreline soils shows less variability among samples, with two comparatively low values being recorded for locations 28 and 38. If these values are removed from the analysis, the concentration of potassium in the Issyk-Kyol shoreline is 5.8 ± 1.1%, in the range of the data of Chang et al. (1974), but about 65% higher than Kohman and Saito's (1954) high value.

Several representative sampling points plotted relative to the lake's shoreline are shown in Figure 15.3. These particular locations are plotted to highlight areas found to have elevated radiation exposure rates (Hamby and Tynybekov, 1999) and areas

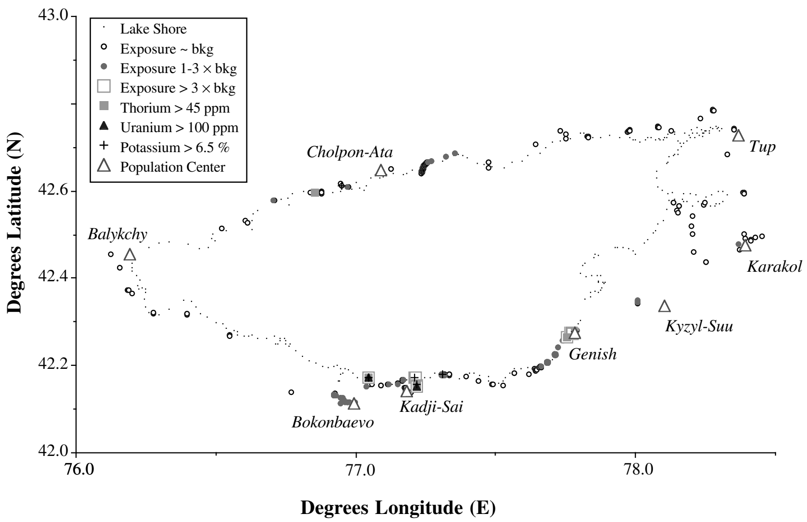


FIGURE 15.3 Relative radiation levels and areas of relatively high thorium, uranium, and potassium concentrations at specific locations on the shoreline of Lake Issyk-Kyol. (From Hamby, D.M., and Tynbekov, A.K. 2002. *Environ Monitor. Assess.*, 73: 101–108. With permission.)

of relatively high uranium, thorium, and potassium concentrations. As expected, locations with high soil concentrations of these radionuclides (locations 37, 38, 87, 90, 95, and 96) are consistently located near areas of the lake previously determined to have high exposure-rate measurements (SCSC, 1990; Karpachov, 1996; Hamby and Tynybekov, 1999).

Measurements by our international team of scientists have confirmed the existence of areas with elevated levels of radiation exposure and high concentrations of naturally occurring radionuclides on the southern shore of Lake Issyk-Kyol. Thorium, uranium, and potassium concentrations in specific areas near the lake are somewhat higher than average concentrations around the world. Visual inspection of Lake Issyk-Kyol's white, sandy beaches near the towns of Bokonbaevo and Kadji-Sai show a distinctive mixture of black sands in very localized areas. Monazite is an insoluble rare-earth mineral that is known to appear with the mineral ilmenite in sands at other locations in the world (Eisenbud, 1987). Monazite contains primarily radionuclides from the ^{232}Th series, and also contains radionuclides in the ^{238}U series. The sands on the Lake Issyk-Kyol beaches very likely contain monazite and ilmenite. These mineral outcroppings are the source of radioactivity along the shoreline of Lake Issyk-Kyol. Historical evidence provides insight into possible other sources of radioactivity in this area of the world; however, our results suggest that shoreline radioactivity is of natural origins.

ACKNOWLEDGMENT

This work was conducted with partial financial support from the U.S. Civilian Research and Development Foundation (Grant No. YB1-121) and the NATO Science Program and Cooperation Partners (Linkage Grant No. 960619).

REFERENCES

Chang, T.Y., Cheng, W.L., and Weng, P.S. 1974. Potassium, uranium and thorium content in building material of Taiwan. *Health Phys.*, 27(4): 385–387.

Chief Administration of Geodesy and Cartography (CAGC). 1987. Atlas of the Kyrghyz Soviet Socialist Republic. Natural conditions and resources. Vol. 1. Moscow (in Russian).

Eisenbud, M. 1987. *Environmental Radioactivity: From Natural, Industrial, and Military Sources*. 3rd ed., Academic Press, New York.

Hamby, D.M. and Tynybekov, A.K. 1999. A screening assessment of external radiation levels on the shore of lake Issyk-Kyol in the Kyrghyz Republic. *Health Phys*., 77(4): 427–430.

Hamby, D.M. and Tynybekov, A.K. 2002. Uranium, thorium, and potassium in soils along the shore of Lake Issyk-Kyol in the Kyrghyz Republic, *Environ. Monitor. Assess*., 73: 101–108.

Karpachov, B.M. 1996. Regional radiological investigations in the Kyrghyz Republic. In: *Proceedings of the 1st Conference on Prospective Planning for Continued Ecological Investigations in the Kyrghyz Republic* (translated), pp. 14–15. Bishkek, Kyrghyzstan (in Russian).

Kohman, T. and Saito, N. 1954. Radioactivity in geology and cosmology. *Annu. Rev. Nucl. Sci.*, 4.
Murith, C., Voelkle, H., and Huber, O. 1986. Radioactivity measurements in the vicinity of Swiss nuclear power plants. *Nucl. Instrum. Methods*, A243: 549–560.
Myrick, T.E., Berven, B.A., and Haywood, F.F. 1983. Determination of concentrations of selected radionuclides in surface soil in the U.S. *Health Phys*., 45(3): 631–642.
SCSC. 1990. Radiation Investigation of Lake Issyk-Kyol. Issyk-Kyol Ecology Branch of the State Committee Scientific Center for the Kyrghyz Republic and the Issyk-Kyol Station of Chemistry Planning and Investigation (SCSC). 1990. Bishkek, Kyrghyzstan (in Russian).

16 Monitoring and Assessment of the Fate and Transport of Contaminants at a Superfund Site

K.T. Valsaraj and W.D. Constant

CONTENTS

16.1 INTRODUCTION

Contamination of soils poses a serious environmental problem in the U.S. The Comprehensive Environmental Response Compensation and Liability Act (CERCLA) of 1980 established the so-called Superfund provisions whereby a trust fund was set up to provide for cleanup of hundreds of abandoned hazardous waste sites. Several of these sites were put on the National Priorities List (NPL) and slated for cleanup. Two such sites are located north of Baton Rouge in Louisiana in the U.S. Environmental Protection Agency (EPA) Region 6 and are called the Petro Processors, Inc. (PPI) sites.

PPI sites comprise two former petrochemical disposal areas situated about 1.5 miles apart near Scotlandville, about 10 miles north of Baton Rouge, the Scenic Highway, and Brooklawn sites, totaling 77 acres. Brooklawn is the larger of the two areas, currently estimated at 60 acres. These sites were operated in the late 1960s and early 1970s to accept petrochemical wastes. During their operation,

1-56670-641-6/04/$0.00+$1.50

TABLE 16.1
Principal Organic and Inorganic Contaminants at the PPI Superfund Site

Volatile Organic Compounds	1,1,2-Trichloroethane
Vinyl chloride	Chlorobenzene
1,1-Dichloroethene	Ethylbenzene
Chloroform	1,1,2,2-Tetrachloroethane
Benzene	1,2-Dichlorobenzene
1,2-Dichloroethane	1,3-Dichlorobenzene
Trichloroethene	1,1,1-Trichloroethane
1,2-Dichloropropane	
Toluene	
Tetrachloroethene	
Semivolatile/Base Neutrals	
Naphthalene	Fluorene
Hexachlorobutadiene	Phenanthrene
Hexachlorobenzene	Anthracene
Diethylphthalate	
Bis-chloroethyl ether	
Chloro-1-propyl ether	
Hexachloroethane	
Isophorone	
1,2,4-Trichlorobenzene	
2,4-Dinitrotoluene	
Metals	
Copper	
Zinc	
Cadmium	
Lead	
Chromium	
Nickel	

approximately 3.2×10^5 tons of refinery and petrochemical wastes were disposed in nonengineered pits on the two sites. Free phase organics are present in buried pits and high permeability soil lenses are found in the proximity of the disposal area. A concerted effort was made in the early 1980s through soil borings and drilling wells to determine the types and nature of contaminants at the site. Table 16.1 lists the major contaminants found at the site. Contaminants at the sites are predominantly chlorinated organic solvents and aromatic hydrocarbons.

A conventional hydraulic containment and recovery system, pump-and-treat (P&T), was initiated in 1989 with a plan for a total of 214 wells at the bigger Brooklawn site. However, this method was shown to require unrealistically long times to make significant reductions in the quantity of organic contaminants.[1] This was primarily attributed to the fact that the removal of hydrophobic organic compounds

(HOCs) from contaminated soils is usually hindered by very low solubility in water. Thus, there was a need for alternative technologies or other methods to enhance the recovery of contaminants. The use of surfactants to enhance the performance of the existing well facilities was suggested, since the P&T wells were already in production, and any additional wells and equipment necessary could easily be incorporated into the system. However, one potential problem that the researchers and the regulators were concerned about was the possibility of downward migration of mobilized contaminants and surfactants into deeper depths of the pristine subsurface soils.

Our research project, which was initiated to support the groundwater geochemical model, MODFLOW®, yielded interesting findings. We observed that a significant fraction of the contaminant was irreversibly bound to the soil. A measure of the desorption-resistant fraction and its bioavailability was not readily available. To fill this knowledge gap, a comprehensive study on the sorption/desorption hysteresis, desorption kinetics, and bioavailability of key contaminants was undertaken. The findings of these studies along with other supporting data eventually resulted in the selection of monitored natural attenuation (MNA) as the current remediation scheme now in place for the two sites. This chapter details the results of our environmental chemodynamic studies at the site. The discussion is selective in that only the chemodynamic data for two of the several contaminants are considered for this chapter as illustrations.

16.2 ASSESSMENT OF CHEMODYNAMIC DATA FOR FIELD SOILS

Fate, transport, and risk assessment models require both equilibrium and kinetic data on desorption of contaminants from soil. Studies suggested a two-stage (biphasic) desorption of organic chemicals from soils and sediments. A rapid release of a loosely bound fraction is followed by the slow release of a tightly bound fraction.[2–4] Quantitative models have been only partly successful in explaining desorption hysteresis, irreversibility, and slowly reversible, nonequilibrium behavior. The bioavailability of a chemical is also controlled by a number of physical–chemical processes such as sorption and desorption, diffusion, and dissolution.[5,6] Several researchers have confirmed that biodegradation can be limited by the slow desorption of organic compounds.[7–10] Long-term persistence in soils of intrinsically biodegradable compounds in field contaminated (aged) soil has also been noted.[2,11]

16.2.1 EQUILIBRIUM DESORPTION FROM SOIL

To test the above hypothesis, we conducted an experiment using field-contaminated soil from the PPI sites.

Though the tests involved several chlorinated organics the discussion here is limited to HCBD as it is one of the most prevalent compounds present at high concentrations throughout the site. The soil was subjected to sequential desorption using distilled water and the aqueous concentration after each desorption step was obtained.[12] The initial concentration in water in equilibrium with the soil was 2120 ± 114 μg/l and declined to about 1200 μg/l after 20 desorption steps (140 d). The mass

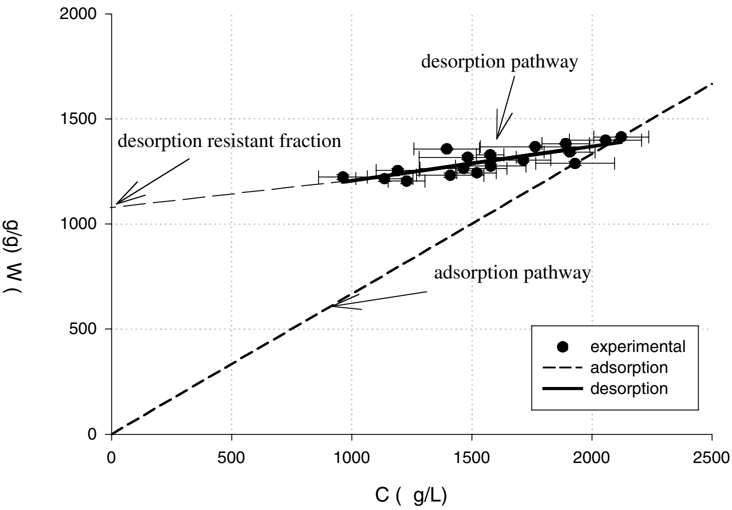


FIGURE 16.1 Equilibrium sequential desorption of hexachlorobutadiene (HCBD) from the site soil. Soil characteristics are given in Reference 12. (From Kommalapati, R.R., Valsaraj, K.T., and Constant, W.D., 2002, Soil-water partitioning and desorption hysteresis of volatile organic compounds from a Louisiana Superfund site soil, *Environ. Monit. Assess.*, 73(3): 289. With permission.)

remaining on the soil after each desorption step was determined through a mass balance and the resulting desorption isotherm relating the aqueous concentration to the equilibrium soil concentration was plotted as shown in Figure 16.1. Each point on the plot represents the mean of quadruplicate samples.

The slope of the straight line (667 l/kg) represents the partition coefficient K_{sw}^{rev} if the partitioning was entirely reversible. It is quite clear that considerable hysteresis in desorption exists. A linear fit of data of the desorption data yields a slope equivalent to a partition coefficient for the loosely bound fraction which we represent as $K_{sw}^{des,1}$ and is 166 l/kg in the present case. Extrapolating back on this slope suggests that approximately 1204 ± 13 μg/g of the HCBD is tightly bound to the soil and may desorb only very slowly (months to years). An attempt was made to obtain the desorption constant for the tightly bound fraction. The derived value thus was 10,458 l/kg. Thus, an estimated ratio of $K_{sw}^{des,2}/K_{sw}^{des,1}$ for HCBD on PPI site soil is 63. This is only meant to illustrate the relative magnitude of the two compartments into which the HCBD partitions within the soil. A number of investigators have also shown that the ratio of the partition coefficients varied from 1 to 33[2] for a variety of other compounds on contaminated soils.

16.2.2 Kinetics of Desorption from Soil

Kinetic studies on desorption were conducted with freshly contaminated soils and aged soils (i.e., soils that had contact time of 3 d to 5 months). Three levels of contamination were used.[13] To illustrate here, data for 1,3-DCB with silty soil from

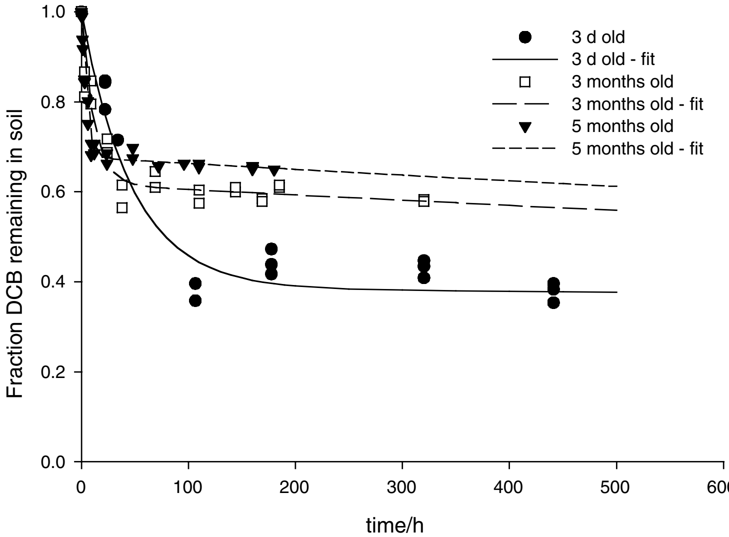


FIGURE 16.2 The rate of desorption of 1,3-dichlorobenzene (DCB) from the site soil. Three soils with varying degrees of contaminant aging in the soils are shown. The lines represent the model fit to the data. Parameters were obtained from Lee S., Kommalapati, R.R., Valsaraj, K.T., Pardue, J.H., and Constant, W.D., 2002, Rate-limited desorption of volatile organic compounds from soils and implications for the remediation of a Louisiana Superfund site, *Environ. Monit. Assess.*, 75(1): 93–111.

the PPI site are plotted with the fraction of the contaminant remaining as a function of time in Figure 16.2. The results for other chemicals are not discussed, as the findings are very similar. A substantial portion of the contaminant is released within the first 20 to 30 h, followed by a very slow release over a very long period. This slow release was observed over the entire duration of the experiment (100 to 450 h).

An empirical model was used to describe the contaminant rate of release (ROR); a relatively rapid release of the chemical followed by a much slower release of the remaining chemical.[14] The nonlinear equation used to describe this biphasic behavior during desorption was given by:

$$1-\frac{S_t}{S_0}=Fe^{-k_1t}+(1-F)e^{-k_2t} \tag{16.1}$$

where t is time, S_t/S_0 is the fraction of chemical released after time t, F is the loosely bound fraction of chemical, and k_1 and k_2 are the first order rate constants describing the desorption of the loosely bound and tightly bound fraction (time^{-1}). The model parameters, F, k_1, and k_2 were determined by fitting the experimental data to the model. The lines in the figures are obtained using Equation (16.1) with the model parameters determined from the nonlinear fit of the composite data. An F value of 0.6 was obtained for the 3-d aged soil. The first order rate coefficients for the loosely (k_1) and tightly (k_2) bound fractions are 0.03 and 6×10^{-4} h^{-1}, respectively. Thus, at least two orders of magnitude difference was found in the values of rate constants for the loosely and tightly bound fractions.

The desorption kinetics data presented in Figure 16.2 also show the effect of the age of contamination for 1,3-DCB. The values of the model parameters, F, k_1, and k_2 are 0.38, 0.09, and 2×10^{-4} for 3-month-old contamination and 0.32, 0.22, and 2×0^{-4} for the 5-month-old contamination. As expected, the loosely bound fraction was reduced from about 0.6 to 0.32 for 1,3-DCB. The longer incubation period in the case of the aged soil allows organic molecules to be sequestered into the soil and exhibit slow desorption kinetics. However, the rate constant, k_1, for the loosely bound fraction was higher for the aged contaminated soil than the freshly contaminated soil. This suggests that the initial release from the aged soil is faster than the initial release from the freshly contaminated soil. Therefore, during adsorption, all the reversible sites on the soil and the organic carbon are first occupied by the contaminant before it starts to be sequestered in the irreversible compartment. For freshly contaminated soil within the short equilibrium time of 3 d, the contaminant should be mainly in the loosely bound sites and diffusion of contaminant into tightly bound sites might still be occurring. Thus, both processes of desorption into water and migration to the tightly bound sites are occurring simultaneously and the overall desorption rate is lowered. However, for aged soils the diffusion of contaminant from the loosely bound sites to tightly bound sites would be nearly complete in the 3- and 5-month periods. Thus, when the desorption process began, the fraction of contaminant in the loosely bound sites, though smaller, desorbs at a faster rate than from the freshly contaminated soil. The aged soil study also further reinforces the need to assess F with k_1 and k_2 when comparing fast and slow release rates, as the rate coefficients also may be misleading if the aging process is not well known.

16.2.3 Bioavailability of the Tightly Bound Fraction in the Soil

A freshly DCB-contaminated soil was used in the first set of experiments. A second set of experiments was conducted using soil which was subjected to sequential desorption to remove the loosely bound fraction; this represented the desorption-resistant fraction. In both cases the soil was amended with nutrients and inoculated with seed cultures. Figure 16.3 shows the percent of 1,3-DCB degraded as a function of incubation time for the freshly contaminated soil and that containing only the desorption-resistant fraction. Again, data points are obtained by averaging the triplicate samples used. The percent degradation presented in the plot accounts for all the losses reported in the control samples which were less than 10%, a reasonable loss for complex systems such as this one.

For the freshly contaminated soil, the percent biodegradation was 23% after the first day of incubation in the microcosm. Thereafter, the biodegradation increased slowly but steadily as indicated by the positive slope to about 55% by the end of a 6-week incubation period. The positive slope of the biodegradation curve indicates that neither oxygen nor nutrients were limiting microbial growth. It is possible then that the degradation of 1,3-DCB will proceed, albeit at a very slow rate, if the experiments

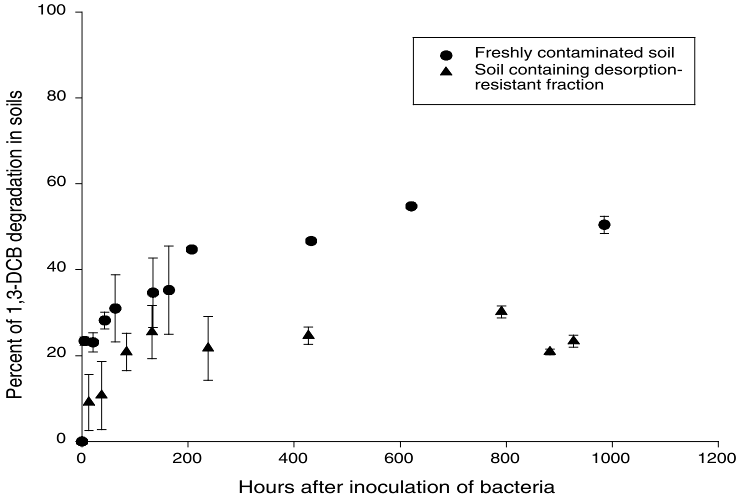


FIGURE 16.3 Biodegradation of 1,3-DCB from a freshly contaminated soil and soil containing only the tightly bound (desorption resistant) fraction. The fraction of DCB degraded was calculated after correcting for the losses in the control vials.

were continued beyond the 6-week period. As we noted earlier, about 60% [F = 0.6 from Equation (16.1)] of the sorbed contaminant 1,3-DCB is reversibly bound to soil. It appears from the figure that bacteria were able to degrade a significant fraction of the readily available fraction of 1,3-DCB.

Figure 16.3 displays percent biodegradation as a function of incubation time for the soil containing only the tightly bound fraction of 1,3-DCB. The biodegradation of 1,3-DCB was monitored with time over a 6-week incubation period. The initial soil concentration was only about 1650 mg/kg soil compared to 18,000 mg/kg soil for freshly contaminated soil experiments. The percent biodegradation was calculated after accounting for the losses from the control samples. The 1,3-DCB degradation was about 22% during the first week with only a total of 33% during the total 6-week incubation period. After the first week, the 1,3-DCB biodegradation rate was very small, as one would expect. Desorption from the tightly bound nonlabile phase is thus limiting the availability of the contaminant for biodegradation. Microorganisms metabolize the substrate present in the aqueous phase rather than direct metabolism from the soil phase. Though there have been some reports that sorbed substrate may be directly available for degradation by attached cells either by direct partitioning to the cell membrane or via degradation by extracellular enzymes,[15] there is still disagreement among the researchers on this point.

Our microcosm batch studies showed that microorganisms readily degraded the easily extractable or loosely bound fraction of the sorbed DCB and that desorption mass transfer is limiting the growth of the microorganisms and thus the rate of biodegradation for the desorption-resistant fraction. Similar arguments are applicable to other compounds including HCBD listed in Table 16.1.

16.3 IMPLICATIONS FOR SITE REMEDIATION

The observed hysteresis in the chemical sorption/desorption process has implications as far as the selection of a remedy for the site is concerned. In order to illustrate this aspect, we used a model to estimate the concentration in groundwater from an aquifer contaminated with a hydrophobic organic compound.[16] The volume of the contaminated zone is divided into n sub-volumes of equal size, and the concentration of water exiting the contaminated zone (withdrawal well) is obtained. The exit concentration of the contaminant C_w in the porewater relative to the initial concentration C_w^0 is then given by

$$\frac{C_w}{C_w^0} = \exp\left(-\frac{nt}{\tau_c}\right)\left[1 + \frac{nt}{\tau_c} + \cdots + \frac{\left[\frac{nt}{\tau_c}\right]^{(n-1)}}{(n-1)!}\right] \tag{16.2}$$

In the above equation t is the time (years) and τ_c is the chemical residence time (years) given by the following equation:

$$\tau_c = [\varepsilon \ + \ K_{sw}\,(1-\varepsilon)\rho_b]\frac{L}{V_x} = R_F \cdot \frac{L}{x} \tag{16.3}$$

For illustrative purposes we choose the most conservative chemical, namely, HCBD (Figure 16.4). For Case I we choose $K_{sw} = K_{sw}^{\mathrm{des},1}$, whereas for Case II we choose $K_{sw} = K_{sw}^{\mathrm{des},2}$. The corresponding retardation factors are R_{F1} = 750 and R_{F2} = 11,250 for the two cases. We choose a 10-m zone length with a soil porosity of 0.25. An n-value of three is chosen in keeping with the fact that the changes in concentrations predicted are only marginally affected for values of n greater than three. L is the length of the zone (10 m) and V_x is the groundwater velocity, which is taken to be approximately 3.3 m/y. For Case I, the initial concentration is given by $\rho_s W^o/R_{F1}$ = 2,120 μg/l as in Figure 16.1. For Case II, the initial aqueous concentration is rescaled to 160 μg/l using the sediment concentration for the irreversible fraction (1043 μg/g) and R_{F2}. Figure 16.4 shows model results of the water concentration at the exit (withdrawal well) of the 10-m zone vs. time for HCBD at the PPI site. Under the assumptions made above, the exit concentration in water from the zone decreased from 2280 μg/l initially to 2178 μg/l after 100 years, 648 μg/l after 500 years and 173 μg/l after 1000 years. Even after 5000 years, the concentration in the water was small, but measurable (17 μg/l), and was determined by the gradual leaching of the tightly bound fraction. Note that the national recommended water quality criterion for HCBD is 0.44 μg/l, which would be reached in about 9600 years. The area under the curve represents the mass removed and in 1000 years was only 17% of the total. However, if the partitioning was entirely reversible, the predicted aqueous concentration decreased to only 1555 μg/l after 1000 years and the mass recovery was 41%. Due to the continuous removal of the reversibly sorbed HCBD, the predicted

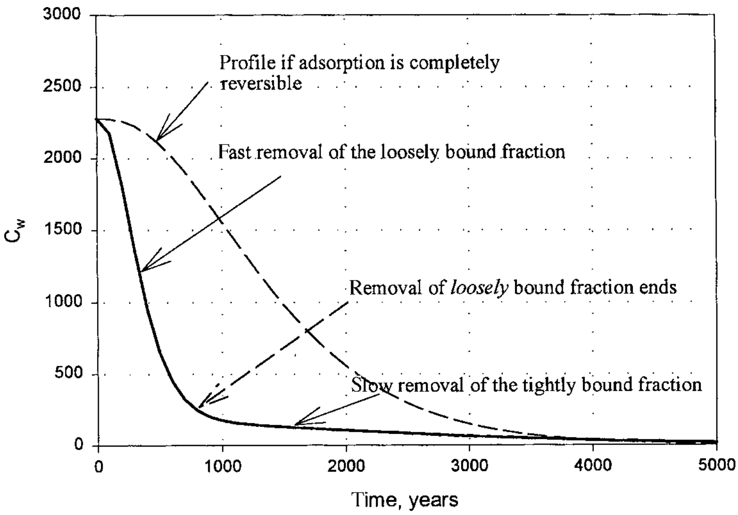


FIGURE 16.4 The predicted concentration of HCBD in the withdrawal well during P&T remediation of a 10-m contaminated zone at the site.

concentration in the aqueous phase after 5000 years was only 6.8 μg/l in this case. As expected, the concentration in the withdrawal well is predicted to be approximately three times as large over the long term when the process is only partly reversible than when the adsorption is completely reversible. Clearly, the total mass recovery of the contaminant (i.e., HCBD) from the aquifer by pump-and-treat is expected to be very small. A large fraction of the mass (67%) is tightly bound to the soil and will take several thousand years to be removed by conventional groundwater-removal technology.

There are several important consequences of the above findings. First, it is clear that satisfactory removal of residual soil-sorbed HCBD (i.e., the fraction left after the free-phase removal) by conventional pump-and-treat is difficult and requires an inordinate and unacceptable time frame. In other words, much of the HCBD is irreversibly bound to the soil. If recovery is the only remedy, other enhanced removal schemes should be considered. Second, since most of the HCBD is bound to the soil and leaches only slowly, and since microbes capable of consuming HCBD are known, monitored natural attenuation or enhanced bioremediation is a better option for site remediation if removal is not required. This would, of course, require a continuous monitoring of the plume so that no offsite migration of the contaminants occurs during the site cleanup, which is required in MNA projects. Third, since HCBD predominantly remains bound to soil particles, its movement offsite with the groundwater is likely not to be significant, and justifies its consideration as a good candidate compound for MNA.

The calculations made above are approximate and neglect many other factors such as the site heterogeneity, biodegradation, and contaminant concentration variations. However, the basic conclusions with respect to the difficulty in extracting HCBD and the benefits it illustrates in MNA should remain the same even if a more complete and detailed model such as MODFLOW is considered.

Experimental results and basic modeling of transport suggest that at most only 60% of the compound is labile and participates in the reversible sorption equilibrium. The half-life for desorption of this fraction is of the order of a few hours (2 to 24 h) for TCE (from Lee et al.[13] and HCBD, and more for DCB. The labile fraction becomes smaller as the contaminant ages within the soil. Our earlier work[12] showed that even when water in contact with the soil is replaced with fresh water at every step, after 24-h of equilibration only a small percentage of the material is recovered from the contaminated soil. Hence, sequential desorption also is incapable of removing the nonlabile fraction from the soil. The projections from the above models based on the 72-h equilibrium data only pertain to the labile fraction of pollutant. Hence, the long-term predictions using the batch K_d values will significantly overpredict movement of contaminants that have been in contact with the soil for decades at the PPI site. On the other hand, we conclude that a significant portion of the contaminant in the aged site soil is inaccessible to water in a conventional pumping scheme and hence remains within the soil to pose no significant threat of migration away from the site. Moreover, the slowly released fraction can probably be managed by the natural assimilative capacity of the soil, given that biodegradation and sorption are common at the site(s). Thus, the desorption resistant concentration may be considered an environmentally acceptable end point (EAE) in the soil that can be managed by MNA procedures currently in practice.

Monitored natural attenuation, as mentioned previously, is currently the remedy of choice for sites at PPI. Implementation of MNA is in part due to results of research summarized in this chapter and is thus an illustration of how monitoring and assessing the fate and transport of contaminants at a site is helpful in choosing a particular remediation choice for the site. If one considers the site history, the remedy has changed over the last decade from very active (excavation) through active (hydraulic containment and recovery or pump-and-treat) to passive methods (MNA). Our work was instrumental in directing the remedy to MNA. First, simple models employed herein, followed by more sophisticated ones (not discussed here), show that it will take many years to obtain remediation of the contamination down to target levels by active (pump-and-treat) technology. The active process may actually work against natural processes with a basis in our findings, i.e., the loosely bound fraction biodegrades at high rates *in-situ* and the tightly bound fraction binds contaminants to the point of rate-limiting degradation. In other words, conditions seem to be favorable for significant contaminant transformation in the groundwater. Continued modeling efforts and long-term monitoring will determine if direction from this work and the remedy are environmentally acceptable over the long term. It is anticipated that if the remedy is found effective in the short term that this highly ranked NPL site will be delisted, with long-term monitoring continuing to ensure protection of human health and the environment.

ACKNOWLEDGMENTS

This work was supported by a grant from the LSU Hazardous Waste Research Center and sponsored by the U.S. District Court, Middle District of Louisiana. We thank the various personnel at NPC Services, Inc., who were instrumental in providing data and collecting and analyzing various samples in this work.

REFERENCES

1. Mackay, D.M. and Cherry, J.A., 1989, Groundwater contamination, pump and treat, *Environ. Sci. Technol.*, 23(6): 630–634.
2. Pignatello, J.J. and Xing, B., 1996, Mechanisms of slow sorption of organic chemicals to natural particles, *Environ. Sci. Technol.*, 30: 1–11.
3. Di Toro, D.M. and Horzempa, L.M., 1982, Reversible and resistant components of PCB adsorption–desorption isotherms, *Environ. Sci. Technol.*, 16: 594–602.
4. Karickoff, S.W. and Morris, K.W., 1985, Impact of tubificid oligochaetes on pollutant transport in bottom sediments, *Environ. Toxicol. Chem.*, 9: 1107–1115.
5. Ogram, A.V., Jessup, R.E., Ou, L.T., and Rao, P.S.C., 1985, Effects of sorption on biological degradation rates of 2,4-dichlorophenoxy acetic acid in soils, *Appl. Environ. Microbiol.*, 49: 582–587.
6. Al-Bashir, B., Hawari, J., Samson, R., and Leduc, R., 1994, Behavior of nitrogen-substituted naphthalenes in flooded soil-part II: effect of bioavailability of biodegradation kinetics, *Water Res.*, 28(8): 1827–1833.
7. Robinson, K.G., Farmer, W.S., and Novak, J.T., 1990, Availability of sorbed toluene in soils for biodegradation by acclimated bacteria, *Water Res.*, 24: 345–350.
8. Steinberg, S.M., Pignatello, J.J., and Sawhney, B.L., 1987, Persistence of 1,2-dibromoethane in soils: Entrapment in intraparticle micropores, *Environ. Sci. Technol.*, 21: 1201–1208.
9. Pignatello, J.J., 1989, Sorption dynamics of organic compounds in soils and sediments, in *Reactions and Movement of Organic Chemicals in Soils*, Sawhney, B.L. and Brown, K., Eds., Soil Science Society of America and American Society of Agronomy, pp. 31–80.
10. Valsaraj, K.T., *Elements of Environmental Engineering*, CRC Press, Boca Raton, FL, 1995.
11. Hatzinger, A.B. and Alexander, M., 1995, Effects of aging of chemicals in soils on their biodegradability and extractability, *Environ. Sci. Technol.*, 29: 537–545.
12. Kommalapati, R.R., Valsaraj, K.T., and Constant, W.D., 2002, Soil–water partitioning and desorption hysteresis of volatile organic compounds from a Louisiana Superfund site soil, *Environ. Monit. Assess.*, 73(3): 275–290.
13. Lee, S., Kommalapati, R.R., Valsaraj, K.T., Pardue, J.H., and Constant, W.D., 2002, Rate-limited desorption of volatile organic compounds from soils and implications for the remediation of a Louisiana Superfund site, *Environ. Monit. Assess.*, 75(1): 93–111.
14. Opdyke, D.R. and Loehr, R.C., 1999, Determination of chemical release rates from soils: experimental design, *Environ. Sci. Technol.*, 33: 1193–1199.
15. Park, J.H., Zhao, X., and Voice, T.C., 2001, Biodegradation of non-desorbable naphthalene in soils, *Environ. Sci. Technol.*, 35(13): 2734–2740.
16. Thibodeaux, L.J., 1996, *Environmental Chemodynamics*, John Wiley & Sons, New York.

17 Statistical Methods for Environmental Monitoring and Assessment

E. Russek-Cohen and M. C. Christman

CONTENTS

17.1 INTRODUCTION

During the last decade, there have been significant advances in statistical methodology for environmental monitoring and assessment. The analysis of environmental data is challenging because data are often collected at multiple locations and multiple time points. Correlation among some, if not all, observations is inevitable, making many of the statistical methods taught in introductory classes inappropriate. In the last decade we have also seen parallel developments in such areas as geographic information systems (GIS) and computer graphics that have enhanced our ability to visualize patterns in data collected in time and space. In turn, these methods have stimulated new statistical research for the analysis of spatially and temporally referenced data.[1,2]

We recognize that writing a chapter that covers all aspects of statistical methodology related to environmental monitoring and assessment would be impossible. So we provide a short overview that points readers to other resources including texts and journals. Then we discuss some of the types of variables one is apt to see in environmental studies and the problems they may pose from a statistical perspective.

1-56670-641-6/04/$0.00+$1.50

We highlight some methodological advances and issues in assessment and provide a similar discussion for environmental monitoring. Finally, we try to note some shortcomings in existing methodology and data collection.

17.2 OVERVIEW

The majority of environmental studies are observational studies rather than controlled experiments. As a result, observational studies are often harder to design and interpret than planned experiments. Careful thought must be given to the study design to avoid multiple interpretations to the same study result because of the potential for confounding variables.[3] Some principles associated with good study design are not unique to environmental data. For example, there are several well-established references on survey methodology.[4,5] There are also texts focused exclusively on environmental studies.[6–8] Both statisticians and environmental scientists have contributed to this body of literature, and one can find methodological papers in journals such as *Ecology*, *Ecological Applications*, and *Environmental Monitoring and Assessment*. In addition, in just this past decade we have seen tremendous growth in journals devoted to statistical methods for the analysis of environmental data including the *Journal of Agricultural, Biological and Environmental Statistics, Environmetrics*, and the journal *Environmental and Ecological Statistics*. Also, recently an encyclopedia devoted to environmental statistics, the *Encyclopedia of Environmetrics*,[9] was published.

In spite of all the sophisticated software tools that exist, some statistical issues continue to plague the scientific community. We feel we would be remiss if we failed to mention some of these. For example, most published research relies on statistical tests of hypotheses. Hurlbert[10] noted the difficulties in testing certain hypotheses in certain types of observational experiments. More recently, McBride,[11] in one of a series of articles defining how scientists view statistics, points to the arbitrariness of null hypotheses, and tests of significance. In classical statistical hypothesis testing, one does not prove a null hypothesis to be true when one fails to reject it. This failure to reject the null could be due to either the null hypothesis being true or too small a sample size with insufficient power to see an effect. Conversely, low power can also result in environmentally relevant impacts being missed because insufficient data are collected. The opposite can also occur. If sample sizes are large enough, small effects may be statistically significant but may or may not have environmental consequences.

McBride[11] argues that interval estimates such as confidence intervals and Bayesian posterior intervals would be more effective. In principle, we agree. However, many users of statistics are not familiar with interval estimates for anything other than the simplest estimators. Simultaneous or multivariate confidence intervals can be difficult to present or interpret,[12] and Bayesian highest posterior density intervals[13] have as yet to gain acceptance by much of the environmental science community.

Bayesian approaches[13,14] involve specification of a prior distribution which is the user's beliefs concerning the likely or appropriate truth about the parameters under study. For example, a user can incorporate the belief that the mean will be between 2 and 4 units. The former is then used in conjunction with the observed data to derive an estimate of a mean that weighs the prior information with the observed data. Many suspect the choice of a prior can be used to manipulate a conclusion. In addition, irrespective of how an interval is calculated, resource managers and environmental decision makers may choose to pick a value in the interval that best suits their purpose and as a result make poor decisions. In resource management it is especially critical that statistical methods be used correctly and with objectivity.

17.3 TYPES OF ENDPOINTS

Different kinds of variables inevitably lead to different types of statistical analyses. In beginning statistics classes we are taught to categorize variables into continuous, discrete, nominal, and ordinal variables. However, many variations exist on these four categories and some variables defy such labels. The methods used will differ depending not only on the type of data but also on assumptions concerning the distribution of our variables. The methods will also differ depending on the questions or hypotheses of interest. Thus if one is trying to predict pesticide residues in the soil as a function of when and where the pesticide has been applied, the analysis may be very different from trying to predict the impact these pesticides will have on the wildlife that reside in the area of application. By far the majority of existing methods assume the variable of interest is continuous, is measured without error, and is Gaussian, i.e., normally distributed. One can see this by looking at the extensive list of available software that exists for regression methods,[15] time series methods,[16,17] for spatially referenced data,[18] and for mixed models often used in shorter term longitudinal studies.[19,20] Unfortunately, many endpoints recorded in environmental studies do not fit this paradigm.

Sometimes the data are quantitative but are not normally distributed and alternative parametric models can be fit to the data. Unfortunately, few of these models have been extended to data that are temporal (i.e., time series data) or spatial in nature as is common in environmental monitoring. A few software procedures exist that accommodate alternative parametric models, e.g., procedures for generalized linear models including GEE extensions for repeated measures[20,21] and procedures such as LIFEREG which is found in SAS.[22] A recent monograph by Kedem and Foikanos[23] suggests a tractable approach for time series data for some discrete and continuous data models.

A common problem in many areas of science is that observations can be censored. For example, when a pesticide level in soils falls below the limit of detection of the assay, we say the observation is left-censored. Left-censoring means we know the value is below a set number but we cannot provide an exact value. On the other hand, when a settling plate is overgrown with algae so an exact enumeration of algae is unobtainable, we have a right-censored value. Right-censored values have received considerably more attention in the statistics community. Statistical methods for survival data or time until failure have been studied extensively by medically oriented statisticians. These methods accommodate patients still alive at the end of a clinical study, so an exact time of death is unknown. In the medical statistics literature there is a heavy emphasis on nonparametric or semiparametric approaches (e.g., the Cox Proportional Hazards Model described in many basic texts[24]). These methods make fewer assumptions than parametric models. Reliability engineers have also been interested in time until a component fails but have developed a number of parametric models to accommodate such data.[25] Left-censored data and doubly censored data (data which can be left- or right-censored) have gotten less attention in the statistics literature, especially when the variable subject to censoring is an explanatory variable rather than a response variable in a model. Only a limited number of models for spatial-referenced data subject to censoring can be found in the statistics literature (e.g., Li and Ryan[26]). El-Shaarawi and Nader[27] discuss some issues associated with censored data in the context of environmental problems whereas Hawkins and Oehlert[28] suggest some simple model formulation and estimation schemes for data subject to censoring, including left-censoring. However, neither paper discusses the complexities of using censored data as response variables in the monitoring setting.

In the U.S., the most common endpoint for monitoring microbial water quality is a measure of fecal coliform values calculated using a most probable number assay.[29] Such assays are based on multiple test tubes at multiple dilution levels. These values are really doubly censored and subject to measurement error (see text below). These issues are often ignored and MPN or log10(MPN) values are frequently used as a response in regression models. The end consequence of ignoring these problems has not been fully explored.

Many of the current parametric models fail to accommodate the number of zeros that are often observed in environmental datasets. For example, insect counts may be zero for half of the traps set but may vary among sites where at least one insect is found. Such variables can be thought of as a mixture of two processes or distributions.[30] Lambert[31] proposed zero-inflated Poisson (ZIP) and zero-inflated negative binomial models for independent observations. Lambert's method has been subsequently generalized to correlated data by Warren[32] and Hall[33] and for time series by Wang.[34] In each of these zero-inflated models, there are two parts to the model: (1) a model that describes the probability of observing a zero event and (2) a model that describes the probability for observing a specified count, given the count is nonzero. These ZIP-type models can incorporate covariates and allow each part of the model to have the same or different covariates. Similar mixture models have been proposed for zeros plus continuous log-normally distributed data.[35] Few of these models have been extended to spatial data or spatial–temporal data context, i.e., allowing for complex correlation structures in the data. Fitting these models using conventional statistics packages is not straightforward.

Many variables are measured inexactly and we often fail to recognize this in the analysis we choose. For example, assays for pollutants in soil or feathers may have errors. Some variables (e.g., number of failed septic tanks in a watershed) may be surveyed infrequently and are therefore an approximation of what is there. When these variables are response variables, this adds noise to the analysis. When variables with measurement error are explanatory variables, the regression coefficients are apt to be biased and the ability to ascertain the importance of these variables in the model is hindered.[36] Spiegelman and colleagues[37] have conducted a series of studies focusing on measurement errors, including substudies, to quantify the degree of measurement error and then incorporating these results into overall study objectives.

While Spiegelman's research has focused on epidemiological applications including those in environmental health, more attention to these issues is needed when analyzing environmental data.

Indices or variables that are composite variables are often seen in large-scale environmental studies. These indices are constructed because they give a value that can be measured over many sites and multiple time points. Examples include the multitude of diversity measures that exist[38] and integrated biota indices or IBIs, such as the fish IBI calculated in the Maryland Biological Stream Survey.[39] These indices may vary with time of year and some caution is needed when comparing sites across time and space. The indices typically vary with the number and type of species present, and they may depend on the relative abundance of each species. Several cautions must be stated with any such index. First of all, these indices rarely fit some nice parametric model. Smith[40] has developed an ANOVA-like approach for diversity measures while almost everyone else transforms the index and applies a method based on the normal distribution. Two locations may have the same value of the index but may appear quite different to the scientist. Solow[41] does an analysis of fish harvested in the Georges Bank. He finds that while diversity has not changed in a substantive fashion over a 10-year period, the composition of the community has changed. Fish with significant commercial value was on the decline over this period while other species rose in relative abundance.

17.4 ASSESSMENT

All sampling for environmental questions involves a type of assessment since the intention is to describe the population under study. What distinguishes assessment from the more general question of characterization of the population is the need for accurate, precise estimators that can assess the impact of a change in the environment and the need to control the costs of committing Type I or II errors during that assessment. For example, the introduction of Bt corn (corn that has been genetically modified to include expression of an endotoxin found in the bacterium *Bacillus thuringiensis*) has generated arguments over whether the use of the corn is destructive to nontarget species such as the monarch butterfly (Lepidoptera: Danainae).[42–44] In this instance committing a type I error has consequences for the environment and committing a type II error is costly to the company in terms of development costs and lost revenue.

In some instances assessment is performed in order to simply characterize the environment that is potentially affected by future changes. In these cases, it is necessary to sample the study region adequately and with sufficient numbers of observations so that most if not all of the variation in the environment is captured. For example, a difficult problem is one of constructing sampling designs to determine the abundance and spatial extent of rare or elusive species. The usual sampling designs such as stratified random sampling, systematic sampling, double sampling, and cluster sampling[4,5] can often miss rare elements of interest and, in fact, depending on the rarity, even result in samples with no rare elements observed. Hence, there have been many alternative designs proffered for characterizing rare populations (for a review see Christman[45]).

There is a vast literature on sampling or monitoring for natural resource estimation.[46–52] Some of the sampling strategies for rare, clustered populations are described.[53–64] They include stratification,[4,5,60] adaptive allocation of samples to strata,[54] adaptive cluster sampling,[61,62] inverse or sequential sampling,[63] and others. In stratification schemes for sampling of rare elements, Kalton and Anderson[60] considered a technique in which the population of interest is first divided into two strata, a small one containing as many of the rare elements as possible and the other stratum containing the remaining population elements. Then the small stratum is disproportionately sampled (relative to its size) in order to obtain accurate estimators of the mean or total. Thompson and Seber[54] describe a method for adaptively allocating samples to different strata based on an initial pilot survey. The allocation of the remaining samples could be based on either the initial estimates of the mean or of the variance in each stratum.

Adaptive cluster sampling[54] is ideally suited for species that are spatially rare, appearing in a few dense clusters. In adaptive cluster sampling, an initial sample is first taken according to some probability-based sampling scheme. Then, if an observation meets some criterion, its 'neighbors' are sampled. The neighbors and definition of neighborhood are fixed prior to sampling and are determined for every element in the population. For example, if the population consists of all small stream watersheds in West Virginia, then a neighborhood for a given watershed, W_i, might be defined to be the set of watersheds that share a boundary with W_i. So, for example, if a rare fern is found in a sampled watershed, an adaptive cluster sampling design would require that the contiguous watersheds also be searched for the rare fern. Should the fern show up in any of those watersheds, their neighbors would also be sampled. As a result, a cluster would be completely sampled. The main disadvantage of this approach is that the final sample size is not controlled. There have been several recommended variations that describe methods for putting at least an upper limit on the total number of observations taken.[55,64]

When sampling is performed in order to characterize the spatial distribution of the variable of interest, the optimal sampling design is often a variant of stratified random sampling or a systematic design.[18] The obvious reason for this approach is to ensure that the study region is adequately spatially covered. Kriging or similar interpolation techniques are then used to create maps showing the distributions of the variables of interest.[65] One main concern about these approaches is that often the scientist ignores the fact that the sample is exactly that: a sample from the population of interest and, as such, has a sampling error associated with it. This sampling error is also valid for the maps as well since a different sample would lead to a different map. As a result, only those interpolation techniques that provide standard errors for the predicted values should be used. These are basically the various kriging methods that are available[18] and include ordinary kriging, universal kriging, indicator kriging, and disjunctive kriging.

What often sets environmental assessment apart from the more general question of environmental condition is that the former is interested in determination of the effect of an impact. In impact studies the researcher is interested in the effect of a particular event on that particular place at that particular time. In environmental research, however, we are interested in the generalization of the effect of an event on an environment (a large population of possible instances) rather than on a particular site. For example, the question of the effect of placing a paper mill on a specific river in Canada, say, is a different research problem from the more general question of the effect of paper mill effluent on benthic species found downstream of mill sites. As a result, the data collection and analyses also differ.

For environmental impact assessment, the standard approach is a variation on the Before/After Control Impact Sampling Design (BACI).[6,66] In these studies, it is known that there will be a future event, such as the building of a pier or offshore dredging or building of a power plant, etc. The effect of the activity is determined by first identifying which variables are likely to indicate an effect, if there is one, and then testing to determine if differences between the mean values at a control site and an impact site change once the impact begins. Hence, samples are taken both before and after at both control and impact locations. Control sites are chosen so as to be free of the influence of the impact yet sufficiently similar to the impact site so as to exhibit the same phenomena of interest. Data are analyzed using analysis of variance (ANOVA) techniques and hence must also meet the assumptions of the ANOVA, including homogeneity of variance, independence, and normality.[16,25]

A common sampling design for a BACI study is as follows. At the control and impact sites, simultaneous samples are taken at several fixed times before and after the impact. If the population mean difference is constant between the two sites during the "before" phase (analogous to having two parallel lines of abundance in time), and if the time series at two locations are realizations of the same phenomenon (if not, it might be possible to transform the data), and if the observations are independent, then one can do a t-test comparing the average "before" difference to the average "after" difference. The design is intended to account for the temporal variability in the process under study in order to distinguish the effect of the disturbance. Not accounted for is the spatial variability (but see Underwood[67]). Note the caveats (assumptions) inherent in the procedure. For the question of nonindependence, one might be able to do an intervention analysis as is done in time series studies.[17,68]

An important consideration is that the BACI design described here assumes that the effect is a change in the mean level.[66] This, in fact, may or may not be true; for example, the effect could be a change in variability within the population under study.[67,69,70] In addition, the design assumes a constant level of effect after the event occurs which also may not be true. That is, it assumes an immediate and constant shift in the mean location. It could be that there is an immediate impact that gradually dissipates or, conversely, an impact that gradually becomes more severe, depending on the type of impact and the variable under study.

Underwood[67] recommends an asymmetrical BACI design in which there is one impact site and several control sites. This is especially important in cases where the response variable exhibits temporal interaction with a site location that cannot be addressed by the more typical single control study. He shows that the design is useful for determining the effect of an impact on (1) temporal variability, (2) short-term (pulse) responses to disturbances, and (3) a combination of a change in the mean as well as heterogeneity in the temporal variability of the process. Another consideration is the variable(s) of interest, i.e., the variable being used to determine if an impact occurs. In an ecosystem, for example, the effect on species richness could be masked by tolerant species displacing intolerant species and hence there would be no perceived difference in diversity.

Conversely, the impact could have an effect opposite to that expected, for example, a prey species could decrease in abundance due to emigration or mortality or it could conversely increase in abundance due to a loss of predation pressure.[70] As a result, some methods rely on multivariate BACI designs in which the information is a result of multivariate dimension reductions techniques, such as canonical correspondence analysis or principal component analysis.[71] As described earlier, indices of biotic integrity[39] are another example of methodology aimed at summarizing the ecosystem or community under assessment.

An alternative approach would be to choose a population parameter that would display large effects, i.e., the difference in the "before" and "after" phase would be large, relative to the standard error of the difference.[72] The problem here is that the ability to detect this effect size depends on both the natural variability in the process under study as well as the number of sampling events. Osenberg et al.[72] recommend using results from other types of studies, such as long-term monitoring studies and after-only studies, to help identify variables and sample sizes that show large differences that can be detected in the presence of background noise or natural variability. In addition, they address the issue of sampling period in order to obtain independent observations.

One of the potential problems with a BACI or similar sampling design is that it does not adhere to the classic experimental design in which treatments are randomized among sampling units. For example, in most environmental assessments, the particular event, such as the building of a power plant or construction of a roadway, is fixed in location and its impact must be compared against a similar site that will not be impacted. As a result, hypothesis testing can be problematic.[8,10,67] This has led to several variations on the basic BACI design first proposed by Green,[6] including multiple independent observations paired in time at both the control and impact sites[66] and including multiple control sites.[67]

Sometimes, the effect in question is not an impact but instead its opposite, the remediation of an impact. In that case, the studies involve a somewhat different question, namely whether remediation or reclamation efforts have been successful. Like impact assessment, the main issues are adequate sampling and hypothesis testing. The main difference is that in remediation studies the classical statistical testing paradigm is inappropriate since it is the null hypothesis of no difference that is of interest. As a result, recent studies have recommended bioequivalence testing.[8,73–75] In the classical testing approach, nonrejection of the null hypothesis is not sufficient proof that the null hypothesis is true. Hence, bioequivalence testing reformulates the question by assuming the null case to be that there has been damage (i.e., prior to remediation there is a difference in means between impact and control sites) and the alternative to be that remediation has been effective (the means of the two sites are now statistically within some percentage of each other). Like BACI designs, issues of sample size, adequate control sites, and whether the means of the two sites under study should be assumed to be equal in the absence of an impact are relevant in remediation studies as well.[76]

17.5 ENVIRONMENTAL MONITORING

Assessment studies can be difficult to execute because a site can encompass a watershed or a river and finding a comparable control site as in the BACI designs discussed above can be a challenge. Monitoring studies are typically of longer term than assessment studies, perhaps spanning over years, and are aimed at determining long-term trends or changes. So monitoring studies can have all the complexities of an assessment study and more. The analyses are apt to vary, depending on the regularity of the data collection periods. In addition, if monitoring occurs over several years, there is also the issue that technology, staff, and lab methods are apt to evolve over the course of the study. A recent National Research Council (NRC) study[77] has noted that the fisheries surveys of the National Marine Fisheries Service (NMFS) have received skepticism by the commercial fisheries community. Because NMFS has been using the same gear for over 20 years while commercial fishing gear has changed considerably over the same time period, the fisherman believe NMFS chronically underestimates the available catch. Similar criticisms are drawn when the gold standard for an assay has changed over the course of a decade.

Objectives of monitoring can vary considerably among studies but almost always include establishing normal ranges of water, soil, or air quality, for example. Because these ranges may vary by season and location and over wet and dry years, it may take several years to get a sense of what is normal. Many government agencies have regular monitoring programs (see Chapters 1, 22, and 27 to 32 this volume) to help them set regulatory standards in line with their mission. Long-term ecological research stations funded by National Science Foundation (NSF) are designed to help scientists understand if and how environments change over a decade or more. Resource planners may want to monitor natural resources such as fish and forest lumber to regulate when and how people harvest. In this context, patterns over time are of interest, but this year's numbers may be immediate cause for action. By understanding normal patterns and being able to define outliers, there is a potential for early warning systems, e.g., to detect a crash in striped bass populations in the Chesapeake Bay or to detect a cornborer outbreak in Nebraska. Outliers are by definition those outside the normal range of values. Sometimes the objective is to detect shifts in mean, e.g., establishing if global sea ice levels are declining over time in spite of seasonal ups and downs.[78] A decline in global sea ice could be indicative of global warming. Looking for shifts in mean in one or more variables over time is often referred to as trend detection.[79–81] Invariably, some form of regression method is employed. When monitoring air and water quality, one may model when particulate matter or some other quantity exceeds a legally mandated cutoff[79] or one can model the elapsed time between observed violations of this cutoff.[82]

In contrast to assessment studies where the data are often spatially referenced, monitoring studies often consist of one or more time series or a set of data that is both temporally and spatially referenced. Statistical models for spatio-temporal referenced data have been developing rapidly in the past few years, but the field is still in its infancy[83] and statistics packages such as SAS and S-plus are limited in providing procedures for these data. The most common approaches to analyzing this kind of data have been model-based (see Gregoire[84]) but a few approaches have been design based.[80,81] In design-based approaches, the analysis is solely determined by when and where one samples. Urquhart and Kincaid[80,81] describe different kinds of designs that vary in whether the same sites are revisited none, all, or only some of the time. Split panel designs allow one to have some sites regularly monitored while augmenting the survey with a random sample during each sampling period. Model-based inference can include covariates, can incorporate equations derived elsewhere, and can incorporate complex parametric model assumptions. The adequacy of these models is somewhat tied to the appropriateness of the assumptions, so design-based methods are often used since they rely on fewer assumptions overall.

Statistical models are apt to vary depending on how data are aggregated in space or time. Data can be spatially indexed point data such as measurements recorded from soil and sediment core samples. Lattice data are spatial data that have been aggregated such as by watershed or by county. In a third class of spatial data (point patterns), the location of an event is the variable of interest and one looks at patterns of such events in space[18] or in time and space.[85] In these instances we may look at clusters of events or regularity of the points in space. So, for example, in environmental health a cluster of leukemia cases that are close in space and in time may suggest a point source for a carcinogen. The majority of monitoring studies associated with air, soil, and water quality consist of either point or lattice data. The presence of a temporal component adds another level of complexity to the analysis.

The most common time series models assume regular collection intervals although several of the newer methods relax such assumptions.[19,20] Most monitoring approaches assume the study sites are selected at the beginning of a study. However, Zidek et al.[86] suggests approaches to adding study sites and Wikle and Royle[87] suggest a more adaptive approach in which study sites are selected each year for monitoring purposes. Lin[65] describes a method for adding and subtracting monitoring sites based on geostatistics and kriging.

17.6 STATISTICAL ASPECTS OF MONITORING AIR QUALITY: AN EXAMPLE

We have chosen to focus some attention on methods specific to the monitoring of air quality. Monitoring of air quality has some unique challenges, though the primary values of interest are continuous variables such as levels of particulate matter or specific gasses (e.g., ozone, sulfur dioxide). As we indicated earlier, models for continuous responses are among the most common, certainly in the context of spatio-temporal data. Many natural resources such as forest inventories or fisheries are monitored on an annual or perhaps a monthly or biweekly schedule and are not always spatially referenced and are therefore easier to evaluate. Air quality needs to be almost continuously monitored with readings collected hourly over many locations.

Suspended particulate matter of a diameter less than 10 μm (PM10) is often used as a measure of air quality. Sustained exposure to elevated PM10 values is thought to be associated with a variety of health problems.[88] Estimating the level of PM10 over a broad geographic range poses statistical problems that include defining adequate spatial coverage and capturing of spatial and temporal variation in PM10 values. Because elevated PM10 values are thought to be associated with specific health problems, regulatory limits for PM10 values have been set.[83] There are statistical methods designed to assess compliance with a regulatory standard (e.g., Polansky and Check[82]). We contrast objectives and approaches of several statistical methods that have focused on air quality. We assume that PM10, log transformed, is the variable of interest.

In order to quantify PM10 values over an area, one needs to interpolate values between monitoring stations. In doing so, one may be able to identify hot spots or areas that require greater attention, e.g., near an industrial area or in an area with high traffic volume. Mapping can be done using conventional geostatistical methods such as construction of a variogram to model the correlation that exists as a function of distance between stations[18,65] and then kriging to calculate predicted values for locations in an area. Some analyses fail to recognize that the outcomes of the evolution of PM10 values over time in a region are probably due to varying spatial and temporal behaviors. As a result, these analyses tend to ignore the effects of either time (on the spatial structure) or space (on the temporal structure). For example, weekends may have different PM10 values than weekdays, and time of day may even be a factor that accounts for observed PM10 values. This can especially be a problem if the number of monitoring stations is limited. Models that consider known patterns of temporal variation should be more efficient in predicting values and should yield smaller standard errors associated with the predicted values.[88]

Holland et al.[89] analyze sulfur dioxide values in a sampling regime in which multiple monitoring stations are monitored at regular intervals. They developed a time series model for each monitoring station using generalized additive models[90] with an explanatory variable corresponding to day of week and week of year, and then used these equations to develop spatial patterns over time.

17.7 SUMMARY

There are many common themes for monitoring and assessment as can be seen in this chapter. Both rely on sampling that covers a region (or regions) and hence have either an implicit or explicit spatial component. What distinguishes monitoring is that it adds the complexity of time, since monitoring is often conducted over long periods. Another major difference is that assessment is often used to determine if a change, either remediation or its opposite, development, has occurred, whereas monitoring is often more related to determining trend.

Design-based and model-based inference procedures exist for both monitoring and assessment. In general, design-based procedures have fewer assumptions but probably are more restrictive from a data analysis perspective in that they are aimed at providing estimates of population quantities but not at hypothesis testing. Hence, they are little used in most environmental assessments since the major emphasis is to determine effect. The design-based approaches are more used in monitoring situations in which the interest is reporting averages or totals such as might be seen in fisheries management.

Both assessment and monitoring are similar in that, like any statistical method, the techniques used are sensitive to sample size issues and to collection techniques. For example, most spatial and temporal statistical methods are best used on data that have been at least somewhat regularly collected. In the case of space, this usually means a systematic sample taken on a grid if one is interested in a map of the region. For time series, this usually means sampling on a regular schedule. Note, though, that the motivation for the data collection will often drive the collection to be done on a different schedule or over specific regions of space in order to capture unusual events. The analytical techniques must therefore be modified.

It should be evident from these discussions that environmental statistics is an active area of research. We will invariably see more research in the analysis of data that varies both spatially and temporally. We expect to see better models for correlated data that do not require an assumption of a normal or Gaussian distribution and more models that accommodate censoring in both explanatory and response variables. More adaptive methods are also on the horizon, allowing for more effective use of resources. Right now the biggest impediment to implementing a number of these algorithms is the availability of user-friendly software. We expect that will change when we see more general methods developed that are applicable in multiple settings.

REFERENCES

1. Burroughs, P. A. GIS and geostatistics: essential partners for spatial analysis. *Environ. Ecol. Stat*. 8: 361–377, 2001.
2. Waller, L. and Gotway-Crawford, C. *Applied Spatial Analysis of Public Health Data*. John Wiley & Sons, New York (in press).
3. Kuehl, R. *Design of Experiments: Statistical Principles of Research Design and Analysis*. 2nd ed. Duxbury Press, Pacific Grove, CA. 2000.
4. Cochran, W. *Sampling Techniques*. 3rd ed. John Wiley & Sons, New York, 1977.
5. Thompson, S. *Sampling*. John Wiley & Sons, New York, 2002.
6. Green, R. *Sampling Design and Statistical Methods for Environmental Biologists*. John Wiley & Sons, New York, 1979.
7. Gilbert, R. O. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York, 1987.
8. Manly, B. F. J. *Statistics for Environmental Science and Management*. Chapman and Hall, London, 2000.
9. El-Shaarawi, A. et al., Eds. *Encyclopedia of Environmetrics*. John Wiley & Sons, London, 2002.
10. Hurlbert, S. J. Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* 54: 187–211, 1984.
11. McBride, G. B. Statistical methods helping and hindering environmental science and management. *J. Agric. Biol. Environ. Stat*. 7: 300–305, 2001.
12. Johnson, R. A. and Wichern, D. *Applied Multivariate Statistical Methods*. Prentice Hall, New York, 1998.
13. Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. *Bayesian Data Analysis*. Chapman and Hall, New York, 1995.
14. Spiegelhalter, D., Thomas, A., and Best, N. G. WINBUGS Version 1.2 User Manual. MRC Biostatistics Unit, Oxford, U.K., 1999 (available on-line).

15. Kleinbaum, D. G., Kupper, L. L., Muller, K. E., and Nizam, A. *Applied Regression Analysis and Other Multivariable Methods*. Duxbury Press, Pacific Grove, CA, 1998.
16. Chatfield, C. *The Analysis of Time Series: An Introduction*. Chapman and Hall, London, 1989.
17. Wei, W. W. S. *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley, Redwood City, CA, 1994.
18. Cressie, N. *Statistics for Spatial Data*. 2nd ed. John Wiley & Sons, New York, 1993.
19. Verbeke, G. and Mohlenberghs, G. *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, Berlin, 2000.
20. Diggle, P., Heagarty, P., Liang, K.-Y., and Zeger, S. *Analysis of Longitudinal Data*. 2nd ed. Oxford University Press, Oxford, U.K., 2002.
21. McCullagh, P. and Nelder, J. A. *Generalized Linear Models*. 2nd ed. Chapman and Hall, London, 1989.
22. SAS. SAS on-line documentation for Version 8.2, SAS Institute. Cary, NC, 2003.
23. Kedem, B. and Foikanos, K. *Regression Models for Time Series*. John Wiley & Sons, New York, 2002.
24. Rosner, B. *Fundamentals of Biostatistics*. 5th ed. Duxbury, Pacific Grove, CA, 2000.
25. Lawless, J. E. *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, New York, 1982.
26. Li, Y. and Ryan, L. Modeling spatial survival data using semiparametric frailty models. *Biometrics* 58: 287–297, 2002.
27. El-Shaarawi, A. and Nader, A. Statistical inference from multiply censored environmental data. *Environ. Monit. Assess*. 17: 339–347, 1991.
28. Hawkins, D. and Oehlert, G. W. Characterization using normal or log-normal data with multiple censoring points. *Environmetrics* 11: 167–182, 2000.
29. Russek-Cohen, E. and Colwell, R. R. Computation of most probable numbers. *Appl. Environ. Microbiol*. 45: 1146–1150, 1983.
30. Everitt, B. S. and Hand, D. J. *Finite Mixture Distributions*. Chapman and Hall, London, 1981.
31. Lambert, D. Zero-inflated Poisson regression, with an application in manufacturing. *Technometrics* 34: 1–14, 1992.
32. Warren, W. G. Changes in the within-survey spatio-temporal structure of the northern cod (*Gadus morhua*) population 1985–1992. *Can. J. Fish. Aquat. Sci*. 54(supplement): 139–148, 1997.
33. Hall, D. B. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics* 56: 1030–1039, 2000.
34. Wang, P. Markov zero-inflated poisson regression models for a time series of counts with excess zeros. *J. Appl. Stat*. 28: 623–632, 2001.
35. Owen, W. J. and DeRouen, T. A. Estimation of the mean for lognormal data containing zeroes and left censored values, with applications to the measurement of worker exposure to air contaminants. *Biometrics* 36: 707–719, 1980.
36. Fuller, W.A. *Measurement Error Models*. John Wiley & Sons, New York, 1987.
37. Thurigan, D., Spiegelman, D., Blettner, M., Heuer, C., and Brenner, H. Measurement error correction using validation data: a review of methods and their applicability in case-control studies. *Stat. Methods Med. Res*. 9: 447–474, 2000.
38. Pielou, E.C. *Ecological Diversity.* John Wiley & Sons, New York, 1975.
39. Roth, N. E., Southerland, M. T., Chaillou, J. C., Klauda, R. J., Kazyak, P. F., Stranko, S. A., Weisberg, S. B., Hall, L. W., Jr., and Morgan, R. P., II. Maryland biological stream survey: development of a fish index of biotic integrity. *Environ. Monit. Assess*. 51: 89–106, 1998.

40. Smith, W. F. ANOVA-like similarity analysis using expected species shared. *Biometrics* 45: 873–881, 1989.
41. Solow, A. R. Detecting changes in the composition of a multispecies community. *Biometrics* 50: 556–565, 1994.
42. Zangerl, A. R., McKenna, D., Wraight, C. L., Carroll, M., Ficarello, P., Warner, R., and Berenbaum, M. R. Effects of exposure to event 176 *Bacillus thuringiensis* corn pollen on monarch and black swallowtail caterpillars under field conditions. *Proc. Natl. Acad. Sci. U.S.A*. 98: 11908–11912, 2001.
43. Stanley-Horn, D. E., Dively, G. P., Hellmich, R. L., Mattila, H. R., Sears, M. K., Rose, R., Jesse, L. C. H., Losey, J. E., Obrycki, J. J., and Lewis, L. Assessing the impact of cry1Ab-expressing corn pollen on monarch butterfly larvae in field studies. *Proc. Natl. Acad. Sci. U.S.A*. 98: 11931–11936, 2001.
44. Sears, M. K., Hellmich, R. L., Stanley-Horn, D. E., Oberhauser, K. S., Pleasants, J. M., Mattila, H. R., Siegfried, B. D., and Dively, G. P. Impact of Bt corn pollen on monarch butterfly populations: a risk assessment. *Proc. Natl. Acad. Sci. U.S.A*. 98: 11937–11942, 2001.
45. Christman, M. C. A review of quadrat-based sampling of rare, geographically clustered populations. *J. Agric. Biol. Environ. Stat*. 5: 168–201, 2000.
46. Seber, G. A. F. *The Estimation of Animal Abundance and Related Parameters*. Charles Griffin, London, 1982.
47. Seber, G. A. F. A review of estimating animal abundance. *Biometrics* 42: 267–292, 1986.
48. Seber, G. A. F. A review of estimating animal abundance II. *Int. Stat. Rev.* 60: 129–166, 1992.
49. Kenkel, N. C., Juhsz-Nagy, P., and Podani, J. On sampling procedures in population and community ecology. *Vegetatio* 83: 195–207, 1989.
50. Seber, G. A. F. and Thompson, S. K. Environmental adaptive sampling, *Handbook for Statistics*, Vol. 12. Patil, G. P. and Rao, C. R., Eds. Elsevier Science, Amsterdam, 1994, pp. 201–220.
51. Stehman, S. V. and Overton, W. S. Environmental sampling and monitoring, in *Handbook of Statistics*, Vol. 12. Patil, G. P. and Rao, C. R., Eds. Elsevier Science, Amsterdam, 1994, pp. 263–306.
52. Patil, G. P. and Taillie, C. Contemporary Challenges and Recent Advances in Ecological and Environmental Sampling, Technical Report 97-0501, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, The Pennsylvania State University, University Park, PA, 1997.
53. Pennington, M. Efficient estimators of abundance, for fish and plankton surveys. *Biometrics* 39: 281–286, 1983.
54. Thompson, S. K. and Seber, G. A. F. *Adaptive Sampling*. John Wiley & Sons, New York, 1996.
55. Brown, J. A. The application of adaptive cluster sampling to ecological studies, in *Statistics in Ecology and Environmental Monitoring*, Otago Conference Series 2. Fletcher, D. J. and Manly, B. F., Eds. University of Otago Press, Dunedin, New Zealand, 1994, pp. 86–97.
56. Smith, D. R., Conroy, M. J., and Brakhage, D. H. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics* 51: 777–788, 1995.
57. Christman, M. C. Efficiency of adaptive sampling designs for spatially clustered populations. *Environmetrics* 8: 145–166, 1997.
58. Pontius, J. S. Strip adaptive cluster sampling: probability proportional to size selection of primary units. *Biometrics* 53: 1092–1095, 1997.

59. Salehi, M. M. and Seber, G. A. F. Adaptive cluster sampling with networks selected without replacement. *Biometrika* 84: 209–220, 1997.
60. Kalton, G. and Anderson, D. W. Sampling rare populations. *J. R. Stat. Soc. Series A* 149: 65–82, 1986.
61. Thompson, S. K. Adaptive cluster sampling. *J. Am. Stat. Assoc.* 85: 1050–1059, 1990.
62. Thompson, S. K. Stratified adaptive cluster sampling. *Biometrika* 78: 389–397, 1991b.
63. Christman, M. C. and Lan, F. Inverse adaptive cluster sampling. *Biometrics* 57: 1096–1105, 2001.
64. Christman, M. C. and Pontius, J. Bootstrap confidence intervals for adaptive cluster sampling. *Biometrics* 56: 503–510, 2003.
65. Lin, Y.-P. Geostatistical approach for optimally adjusting a monitoring network, in *Environmental Monitoring*. Wiersma, G. B., Ed. CRC Press, Boca Raton, FL (in press).
66. Stewart-Oaten, A., Murdoch, W. W., and Parker, K. R. Environmental impact assessment: pseudoreplication in time? *Ecology* 67: 929–940, 1986.
67. Underwood, A. J. On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecol. Appl.* 4: 3–15, 1994.
68. Box, G. E. P. and Tiao, G. C. Intervention analysis with applications to economic and environmental problems. *J. Am. Stat. Assoc.* 70: 70–79, 1975.
69. Osenberg, C. W. and Schmitt, R. J. Detecting environmental impacts: detecting human impacts in marine habitats. *Ecol. Appl.* 4: 1–2, 1994.
70. Smith, E. P. BACI design in *The Encyclopedia of Environmetrics*, Vol. 1. El-Shaarawi, A. H. and Piegorsch, W. W., Eds. John Wiley & Sons, Chichester, U.K., 2002.
71. Kedwards, T. J., Maund, S. J., and Chapman, P. F. Community-level analysis of ecotoxicological field studies: II. Replicated design studies. *Environ. Toxicol. Chem.* 18: 158–166, 1999.
72. Osenberg, C. W., Schmitt, R. J., Holbrook, S. J., Abu-Saba, K. E., and Flegal, A. R. Detection of environmental impacts: natural variability, effect size, and power analysis. *Ecol. Appl.* 4: 16–30, 1994.
73. McBride, G. B., Loftis, J. C., and Adkins, N. C. What do significance tests really tell us about the environment? *Environ. Manage.* 17: 423–432, 1993.
74. McDonald, L. L. and Erickson, W. P. Testing for bioequivalence in field studies: has a disturbed site been adequately reclaimed? in *Statistics in Ecology and Environmental Monitoring*. Fletcher, D. J. and Manly, B. F. J., Eds. University of Otego Press, Dunedin, New Zealand, 1994, pp. 183–197.
75. McBride, G. B. Equivalence tests can enhance environmental science and management. *Aust. N. Z. J. Stat.* 41: 19–29, 1999.
76. Grayson, J. E., Chapman, M. G., and Underwood, A. J. The assessment of restoration of habitat in urban wetlands. *Landscape Urban Plann*. 43: 227–236, 1999.
77. Sullivan, P. K., Able, K., Jones, C. et al. *Improving the Collection, Management and Use of Marine Fisheries Data.* National Academy Press, Washington, D.C., 2000.
78. Parkinson, C. L., Cavalieri, D. J., Gloersen, P., Zwally, H. J., and Comiso, J. C. Arctic sea ice extents, areas and trends. *J. Geophys. Res.* 104: 20837–20856, 1999.
79. Piegorsch W. W., Smith, E. P., Edwards, D., and Smith R. L. Statistical advances in environmental science. *Stat. Sci*. 13: 186–208, 1998.
80. Urquhart, N. S. and Kincaid, T. M. Monitoring for policy-relevant regional trends over time. *Ecol. Appl*. 8: 246–257, 1998.
81. Urquhart, N. S. and Kincaid, T. M. Designs for detecting trend from repeated surveys of ecological resources. *J. Agric. Biol. Environ. Stat*. 4: 404–414, 1999.
82. Polansky, A. M. and Check, C. E. Testing trends in environmental compliance. *J. Agric. Biol. Environ. Stat*. 7: 452–468. 2002.

83. Cox, L. H. Statistical issues in the study of air pollution involving airborne particulate matter. *Environmetrics* 11: 611–626, 2000.
84. Gregoire, T. G. Design-based and model-based inference in survey sampling: appreciating the difference. *Can. J. For. Res*. 28: 1429–1447, 1998.
85. Kulldorf, M. and Hjalmars, U. The knox method and other tests for space–time interaction. *Biometrics* 55: 544–552. 1999.
86. Zidek, J. V., Sun, W., Le, N. D. Designing and integrating composite networks for monitoring multivariate Gaussian pollution fields. *Appl. Stat*. 49: 63–79, 2000.
87. Wikle, C. K. and Royle, J. A. Space–time dynamic design of environmental monitoring networks. *J. Agric. Biol. Environ. Stat*. 4: 489–507, 1999.
88. Sun, L., Zidek, J. V., Le, N. D., and Ozkaynak, H. Interpolating Vancouver's daily ambient PM10 field. *Environmetrics* 11: 651–664, 2000.
89. Holland, D. M., DeOliveira, V., Cox, L. H., and Smith, R. L. Estimation of regional trends over the eastern United States. *Environmetrics* 11: 373–394, 2000.
90. Hastie, T. J. and Tibshirani, R. J. *Generalized Additive Models*. Chapman and Hall, London, 1990.

18 Geostatistical Approach for Optimally Adjusting a Monitoring Network

Y.-P. Lin

CONTENTS

18.1 INTRODUCTION

Designing an environmental monitoring network involves selecting sampling sites and frequencies.[1] However, an optimal information-effective monitoring network should provide sufficient but no redundant information of monitoring variables. The information generated by such monitoring networks may be used to characterize natural resources and to delineate polluted area. Given the high cost and risks associated with such investigations, the development of efficient procedures for designing or adjusting monitoring networks is crucial. In such investigations, the collected data may include significant uncertainty, including complex (unexplainable) or extremely complicated variations in observed values of measurable characteristics of the investigated medium in time and space. Accordingly, several authors have used statistical procedures to model the spatial structures of investigated variables.[2–16]

Geostatistics, a spatial statistical technique, is widely applied to analyze environmental monitoring data in space and time. Geostatistics can characterize and quantify spatial variability, perform rational interpolation, and estimate the variance

1-56670-641-6/04/$0.00+$1.50

in the interpolated values. Kriging, a geostatistical method, is a linear interpolation procedure that provides a best linear unbiased estimator (BLUE) for quantities that vary spatially. Recently, kriging has been widely used to analyze the spatial variability of monitoring data on environmental, hydro-geological, and ecological studies.[11,12,17–39]

Statistical approaches have also been taken to designing or adjusting environmental monitoring systems and quantifying the informational value of monitoring data. Among the proposed statistical methods, some treat designing a monitoring network as an optimization problem. For example, James and Gorelick[40] provided an approach, including Risk–Cost–Benefit and Data Worth analyses, to determine the optimum number of and the best locations of observation wells. Christakos and Olea[41] maximized the accuracy of estimates and minimized the cost of exploration to determine the best sampling strategy. Rouhani[42] and Rouhani and Hall[43] discretized the field into a grid of potential estimation points, which were then sequentially ranked according to their efficiency in reducing the site-wide estimation variance or other risk-driven criteria. Loaiciga[44] added a binary variable to a linear combination estimator and used mixed integer programming (MIP) to minimize the estimated variance. Hudak and Loaiciga[45] selected monitoring sites from a pool of potential nodes with various different weights. Mckinney and Loucks[46] used first-order uncertainty analysis to compute the prediction variance and select new locations.

Groenigen and Stein[47] used spatial simulated annealing to optimize spatial environmental schemes at the point-level, accounting for sampling constraints and preliminary observations. Groenigen et al.[48] extended spatial simulated annealing to optimize spatial sampling schemes for obtaining minimal kriging variance. Like variance reduction analysis,[42] these schemes reduced the high values of kriging variance near the boundaries of the study area. The objective of these schemes was to minimize the prediction variance of the simulation model by selecting new measurement locations. James and Gorelick[40] developed risk–benefit data worth analyses to determine the optimum number and the best locations of observation wells, and to minimize the expected cost of remediation plus sampling. One of the main deficiencies of these latter algorithms is related to the fact that they typically require potential sampling locations to be defined. They are commonly defined by spatial discretization, which confines the selected sites to a few predefined locations. Moreover, the sequential ranking of sampling sites imposes additional constraints on the optimality of the selected points.

Lin and Rouhani[1] have developed multiple-point variance analysis (MPV),[1] including Multiple-Point Variance Reduction Analysis (MPVR) and Multiple-Point Variance Increase Analysis (MPVI), to expand on the foregoing works by providing automatic procedures for simultaneously identifying groups of sites without any need for spatial discretization or sequential selection. The effect of any set of additional sampling locations can be determined through MPVR. The effect of deleting redundant sampling points can be estimated by MPVI. The goal of MPVR[1] is to develop a framework for the optimal simultaneous selection of additional or redundant sampling locations. Lin and Rouhani[1] successfully improved and applied optimal MPV to a field case. This chapter introduces optimal MPV and its applications in a more realistic hypothetical case.

18.2 MULTIPLE-POINT VARIANCE ANALYSIS (MPV)

18.2.1 Geostatistics

Geostatistical methods are based on the regionalized variables theory, which states that variables in an area have both random and spatial properties.[49] A geostatistical variogram of data must first be determined. A variogram quantifies the commonly observed relationship between the values of data pertaining to the samples and the samples' proximity. Geostatistics provide a variogram of data within a statistical framework, including spatial and temporal covariance functions. These variogram models are termed spatial or temporal structures and are defined in terms of the correlation between any two points separated either spatially or temporally. The variogram (h) is defined as:

$$\gamma(h) = (1/2)Var[Z(x) - Z(x+h)] \tag{18.1}$$

An experimental variogram, (h), is given by

$$\gamma(h) = 1/[2n(h)]\sum_{i=1}^{n(h)}[Z(x_i + h) - Z(x_i)]^2 \tag{18.2}$$

Ordinary kriging, as applied within moving data neighborhoods, is a nonstationary algorithm that corresponds to a nonstationary random function model with varying mean but stationary covariance.[50] Kriging estimates are determined as weighted sums of the adjacent sampled concentrations. These weights depend on the exhibited correlation structure. For illustration, if data appear to be highly continuous in space, the points closer to the estimated points are more highly weighted than those farther away. The criterion for determining these weights is the minimization of the estimation variance. In this context, kriging estimates (best linear unbiased estimator) can be regarded as the most accurate among all linear estimators. Consequently, at an unsampled location and for a given variogram, a kriging estimate can be considered simply to be an optimally weighted average of the data sampled at surrounding locations).[51] Accordingly, kriging estimates the value of the random variable at an unsampled location X_0, based on the measured values in linear form.[42]

$$Z^*(X_0) = \sum^{N} \lambda_{i0} Z(X_i) \tag{18.3}$$

Two criteria are imposed on the preceding kriging estimation. The first criterion requires the estimator $Z^*(X_0)$ to be unbiased.

$$E[Z^*(X_0) - Z(X_0)] \tag{18.4}$$

The second criterion requires the estimator $Z^*(X_0)$ to yield estimates with minimum variance, which is

$$Var[Z^*(X_0) - Z(X_0)] \tag{18.5}$$

In ordinary kriging, the mean is assumed to be constant but known. The minimization process in this kriging method can be written as

$$\frac{\partial[Z^*(X_0) - Z(X_0)]}{\partial\lambda_i} - 2\mu = 0 \qquad \text{for i = 1, 2, …, n}$$

where μ is the Lagrange multiplier for the unbiasedness constraint.

The estimation variance of ordinary kriging is

$$\begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & \gamma_{I,I} & \gamma_{I,II} & \cdots & \gamma_{I,N} \\ 1 & \gamma_{II,I} & \gamma_{II,II} & \cdots & \gamma_{II,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \gamma_{N,I} & \gamma_{N,II} & \cdots & \gamma_{N,N} \end{bmatrix} \begin{bmatrix} \mu \\ \lambda_{I,0} \\ \lambda_{II,0} \\ \vdots \\ \lambda_{N,0} \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{I,0} \\ \gamma_{II,0} \\ \vdots \\ \gamma_{N,0} \end{bmatrix} \tag{18.6}$$

where $\gamma_{i,j}$ = (semi-) variogram between Z_i and Z_j, $\gamma(|x_i - x_j|)$; Z_i = the random variable at x_i; $\gamma_{\text{ i,0}}$ = the kriging weight of Z_i in the estimation Z_0^*; μ = Lagrange multiplier in the kriging system for estimation; z_i = measured value at x_i; Z_0^* = estimated value at x_o; $i = I, \ldots, N$; N = the number of existing data points used.

Three main features of a typical variogram are (1) range, (2) sill, and (3) nugget effect. Range is the distance at which the variogram reaches its maximum value. Paired samples separated by a distance greater than the range are uncorrelated.[52] The sill is the upper limit of any variogram, to which it tends at a large distance. The sill is a measure of the population variability of investigated variable; a higher sill corresponds to greater variability in the population.[52] The nugget effect is exhibited by the apparent jump of the variogram at the origin, which fact may be attributed to the small-scale variability of the investigated process and/or measured errors.[53]

18.2.2 Multiple-Point Variance Reduction Analysis (MPVR)

The general MPVR equation is initially developed by computing the effect of one additional sampling point on the kriging estimation variance at another location.[1] All developments of MPVR are presented in the context of ordinary kriging and variance reduction analysis.[42] The kriging system can be written as

$$Aw_0 = a_0 \tag{18.7}$$

where $a_0 = \begin{bmatrix} 1 \\ \gamma_{I,0} \\ \gamma_{II,0} \\ \vdots \\ \gamma_{N,0} \end{bmatrix}$; $w_0 = \begin{bmatrix} \mu \\ \lambda_{I,0} \\ \lambda_{II,0} \\ \vdots \\ \lambda_{N,0} \end{bmatrix}$

If a new measurement point at x_I is added, the matrix A in Equation 18.7 should be expanded into a matrix A_I. A_I is matrix A with a new bottom row and a new right-hand side column. Based on the above kriging system, Rouhani[42] developed a variance reduction equation

$$VR = 1/V_1(N)[\gamma_{10} - a_0^T\, w_1]^2 \tag{18.8}$$

where $V_1(N)$ = variance of estimation at x_1 prior to any sampling at that point;

$$a_0^T = [1 \;\; \gamma_{I,0} \;\; \gamma_{II,0} \;\; \cdots \;\; \gamma_{N,0}];\;\; w_1^T = [\mu \;\; \gamma_{I,1} \;\; \gamma_{II,1} \;\; \cdots \;\; \gamma_{N,1}]\,.$$

When two new measurement points at x_1 and x_2 are added, the matrix A_1 in Equation 18.8 is expanded into a bordered matrix A_2. By an approach similar to that of variance reduction, $VR_0(x_1, x_2)$ can be written as

$$\begin{aligned} VR_0(x_1, x_2) &= V_0(N) - V_0(N+2) \\ &= \alpha_1\left(\gamma_{1,0} - a_0^T w_1\right)^2 + \alpha_2\left(\gamma_{2,0} - \left[a_0^T \quad \gamma_{1,0}\right] w_2\right)^2 \end{aligned} \tag{18.9}$$

where

$a_{1,0}^T = [1 \;\; \gamma_{I,0} \;\; \gamma_{II,0} \;\; \cdots \;\; \gamma_{N,0} \;\; \gamma_{1,0}];$

$\alpha_2\left[\gamma_{2,2} - a_2^T A_1^{-1} a_2\right]^{-1} = [V_2(N+1)]^{-1};\; w_2 = A_I^{-1} \;\; a_2;$

$A_1 = \begin{bmatrix} A & a_1 \\ a_1^T & \gamma_{1,1} \end{bmatrix};\;\; a_1^T = [1\gamma_{I,1} \;\; \gamma_{II,1} \;\; \cdots \;\; \gamma_{N,1}].$

Similarly, when m new measurement points at $x_1, x_2,\ldots, x_m$ are added, the matrix A_m^1 can be expanded to be a bordered matrix A_m. Consequently,

$$\begin{aligned} VR_0(x_1, x_2, \ldots, x_m) &= V_0(N) - V_a(N+m) \\ &= \alpha_1\left(\gamma_{1,0} - a_0^T w_1\right)^2 + \alpha_2\left(\gamma_{2,0} - \left[a_0^T \quad \gamma_{1,0}\right] w_2\right)^2 \\ &\quad + \cdots + \alpha_m\left(\gamma_{m,0} - \left[a_0^T \quad \gamma_{1,0} \quad \ldots \quad \gamma_{m-1,0}\right] w_m\right)^2 \end{aligned} \tag{18.10}$$

where $\alpha_m = \left[\gamma_{m,m} - a_m^T A_{m-1}^{-1} a_m\right]^{-1} [V_m(N+m-1)^{-1}]; w_m = A_{m-1}^{-1} a_m$.

$VR_0(x_1, x_2, \ldots, x_m)$ is defined as the variance reduction at location x_0 due to the adding of m sampling points.

In many instances, the estimated value has a block support, while measurements are point values.[1] Under such conditions, the given MPVR equations can be modified to cover block-estimation cases. If m new measurement points are added, then Equation 18.10 is modified, using block-points and block-block variograms, as:

$$\begin{aligned} VR_B(x_1, x_2, \ldots, x_m) &= \alpha_1\left(\gamma_{1,B} - a_B^T w_1\right)^2 \alpha_2\left(\gamma_{2,B} - \left[a_B^T \quad \gamma_{1,B}\right] w_2\right)^2 \\ &\quad + \cdots + \alpha_m\left(\gamma_{m,B} - \left[a_B^T \quad \gamma_{1,B} \quad \ldots \quad \gamma_{m-1,B}\right] w_m\right)^2 \end{aligned} \tag{18.11}$$

$$\text{where} \quad a_B = \begin{bmatrix} 1 \\ \gamma_{I,B} \\ \gamma_{II,B} \\ \vdots \\ \gamma_{N,B} \end{bmatrix}; \; w_B = \begin{bmatrix} \mu \\ \lambda_{I,0} \\ \lambda_{II,0} \\ \vdots \\ \lambda_{N,0} \end{bmatrix}$$

N is the number of data points used in kriging, the block-point variogram

$$\gamma_{I,B} = \frac{1}{V}\int_V \gamma\left(|x_I - x_j|\right) dx_j \quad \text{and} \quad \gamma_{B,B} = \frac{1}{V}\int_V \gamma\left(|x_I - x_j|\right) dx_i\, dx_j.$$

V represents the block volume and subscript B denotes the block. Values α_m and w_m are the same as those in Equation (18.10). Other definitions are the same as those determined in Equation 18.6.

18.2.3 Multiple-Point Variance Increase Analysis (MPVI)

Like variance reduction analysis, the first step of MPVI is to establish a relationship between the increase in kriging variance at an arbitrary point due to the deletion of an existing monitoring point.[1] The ordinary kriging procedure is used in this development. However, it is general and can be expanded into other forms. If an existing measurement point, x_N, is dropped, then the matrix A_{NI} can be written in the following matrix form:

$$A_{N-1} = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & \gamma_{I,I} & \gamma_{I,II} & \cdots & \gamma_{I,N-1} \\ 1 & \gamma_{II,I} & \gamma_{II,II} & \cdots & \gamma_{II,N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \gamma_{N-1,I} & \gamma_{N-1,II} & \cdots & \gamma_{N-1,N-1} \end{bmatrix} \tag{18.12}$$

where $N-1$ = the number of the remaining data points used in kriging. The relationship between matrix A and A_{N-1} is written as

$$A = \begin{bmatrix} A_{N-1} & a_N \\ a_N^T & \gamma_{N,N} \end{bmatrix} \tag{18.13}$$

The bordered matrices in matrix A^{-1} can be expressed as:

$$A^{-1} = \begin{bmatrix} A_{N-1} & a_N \\ a_N^T & \gamma_{N,N} \end{bmatrix}^{-1} = \begin{bmatrix} F_N & p_N \\ p_N^T & \alpha_N \end{bmatrix} \tag{18.14}$$

where $F_N = A_{N-1}^{-1} + \alpha_N A_{N-1}^{-1} a_N a_N^T A_{N-1}^{-1}$; $\alpha_N = \left[\gamma_{NN} - a_N^T A_{N-1}^{-1} a_N\right]^{-1}$; $p_N = -\alpha_N A_{N-1}^{-1} a_N$; $w_N = A_{N-1}^{-1} a_N$; $\alpha_N = \left[\gamma_{N,N} - a_N^T A_{N-1}^{-1} a_N\right]^{-1}$.

With an approach similar to that of variance reduction, $VI_0(x_N)$ can be written as:

$$VI_0(x_N) = \alpha_N \left(\gamma_{N,0} - a_{N-1,0}^T w_N\right)^2 \tag{18.15}$$

Similarly, when m measurement points at $x_N, x_{N-1}, x_{N-2}, \ldots, x_{N-m+1}$ are dropped, the relationship between matrix A_{N-m+1} and A_{N-m} becomes:

$$A_{N-m+1} = \begin{bmatrix} A_{N-m} & a_{N-m+1} \\ a_{N-m+1}^T & \gamma_{N-m+1,\,N-m+1} \end{bmatrix} \tag{18.16}$$

where

$$a_{N-m}^{T} = [1 \quad \gamma_{I,N-m} \quad \gamma_{II,N-m} \quad \cdots \quad \gamma_{N-m+1,N-m}];$$

$$a_{N-m+1}^{T} = [1 \quad \gamma_{I,N-m+1} \quad \gamma_{II,N-m+1} \quad \cdots \quad \gamma_{N-m+1,N-m+1}].$$

The A_{N-m+1}^{-1} can be written as:

$$A_{N-m+1}^{-1} = \begin{bmatrix} A_{N-m} & a_{N-m+1} \\ a_{N-m+1}^{T} & \gamma_{N-m+1,N-m+1} \end{bmatrix}^{-1} = \begin{bmatrix} F_{N-m} & P_{N-m+1} \\ P_{N-m+1}^{T} & \alpha_{N-m+1,N-m+1} \end{bmatrix} \tag{18.17}$$

where $\alpha_{N-m+1} = \left[\gamma_{N-m+1,N-m+1} - a_{N-m+1}^{T} A_{N-m}^{-1} a_{N-m+1}\right]^{-1}$.

The increase in variance at x_0 caused by dropping m points at x_N, x_{N-1}, …, x_{N-m+1}, $V_0(m)$, can be written as:

$$VI_0(x_N, x_{N-1}, \ldots, x_{N-m+1}) = \alpha_{N-m+1}\left(\gamma_{N-m+1,0} - a_{N-m,0}^{T} w_{N-m+1}\right)^2 + \cdots$$
$$+ \alpha_{N-1}\left(\gamma_{N-1,0} - a_{N-2,0}^{T} w_{N-1}\right)^2 + \alpha_N\left(\gamma_{N,0} - a_{N-1,0}^{T} w_N\right)^2 \tag{18.18}$$

where $a_{N-m,0}^{T} = [1 \quad \gamma_{I,0} \quad \ldots \quad \gamma_{N-m,0}]$; $w_{N-m+1} = A_{N-m}^{-1} a_{N-m+1}$.

If the estimated value has a block support, when measurements are point values, the above-point MPVI can be expanded to the block estimation case. In this case, the increase in variance at block B due to the deletion of *m* sampling points, at x_N, x_{N-1},…, x_{N-m+1} $VI_B(x_N, x_{N-1},\ldots, x_{N-m+1})$, can be defined as

$$VI_B(x_N, x_{N-1}, \ldots, x_{N-m+1}) = \alpha_{N-m+1}\left(\gamma_{N-m+1,B} - a_{N-m,B}^{T} w_{N-m+1}\right)^2 + \cdots$$
$$+ \alpha_{N-1}\left(\gamma_{N-1,B} - a_{N-2,B}^{T} w_{N-1}\right)^2 + \alpha_N\left(\gamma_{N,B} - a_{N-1,B}^{T} w_N\right)^2 \tag{18.19}$$

18.2.4 Optimal MPVR and MPVI

Most network design and adjustment procedures have focused on adding new locations for getting more information on monitoring variables. In MPVR, the selection of *m* additional points can be defined as an optimization problem. The objective function of this optimization problem is provided by the MPVR equation and so can be formulated as a nonlinear optimization problem.[1] This objective function is in terms of the reduction in the estimation variance over a targeted block of study sites. The existing sampling points are denoted as $i = I, II, \ldots, N$.

The locations of the new m points are selected simultaneously by maximally reducing the estimation variance:

$$\max\ VR_B(x_1, x_2, ..., x_m) = \alpha_1\left(\gamma_{1,B} - a_B^T A_{-0}^{-1}\, a_1\right)^2 + \alpha_2\left(\gamma_{2,B} - \left[a_B^T \quad \gamma_{1,B}\right] A_1^{-1}\, a_2\right)^2$$
$$+ \cdots + \alpha_m\left(\gamma_{m,B} - \left[a_B^T \quad \gamma_{1,B}\ \ldots\ \gamma_{m-1,B}\right] A_{m-1}^{-1}\right)^2 \tag{18.20}$$

where $VR_B(x_1, x_2, \ldots, x_m)$ = Estimation variance reduction at block B due to additional measurements at $x_1, x_2, \ldots, x_m$; m = number of additional sampling points; N = number of existing sampling points.

$$\gamma_{i,j} = \frac{1}{2}E\{[x_1 - x_j]^2\} = \gamma(|x_i - x_j|);\ \ \gamma_{k,B} \cong \frac{1}{V}\int_V \gamma(|x_k - x_j|)\,dx_j;$$

$$\alpha_k = \left[\gamma_{k,k} - a_k^T\, A_{k-1}^{-1}\, a_k\right]^{-1};\ \ A_k = \begin{bmatrix} A_{k-1} & a_k \\ a_k^T & \gamma_{k,k} \end{bmatrix};\ \ a_B^T = [1 \quad \gamma_{I,B} \quad \gamma_{II,B}\ \ldots\ \gamma_{N,B}];$$

$$a_k^T = [1 \quad \gamma_{I,k} \quad \gamma_{II,k}\ \ldots\ \gamma_{N,k}];\ \ i,j = I, II, ..., N;\ k = 1,2,3..., m;$$

V = Volume or area of the investigated block.

This nonlinear optimization problem is solved using the steepest ascent technique with the Bolzano search plan.[54] This method is an iterative process based on a search for the largest feasible gradient of the objective function at each iteration, followed by the optimal step size along this gradient. The process is repeated until the gradient becomes small enough.

In summary, Equation 18.20 is optimized by considering $X = (x_1, x_2, \ldots, x_m)$ to be the vector of decision variables, giving the coordinates of the selected additional sampling points. The steepest ascent requires an initial vector X^0 to be defined as the initial solution. Let X^k be the solution at the k-th iteration. X^{k+1} is generated and defined as

$$X^{k+1} = X^k + S^k \nabla f(X^k) \tag{18.21}$$

where $f(X^{k+1}) > f(X^k); if\ \nabla f(X^k) \neq 0$ and S^k is the optimal step size.

Determining the greatest feasible gradient requires ∇f to be computed at each iteration. This determination is made by the central-difference approximation. S^k is determined by maximizing $f(X^k + S^k \nabla f(X^k))$ through a one-dimensional search, as discussed by Hillier and Lieberman.[54] This procedure is conducted by determining whether $\partial f/\partial S^k$ is positive or negative in a trial solution. If $\partial f/\partial S^k$ is positive, then the optimal S must exceed the currently determined S. This latter S value becomes the lower bound. If, however, $\partial f/\partial S^k$ is negative, then the optimal S must be smaller than the currently determined S, which becomes the upper bound. At each iteration, the bounds become tighter. The resulting sequence of iterated solutions causes the upper and the lower bounds to converge to the optimal S. In practice, the iterations continue until the distance between the bounds is small enough. Several rules can be applied to select a new trial solution at each iteration. The one used herein is the Bolzano search plan,[54] which simply selects the midpoint between the upper and the lower bounds as the solution at the end of each iteration. The iterative process of steepest ascent will continue until two successive solutions X^k and X^{k+1} are approximately equal. In this case, $S^k\ \nabla f(X^k) \cong 0$, which can occur if the gradient vector, $\nabla f(X^k)$, is nearly null. Figure 18.1 shows the flowchart of this procedure.[1]

Meanwhile, in many instances, however, an existing sampling network may suffer from redundancy of information. The proposed MPVI procedure provides a mechanism for optimally identifying the m most redundant sampling locations. Accordingly, the objective function is defined in terms of increases in estimation variance of the designated block B, associated with the simultaneous deletion of m existing points. This minimization problem is defined as:

$$\begin{aligned}\text{Min}\ \ VI_B(x_N, x_{N-1}, ..., x_{N-m+1}) &= \alpha_{N-m+1}\left(\gamma_{N-m+1,B} - a^T_{N-m,B}\, w_{N-m+1}\right)^2 + \cdots \\ &+ \alpha_{N-1}\left(\gamma_{N-1,B} - a^T_{N-2,B} w_{n-1}\right)^2 + \alpha_N\left(\gamma_{N,B} - a^T_{N-1,B}\, w_N\right)^2\end{aligned} \tag{18.22}$$

where $VI_B\,(x_N, x_{N-1}, \ldots, x_{N-m+1})$ = Estimation variance increase at Block B due to the deletion of points at $x_N, x_{N-1}, x_{N-2}, \ldots, x_{N-m+1}$; m = number of deleted sampling points; N = number of existing sampling points;

$$\alpha_{N-k+1} = \left[\gamma_{N-k+1,N-k+1} - a^T_{N-k+1}\, A^{-1}_{N-k+1}\, a_{N-1}\right]^{-1};\ \ A_{N-k+1} = \begin{bmatrix} A_{N-k} & a_{N-k+1} \\ a^T_{N-k+1} & \gamma_{N-k+1,N-k+1} \end{bmatrix};$$

$$a^T_N = [1 \quad \gamma_{I,N} \quad \gamma_{II,N} \ \cdots \ \gamma_{N-1,N}];\ \ a^T_{N-k} = [1 \quad \gamma_{I,N-k} \quad \gamma_{II,N-k} \ \cdots \ \gamma_{N-k+1,N-k}];$$

$$a^T_{N-m,B} = [1 \quad \gamma_{I,B} \quad \gamma_{II,B} \ \cdots \ \gamma_{N-m,B}];\ \ a^T_{N-k+1,B} = \left[a^T_{N-k} \quad \gamma_{N-k+1,B}\right];\ \ a^T_{N,B} = \left[a^T_{N-1} \quad \gamma_{N,B}\right];$$

$$k = 1, 2, 3 ..., m.$$

The optimal MPVI attempts to drop the redundant sampling points from the set of existing monitoring points, which can be obtained through a rather limited search.[1] Consequently, a repetitive procedure is used to evaluate the effect of deleting all possible combinations of m out of N existing sampling points. The combination that yields the smallest increase in variance is then selected as the optimal set of deleted sampling points. In these analyses, the optimal MPVI is used to identify the best set of sampling points for deletion, resulting in the smallest increase in estimation variance of the investigated block. Moreover, the verifications of MPVR and MPVI under two hypothetical cases and a field case have been done by Lin and Rouhani.[1]

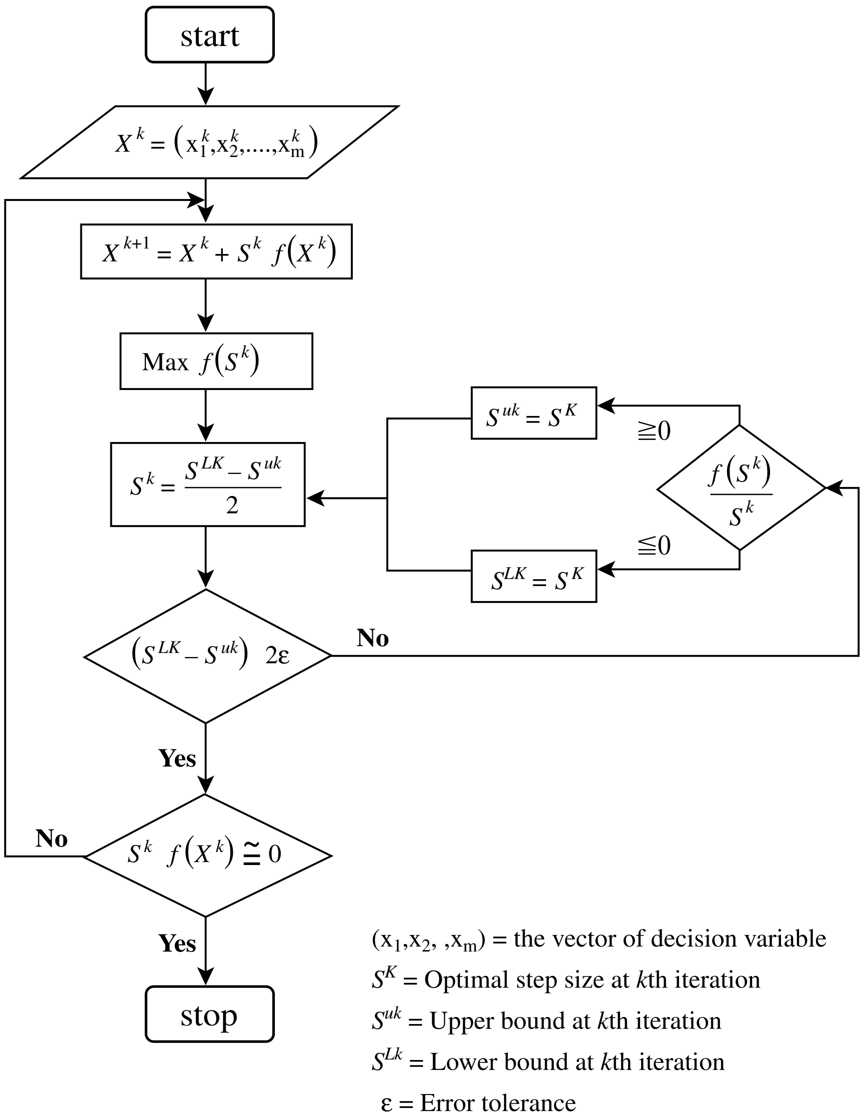


FIGURE 18.1 Steepest ascent procedure with Bolzano search plan. (From Lin, Y.P. and Rouhani, S., Multiple-point variance analysis for optimal adjustment of a monitoring network, *Environ. Monit. Assess.*, 69, 239, 2001. With permission from Kluwer Academic Publishers.)

18.3 CASE STUDY OF AN OPTIMAL ADJUSTMENT

In this chapter optimal MPVR is applied under a realistic hypothetical groundwater monitoring network. The surficial groundwater monitoring system at this site includes 14 wells, as shown in Figure 18.2. In this hypothetical case study, the surficial aquifer is contaminated and the plume is assumed to be time-stationary. The time-stationary

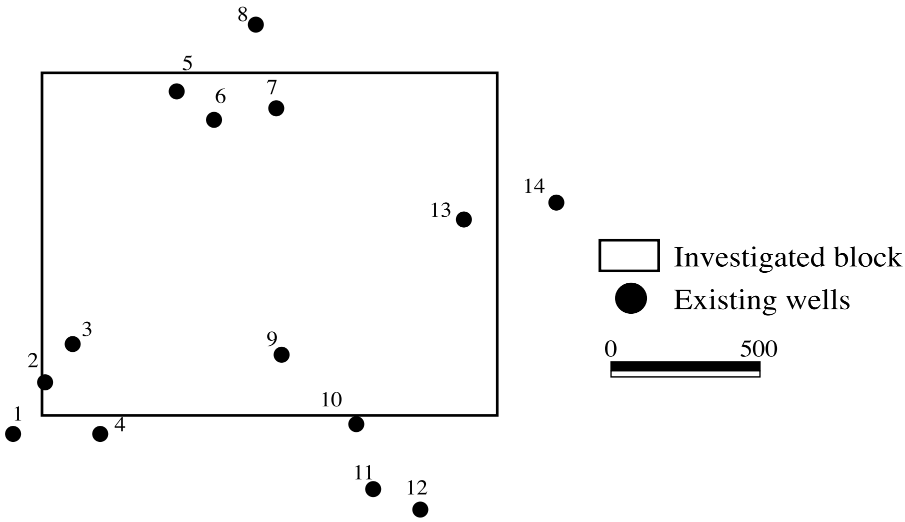


FIGURE 18.2 Existing groundwater monitoring network at the hypothetical site.

characteristic of the investigated contaminants was confirmed during analyses of available time series at various monitoring wells. The apparent stationary characteristic of plumes is assumed to be primarily attributable to the balance between advective and natural attenuation processes that control the fate and transport of solvent plumes. Groundwater in the water table aquifer flows in a southwestward direction and the horizontal gradient of this flow is assumed less than 0.002 m/m.

The use of optimal MPVR and MPVI requires the following information: (1) positions of existing monitoring locations, (2) geometry of the targeted block, and (3) a variogram of the investigated variable. Consequently, the variogram of the investigated variable is assumed to be a spherical model (Equation 18.23) with sill = 2.3 $(\mu g/l)^2$ and range = 720 m, and the size of the interesting area (Investigated block) is assumed to be 1600 m × 1200 m, as shown in Figure 18.2.

$$\gamma(h) = 2.3[1.5(h/720) - 0.5(h/720)^3] \tag{18.23}$$

18.3.1 MPVR and MPVI Applications

Using the variogram of the investigated variable, the location of an additional well with the highest variance reduction is at the east of the investigated block, as shown in Figure 18.3(a). The optimal MPVR is also applied to cases of multiple additional points. Figure 18.3(b) to Figure 18.3(e) show locations of two, three, four, and five additional monitoring wells. Importantly, the selected locations of one set do not include locations of other sets, obviously in Figure 18.3(c–e). For a notable example, neither of the five selected wells in Figure 18.3(d) are at the position of the four selected wells in Figure 18.3(e), implying that sequentially selecting monitoring locations (see, for example, Rouhani,[42] Rouhani and Hall,[43]

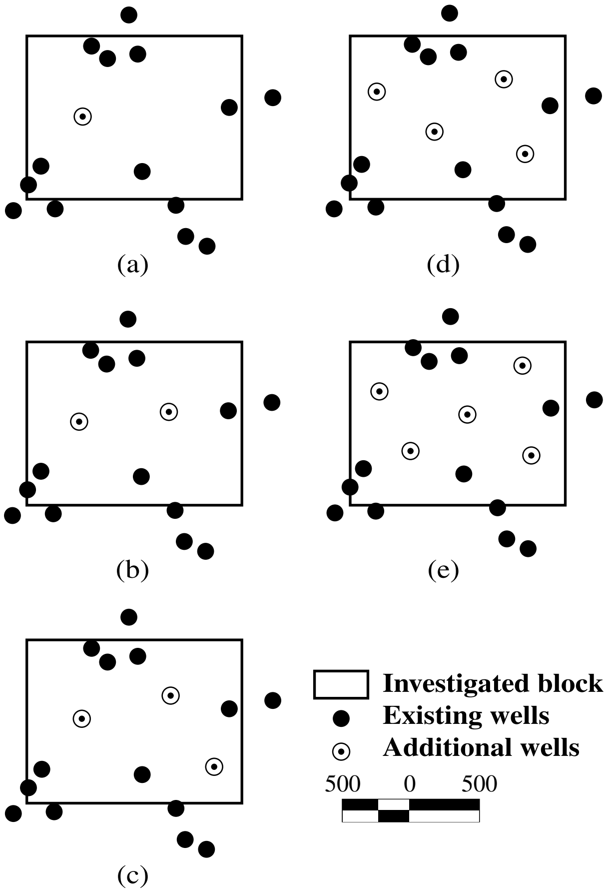


FIGURE 18.3 MPVR results: (a) location of one additional well; (b) locations of two additional wells; (c) locations of three additional wells; (d) locations of four additional wells; (e) locations of five additional wells.

and Lin and Rouhani[1]) does not necessarily yield optimal solution. Moreover, as expected, additional monitoring locations have a diminishing effect on the targeted estimation variance. Figure 18.4(a) shows the percentage of variance reduction as a function of the number of additional wells. As shown in Figure 18.4(b), the marginal variance reduction decreases as the number of additional wells increase.

MPVI is also applied to select the redundant monitoring wells in the hypothetical case. Using the variogram of the investigated variable, various sets of the five most redundant monitoring wells are Wells No. 2, No. 8, No. 11, No. 12, and No. 14, respectively. Figure 18.5 presents the percentage of variance increases as functions of the number of deleted wells. The deletions of Wells No. 2 and No. 8 do not increase the estimation variance of the investigated block, as displayed in Figure 18.5.

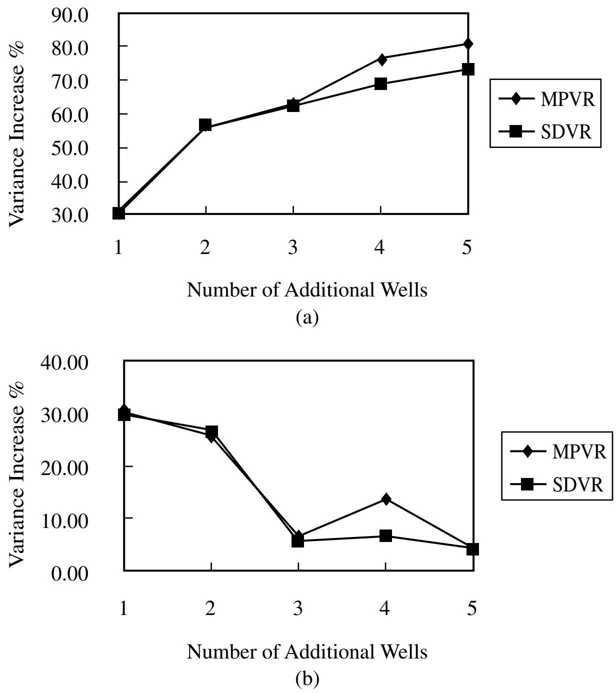


FIGURE 18.4 Optimal variance reduction results: (a) variance reduction percentage; (b) marginal variance reduction percentage.

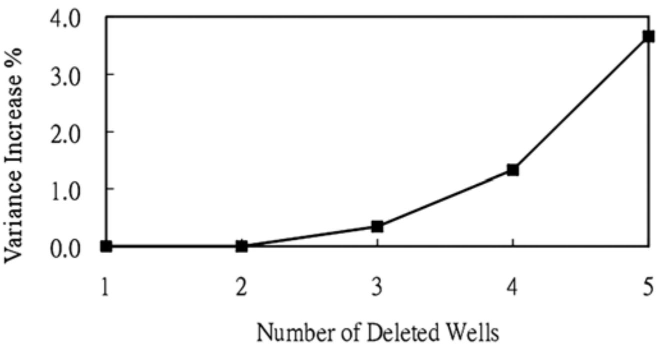


FIGURE 18.5 Optimal variance increase through deletion of redundant wells.

18.3.2 Information Efficiency

To compare the efficiency of the MPVR procedures, their results were compared to those of SDVR (sequential discretized variance reduction). SDVR[42] is a sequential procedure for selecting additional sampling points from a pool of potential sites. In this procedure, at each round, the location associated with the highest variance reduction is selected and added to the list of existing points. After the expanded data set is compiled, the process is repeated. All the additional points are taken from a predefined set of potential sampling sites which was a square 32 column × 24 row comprising 768 points at the investigated block. In SDVR process, the selected point can be identified from any arbitrary site in the investigated area. The selected point is then added to the list of existing locations and the process is repeated. Figure 18.6 shows the five additional wells selected using SDVR. The locations of additional wells selected using SDVR are different from those selected using the optimal MPVR, obviously in the cases of adding four and five additional wells, as illustrated in Figure 18.3 and Figure 18.6. However, Figure 18.4 clearly presents that the optimal MPVR always reduces variance more than SDVR. This advantage is related to the fact that the optimal MPVR requires neither artificial discretizations nor sequential selection.

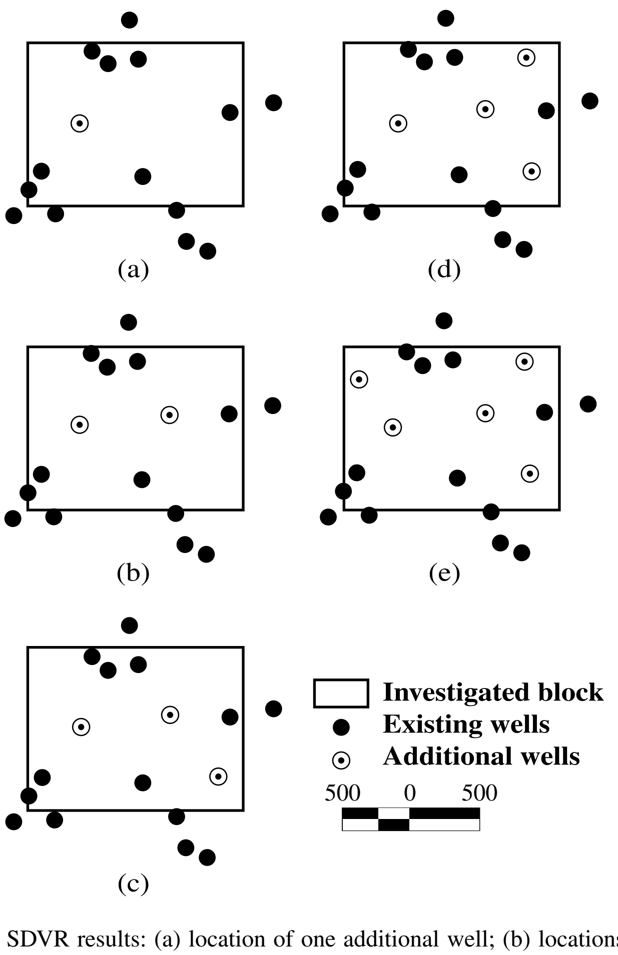


FIGURE 18.6 SDVR results: (a) location of one additional well; (b) locations of two additional wells; (c) locations of three additional wells; (d) locations of four additional wells; (e) locations of five additional wells.

18.3.3 Combined Optimal MPVI and MPVR

The optimal MPVR and MPVI should be applied together to improve the information provided by and cost-effectiveness of a monitoring network. This combined procedure allows the decision maker to locate redundant wells while determining data-gap locations.[1] The combined approach can be implemented in two stages: (1) deleting the redundant monitoring points through the optimal MPVI, and then (2) determining additional locations through the optimal MPVR.[1]

The above approach is applied to the hypothetical case. In the case based on the variogram, Wells No. 2 and No. 8 are the most redundant and are removed. These results imply that the estimation variance of the investigated block remains the same when Wells No. 2 and No. 8 are deleted. However, adding new monitoring wells easily compensates for the resulting reduction in the estimation variance of the investigated block. The locations of the wells are determined by the optimal MPVR. Figure 18.7(a) shows the location of one optimal additional well selected by the optimal MPVR after removing Well No. 2. In this case, the remaining 14 monitoring wells, the combined optimal MPVI and MPVR provide a new monitoring well

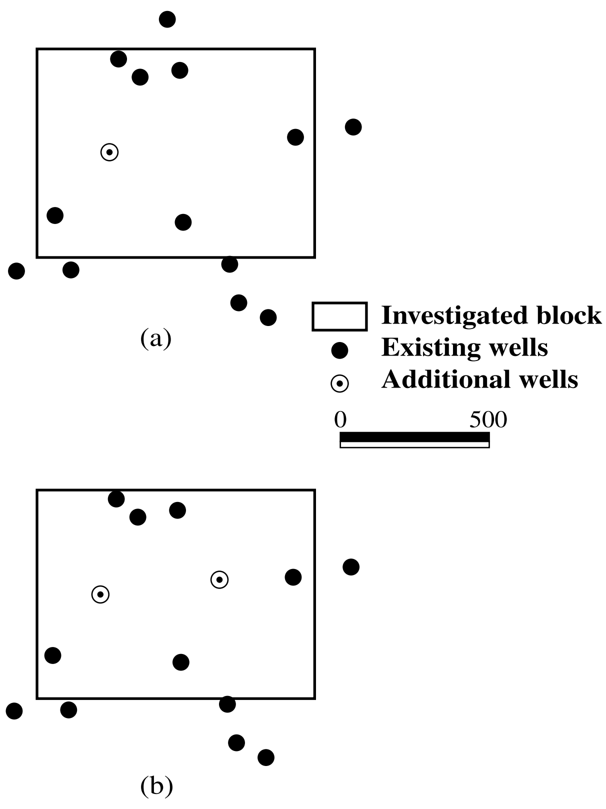


FIGURE 18.7 Locations of the wells: (a) deleting one redundant well and adding one additional well; (b) deleting two redundant wells and adding two additional wells.

system that can reduce the estimation variance of the investigated block by over 20.9% compared to the original monitoring well system.

Figure 18.7(b) shows the locations of two optimal additional monitoring wells selected by the optimal MPVR after deleting Wells No. 2 and No. 8. In this case, the remaining 14 monitoring wells, the combined optimal MPVI and MPVR, offer a new monitoring well system that can reduce the estimation variance of the investigated block by more than 56.0% compared to the original monitoring well system. However, using the same number of monitoring wells, the combined optimal MPVI and MPVR can provide efficient information.

18.4 SUMMARY AND CONCLUSION

The presented algorithms are geostatistical multiple-point selection procedures for maximizing variance reduction and minimizing variance increase for optimal adjusting monitoring network. The aim of the optimal MPVR is to select a set of additional sampling locations that produce the largest reduction in the estimation variance of a targeted block. Unlike most similar network design or adjustment algorithms, this nonsequential algorithm does not depend on any prior definition of the sampling locations, improving the information effectiveness of its results.

As a complementary component of the MPV, the optimal MPVI is designed to identify the set of redundant sampling points whose deletion yields the minimum increase in variance. The selection process is based on an exhaustive assessment of all possible combinations to determine the set of redundant points with the smallest variance. Such deletions can yield a cost-effective monitoring network while respecting appropriate information criteria. The combined optimal MPVR and MPVI are introduced to provide a tool for simultaneously adding and/or deleting monitoring points. This procedure generates optimal adjustments, leading to an information-effective/cost-effective monitoring plan.

REFERENCES

1. Lin, Y.P. and Rouhani, S., Multiple-point variance analysis for optimal adjustment of a monitoring network, *Environ. Monit. Assess.*, 69, 239, 2001.
2. Clifton, P.M. and Neuman, S.P., Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in Southern Arizona, *Water Resour. Res.*, 18, 1251, 1982.
3. Anderson, J. and Shapiro, A.M., Stochastic analysis of one-dimensional steady state unsaturated flow: a comparison of Monte Carlo and perturbation methods, *Water Resour. Res.*, 19, 121, 1983.
4. Neuman, S.P., Winter, C.L., and Newman, C.M., Stochastic theory of field-scale fickian dispersion in anisotropic porous media, *Water Resour. Res.*, 23(3), 453, 1987.
5. Hoeksema, R.J. and Kitanidis, P.K., Analysis of spatial structure of properties of selected aquifers, *Water Resour. Res.*, 21, 563, 1985.
6. Rubin, Y., Stochastic modeling of microdispersion in heterogeneous porous media, *Water Resour. Res.*, 26, 133, 1990.

7. Bjerg, P.L., Hinsby, K., Christensen, T.H., and Gravesen, P., Spatial variability of hydraulic of an unconfined sandy aquifer determined by a mini slug test, *J. Hydrol.*, 136, 107–122, 1992.
8. Federico, V.D., Flow in multiscale log conductivity fields with truncated power variogram, *Water Resour. Res*., 34, 975, 1998.
9. Salandin, P. and Fiorotto, V., Solute transport in highly heterogeneous aquifers, *Water Resour. Res.*, 34, 949, 1998.
10. Einax, J.W. and Soldt, U., Geostatistical and multivariate statistical methods for the assessment of polluted soils — merits and limitations, *Chemometr. Intell. Lab. Syst*., 46, 13, 1999.
11. Lin, Y.P. and Chang, T.K., Geostatistical simulation and estimation of the spatial variability of soil zinc, *J. Environ. Sci. Health A*., 35, 327, 2000.
12. Lin, Y.P., Chang, T.K., and Teng, T.P., Characterization of soil lead by comparing sequential Gaussian simulation, simulated annealing simulation and kriging methods, *Environ. Geol.*, 41, 189, 2001.
13. Pelissier, R. and Goreaud, F.A., Practical approach to the study of spatial structure in simple cases of heterogeneous vegetation, *J. Veg. Sci*., 12, 10, 2001.
14. Bergstrom, U., Englund, G., and Bonsdorff, E., Small-scale spatial structure of Baltic Sea zoobenthos-inferring processes from patterns, *J. Exp. Mar. Biol. Ecol*., 281, 14, 2002.
15. Battaglia, M.A., Mou, P., Palik, B., and Mitchell, R.J., The effect of spatially variable overstory on the understory light environment of an open-canopied longleaf pine forest, *Can. J. For. Res. — Rev. Can. Rech. For*., 32, 8, 2002.
16. Legendre, P., Dale, M.R.T., Fortin, M.J., Gurevitch, J., Hohn, M., and Myers, D., The consequences of spatial structure for the design and analysis of ecological field surveys, *Ecography*, 25, 15, 2002.
17. Lin, Y.P., Li, C.G., and Tan, Y.C., Geostatistical approach for identification of transmissivity structure at Dulliu area in Taiwan, *Environ. Geol.*, 40(1/2), 111, 2000.
18. Zhang, R., Myers, D.E., and Warrick, A.W., Estimation of the spatial distribution for soil chemicals using pseudo-cross-variograms, *Soil Sci. Soc. Am. J.*, 56, 1444, 1992.
19. Keck, T.J., Quimby, W.F., and Nielson, G.A., Spatial distribution of soil attributes on reconstructed mine soils, *Soil Sci. Soc. Am. J.*, 57, 782, 1993.
20. Samra, J.S. and Gill, H.S., Modelling of variation in a sodium-contaminated soil and associated tree growth, *Soil Sci. Soc. Am. J.*, 52, 1418, 1993.
21. Litaor, M.I., Spatial analysis of plutonium-239+240 and americium-241 in soils around Rocky Faults, Colorado, *J. Environ. Qual.*, 24, 506, 1995.
22. Zhang, R., Rahmani, S., Vance, G.F., and Munn, L.C., Geostatistical analyses of trace elements in soils and plants, *Soil Sci.*, 159, 383, 1995.
23. Steiger, B.V., Webster, R., Schulin, R., and Lehmann, R., Mapping heavy metals polluted soil by disjuntive kriging, *Environ. Pollut.*, 94, 205, 1996.
24. Couto, E.G., Stein, A., and Klamt, E., Large area spatial variability of soil chemical properties in central Brazil agriculture, *Ecosyst. Environ.*, 66, 139, 1997.
25. Juang, K.W. and Lee, D.Y., A comparison of 3 kriging methods using auxiliary variables in heavy-metal contaminated soils, *J. Environ. Qual.*, 27, 355, 1998.
26. Wang, X.J., Kriging and heavy-metal pollution assessment in waste water irrigated agricultural soil of Beijing Eastern Farming Regions, *J. Environ. Sci. Health A*., 33, 1057, 1998.
27. Wang, X.J. and Zhang, Z.P., A comparison of conditional simulation, kriging and trend surface analysis for soil heavy metal pollution pattern analysis, *J. Environ. Sci. Health A*., 34, 73, 1999.

28. Zhang, R., Shouse, P., and Yates, S., Estimates of soil nitrate distributions using cokriging with pseudo-crossvariogram, *J. Environ. Qual.*, 28, 424, 1999.
29. White, J.G., Welch, R.M., and Norvell, W.A., Soil zinc map of the USA using geostatistics and geographic information systems, *Soil Sci. Soc. Am. J.*, 61, 185, 1997.
30. Chang, T.K., Shyu, G.S., Lin, Y.P., and Chang, N.C., Geostatistical analysis of soil arsenic content in Taiwan, *J. Environ. Sci. Health A.*, 34, 1485, 1999.
31. Lin, Y.P., Chang, T.K., Shih, C.W., and Tseng, C.H., Factorial and indicator kriging methods using a geographic information system to delineate spatial variation and pollution sources of soil heavy metals, *Environ. Geol.*, 42, 900, 2002.
32. Dille, J.A., Milner, M., Groeteke, J.J., Mortensen, D.A., and Williams, M.M., How good is your weed map? A comparison of spatial interpolators, *Weed Sci*., 51, 44, 2003.
33. Monestiez, P., Courault, D., Allard, D., and Ruget, F., Spatial interpolation of air temperature using environmental context: application to a crop model, *Environ. Ecol. Stat.*, 8, 297, 2001.
34. Wang, Y., Ma, T., and Luo, Z., Geostatistical and geochemical analysis of surface water leakage into groundwater on a regional scale: a case study in the Liulin karst system, northwestern China, *J. Hydrol.*, 246, 223, 2001.
35. Hassen, M.B. and Prou, J., A GIS-based assessment of potential aquacultural nonpoint source loading in an Atlantic bay (France), *Ecol. Appl.*, 11, 800, 2001.
36. Mahapatra, S., Mahapatra, M., and Mondal, B., Phenol — an indicator of groundwater pollution by industrial effluents in Durgapur, West Bengal, *J. Geol. Soc. India.*, 59, 259, 2002.
37. Viau, A.E. and Gajewski, K., Holocene variations in the global hydrological cycle quantified by objective gridding of lake level databases, *J. Geophys. Res. — Atmos.*, 106, 31703, 2001.
38. Miller, J. and Franklin, J., Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence, *Ecol. Model.*, 157, 227, 2002.
39. Wylie, B.K., Meyer, D.J., Tieszen, L.L., and Mannel, S., Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands — a case study, *Remote Sens. Environ.*, 79, 266, 2002.
40. James, B.R. and Gorelick, S.M., When enough is enough: the worth of monitoring data in aquifer remediation design, *Water Resour. Res*., 30, 3499, 1994.
41. Christakos, G. and Olea, R., A multiple-objective optimal exploration strategy, *Math. Comput. Model.*, 11, 413, 1988.
42. Rouhani, S., Variance reduction analysis, *Water Resour. Res*., 21, 837, 1985.
43. Rouhani, S. and Hall, T.J., Geostatistical schemes for groundwater sampling, *J. Hydrol.*, 103, 85, 1988.
44. Loaiciga, H.A., An optimization approach for groundwater quality monitoring network design, *Water Resour. Res*., 25, 1771, 1989.
45. Hudak, P.F. and Loaiciga, H.A., An optimization method for monitoring network design in multilayered groundwater flow systems, *Water Resour. Res*., 29(8), 2835, 1993.
46. Mckinney, D.C. and Loucks, D.P., Network design for predicting groundwater contamination, *Water Resour. Res*., 28, 133, 1992.
47. Groenigen, J.W. and Stein, A., Constrained optimal of sampling using continuous simulated annealing, *J. Environ. Qual*., 27, 1078, 1998.
48. Groenigen, J.W., Siderius, W., and Stein, A., Constrained optimal of soil sampling for minimization of the kriging variance, *Geoderma*, 87, 239, 1999.

49. Journel, A.G. and Huijbregts, C.J., *Mining Geostatistics*, 5th ed., Academic Press, San Diego, CA, 1991, 600.
50. Deutsch, C.V. and Journel, A.G., *Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 1992, 340.
51. Cressie, C., The origins of kriging, *Math. Geol.*, 22, 239, 1990.
52. Rouhani, S., Lin, Y.P., Majid, A., and Shi, Y., *Geostatistical Analysis and Evaluation of Groundwater Quality Monitoring at Silverton Road Waste Unit*, EPDA Project 93029, Georgia Institute of Technology, Atlanta, GA, 1993.
53. Rouhani, S., Course Notes in Geostatistics: Theory, Practice, and Personal Computer Applications, Geostatistics Seminar, July 20–22, Augusta, GA, 1993.
54. Hillier, F.S. and Lieberman, G.J., *Introduction to Operations Research*, Holden-Day, Oakland, CA, 1986, 888.

19 The Variability of Estimates of Variance: How It Can Affect Power Analysis in Monitoring Design*

J.M. Carey

CONTENTS

19.1 INTRODUCTION

When designing a sampling program for a monitoring study in which frequentist hypothesis testing will be undertaken, it is essential that the program provides adequate statistical power to detect the effect of interest, should it exist. Conversely, it is also desirable to avoid wasting resources in excessive sampling which returns only minimal or no further improvement in sensitivity. Power analysis is a statistical tool which enables the determination of the number of samples needed to detect a certain effect size with a given confidence.[1] (For readers unfamiliar with the concept of statistical power, a brief introduction is provided as an appendix to this chapter.) A necessary input for power analysis is an estimate of the error variance appropriate for the statistical test to be used.[1] While it is recognized that "one must somehow guess at the variance, … unfortunately the results depend on having a fairly good

* This chapter is based on an article previously published in *Environmental Monitoring and Assessment*, 74: 225–241, 2002. Copyright 2002 Kluwer Academic Publishers, with permission.

1-56670-641-6/04/$0.00+$1.50

guess as to the actual value."[2] There are two recommended sources of variance estimates for use in power analysis:

- Pilot studies[3–5]
- Existing studies[6–7]

Whatever their origin, an important question underlying the use of variances in survey design is how reliable those estimates actually are. Are they close enough to the final value from the study to produce a satisfactory design? In a naturally variable system, it may be that variance estimates themselves are so variable as to be of limited use in survey design. Where the effect size is defined as an absolute value or as a percentage of the mean, rather than as a percentage of variance[7] or in the form of a signal-to-noise-ratio,[8] a poor estimate of variance can have a major effect on the calculation of the number of samples in the main study.

Variances (or standard deviations) of monitoring variables have been of interest for some time in ecology (e.g., freshwater biota,[9] zooplankton[10]), in chemistry (e.g., acid-volatile sulfide in sediments,[11] ion concentrations in soil solutions[12]), and in physical sciences (streamflows[13]). However, interest has often focused on precision or the number of samples needed to obtain a desired level of reliability. More recently, there have been considerations of variance specifically in relation to power analysis and sampling design, including studies on seagrass[14] and mercury contamination in fish and shellfish.[15]

19.2 CHOOSING THE CORRECT VARIANCE

In order to use power analysis as an aid to monitoring design, it is essential to decide what statistical test will be used to analyze the data and then identify the correct variance for use in the power calculations. For example, in an analysis of variance (ANOVA), the correct variance for calculating the power of an individual test is the mean square term used in the denominator of that particular F-ratio[16] (but there are additional considerations when the error variance is composed of more than one source of variation[7,17,18]). In the case of an orthogonal two-factor ANOVA, with factor A fixed and factor B random, factor B and the interaction would be tested over the residual mean square, but factor A would be tested over the interaction mean square. So power calculations for factor B and the interaction would be made using the residual mean square, while those for factor A would use the interaction mean square. In simple linear regression, the power of a t-test to detect a slope significantly different to zero may be determined using the variance of the slope.[19]

19.3 PILOT STUDIES AS A SOURCE OF AN ESTIMATE OF VARIANCE

Pilot studies preceding a main study are usually small scale programs, frequently with sampling undertaken at only single spatial or temporal scales, and often with minimal replication. Thus the variance estimates drawn from them may well be biased.[20] Green suggested that estimates of variance from pilot studies will tend to be large unless the size of the pilot study approximates that of the final study,[4] but it has also been suggested that pilot data from environmental impact designs such as the Before/After Control/Impact (BACI) designs[21–23] will probably underestimate variances,[7] presumably because inclusion of temporal variation is minimal or nonexistent.

In a pilot study, we know we are unlikely to obtain an estimate actually equal to the final value in the study proper.[24] What we do not know is where within the range of possible estimates our pilot value lies; that is, how close is our pilot estimate to the true or final value for our main study. If the estimate from the pilot study is substantially higher than the final value from the main study, then the sampling design may well have sufficient power to detect the effect size of interest. However, it may also prove to be more expensive and time-consuming than necessary, perhaps even prohibitively so. If the estimate of variance proves to be much lower than the final value, there will be a serious problem, with the final survey design having insufficient power to detect the effect size of interest.

19.3.1 A Case Study: Marine Infauna

A retrospective analysis of data from monitoring programs for two ocean outfalls in southern Australia explored the effects of variability in estimates of variance on monitoring design.[25] The two outfalls discharge wastewater off the Ninety Mile Beach in eastern Victoria, and abundance data for infauna (organisms living within the soft sediments) from the following sampling events were used in the study:

Saline Wastewater Outfall Program (SWOP)
- Five annual surveys from 1982 to 1986
- Three locations, spanning approximately 15 km of coastline
- Five replicate sites at each location
- Water depth: 3 to 15 m

Latrobe Valley Ocean Outfall (LVOO)
- Four annual surveys between 1989 and 1994
- Three locations, spanning nearly 50 km of coastline
- Eight replicate sites at each location
- Water depth: 13 to 18 m

Sampling was undertaken at the same time of year in both programs, using identical sampling methods, and there was spatial overlap of sampling locations for the two programs.

Soft-sediment infauna is noted for its patchy distribution in both space and time.[26–29] While the Ninety Mile Beach is an unusually uniform stretch of coastline for southern Australia, there were nonetheless some large-scale variations in sediment grain size within the study area, as well as the small-scale patchiness common in such habitats.[27] Major temporal fluctuations in the abundance of some species were recorded during the programs. For example, the bivalve *Mysella donaciformis* underwent an order of magnitude increase in abundance, from a mean of less than 300 individuals per m^2 in the years 1982 to 1986 to a peak of over 2500 per m^2 in 1987, and remained at high levels for several years before gradually falling. In contrast, the gastropod *Bankivia fasciata* underwent a relatively short-term bloom. From mean abundances of generally less than 15 individuals per m^2, it exceeded 8000 per m^2 at two locations during the designated monitoring period in the 1993 survey, but the blooms were no longer apparent on subsequent visits some 6 weeks later. Some temporal variations may be related to known cycles such as seasonality, and could thus be accommodated in sampling programs. However, others, such as the examples here, are apparently stochastic and thus unpredictable with our present knowledge of the marine biota. The case study data might thus be considered to represent a challenge in terms of designing a monitoring program.

For the purpose of examining the effects of variability of variances, two forms of ANOVA were used. The first is the MBACI design,[23] where fixed sampling times, surveys, are nested in the factor *BA* (i.e., before or after impact), and randomly selected locations are nested in the factor *CI* (i.e., control or impact). The test statistic for long-term impact is

$$F_{1,Locations-2} = \frac{MS_{BA\times CI}}{MS_{Locations(CI)\times BA}} \tag{19.1}$$

Thus the correct variance for term of interest for power calculations is the *Location* (*CI*) × *BA* interaction mean square. The second test was a simple two-way fixed-factor ANOVA, where the error term for the *Survey* × *Location* interaction was the residual mean square. Both ANOVAs were run using both raw and $\log_{10}$ transformed data, again for comparative purposes because it is unlikely that any single scale of measurement would be appropriate for all taxa.

Power calculations were made for 148 taxa (55 for SWOP taxa and 93 for LVOO), ranging from individual species to pooled totals for higher taxonomic levels. For each of the 148 cases, variances were calculated for all possible combinations of surveys — that is, for each individual survey, each pair of surveys, each set of three surveys, and so on, through to the full number of surveys which was considered to be the best estimate for the set. For the five surveys of the SWOP program, this produced a set of 31 estimates of variance, and for the four LVOO surveys, a set of 15 estimates.

The worst estimates of variance were usually obtained from single surveys, with the highest estimates in 92% of cases based on datasets which had no component of temporal variation. Scatter plots of error term against the number of contributing surveys were consistently tapered left to right, from a large range of estimates based on a single survey to the single value for the best estimate obtained with the maximum number of surveys (Figure 19.1). Plots for fixed-factor ANOVAs were generally wedge shaped, with estimates distributed fairly evenly above and below the best estimate (Figure 19.1a). For the MBACI ANOVAs, points on the upper half of the scatter plots formed a more concave pattern, while the lower halves were frequently compressed vertically (Figure 19.1b).

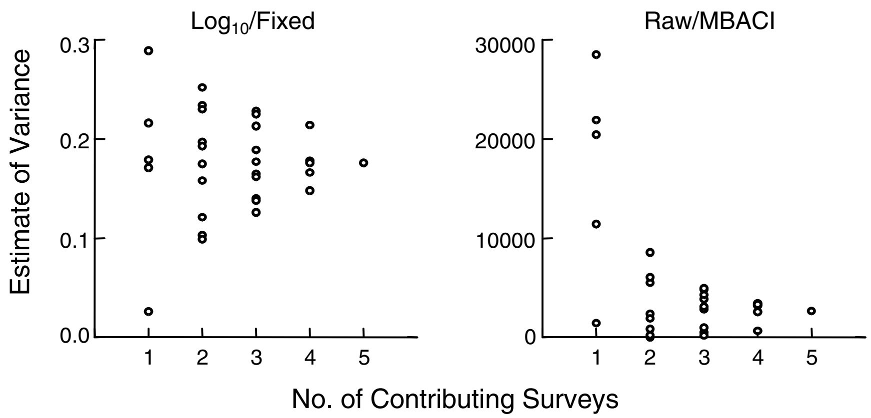


FIGURE 19.1 Estimates of variance vs. number of surveys contributing to calculation of the estimates. An example is the amphipod *Tipimegus thalerus* from the SWOP program. (From Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess*., 74, 225, 2002. With permission from Kluwer Academic Publishers.)

Error terms were grouped into five categories according to their magnitude relative to the best estimate:

VL (very low):	50% or less of the best estimate
L (low):	Between the best estimate and 50% less than it
E (equal):	Equal to the best estimate
H (high):	Between the best estimate and 50% greater than it
VH (very high):	50% or greater than the best estimate

This grouping revealed that in 43% of cases, error terms differed from their appropriate best estimate by 50% or more. There were generally more values lower than the best estimate rather than higher (Table 19.1).

The difference in the behavior of sets of error terms with analysis and transformation can also be seen in Table 19.1, where smaller percentages indicate better behaved sets of variances, that is, less variable sets with more estimates closer to the best estimate. The fixed-factor ANOVA on $\log_{10}$ transformed data clearly gave consistently better results than the other forms of analysis (4 to 7% of estimates), while the MBACI ANOVA on untransformed data produced most error terms furthest from the best estimate.

A typical distribution of error terms by category is shown in Figure 19.2. A good plot, with regard to use of the error terms for power analysis and sampling design, is one in which the majority of the error terms fall in the innermost categories, that is L, E, and H. One set of variances was considered better than another when a greater proportion of its error terms lay closer to the best estimate in category E, that is when the category score L+E+H-VL-VH was higher (Table 19.2). Some general patterns in the distribution of the error terms were evident from the category scores.

TABLE 19.1
Percentage of Estimates Which Differ from Their Best Estimate by 50% or More

Sampling Program	Transformation	ANOVA	VL	VH
SWOP	None	Fixed-factor	26	19
		MBACI	36	25
	Log_{10}	Fixed-factor	4	4
		MBACI	30	25
LVOO	None	Fixed-factor	27	18
		MBACI	36	23
	Log_{10}	Fixed-factor	7	6
		MBACI	30	24
Combined	Both	Both	24	18

Note: VL = 50% or less than best estimate. VH = 150% or more than best estimate. For each SWOP analysis, N = 55 taxa × 31 estimates = 1705. For each LVOO analysis, N = 93 taxa × 15 estimates = 1395.

Source: Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess*., 74, 225, 2002. With permission from Kluwer Academic Publishers.

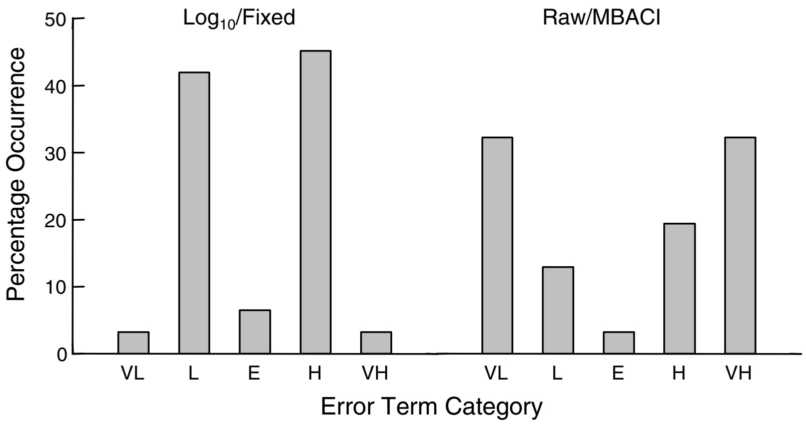


FIGURE 19.2 Percentage of error terms by category. An example is the amphipod *Tipimegus thalerus* from the SWOP program. VL = 50% or less of the best estimate, L = between best estimate and 50% less than it, E = equal to the best estimate, H = between the best estimate and 50% greater than it, VH = 50% or greater than the best estimate. (From Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ, Monit. Assess*., 74, 225, 2002. With permission from Kluwer Academic Publishers.)

TABLE 19.2
Percentage of Cases Where Analysis 1 > Analysis 2 for Category Score = (L + E + H − VL − VH)

	Fixed-Factor ≥ MBACI		Log_{10} ≥ Raw Data	
Sampling Program	**Raw Data**	**Log_{10}**	**Fixed-Factor**	**MBACI**
SWOP (N = 55)	78	100	95	58
LVOO (N = 93)	56	94	83	51
Combined (N = 148)	64	96	87	53

Source: Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess.*, 74, 225, 2002. With permission from Kluwer Academic Publishers.

Error terms from $\log_{10}$ transformed data gave better plots than those based on raw data in 70% of cases, presumably due to the dampening effect of the transformation on population peaks. This pattern was more marked for residuals from the fixed-factor ANOVA than for interactions from the MBACI ANOVA. The fixed-factor ANOVA gave better plots of error terms than the MBACI ANOVA in 80% of cases, a pattern most evident in the $\log_{10}$ transformed data. The difference in the quality of estimates is probably related to the nature of the calculation of the error terms. Irregular peaks in populations may be more evident among the locations and surveys incorporated in the MBACI error term than within them, as used for the fixed-factor calculation. Patterns such as these should be considered when assessing the confidence which should be placed in pilot estimates.

To allow quantitative comparisons of the variances for different taxa, error terms were also standardized for each transformation and analysis by dividing by the appropriate best estimate. When considering all standardized error terms from SWOP taxa as a single set for each taxon and analysis (N = 31), the taxa with the smallest variation among error terms, as measured by standard deviations, were not the same for all analyses (Table 19.3). Thus low variance for a taxon under one analytical scenario should not be taken to indicate that variance will also be low under another.

Although it might reasonably be expected that peaks in abundances of individual species may be smoothed out by pooling to higher taxonomic levels, changing the level of taxonomic resolution made no general difference to the closeness with which the calculated variances for infauna approximated their best estimates. In terms of power analysis, it would appear that choosing a particular level of taxonomic resolution would have had little intrinsic effect on a final survey design.

Other possible criteria for selecting variables for ecological monitoring, such as a higher taxonomic group (e.g., Mollusca, Polychaeta, Crustacea, Echinodermata), broad abundance categories ("abundant" = 64 or more organisms per m^2; "moderately abundant" = 16 to 63 organisms per m^2),[30] and miscellaneous criteria, such as ecological importance (e.g., with respect to the food chain) or ease of sorting and identification in the laboratory, produced no consistent patterns in the estimates of variance.

TABLE 19.3
Relative Rankings of Standard Deviations of Standardized Error Terms for Selected SWOP Taxa, Grouped by Analysis and Transformation

		Rank Order	Corresponding Rank in Other Analyses		
		Fixed-Factor	Fixed-Factor	MBACI	
Taxon		Raw	Log_{10}	Raw	Log_{10}
Anthuridae	(C)	1	30	52	53
Apanthura isotoma	(C)	2	24	51	54
Chaetozone platycerca	(P)	3	48	46	37
Amphipoda	(C)	4	8	49	46
Marginellidae	(M)	5	26	12	13
Corophiidae	(C)	51	1	38	32
Mollusca		52	47	24	3
Bivalvia	(M)	53	53	28	10
Mysella donaciformis	(M)	54	21	32	36
Cirolana woodjonesi	(C)	55	55	30	9

Note: Ranks are in order of increasing standard deviation for fixed-factor ANOVAs on raw data, while those for the other analyses indicate the rank order of the taxon within that transformation/analysis; M = Mollusca, P = Polychaeta, C = Crustacea

Source: Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess.*, 74, 225, 2002. With permission from Kluwer Academic Publishers.

Overall, the estimates of variance calculated from the case study data proved to be very variable. While they showed some pattern with the type of analysis used and the data transformation applied, they were not readily predictable *a priori* on the basis of characteristics such as taxonomic resolution or higher taxonomic group.

The effect on calculations of sample size of using different estimates of variance is well known in general terms,[1,16] but it is perhaps worthwhile considering a quantitative example to emphasize the consequences of poor estimates. In Table 19.4, power calculations for an MBACI ANOVA on a species of moderate abundance are shown. Use of an estimate of variance half the size of the best estimate resulted in a calculated sample size, in this case the number of locations — 40% less than would really be required to detect the nominated effect with the desired confidence. At the end of the monitoring program, the design would be found to be inadequate for its purpose. Conversely, the use of a larger estimate resulted in more locations than necessary, meaning 60% more field sampling and laboratory processing, both of which can be very time consuming and/or expensive activities.

TABLE 19.4
Effect of Variance on Sample Size, Using an MBACI ANOVA

Estimate of Variance ($MS_{(Location(CI) \times BA)}$) Description	Value	Sample Size (No. of Locations)	Change in Sample Size
Best estimate	946	8	—
Best est. × 0.5	473	5	× 0.6
Best est. × 1.5	1419	10	× 1.3
Best est. × 2.0	1892	13	× 1.6

Note: Analysis is to detect an absolute difference of 50% of the mean number of individuals per site, with 90% confidence and alpha = 0.10, and 5 sites sampled at each location. Sample size is rounded to the next highest integer. Example is the isopod *Apanthura isotoma* from SWOP program, with mean abundance = 34 and best estimate of variance = 946.

Source: Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess*., 74, 225, 2002. With permission from Kluwer Academic Publishers.

19.4 EXISTING STUDIES AS A SOURCE OF AN ESTIMATE OF VARIANCE

Existing studies may also be a source of estimates of variance, possibly overcoming the problems of scale often evident in pilot studies. However, an existing study, in a comparable habitat and using similar variables, may not always be conveniently available.

19.4.1 A CASE STUDY

If variance estimates from an existing study were to be used in designing a new program, then the Ninety Mile Beach situation would seem almost ideal. If we were at the planning stage of the LVOO program, with access to data from the earlier SWOP program, we would find an existing study of the same habitat within the intended study area undertaken 3 to 7 years earlier in only slightly shallower water and using identical sampling methods.

It was possible to compare the best estimates of variance from the two sampling programs for 43 taxa common to both. Correlations between SWOP and LVOO taxa were statistically significant for both analyses and transformations (Table 19.5). However, while the estimates from the earlier SWOP study proved to be reasonable ones for some LVOO taxa, for 25% of taxa the SWOP estimates differed from those of LVOO by more than one order of magnitude, and in a few cases by as much as four (Figure 19.3). Differences of this magnitude would have a major effect on calculations of sample size for the intended program. The worst differences resulted from large natural increases in abundance in a single survey or location, such as described earlier for the gastropod *Bankivia fasciata*. Thus, it is clear that even a relatively large scale existing study, well matched to the study being planned, will not necessarily provide accurate estimates of variance.

TABLE 19.5
Correlation between Best Estimates of Variance for SWOP and LVOO Monitoring Programs (N = 43)

Transformation	ANOVA	r	p
Raw data	Fixed-factor	0.734	<.001
	MBACI	0.336	0.028
Log_{10} transformed	Fixed-factor	0.385	0.011
	MBACI	0.506	0.001

Source: Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess.*, 74, 225, 2002. With permission from Kluwer Academic Publishers.

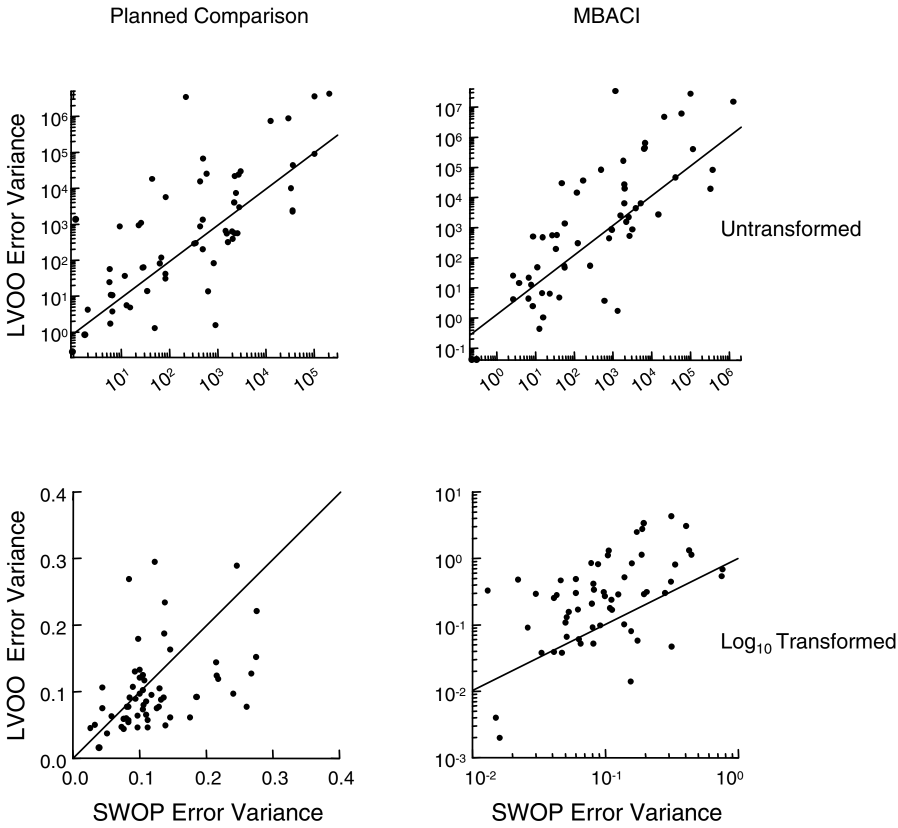


FIGURE 19.3 Best estimates of error variances for SWOP and LVOO programs by data transformation and type of analysis. The line $y = x$ on each graph indicates equivalence of the SWOP and LVOO estimates. (From Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess.*, 74, 225, 2002. With permission from Kluwer Academic Publishers.)

19.5 STRATEGIES FOR REDUCING VARIATION AND DEALING WITH UNCERTAINTY

Variability in variances is an unfortunate fact that designers of monitoring programs must face. Variances may be large and/or unpredictable, not only because of natural fluctuations of the variable through time and space but also as a result of other sources of epistemic uncertainty such as measurement error, systematic error, and uncertainty arising from the use of statistical models.[31,32] The following strategies are suggested as possible means of reducing variance due to the latter causes or when no further reduction is possible, confronting the problem rather than simply ignoring it and keeping one's fingers crossed!

Reducing variation:

1. When discussing the philosophy and rationale of power analysis in environmental monitoring, Green recommended employing sampling strategies that can minimize variation.[31] For example, using a stratified sampling design where natural strata exist in the population to be sampled will result in smaller sample variances than would be the case if simple random sampling was applied across the entire study area.[33] Where data are to be tested using ANOVA, a variance-minimizing strategy may involve the incorporation of additional factors that further reduce the unexplained variation or, in the case of regression analysis, the incorporation of additional explanatory or predictor variables.[16]
2. It is important to match the sampling scale of the monitoring program with the expected scale of potential environmental change because mismatches can result in larger variances than would otherwise be the case.[34]
3. When sampling ecological response variables, use a size and number of spatial sampling units appropriate to the spatial aggregation of the organisms.[3,35]
4. When an estimate of variation is to be obtained from a pilot study, it is desirable to undertake the pilot study on the same spatial and temporal scales as the proposed monitoring program to obtain a more appropriate measure of variation.[23,36] On the other hand, pilot studies are intended as smaller, cheaper versions of the main study to follow and should not be allowed to drain resources from that study.[3] The results of the case study suggest that a compromise position may be worthwhile, adding at least a second sampling time to a pilot program to hopefully remove the most extreme of the estimates of variance.[25]
5. Given a choice of otherwise equally desirable response variables for monitoring, select those with small variances[31,34] that are relatively stable.

Dealing with remaining variation:

1. In a study of cost benefit analyses using simulated data, McArdle and Pawley identified circumstances where pilot study variances were likely

to be unreliable, and recommended that a little critical judgment be used when applying them.[37] If a researcher has a good working knowledge of a particular system and obtains pilot data which seem extreme or atypical, modification of the inputs to power analysis for design purposes in the light of that researcher's experience would seem a sensible approach,[25] notwithstanding the additional uncertainty that arises with the application of such subjective judgment.[32]

2. Even better than relying on a single estimate of variance for a statistical test would be the use of plausible lower and upper values for the variance to provide a range of sample sizes and resulting power estimates as an aid to informed decision making.[36,38]

Finally, a long-term strategy for improving the performance of monitoring programs is to improve our understanding of the natural variability of variances associated with monitoring variables, which will in turn aid in the critical evaluation of estimates of variance before applying them to power analysis for the purpose of survey design.

ACKNOWLEDGMENTS

The author thanks Mick Keough for his collaboration on the paper on which this chapter is based, Gippsland Water for their permission to use the case study data in that paper, and Kluwer Academic Publishers for permission to use previously published material.

REFERENCES

1. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.
2. Eberhardt, L.L., Appraising variability in population studies, *J. Wildl. Manage.*, 42, 207, 1978.
3. Andrew, N.L. and Mapstone, B.D., Sampling and the description of spatial pattern in marine ecology, *Oceanogr. Mar. Biol. Annu. Rev.*, 25, 39, 1987.
4. Green, R.H., Power analysis and practical strategies for environmental monitoring, *Environ. Res.*, 50, 195, 1989.
5. Silsbee, D.G. and Peterson, D.L., Planing for implementation of long-term resource monitoring programs, *Environ. Monit. Assess.*, 26, 177, 1993.
6. Osenberg, C.W. et al., Detection of environmental impacts: natural variability, effect size and power analysis, *Ecol. Applic.*, 4, 16, 1994.
7. Keough, M.J. and Mapstone, B.D., Protocols for designing marine ecological monitoring programs associated with BEK Mills, National Pulp Mills Research Program, Technical Report No. 11. CSIRO, Canberra, Aust., 1995, chap. 6.
8. Page, D.S. et al., Shoreline Ecology Program for Prince William Sound, Alaska, following the Exxon Valdez oil spill: Part 1 — study design and methods, in *Exxon Valdez Oil Spill: Fate and Effects in Alaskan Waters*, Wells, P.G., Butler, J.N., and Hughes, J.S., Eds., ASTM STP 1219, American Society for Testing and Materials, Philadelphia, PA, 1995, 263.

9. Cassie, R.M., Sampling and statistics, in *A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters*, Edmondson, W.T. and Winberg, G.G., Eds., International Biological Programme Handbook No. 17, Blackwell Scientific Publications, Oxford, 1971, chap. 4.
10. Alden, R.W., III, Dahiya, R.C., and Young, R.J., Jr., A method for the enumeration of zooplankton subsamples, *J. Exp. Mar. Biol. Ecol.*, 59, 185, 1982.
11. Mackey, A.P. and Mackay, S., Spatial distribution of acid–volatile sulphide concentration and metal bioavailability in mangrove sediments from the Brisbane River, Australia, *Environ. Pollut.*, 93, 205, 1996.
12. Manderscheid, B. and Matzner, E., Spatial heterogeneity of soil solution chemistry in a mature Norway spruce (*Picea abies* (L.) Karst.) stand, *Water Air Soil Pollut.*, 85, 1185, 1995.
13. Poff, N.L. and Ward, J.V., Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns, *Can. J. Fish. Aquatic Sci.*, 46, 1805, 1989.
14. Heidelbaugh, W.S. and Nelson, W.G., A power analysis of methods for assessment of change in seagrass cover, *Aquat. Bot.*, 53, 227, 1996.
15. Nicholson, M.D., Fryer, R.J., and Ross, C.A., Designing monitoring programmes for detecting temporal trends in contaminants in fish and shellfish, *Mar. Pollution Bull.*, 34, 821, 1997.
16. Quinn, G.P. and Keough, M.J., *Experimental Design and Data Analysis for Biologists.* Cambridge University Press, Cambridge, U.K., 2002, chap. 7.
17. Winer, B.J., Brown, D.R., and Michels, K.M., *Statistical Principles in Experimental Design*, 3rd ed., McGraw-Hill, New York, 1991, chap. 5.
18. Tiku, M.L., A note on the distribution of the doubly non-central *F*-distribution, *Aust. J. Stat*., 14, 37, 1972.
19. Neter, J., Wasserman, W., and Kutner, M.H., *Applied Linear Statistical Models: Regression*, *Analysis of Variance*, *and Experimental Designs*, 3rd ed., Richard D. Irwin, Boston, MA, 1990, chap. 3.
20. Fairweather, P.G., Statistical power and design requirements for environmental monitoring, *Aust. J. Mar. Freshwater Res.*, 42, 555, 1991.
21. Green, R.H., *Sampling Design and Statistical Methods for Environmental Biologists*, Wiley-Interscience, New York, 1979, p. 30.
22. Underwood, A.J., On beyond BACI: sampling designs that might reliably detect environmental disturbances, *Ecol. Appl.*, 4, 3, 1994.
23. Keough, M.J. and Mapstone, B.D., Designing environmental monitoring for pulp mills in Australia, *Water Sci. Technol.*, 35(2–3), 397, 1997.
24. Mapstone, B.D., Scalable decision rules for environmental impact studies: effect size, Type I, and Type II errors, *Ecol. Appl.*, 5, 401, 1995.
25. Carey, J.M. and Keough, M.J., The variability of estimates of variance, and its effect on power analysis in monitoring design, *Environ. Monit. Assess*., 74, 225, 2002.
26. Buchanan, J.B. and Moore, J.J., A broad view of variability and persistence in the Northumberland benthic fauna — 1971–85, *J. Mar. Biol. Assoc. U.K.*, 66, 641, 1986.
27. Thrush, S.F., Spatial patterns in soft-bottom communities, *Trends Ecol. Evol.*, 6, 75, 1991.
28. Morrisey, D.J. et al. Spatial variation in soft-sediment benthos, *Mar. Ecol. Prog. Ser.*, 81, 197, 1992.
29. Morrisey, D.J. et al. Temporal variation in soft-sediment benthos, *J. Exp. Mar. Biol. Ecol.*, 164, 233, 1992.

30. Pearson, T.H., Gray, J.S., and Johannessen, P.J., Objective selection of sensitive species indicative of pollution-induced change in benthic communities. 2. Data analyses, *Mar. Ecol. Prog. Ser.*, 12, 237, 1983.
31. Green, R.H., Aspects of power analysis in environmental monitoring, in *Statistics in Ecology and Environmental Monitoring*, Fletcher, D.J. and Manly, B.F.J., Eds., Otago Conference Series 2, University of Otago Press, Dunedin, NZ, 1994, p. 173.
32. Regan, H.M., Colyvan, M., and Burgman, M.A., A taxonomy and treatment of uncertainty for ecology and conservation biology, *Ecol. Appl.*, 12, 618, 2002.
33. Gilbert, R.O., *Statistical Methods for Environmental Pollution Monitoring*, John Wiley & Sons, New York, 1987, chap. 5.
34. Downes, B.J. et al., *Assessing Ecological Impacts: Applications in Flowing Waters*, Cambridge University Press, Cambridge, U.K., 2002, chap. 6.
35. Schneider, D.C., *Quantitative Ecology: Spatial and Temporal Scaling*, Academic Press, San Diego, CA, 1994, chap. 10.
36. Downes, B.J. et al., *Assessing Ecological Impacts: Applications in Flowing Waters*, Cambridge University Press, Cambridge, U.K., 2002, chap. 13.
37. McArdle, B.H. and Pawley, M.D.M., Cost benefit analysis in the design of biological monitoring programs: is it worth the effort?, in *Statistics in Ecology and Environmental Monitoring*, Fletcher, D.J. and Manly, B.F.J., Eds., Otago Conference Series 2, University of Otago Press, Dunedin, New Zealand, 1994, 239.
38. Bernstein, B.B. and Zalinski, J., An optimum sampling design and power tests for environmental biologists, *J. Environ. Manage*., 16, 35, 1983.
39. Neyman, J. and Pearson, E.S., On the use and interpretation of certain test criteria for purposes of statistical inference, Part I, *Biometrika*, 20A, 175, 1928.
40. Toft, C.A. and Shea, P.J., Detecting community-wide patterns: estimating power strengthens statistical inference, *Am. Nat.*, 122, 618, 1983.
41. Suter, G.W., II, Abuse of hypothesis testing statistics in ecological risk assessment, *Human Ecol. Risk Assess.*, 2, 331, 1996.
42. Zar, J.H., *Biostatistical Analysis*, 3rd ed., Prentice-Hall International, Upper Saddle River, NJ, 1996.
43. Downes, B.J. et al., *Assessing Ecological Impacts: Applications in Flowing Waters*, Cambridge University Press, Cambridge, U.K., 2002, p. 91.
44. Fairweather, P.G., Links between ecology and ecophilosophy, ethics, and the requirements of environmental management, *Aust. J. Ecol.*, 18, 3, 1993.
45. Cowles, M. and Davis, C., On the origins of the .05 level of statistical significance, *Am. Psychol.*, 37, 553, 1982.
46. Fisher, R.A., The arrangement of field experiments, *J. Min. Agric.*, 33, 503, 1926.
47. Peterman, R.M., Statistical power analysis can improve fisheries research and management, *Can. J. Fish. Aquat. Sci.*, 47, 2, 1990.
48. Murphy, K.R. and Myors, B., *Statistical Power Analysis: a Simple and General Model for Traditional and Modern Hypothesis Tests*, Lawrence Erlbaum & Assoc., Mahwah, NJ, 1998.
49. Thomas, L. and Krebs, C.J., A review of statistical power analysis software, *Bull. Ecol. Soc. Am.*, 78(2), 126, 1997.

APPENDIX: AN INTRODUCTION TO STATISTICAL POWER ANALYSIS*

For any statistical test, there always exists some probability of the test not correctly identifying the true situation, and the two ways in which this might occur are known as Type I and Type II errors[39] (Table 19.6). The Type I error is the probability of claiming to detect a difference when one does not exist. It has traditionally been the focus of attention in hypothesis testing.[31,40,41] Tabulations of central distributions for common test statistics to determine the probability of the statistic occurring when the null hypothesis is true are readily available in statistical texts such as that by Zar.[42] In contrast, a Type II error is the probability of failing to detect a true difference. Historically, Type II errors have been given little attention, largely because of the focus on the Type I error rate, but also because the relevant noncentral distributions have only become widely accessible with the advent of personal computers.[43] It is only in the last decade or two that Type II errors have become of interest to ecologists[4] and others involved in environmental science.

Statistical power is the probability of detecting a true difference and is defined as the complement of the Type II error rate; that is, $power = 1 - \beta$. The power of a test is dependent on the variation in the datasets (s), the sample size (n), the chosen Type I error rate (α), and the magnitude of difference to be detected (ES, the effect size) as follows[16]:

$$power \propto \frac{ES\alpha\sqrt{n}}{s} \tag{19.2}$$

This relationship can be rearranged to solve for different terms,[20] such as

1. The sample size required to detect a certain effect size with known confidence
2. The power of a completed test to detect an effect of a nominated size
3. The minimum detectable effect size given sample size and variance

TABLE 19.6
Possible Errors in Hypothesis Testing

		Test Result	
		Significant (H_0 Rejected)	**Not Significant (H_0 Not Rejected)**
Reality	**Difference (H_0 false)**	Correct	Type II error (β)
	No difference (H_0 true)	Type I error (α)	Correct

* References cited in this appendix are included in the main list for the chapter.

This last use effectively places bounds on how large an effect could go undetected.[44] While α is conventionally set at 0.05,[16] the value proposed by Fisher in 1925 as a convenient limit,[45] it can be varied *a priori* to suit the particular situation.[39,46] This flexibility creates the potential for relative weighting of the two error rates with regard to the consequences arising from each type of error.[24,40,47]

The calculations of power analysis vary with the type of statistical test under consideration, and examples for common tests such as t-tests, analysis of variance, and simple linear regression can be found in more recent editions of standard statistical text books such as that by Zar,[42] as well as in texts dealing specifically with power analysis, for example, those by Cohen[1] and Murphy and Myors.[48] Power analysis is also available in statistical software packages, as add-ins to spreadsheets, and in specialized software for power analysis. Thomas and Krebs reviewed statistical power analysis software in 1997,[49] and their review is also available online at http://www.zoology.ubc.ca/~krebs/power.html.

20 Discriminating between the Good and the Bad: Quality Assurance Is Central in Biomonitoring Studies

G. Brunialti, P. Giordani,
and M. Ferretti

CONTENTS

1-56670-641-6/04/$0.00+$1.50

20.1 INTRODUCTION

Variability is an inherent property of ecological systems and every attempt to measure and interpret the environment should consider it. Unfortunately, not only variability in the system is of concern for those involved in environmental monitoring, an entire range of actual and potential sources of variability connected to the survey design, methods, and operators should be taken into account.[1–3] Differences between methods, difference in the application of the same method, measurement error, sampling and nonsampling error, and errors related to model applications are all terms of the whole error budget that inevitably affects environmental surveys.[4] In this perspective, the extent to which the objective of the survey is matched depends very much on the ability to manage the various sources of variability.[1,5] While such a management is complex, it always depends on the documentation of the various steps of the investigations. Documentation allows tracking all the steps undertaken to carry out the investigation of concern and helps in identifying when and where problems occur. However, documentation can be properly achieved only by adequate quality assurance (QA).[1,6–8] Politicians, administrators, and decision makers may be not very interested to know with what degree of confidence a certain population parameter was estimated by the survey they are presenting to the public; for example, colored maps showing the spatial variation of lichen diversity as indicator of air pollution are usually much more attractive than statistical details. However, their attitude may change considerably when the survey results are used, for example, to stop a (supposed) harmful power plant and the owners of the plant challenge (in terms of money) such a decision in the court. In this case, every statistical detail about the accuracy and precision of the survey data will be very much welcome. This example suggests that, if biomonitoring should be taken as a serious basis for decision making, it needs to produce robust, defensible data of documented quality. In short, biomonitoring needs QA and it should make the differences between "good" (e.g., of documented and therefore known quality) and "bad" (e.g., undocumented and therefore of unknown quality) monitoring programs.

The aim of this chapter is to recall the basic QA procedures, emphasize the need for a formal design in biomonitoring studies, and provide examples of data quality control in various fields of environmental biomonitoring, with special reference to air pollution monitoring by means of lichens, sensitive tobacco plants, and spontaneous vegetation.

20.2 QUALITY ASSURANCE

Some definitions of the main activities that have to be carried out in all phases of a biomonitoring program in order to assess the quality of the data are reported here. Reference will be made to these phases in the rest of the section using the abbreviations given here.

QA is an organized group of activities defining the way in which tasks are to be performed to ensure an expressed level of quality.[9] The main benefit of a QA plan is the improved consistency, reliability, and cost-effectiveness of a program through time.[8] A QA plan is essential since it forces program managers to identify and evaluate most of the factors involved in the program. In addition, the assessment of data quality enables mathematical management of uncertainty due to the method used.[8] Cline and Burkman[6] consider four main activities in a QA program which take all the steps of the monitoring survey into account:

1. Quality Management (QM). This concerns the proper design of the project and its documentation. It ensures that the proper activities are performed in the proper way. QM activities include, for example, the choice of the proper sampling strategy to be adopted (i.e., where and how the sampling stations have to be located).
2. Quality Assurance (QA). This concerns the first steps of evaluation of the quality of the data. It includes the use and documentation of standard operating procedures. All the activities defined in the sampling protocol are examples of QA procedures. In the case of lichen monitoring, these activities include the selection of sampling subplots, the selection of standard trees, and the positioning of the sampling grid on the tree.
3. Quality Control (QC). This concerns mostly the training, calibration, and control phases. It ensures that data are collected appropriately and that QA is carried out.
4. Quality Evaluation (QE). This concerns mainly the statistical evaluation of the data quality. These activities enable the precision and accuracy of the data collected by the operators to be evaluated, providing the basis for comparability of the data.

20.3 ERRORS

Environmental data for large areas are generally assessed by sample-based methods. The objective of a sample-based survey is to select a subset — the sample — from the population of interest and to estimate population parameters based on probability theory.[10] Obviously, these parameter estimates differ from the true population as they are subject to different sources of errors.[5]

Errors can be classified into four major categories:[2,5,11]

1. Sampling errors: These errors are generated by the nature of the sampling itself and by the degree of data variability. In general, sampling errors can be reduced by increasing the sample size and by introducing a more cost-efficient sampling design.[2,5,11]
2. Assessment errors: These errors incorporate measurement and classification errors. They can occur when the methodology is poorly standardized, when insufficient care is devoted to its application, or when there are problems with instrument calibration.[2]
3. Prediction errors: Many attributes in environmental resources assessment are not directly assessed but derived by models. In this case it is assumed that based on the input values the true population value is derived. Models

and functions, however, are subject to errors, which are defined prediction errors (see Kohl et al.[5] for a complete review).

4. Nonstatistical errors: These kinds of error include human errors and may affect the data quality in all the phases of the sampling. They are frequent, ubiquitous, and can be very serious. They usually originate from errors in measurement, sampling, and/or data processing. Examples are mistakes in data entry, programming errors, and errors in defining the sample frame.[2]

In order to improve the interpretation of survey results and to review the benefit of the retrieved information, the total error of estimates has to be quantified. Some authors introduced the terms "error budget"[5,12,13] or "total sampling design"[14] to define this parameter. The error budget provides a calculation of the total error affecting the survey estimates, which can be achieved by a mathematical model that accounts for the various error sources. A general parameter of the studied population adopted to calculate the total survey error is the mean square error of an estimate. Köhl et al.[5] report the following formula by Kish[12]:

$$\mathrm{MSE}(\bar{y}) = \sum_r S_r^2 + \left(\sum_r B_r\right)^2 \tag{20.1}$$

where $\mathrm{MSE}(\bar{y})$ is the mean square error, $\sum_r S_r^2$ is the sum of all variance terms (S_r) from multiple error sources, and ($\sum_r B_r$) the squared sum of the biases (B_r)

20.4 MONITORING IN A VARIABLE ENVIRONMENT NEEDS PROPER SAMPLING DESIGN

According to Yoccoz et al.,[15] we can define monitoring as the process of gathering information about some system-state variables at different points in time for the purpose of assessing the state of the system and making inferences about changes in state over time. If our focus is on the monitoring of biological diversity, the systems of interest to us are typically ecosystems or components of such systems (e.g., communities and populations), and the variables of interest include quantities such as species richness, species diversity, biomass, and population size.[15]

In the assessment of environmental quality by means of biomonitors, it is important to control the variability of biological data, which often affects the forecasting precision of these techniques.[16] According to Kovacs,[3] the quality of the data originating from biological measurements depends heavily on at least three factors: (1) variability of the biomonitoring organisms (interactions between the organisms and environmental factors), (2) operators involved in data collection (especially for methods requiring taxonomic knowledge), and (3) type of sampling (sampling design, density of sampling points).

The selection of a proper (suitable) sampling design represents the first step to reduce data variability due to sampling error. When selecting the proper design, the objectives of the survey and environmental variability should be taken carefully into account.

20.4.1 Environmental Factors as Source of Noise in Biomonitoring Data

Environmental factors such as geomorphology, climatic variables, and substrate could have a great impact on the ecosystem property being studied in order to assess environmental quality, such as the rate of indicator species, the biodiversity of a community, or the presence of injuries on organisms. For this reason, it is important to understand the environmental processes and the interactions underlying the changes they undergo.

Large-scale monitoring programs cover large geographical regions and raise the question of how to deal with the great differences among ecosystems found in the various areas.[17]

Different species show various patterns of geographical distribution irrespective of anthropogenic effects on the environment; neither is it reasonable to expect their ranges of distribution to be constant over time even in the absence of human activities. The factors that determine the ranges of distribution and geographical patterns in species diversity are not well understood, and, in some cases, it is therefore difficult to separate the natural pattern of variations from the effects of human activities.[17]

Due to this variability, study of environmental properties with known cause–effect relationship is to be preferred. Indeed, the higher the potential to isolate the cause, the lesser is the error in interpreting the data. For some methodologies the cause–effect relationship is more obvious, for example, in the case of ecotoxicological experiments or in the case of ozone-sensitive tobacco, in which noise is easily identified since the cultivar Bel W3 is preferentially sensitive to ozone-related injuries. For other types of biomonitoring, it is more difficult to discriminate between the influence of a single variable and that of the others.

With reference, above all, to biodiversity studies, we should remember that environmental processes are dynamic: populations of organisms are in a constant state of flux. Intensive small-scale studies to calibrate the response of subsets of all species will enhance understanding of the generality and predictability of the trends indicated by the response of subsets.[18] This type of approach enables a valid model for interpreting data obtained on a large scale to be developed.[18] Several examples of such an approach are present in the literature (see, for example, References 19, 20, and 21). Giordani et al.[19] carried out a study aimed at standardizing Lichen Biodiversity monitoring in a Mediterranean region. In particular, they analyzed the influence of the great geomorphological variability and of substrate characteristics on epiphytic lichen vegetation. The results obtained in a small area suggested the use of less restrictive parameters in the sampling protocol (namely, olive trees with an inclination of the trunk >30° were also suitable for biomonitoring relevés). Further information on the applicability of this method was obtained by means of an in-depth study on three tree species monitored in the same microclimatic conditions, showing that the results obtained in the sample area could be extended to vaster areas.[20] In general, these investigations are useful in the preliminary stages of preparation of the sampling protocol and for developing models enabling the information obtained to be extended on a large scale. A wide array of models has been developed to cover aspects as diverse as biogeography, conservation biology, climate change research, and habitat or species management. Conceptual considerations should relate to selecting appropriate procedure for model selection. Testing the model in a wider range of situations (in space and time) will enable the range of applications for which the model predictions are suitable to be defined.[22]

An important issue is selectivity which seems particularly important in ecological measurement. A protocol is selective if the response provided as a measurement depends only on the intended ecosystem property.[17] Regarding this aspect, Yoccoz et al.[15] suggest that quantitative state variables characterizing the system well should be privileged. For example, when defining management objectives in terms of changes of densities of indicator species, the program should incorporate tests to ensure that selected species are indeed indicators of the process and variables of interest.[15]

From the point of view of application, some examples of processing for limiting errors due to environmental variability in interpreting ILB values (Index of Lichen Biodiversity) are reported by Ferretti and Erhardt[2] and by Loppi et al.[23] According to this method, the ILB values are calculated as percentage deviations from natural/normal conditions (Table 20.1), i.e., from a maximum ILB potentially measurable in a natural area. In Italy, ILB values are interpreted according to different scales[21,23] depending on the bioclimatic regions,[24] determined on the basis of the distribution of indicator lichen species and of the main meteorological and climatic parameters (rainfall, altitude, and temperature). This approach was also used in other types of biomonitoring, as, for example, in the case of bioaccumulation of trace elements in lichens.[25–27] The advantage of this interpretation is that it enables data measured on

TABLE 20.1
Interpretative Scales for Index of Lichen Biodiversity Values Scored in the Humid Sub-Mediterranean Bioclimatic Region (Thyrrenian Italy) and in the Dry Sub-Mediterranean Bioclimatic Region (Adriatic) in Italy

% Deviation from Natural Condition	IBL Score (Thyrrenian Italy)	IBL Score (Adriatic Italy)	Naturality/Alteration Classes
100	0	0	Lichen desert
76–100	1–25	1–20	Alteration
51–75	26–50	21–40	Semi-alteration
26–50	51–75	41–70	Semi-naturality
25	>75	>70	Naturality

Note: The scales are based on percentage deviation from maximum score potentially assessed in background natural conditions.

Source: Modified from Loppi, S. et al., A new scale for the interpretation of lichen biodiversity values in the Thyrrenian side of Italy, in *Progress and Problems in Lichenology at the Turn of the Millennium, 4th IAL Symposium (IAL 4)*, Llimona, X., Lumbsh, H.T., and Ott, S., Eds., *Bibliotheca Lichenologica*, 82, J. Cramer in der Gebruder Borntraeger Verlagsbuchhandlung, Berlin, Stuttgart, 2002, 235 and unpublished data.[21]

a regional scale to be compared with data on a national scale. Inevitably, however, these developments are only approximations that do not always lead to a reduction in error, since they tend to simplify the effects of the interactions between organisms and the environment. For this reason, it is important that the quality level of the data that can actually be achieved should be consistent with the level of predictability suggested by the interpretation of the results. Some authors,[28,29] for example, have observed different bioaccumulation rates in different lichen species collected under the same environmental conditions. As a consequence, the combined use of different species in the same bioaccumulation survey has to be verified previously to check the correlation among elemental concentration in the accumulator species. Examples of calibration of interpretation scales in different bioclimatic regions are also reported in other field of environmental monitoring such as in bioaccumulation in mosses[26] or in macroinvertebrates for the assessment of fresh water quality.[30]

20.4.2 Inherent Variability Requires Unambiguous Objectives

An explicit and well-defined objective is the major driver of the whole design process.[2] Once the nature of the study is identified, the unambiguous definition of the objectives involves the explicit identification of:

- Assessment question: Careful attention should be paid in the phase of definition of the objectives of the program. Different objectives require different monitoring designs.[31] As a consequence, the scope of inference of the study and the data collected depend on the aim of the study. If the monitoring objectives are clearly stated, it will be easier to describe the statistical methods to be used to analyze the data.[17]
- Target population: Unfortunately, monitoring programs do not always define their target population in an explicit statement. In many cases, the statement is insufficiently clear to determine whether a potential sample unit is included or excluded from the target population.[17]
- Geographical coverage: The area to be considered by the investigation, as also the characteristic of the area, are important when considering the proper sampling design; e.g., large vs. small survey areas or flat vs. geomorphological complex areas.

20.5 ENVIRONMENTAL FACTORS AND SAMPLING DESIGN

As previously reported, depending on the objectives of the investigation, environmental data for large geographical units are generally collected by sample surveys. The objective is therefore to select a subset from the population of interest (the sample) that allows inferences about the whole population.

A good sampling design is essential to collect data amenable to statistical analyses and to control errors in relation to the costs.

Many environmental programs address the assessment of abundance and richness as state variables of interest.[15] However, when these results have to be extended to a vast geographical range, it is important to estimate the detection probabilities associated with the selected count statistics and survey methods.

McCune and Lesica[32] found tradeoffs between species capture and accuracy of cover estimates for three different within-site sample designs for inventory of macrolichen communities in forest plots. On average, whole-plot surveys captured a higher proportion of species than did multiple microplots, while giving less accurate cover estimates for species. The reverse was true for microplots, with lower species captures and much better cover estimates for common species. Belt transects fell between the over two sample designs.

Similar results were obtained by Humphrey et al.[33] in a study to assess the differences in the species richness of lichen and bryophyte communities between planted and seminatural stands. A high percentage of species was recorded only once and very few species were common to more than half the plots. This "local rarity" phenomenon has been noted in other studies[34–36] and is partially related to sampling area. The authors of that study observe that it is possible that a 1-ha sampling plot used is too small to capture a representative sample of lower plant diversity in forest stands. For example, Rose[37] recommends a minimum sampling area of 1 km^2, but again, this depends on the objective of the survey.

Recently, the influence of different sampling tactics in the evaluation of lichen biodiversity was performed in a test study in Italy. In the first sampling method (Table 20.2), the five trees nearest to the center of the sampling unit were selected, within a Primary Sampling Unit of 1 km^2. In the second sample, the operator moved from the center of the square in all of the cardinal directions, took a circular plot with radius of 56 m (Secondary Sampling Units — SSUs) and selected (if there were)

TABLE 20.2
Sampling Procedures in Four Methods for Assessing Lichen Biodiversity

Sampling Tactic	Plot Dimension and Shape	No. of Trees and Selection Procedures
Method 1	1 km^2 primary unit	The 5 trees nearest to the center of the primary unit
Method 2	4 circular subplots (secondary sampling units — SSU) with 56-m radius, within 1 km^2 primary unit	The 3 trees nearest to the center of each plot
Method 3	4 circular subplots (secondary sampling units — SSU) with 125-m radius, within 1 km^2 primary unit	The 3 trees nearest to the center of each plot
Method 4	10 circular random plots with 30-m radius, within 1 km^2 primary unit	All the trees within the plots

Note: Details in the text.

the three trees nearest to the center of each plot. In the third sampling tactic, the SSUs had a radius of 125 m. Finally, in the fourth method, used to obtain the true average ILB value, a random selection of 30-m-radius SSU was adopted. In each secondary unit a census of the trees within the plot was carried out.

Significant differences in average ILB values were found between the checked tactics. As a result, the third tactic gives the better estimation of the average ILB of the sampling units and it can find a sufficient number of sampling trees, whereas a 56-m-radius SSU is too small to find a proper number of trees. The first tactic (five trees nearest the center) often takes to a "clustering error," i.e., the trees selected are grouped in a small portion of the sampling units and are not representative of the whole area.

Another important aspect to consider is the sampling density, which needs to be defined in relation to the objectives of the study and to its spatial scale. Ferretti et al.[38] used two datasets of lichen diversity (LD) surveys at the subnational level in Italy for establishing the sampling density that can be cost-effectively adopted in medium- to large-scale biomonitoring studies. As expected, in both cases the relative error on the mean LD values and the error associated with the interpolation of LD values for (unmeasured) grid cells increase with decreasing sampling density. However, it was possible to identify sampling density able to provide acceptable errors with quite a strong reduction of sampling efforts. This is important, as reduction of sampling effort can result in a considerable saving of resources that can be used for more detailed investigation in potentially problematic areas.

20.6 INDICATOR DEVELOPMENT

We can define "indicator" as a character or an entity that can be measured to estimate status and trends of the target environmental resource.[39] Further, an "index" is a characteristic, usually expressed as a score, that describes the status of an indicator.[8]

Response indicators should demonstrate the following features:[2,39–41]

- Correlate with changes in processes or other unmeasured components such as the stressor of concern
- Have a broad application, both at local and at large scale
- Integrate effects over time
- Provide early warning on future changes in ecosystem condition
- Provide distinctive information, e.g., cause–effect
- Be related to the overall structure and function of ecosystems
- Have a low and standard measurement error
- Have a sufficient reference on the effective applicability on the field
- Be cost effective

The development of indicator and indices is important in environmental monitoring[39] above all to obtain concise information from a complex environment. The process for indicator development should be taken into account including all the phases of the sampling, from the assessment question to the selection of core indicators and the evaluation of the performance of the indicators adopted.

For this reason it is important to establish *a priori* the variables of interest in a sampling protocol.

An example of a rigorous selection of the variables to be measured is given by EMAN, the Ecological Monitoring and Assessment Network implemented in Canada.[42] In this program a core of suitable variables capable of identifying departures from normal ranges of fluctuations in key ecosystem parameters were selected to detect early warning of ecosystem change. The criteria for this selection were based on data quality, applicability, data collection, repeatability, data analysis and interpretation, and cost-effectiveness.

20.7 OBSERVER AND MEASUREMENT ERRORS

In implementing the guidelines for biomonitoring methods, the documentation of standard operating procedures and a proper sampling design are only the first steps towards meeting Quality Assurance Objectives.[6]

To assess the reliability and consistency of the data, two activities above all are fundamental: training of the personnel involved in data collection and field checks on reproducibility of data.[43] Observer and measurement errors are important issues, especially in large-scale and long-term studies involving many surveyors and subjective estimates of a given attribute.[43–45] Metric measurements, often considered in forest health assessment (dbh, distance between trees, etc.), depend closely on the precision of the instrument used and they are generally easily repeatable and reproducible. When considering methods based on visual estimation, the instrument is the human eye. Visual assessments are quickly made and do not require expensive equipment, chemical tests, or highly trained personnel, but their subjective nature is a matter for concern.[46] Measurements based on visual estimation are consequently less precise and repeatable and less accurately reproducible.

Many different factors may contribute to the variability of application of a visually based assessment method: the operational manual used and the accuracy and precision of the operator, which in turn define his/her position on the calibration curve.

The main causes of error due to the operators involved in biomonitoring surveys are reported in Table 20.3. It is possible to distinguish errors relating to the experimental protocol used: the wrong type of measurement and imprecision of the instruments, the need for more experience in taxonomy, and faulty timing of the various phases.

20.7.1 Too Complex or Too Lax Sampling Protocols May Induce Relevant Observer Errors

A sampling protocol that calls for excessively complex procedures might not be easily applicable on a large scale and by inexperienced operators. The applicability of a protocol can be assessed by measuring precision and reproducibility in the framework of resampling surveys. The development of preparatory studies, conducted perhaps on a small scale and in controlled conditions, is a good method for assessing both the applicability and the repeatability of an experimental protocol.[47]

TABLE 20.3
Main Observer Errors and Procedures to Improve Data Quality

Factors of Errors Due to Operators	Evaluation Criteria	Procedures to Improve Data Quality
Too complex sampling protocol (scarce applicability)	Precision Reproducibility	Propedeutic studies Field checks
Too rough sampling protocol (high level of subjectivity)	Accuracy Precision Reproducibility	Propedeutic studies Field checks
Kind of measures (e.g., nonquantitative measures, inadequate instruments)	Precision Reproducibility	Rigorous protocol Application in the field
Taxonomic knowledge	Intercalibration tests Accuracy Bias estimation Audit certification	Training Certification Debriefing Harmonization
Learning factor	Accuracy	Remeasuring Harmonization
Time for each phase of the sampling protocol	Precision Reproducibility	Rigorous protocol Time restrictions

Note: See text for details.

An excessively lax sampling protocol leaving too much to the operator's subjectivity, would, on the other hand, contribute towards increasing the sampling error and make it difficult to control data quality. Again, in this case, the error induced by the operator's subjectivity can be assessed by repeating every phase of the sampling in order to identify which phase is most influenced by the error.

20.7.2 Type of Sampling Measurements

Efficient field sampling protocols strive to achieve high metric precision because this ensures the repeatability of observations among crews.[47]

Simplifying, we can distinguish between direct and indirect measurements such as visual estimates. This second type of measurement may be desirable in some cases because of its rapidity, but low precision or loss of ecological information may limit its application.[47]

As already mentioned, the use of quantitative state variables is recommended in order to reduce the error in data collecting due to the subjectivity of the operators. However, this is not always possible in environmental studies because of the great variety of parameters involved.

The evaluation of metric precision for a study of the riparian forest surveys suggested that objectively assessed metrics such as tree/snag DBH have high precision, while estimates such as woody debris or tree-cover metrics are less precise. Barker[47] showed that the tree DBH measurements were very precise because they were measured with a diameter tape, while crown ratio measures were less precise because they were visually assessed.[48]

The advantage of visual assessments is that they are quick to evaluate and do not require expensive equipment, chemical tests, or highly trained personnel, but the trade-off is frequently loss of ecological detail.[47] Another example is provided by Englund et al.[49] A comparison among different canopy-density measurement techniques showed that hemispherical photography is a more versatile technique than the spherical densiometer, but the equipment is more expensive, analysis time is not trivial, and waiting for appropriate conditions for taking pictures is a significant limitation. Finally, we must consider that the ultimate use of the data should dictate whether the increased precision that would result from more exacting measures is an acceptable trade-off for the investment in time and required degree of ecological information.[47]

20.7.3 Taxonomic Skill of the Operators

The personnel involved in data collection can influence the quality of the data above all in methods where a high taxonomic knowledge is required. This influence depends on the measure to which taxonomic accuracy is needed. Wilkie et al.[50] observed that in large-scale invertebrate biodiversity assessment, the postcollecting processing requires considerable effort for sorting, identification, and cataloging of the material. The time demand on specialists to process this material would have severe financial implications for any project.[51] It is possible to reduce costs in this phase of the project by reducing the role of the specialists. A variety of strategies[52,53] have been developed for rapid postfield processing of data, such as sorting to higher taxonomic level only ("taxonomic sufficiency") or employing nonspecialist technicians to separate specimens into informal groups based on obvious external characteristics (morphospecies). Oliver and Beattie[54] presented the results of studies in which technicians were used, after a one-off training period, to separate specimens into morphospecies. Samples were sent to experts for checking and a multivariate analysis was performed on both datasets. They found that the results using this simplified approach were very similar to those obtained using species data and concluded that identification errors were insufficient to affect results and conclusions.

The use of guilds or morphological groups as indicators for monitoring changes in ecosystem function has been considered by several authors[55,56] as a good compromise between the need for specialized knowledge and rapid field procedures employing nonspecialist technicians. In particular, McCune et al.[57] observed that random subsamples of species (community level) tended to produce the same patterns as the complete dataset. If, in fact, groups of species are defined by their growth forms (for example foliose, fruticose, and crustose lichens) or their ecology (epiphytic and epiphytic lichens), then different patterns of biodiversity and compositional gradients will result.[58]

Notwithstanding these results, perhaps the simplest and most readily communicated descriptor of diversity is species richness. This parameter, however, depends closely on the skill, the taxonomic knowledge, and the willingness of the trainee to learn. Although some trainees are strongly motivated, most of them need to feel that biodiversity assessment is important.[57] It is obvious that improving the taxonomic training of field observers can significantly reduce the negative bias in estimating species richness.

20.7.4 Observer Error in Lichen Diversity Monitoring

The assessment of lichen diversity is affected by various kinds of errors, mainly regarding the sampling procedures and the influence of environmental factors (see previous paragraphs). In regard to the observer error, important sources of data variability are represented by:

- Subjectivity in the positioning of the sampling grid
- Subjectivity in the selection of trees (sampling tactic) on which the relevés are carried out
- Taxonomic error in the identification of the species in trees
- Count of the frequency of the species appearing in the grid

In order to reduce the influence of the subjectivity of the operators, Italian guidelines for assessing air pollution using lichens[59] recently were changed to adopt a new sampling grid calling for monitoring of the exposure of the whole trunk.

Furthermore, standardization procedures were implemented, also in order to assess the repeatability of different sampling tactics in selecting the trees to be relevéd. The results showed that a very simple tactic is not always repeatable as it is too subjective, while an excessively complex tactic, although it is more objective, is difficult to apply in the field, particularly by inexperienced personnel. This problem can be solved through personnel training and harmonization procedures.

A similar approach was suggested in the American guidelines for lichen diversity monitoring,[60] where only macrolichens are collected and crustose lichens are excluded. As an obvious conclusion, experienced observers found, in general, more species than intermediate observers and beginners. However, McCune et al.[57] noted that during the intercalibration ring tests with operators from different regions, familiarity with the local flora could affect the results.

A method requiring a high level of taxonomic knowledge is in the Italian guidelines for monitoring effects of atmospheric pollution with lichens. The sampling design of the protocol is based on the assessment of the frequency and abundance of all lichen species, including groups of lichens that are difficult to identify in the field, such as crustose lichens.[59]

More detailed biomonitoring data may be obtained using this approach but, on the other hand, the sources of error due to insufficient taxonomic knowledge are more significant and should be carefully assessed by means of intercalibration tests and quality control procedures. Table 20.4 shows an intercalibration ring test conducted in Italy[45] which confirms the influence of operator experience (taxonomic knowledge) on the results. The operators were grouped into three classes based on their lichenological (taxonomic) experience: low, medium, and high. Accuracy of taxonomic identification was strongly affected by this parameter, ranging among the

TABLE 20.4
Percentage Accuracy in Lichen Biomonitoring Studies, in Relation to the Level of Experience of the Operators

Level of Lichenological Experience	Quantitative Accuracy (%)	Qualitative Accuracy (%)	Taxonomic Accuracy (%)
Low	68.6 ± 14.8	67.9 ± 9.1	28.9 ± 14.9
Medium	81.4 ± 11.2	74.6 ± 10.1	48.1 ± 14.1
High	83.8 ± 11.6	83.6 ± 11.3	61.3 ± 17.4

Source: Adapted from Brunialti, G. et al., Evaluation of data quality in lichen biomonitoring studies: the Italian experience, *Environ. Monit. Assess.*, 75, 271, 2002.[45]

TABLE 20.5
Percentage Accuracy Scored by the Operators in Lichen Biomonitoring Surveys Before and After a 7-Month Training Period

	Before	After
Quantitative ILB	72.4 ± 14.4	84.6 ± 10.1
Number of species	64.6 ± 14.3	81.1 ± 8.6
Taxononomic identification	34.5 ± 13.01	56.2 ± 17.9

Source: Adapted from Brunialti, G. et al., Evaluation of data quality in lichen biomonitoring studies: the Italian experience, *Environ. Monit. Assess.*, 75, 271, 2002.[45]

groups from 28.9% (low experience) to 61.2% (high experience). It is noteworthy that the Measurement Quality Objective (MQO) for accuracy of taxonomic identification, stated at 75%, was reached only by a small number of "very experienced" operators, while the MQO for quantitative (frequency of the species) and qualitative (number of species) accuracy were only reached by the groups with average, and much lichenological, experience, suggesting that the average quality of the data produced by the operators must be improved.[45]

The influence of the learning factor on the results of the survey was examined by Brunialti et al.[45] who reported the results of two intercalibration tests of the operators involved in a lichen monitoring program in Italy carried out before and after a 7-months' training period. The operators showed a great improvement in data quality (Table 20.5), in terms of both quantitative lichen biodiversity accuracy (from 72.4 to 84.6%) and accuracy of taxonomic identification (from 32.1 to 56.1%). Furthermore, operator accuracy often improved during the same test, and the results showed that accuracy improved with taxonomic training and, above all, continuous fieldwork and harmonization procedures. Since the task of long-term monitoring is to measure trends of environmental variables over time, the learning factor may cause confusion between the trend of variation of the variable of interest and the trend of improvement of the crews. As observed by McCune et al.,[57] it is possible that observer error will change over a period of time with improvement of skill and changes in motivation. Another aspect to consider is also familiarity with the regional flora. As a consequence, we should also consider that perception depends on preconception, in that we tend to see the species we expect to see.[57] To minimize this error, it is advisable to check particularly carefully the early phase of the survey, planning remeasuring procedures or field assistance of specialists (harmonization). For example, in the case of lichen biomonitoring, crustose lichen species are often overlooked by nonspecialists, even if in many habitats they account for most of the lichens occurring in the relevé. Furthermore, crustose lichens are often the pioneering phase of colonization on trees, and in many situations they indicate a significant improvement of air quality. When considering long-term monitoring with surveys every 2 years, if a crew, after a training period, relevés crustose lichens only in the second survey, these data may be interpreted as improved environmental quality. But who can say whether crustose lichens were already there also during the first survey? This question can only be answered by remeasurement in the early phase of the survey by an expert. Some simple recommendations may also minimize this kind of error. For example, the American Field Method Guide[60] suggests that while collecting the data, the crew should consider everything looking like a lichen as a lichen. This could result in an overestimation but could balance the underestimation caused by the crew's low taxonomic knowledge.

20.7.5 Observer Error in Ozone Monitoring with Tobacco Plants

There are several biomonitoring techniques that do not require taxonomical knowledge or biodiversity assessment. This is the case of biomonitoring of tropospheric ozone using sensitive tobacco (*Nicotiana tabacum* Bel W3) plants (e.g., Heggestad[61]). With this method operators are asked to score the amount of leaf area injured by typical ozone-induced necrosis. Scoring is carried out according to a categorical rating system after a visual examination of the leaves. Given the subjectivity of the visual estimates, differences between observers are likely to occur. Intercalibration exercises carried out in Italy show different levels of agreement among observers. Lorenzini et al.[46] report a high average repeatability (65.4%). As expected,[41] they also saw that agreement on extreme classes of injury is easy, while central classes were evaluated less consistently. Differences between observers may be related to the distribution of injury on the leaf and by the extent of injury. As a rule, when injury covers less than 50% of total leaf area the eye would focus on the diseased tissue and vice versa.[62]

A series of intercalibration exercises were carried out within a more formal quality assurance program developed for a biomonitoring project in Florence, Italy.[63] QA was based on established measured quality objectives (MQOs) and data quality limits (DQLs). MQOs allow an operator to have a ±1 category score; DQLs require observers to match MQOs in 90% of cases. Results obtained after 4350 comparisons between different observers show that a complete agreement (no difference) was reached in 49% of cases, while deviations ±1 class were obtained in 35% of cases

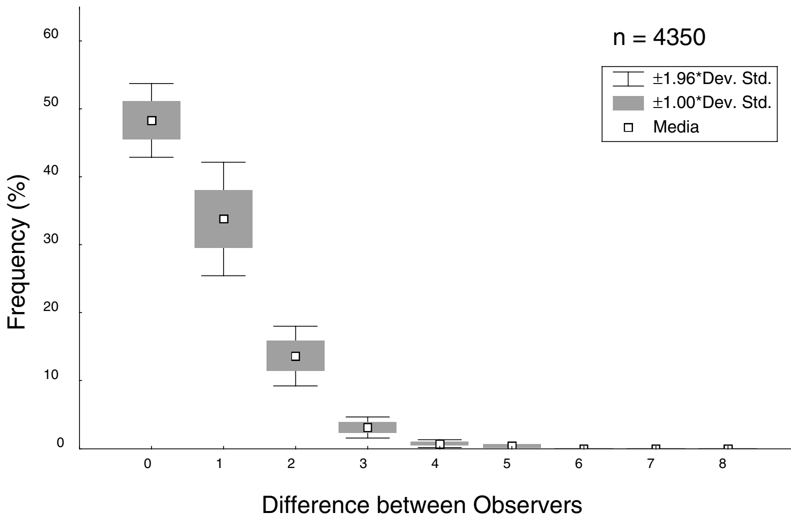


FIGURE 20.1 Difference between observers: 0 means no difference (complete agreement); 1 means 1 class difference, etc. The rating system was based on nine classes (0, 1, 2, 3, 4, 5, 6, 7, 8). Differences up to ±1 class were considered within the MQOs. (From Ferretti, unpublished data.)

(Figure 20.1). Thus, MQOs were actually matched in 84% of cases which is a good result, although below the desired DQLs. These examples show that biomonitoring data can be affected by observer bias. Documentation of differences between observers is essential in order to keep track of their performance and to allow further cross-evaluation of the data. If documentation is available, unusual records may be explained by observer bias, or, in a different perspective, systematic deviations may be managed mathematically. The need for a continuous effort toward more stringent protocols and harmonization procedures among operators and experts is, therefore, obvious, and when adequate documentation is available, it will be possible to track the improvements obtained in data consistency.

20.7.6 Observer Error in Tree Condition Surveys

The condition of the European forests has been monitored since 1986/1987 within the UN/ECE International Cooperative Programme on Assessment and Monitoring of Air Pollutant Effects on Forests (ICP Forests) and the European Union Scheme on the Protection of Forests against Atmospheric Pollution. The basic assessment method is a visual assessment of defoliation of a number of sample trees located at the intersections of a nominal 16 ∞ 16 km grid. In 2002, 5,929 plots with more than 131,000 trees were assessed in 30 European countries.[64] The survey is intended to provide data about spatial and temporal variation of tree defoliation in Europe in relation to air pollution and other stress factors. Obviously, the potential for data comparisons in space and time is fundamental. A number of problems have been identified with the design of this program (e.g., Innes[65]): however, the most serious one concerns the comparability of the assessment done by different surveyors operating in different countries.[66] A series of papers have addressed this issue (e.g., see References [10, 43, 67–73]). All papers agree that comparison between observers and between countries is problematic. The problem is mainly rooted in different methods and reference standards applied by the various countries participating in the program and can be exacerbated by peculiar species, individuals, and site condition.[68] As these problems were not adequately addressed at the beginning of the program, they now have resulted in a strong impact on spatial comparability. For example, the maps published in the UN/ECE reports now always have a caption warning about this problem. Unfortunately, a similar problem is likely to occur also for temporal comparisons. Consider that long-term monitoring programs may encounter considerable turnover in the personnel involved and shift and changes in the methods applied. Actually, there is some evidence that such changes have occurred, and this may weaken the reliability of the time series.[74] In this situation, proper QA procedures are essential; recently, actions were undertaken for the design of a series on international cross-calibration courses[63] as well as to develop image-analysis assessment techniques (e.g., Mizoue[75]). These initiatives will not solve the problem of comparability; however, they will allow proper documentation and QC on the data series.

20.7.7 Management of the Learning Factor

The learning factor — the acquisition of experience while conducting surveys — should not be underestimated, particularly in long-term monitoring. This parameter is controllable by means of remeasuring and harmonization, especially during the early monitoring periods. With this latter procedure, the experience of an expert can be transmitted directly to the rest of the crew and offset the assessment or underestimation errors that could be made in the earlier stages of the work. Variability due to the learning factor may be prevented by employing for the survey only people who have successfully completed training and certification and who meet the MQO requisites. However, this may be difficult and limiting in a cost-effectiveness perspective in the case of large-scale monitoring for which a large number of operators are needed.[42]

20.7.8 Time Required for Each Sampling Phase

In order to save financial resources when planning biomonitoring networks, it is important to know how much time will be required for fieldwork. In order to estimate the applicability of the different tactics, the time required for each sampling procedure will be indicated and the main difficulties found during the fieldwork taken into account. The influence on data quality of the times required for each phase of the sampling protocol has been considered by several authors.[32,50,57] It is obvious that the more time we spend on the sampling procedures, the better results we obtain in terms of data quality. However, we must also consider lowering the costs of the procedures. An example is provided by the issue of site access. A widespread ecological survey will also include visits to remote sites which may be difficult to reach due to time, cost, and safety concerns.[17] Some monitoring informations can be obtained through remote sensing using aerial photography, videogeography, or satellite imagery, but many ecological variables require actual field visits. In these cases, probability sampling can be used to reduce the number of "difficult" sites to be sampled.

Particularly for biological diversity monitoring, the time at disposal for each plot sampling will influence the results considerably. In the case of the Forest Health Monitoring program,[57] the sampling is time-constrained to help standardize effort across crews and to facilitate scheduling crew activities.

20.8 CONCLUSIONS

Biomonitoring is a powerful approach in various environmental studies. However, biomonitoring investigations are subjected to a variety of error sources that need to be acknowledged and documented in order to be managed properly. Error sources encompass a variety of subjects, from the design of the investigation to indicators development and data collection. As decision makers need robust and defensible data, environmental scientists need to consider proper procedures for reassurance about the value of their data. QA is essential in this respect as it allows the identification of the various error sources and forces investigators to address and solve problems. It is therefore important that environmental biologists and field ecologists consider QA as a key attribute of their work. In agreement with References 1, 5, and 7, we suggest that QA and related subjects should be considered an integral part of environmental surveys.

Various examples support our suggestion: biomonitoring surveys using lichens, specific indicator plants, and spontaneous vegetation have reported problems in data comparability at a variety of spatial and temporal scales. While a series of inherent properties of the biological systems will make it almost impossible to achieve a full comparability of biomonitoring data, even the simple documentation of the errors associated with the various investigation steps will provide many benefits. In particular, it will allow investigators, end-users, and the public to know the confidence to be placed on the data. As environmental studies are almost always based on funds from public agencies, we think that the extent to which the quality of the data can be documented should be a criterion that may help in distinguishing the value of biomonitoring (and other) studies.

REFERENCES

1. Wagner, G., Basic approaches and methods for quality assurance and quality control in sample collection and storage for environmental monitoring, *Sci. Total Environ.*, 264, 3, 1995.
2. Ferretti, M. and Erhardt, W., Key issues in designing biomonitoring programmes. Monitoring scenarios, sampling strategies and quality assurance, in *Monitoring with Lichens — Monitoring Lichens*, Nimis, P.L., Scheidegger, C., and Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, chap. 9.
3. Kovacs, M., *Biological Indicators in Environmental Protection*, Horwood, New York, 1992, 207 pp.

4. Gertner, G. et al., Mapping and uncertainty of predictions based on multiple primary variables from joint co-simulation with Landsat TM image and polynomial regression, *Remote Sens. Environ.*, 83, 498, 2002.
5. Kohl, M., Traub, B., and Paivinen R., Harmonisation and standardisation in multinational environmental statistics — mission impossible?, *Environ. Monit. Assess.*, 63, 361, 2000.
6. Cline, S.P. and Burkman, W.G., The role of quality assurance in ecological programs, in *Air Pollution and Forest Decline*, Bucher, J.B. and Bucher-Wallin, J., Eds., IUFRO, Birmensdorf, 1989, p. 361.
7. Shampine, W.J., Quality assurance and quality control in monitoring programs, *Environ. Monit. Assess.*, 26, 143, 1993.
8. Ferretti, M., Forest health assessment and monitoring. Issues for consideration, *Environ. Monit. Assess.*, 48, 45, 1997.
9. Millers, I. et al., North America Sugar Maple Project: Cooperative Field Manual, USDA, Canadian Forest Service, 1994, 51 pp.
10. Köhl, M., Waldschadensinventuren: Mögliche ursachen der Variation der Nadel-/Blattverlustschätzung zwischen Beobachtern und Folgerungen für Kontrollaufnahmen, *Allg. Forst Jagdztg.*, 162, 210, 1991.
11. Cochran, W.G., *Sampling Techniques*, 3rd ed., John Wiley & Sons, New York, 1977, 643 pp.
12. Kish, L., Survey sampling, John Wiley & Sons, New York, 1965, 643 pp.
13. Gertner, G. and Kohl, M., An assessment of some non-sampling errors in national survey using an error budget, *Forest Sci.*, 41(4), 758, 1992.
14. Lessler, J. and Kalsbeek, W., *Nonsampling Error in Surveys*, John Wiley & Sons, New York, 1992, 412 pp.
15. Yoccoz, N.G., Nichols, J.D., and Boulinier, T., Monitoring of biological diversity in space and time, *Trends Ecol. Evol.*, 16, 446, 2001.
16. Laskowsky, S.L. and Kutz, F.W., Environmental data in decision making in EPA regional offices, *Environ. Monit. Assess.*, 51, 15, 1998.
17. Olsen, A.R. et al., Statistical issues for monitoring ecological and natural resources in the United States, *Environ. Monit. Assess.*, 54, 1, 1999.
18. Will-Wolf, S. and Scheidegger, C., Monitoring lichen diversity and ecosystem function, in *Monitoring with Lichens — Monitoring Lichens*, Nimis, P.L., Scheidegger, C., and Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 10, 2002.
19. Giordani, P., Brunialti, G., and Modenesi, P., Applicability of the lichen biodiversity method (L.B.) to a Mediterranean area (Liguria, NW Italy), *Cryptogamie Mycol.*, 22, 193, 2001.
20. Brunialti, G. and Giordani, P., Variability of lichen diversity in a climatically heterogeneous area (Liguria, NW Italy), *Lichenologist*, 35, 55, 2003.
21. Loppi, S. et al., A new scale for the interpretation of lichen biodiversity values in the Thyrrenian side of Italy, in *Progress and Problems in Lichenology at the Turn of the Millennium, 4th IAL Symposium (IAL 4)*, Llimona, X., Lumbsh, H.T., and Ott, S., Eds., *Bibliotheca Lichenologica*, 82, Gebruder Borntraeger Verlagsbuchhandlung, Stuttgart, 2002, 235.
22. Guisan, A. and Zimmermann, N.E., Predictive habitat distribution models in ecology, *Ecol. Model.*, 135, 147, 2000.
23. Loppi, S. et al., Identifying deviation from naturality of lichen diversity in heterogeneous situations, in *Monitoring with Lichens — Monitoring Lichens*, Nimis, P.L., Scheidegger, C., and Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, 281.

24. Nimis, P.L. and Martellos, S., Testing the predictivity of ecological indicator values: A comparison of real and 'virtual' relevés of lichen vegetation, *Plant Ecol.*, 157, 165, 2001.
25. Nimis, P.L. et al., Biomonitoring of trace elements with lichens in Veneto (NE Italy), *Sci. Total. Environ.*, 255, 97, 2000.
26. Bargagli, R., *Trace Elements in Terrestrial Plants: An Ecophysiological Approach to Biomonitoring and Biorecovery*, Springer-Verlag & R.G. Landes, Berlin, 1998, chap. 8 and 9.
27. Bargagli, R. and Mikhailova, I., Accumulation of inorganic contaminants, in *Monitoring with Lichens — Monitoring Lichens*, Nimis, P.L., Scheidegger, C., and Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, chap. 6.
28. Minganti, V. et al., Biomonitoring of trace metals by different species of lichens (*Parmelia*) in North-West Italy, *J. Atmos. Chem.*, 45, 219, 2003.
29. Nimis, P.L., Andreussi, S., and Pittao, E., The performance of two lichen species as bioaccumulators of trace metals, *Sci. Total. Environ.*, 275, 43, 2001.
30. Houston, L. et al., A multi-agency comparison of aquatic macroinvertebrate-based stream bioassessment methodologies, *Ecol. Indicat.*, 1, 279, 2002.
31. Knopman, D.S. and Voss, C.I., Multi-objective sampling design for parameter estimation and model discrimination in groundwater solute transport, *Water Resour. Res.*, 25, 2245, 1989.
32. McCune, B. and Lesica, P., The trade-off between species capture and quantitative accuracy in ecological inventory of lichens and bryophytes in forests in Montana, *Bryologist*, 95, 296, 1992.
33. Humphrey, J.W. et al., The importance of conifers plantations in northern Britain as a habitat for native fungi, *Biol. Conserv.*, 96, 241, 2000.
34. Vitt, D.H., Yenhung, L., and Belland, R.J., Patterns of bryophyte diversity in peatlands of western Canada, *Bryologist*, 98, 218, 1995.
35. Collins, S.L. and Glenn, S.M., Effects of organismal and distance scaling on analysis of species distribution and abundance, *Ecol. Appl.*, 7, 543, 1997.
36. Qian, H., Klinka, K., and Song, X., Cryptograms on decaying wood in old-growth forests of southern coastal British Columbia, *J. Veg. Sci.*, 10, 883, 1999.
37. Rose, F., Ancient British woodlands and their epiphytes, *Br. Wildl.*, 5, 83, 1993.
38. Ferretti, M. et al., Reliability of different sampling densities for estimating and mapping lichen diversity in biomonitoring studies, *Environ. Pollut.*, 127, 249, 2004.
39. Hunsaker, C.T., New concepts in environmental monitoring: the question of indicators, *Sci. Total. Environ.*, Suppl., 77, 1993.
40. Breckenridge, R.P., Kepner, W.G., and Mouat, D.A., A process to select indicators for monitoring of rangeland health, *Environ. Monit. Assess.*, 36, 45, 1995.
41. Muir, P.S. and Mc Cune, B., Index construction for foliar symptoms of air pollution injury, *Plant Dis.*, 71, 558, 1987.
42. Tegler, B., Sharp, M., and Johnson, M.A., Ecological monitoring and assessment network's proposed core monitoring variables: an early warning of environmental change, *Environ. Monit. Assess.*, 67, 29, 2001.
43. Innes, J.L., Landmann, G., and Mettendorf, B., Consistency of observations of defoliation amongst three different European countries, *Environ. Monit. Assess.*, 25, 29, 1993.
44. Ferretti, M., Potential and limitations of visual indices of tree condition, *Chemosphere*, 36, 1031, 1998.
45. Brunialti, G. et al., Evaluation of data quality in lichen biomonitoring studies: the Italian experience, *Environ. Monit. Assess.*, 75, 271, 2002.

46. Lorenzini, G. et al., Visual assessment of foliar injury induced by ozone on indicator tobacco plants: a data quality evaluation, *Environ. Monit. Assess.*, 62, 175, 2000.
47. Barker, J.R. et al., Evaluation of metric precision for a riparian forest survey, *Environ. Monit. Assess.*, 75, 51, 2002.
48. McRoberts, R.E. et al., Variation in forest inventory field measurements, *Can. J. For. Res.*, 24, 1766, 1994.
49. Englund, S.R., O'Brien, J.J., and Clark, D.B., Evaluation of digital and film hemispherical photography and spherical densiometry for measuring forest light environments, *Can. J. For. Res.*, 30, 1999, 2000.
50. Wilkie, L., Cassis, G., and Gray, M., A quality control protocol for terrestrial invertebrate biodiversity assessment, *Biodivers. Conserv.*, 12, 121, 2003.
51. Marshall, S.A. et al., Terrestrial arthropod biodiversity: planning a study and recommended sampling techniques, *Bull. Entomol. Soc. Can.*, 26, 1, 1994.
52. Oliver, I. and Beattie, A.J., Invertebrate morphospecies as surrogates for species: a case study, *Conserv. Biol.*, 10, 99, 1996.
53. Beattie, A.J. and Oliver, I., Taxonomic minimalism, *Trends Ecol. Evol.,* 9, 488, 1994.
54. Oliver, I. and Beattie, A.J., A possible method for the rapid assessment of biodiversity, *Conserv. Biol.*, 7, 562, 1993.
55. Will-Wolf, S., Esseen, P.-A., and Neitlich, P., Monitoring biodiversity and ecosystem function: forests, in *Monitoring with Lichens—Monitoring Lichens*, Nimis, P.L., Scheidegger, C., Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, chap. 14.
56. Will-Wolf, S., Scheidegger, C., and McCune, B., Methods for monitoring biodiversity and ecosystem function, in *Monitoring with Lichens — Monitoring Lichens*, Nimis, P.L., Scheidegger, C., and Wolseley, P., Eds., Kluwer Academic, Dordrecht, Netherlands, 2002, p. 11.
57. McCune, B. et al., Repeatability of community data: species richness versus gradient scores in large-scale lichen studies, *Bryologist*, 100, 40, 1997.
58. McCune, B. and Antos, J.A., Correlations between forest layers in the Swan Valley, Montana, *Ecology*, 62, 1196, 1981.
59. ANPA, I.B.L. Indice di Biodiversità Lichenica: manuale. ANPA, Agenzia Nazionale per la Protezione dell'Ambiente, Roma, 2001, 185 pp.
60. USDA Forest Service, *Forest Health Monitoring 1998: Field Method Guide*, USDA Forest Service, 1998.
61. Heggestad, H.E., Origin of Bel W3, Bel C and Bel B tobacco varieties and their use as indicators of ozone, *Environ. Pollut.*, 74, 264,1991.
62. Horsfall, J.G. and Cowling, E.B., Pathometry: the measurement of plant disease, in *Plant Disease: An Advanced Treatise*, Vol. II, Academic Press, New York, 1978.
63. Ferretti, M. et al., A biomonitoring network to upscale site-related measurements of ground level ozone in the area of Florence, Italy, *EuroBionet2002 Conference on Urban Air Pollution, Bioindication and Environmental Awareness*, University of Hohenheim, November 5–6, Summary 21, 2002.
64. Lorenz, M. et al., Forest Condition in Europe: Results of the 2002 Large-scale Survey, *UN/ECE and EC*, Geneva, 116 pp. plus Annexes, in press.
65. Innes, J. L., Forest health surveys — a critique, *Environ. Pollut.*, 54, 1, 1988.
66. Cozzi, A., Ferretti, M., and Lorenz, M., Quality Assurance for Crown Condition Assessment in Europe, *UN/ECE*, Geneva, 2002, 111 pp.
67. Neumann, M. and Stowasser, S., Waldzustandinventur: zur Objektivitat von Kronenklassifizierung, in *Fortsliche Bundesversuchsanstalt Wien, Jahresbericht 1986*, Vienna: fortsliche Bundesversuchsanstalt, 1986, pp. 101–108.

68. Innes, J. L., Forest health surveys: problems in assessing observer objectivity, *Can. J. For. Res.*, 18, 560, 1988.
69. Schadauer, K., Die Ermittlung von Genauigkeitsmassen terrestrischer Kronenzustandsinventuren im Rahmen der ostereichischen Waldzustandsinventur, 1991.
70. Köhl, M., Quantifizierung der Beobachterfehler bei der Nadel-/Blattverluschatzung, *Allg. Forst Jagdztg.*, 164, 83, 1992.
71. Ghosh, S., Innes, J.L., and Hoffman, C., Observer variation as a source of error in assessment of crown condition through time, *Forest Sci.*, 41, 235, 1995.
72. Ferretti, M., Cenni, E., and Cozzi, A., Indagini sulle condizioni dei boschi. Coerenza econfrontabilità dei dati sulla trasparenza delle chiome degli alberi in Italia, *Montie Boschi*, 2, 5, 1994.
73. Ferretti, M., Cenni, E., and Cozzi, A., Assessment of tree crown transparency and crown discoloration as performed by surveyors from six European countries during the 7th Intercountry EC/ECE Mediterranean Intercalibration Course, Cagliari, Italy, 1994, Regione Autonoma della Sardegna, 1995, 32 pp. plus Annexes.
74. Landmann, G., Nageleisen, M., and Ulrich, E., New evidence for a methodological shift in the visual assessment of the crown condition of broadleaves, *Cah. DSF*, 1, 74, 1998.
75. Mizoue, N., CROCO: semiautomatic image analysis system for crown condition assessment in forest health monitoring, *J. For. Plann.*, 8, 17, 2002.

21 Patchy Distribution Fields: Acoustic Survey Design and Reconstruction Adequacy

I. Kalikhman

CONTENTS

21.1 INTRODUCTION

Patchiness is a fundamental attribute of ecosystems.[1] Each ecosystem component has its typical dimension of patches. The scale of the patchiness for fish concentrations ranges from several feet to hundreds of miles.[2] For plankton, it is usually smaller than that for fish and ranges from one foot to several miles or even dozens of miles.[3] Gaps are a special case of patchiness, corresponding to areas of habitable space in which organisms are noticeably reduced in abundance relative to background levels.[4] Because of the patchiness of ecosystem component distribution, the parameters of an acoustic survey should be chosen on the basis of statistical characteristics of such fields.

One major goal of any acoustic survey is to reconstruct the distribution field studied. In this case, the survey design is considered efficient enough if it ensures the required adequacy of a reconstructed field to the actual one with minimal expenditures.[5] It is assumed in this consideration that the size and structure of the actual distribution field, under natural conditions, should be known. However, in a real situation, the parameters of distribution fields are never known. Therefore, the mathematical simulation method is used to determine the efficiency of survey design and improve the algorithm of data analysis.

1-56670-641-6/04/$0.00+$1.50

The efficiency of an acoustic survey mainly depends on its pattern, the distance between transects, and the unit of sampling distance. The survey pattern represents either zigzag or parallel transects.[6] Zigzag transects are the more economical survey pattern. In the case of zigzag transects, the survey path is formed entirely by transects; for parallel ones, it consists of transects and connecting tracks. Parallel transects are mainly used only if the borders of the region are known.[7] In this case, the collection of data is possible not only on transects but also on connecting tracks. The position of the starting point may, to some extent, affect the efficiency of surveys.

The choice of the distance between transects usually depends on the survey scale. In carrying out large-scale surveys, the distance between transects may reach 50 to 60 mi or more; with small-scale surveys, it may be from 20–30 down to 4–5 mi or less. It was recommended to locate transects at a distance close to the autocorrelation radius for the field.[7] However, no relationships for determining the accuracy of surveys were established. In studies aimed at the revelation of these relationships, the relative error of the abundance estimate (the bias of the survey) was used as a measure of the survey quality.[8,9]

The unit of sampling distance is the length of a survey transect along which the acoustic measurements are averaged to receive one sample. If the sampling distance unit is too large, potentially useful information about the distribution of the stock will be lost. If it is too small, successive samples will be correlated, in which case it will be difficult to determine the confidence interval for the stock estimate.[6] The unit of sampling distance may be as short as 0.1 mi to distinguish dense schools or as long as 10 mi in the case of species widely distributed over large areas of the ocean; usually, the sampling distance unit may be within the range of 1 to 5 mi.[6]

The aim of the present study is to determine survey design parameters (the survey pattern, the distance between transects, and the unit of sampling distance), allowing one to obtain a realistic image of a patchy distribution field. The bias of acoustic surveys is important in studies aimed at estimation of the fish abundance. However, since acoustic surveys began to be used for ecological purposes (e.g., Reference 10), it is often necessary to estimate interrelationships in the organization of populations without considering the bias. Therefore, herein, survey design parameters are determined allowing the unbiased reconstruction of an original distribution field. This means that the bias in the survey results is neglected; the parameters of the reconstructed field and of that originally generated are compared by correlation analysis. Fish and zooplankton distribution in Lake Kinneret (Israel) are used as prototypes in simulating patchy fields.

If the efficiencies of various survey patterns are compared under the assumption of fixed transect spacing, it means that the numbers of parallel or zigzag transects are equal to each other. However, as the length of a zigzag transect is always smaller than that of a parallel transect plus connecting track, the sampling efforts (the overall survey path or, in other words, the time for conducting a survey) for both patterns are not the same. Therefore, the efficiencies of various patterns are often compared under the assumption of fixed sampling effort, when the overall survey paths for both patterns are equal to each other. In this case, a greater number of zigzag transects compared to parallel ones may be carried out. As a result, the adequacy of reconstructing a field from data of a zigzag survey is expected to increase; determination of the extent of this increase is one of the objectives of the present study.

21.2 MATHEMATICAL MODEL DESIGN

The mathematical model is based on the approach described in References 11 and 12. The surveyed area is a rectangular array of numbers making up a square matrix of 50 lines and 50 columns. Each of these 2500 numbers represents the value of a field in an elementary area (node). The distance between the nodes is equal to one, so the grid size is 50×50. To simulate a distribution field, patches (or gaps) are generated and placed into the surveyed area. Outside patches the field variable is zero, while outside gaps it is the constant background. Inside a patch (gap) the variable increases (decreases) from the outer border towards the center. The probabilistic distribution low of the appearance of various field values in a patch (gap) is Poisson's (or the logarithmical) normal; this is in agreement with the results obtained by analyzing data from real acoustic microsurveys.[13]

Patches (gaps) may be isotropic (circular or nondirectional) or anisotropic (elliptical or directional) with various sizes and spatial orientation; or separate or overlapping (partially or fully), forming larger patches (gaps) with more complicated shapes. Patches (gaps) are either immovable (static) or movable (dynamic); moving patches (gaps) have a realistic shape like a comet.[2] Random fluctuations in density can be superimposed on the simulated patchy field.

A resulting (constructed) patchy distribution field is then sampled by simulating an acoustic survey. The starting point (the beginning of the initial transect) is always located either on the abscissa or ordinate axis. We distinguish the general direction of a survey and the survey path. The survey direction coincides either with the abscissa or ordinate axis. The survey path with an overall length S is formed by transects disposed as parallel lines or a zigzag. Transects, as parallel lines, are disposed perpendicularly to the general survey direction. Thus, the overall survey path in this case consists of parallel transects and connecting tracks perpendicular to transects. The distances between parallel transects are regular and equal to the length of connecting tracks (D).

A zigzag is characterized by the shift between neighboring parallel transects in the general survey direction. The shift is a move of a survey when passing from one transect to another. The shift between neighboring parallel transects is regular and twice as large as the distance between corresponding parallel transects. Throughout this study, half a shift between neighboring parallel transects is called the distance between zigzag transects (D). Along transects, a unit of sampling distance can be set. As mentioned above, this distance is the length of a transect over which the values of a field are integrated and averaged. In the model, the sampling distance unit (d) is equal to a specific part of the length of a transect (the minimal value $d = 1/50$).

The values of a field along the survey path over each sampling distance are considered to be measured without error and are used to reconstruct a field. The reconstructed field consists again of the full set of nodes of the array representing the surveyed area. The corresponding values of the reconstructed field and that originally generated are then compared for all nodes. Their adequacy is evaluated by the coefficient of determination representing the square of the standard Pearson correlation coefficient (r).[14]

Further on, the reconstruction of distribution fields will be taken into consideration. The gridding methods use weighted interpolation algorithms. The kriging method widely used in reconstructing distribution fields is tested. This algorithm assumes an underlying variogram, which is a measure of how quickly things change on the average.[15] The method of gridding is chosen so that the maximal correlation between the reconstructed field and that originally generated is attained. Preliminary simulations indicate that in most cases, the maximal correlation is attained when using kriging with the linear variogram model. Application of various gridding methods may have an impact on relationships of interest; therefore, to exclude this impact, the indicated method is used alone. Search option controls which data points are considered by the gridding operation when interpolating grid nodes.[15]

The autocorrelation radius is used as a characteristic of a field in choosing the parameters of a survey.[16] The autocorrelation radius (R) is determined from a plot of the autocorrelation function as the distance from the coordinate center to the point where the lag (autocorrelation) is zero. An important property of the autocorrelation function consists of the fact that it is sensitive only to variable deviations, while constant levels have no effect.[17] This enables us to use the autocorrelation function in seeking the autocorrelation radii of distribution fields that include gaps. Throughout the study, R, S, and D are given in proportion to the size of the square representing a surveyed area.

21.3 MATHEMATICAL EXPERIMENTS

In this section, the results of simulations reflecting the reconstruction of static and dynamic distribution fields are considered.

Static fields. The set of relevant problems can be divided into three groups, where: the transect spacing is fixed, a given sampling effort is allocated to a survey, and a choice as made for the unit of sampling distance.

Fixed transect spacing. Let us consider the set of problems in application to the parallel survey pattern. Examples of the original isotropic field are given in Figure 21.1, first row. The autocorrelation radius for such a field corresponds to the size of the patches and is equal to R = 0.15 (left), R = 0.09 (middle), or R = 0.10 (right), independent of the direction chosen (second row).[18] The adequacy of reconstructing distribution fields is also independent of the direction of the survey performed. For example, the simulations show that the surveys carried out at D = 0.20 in two perpendicular directions give similar results of $r^2 = 0.81$ and $r^2 = 0.85$ (left), $r^2 = 0.20$ and $r^2 = 0.28$ (middle), or $r^2 = 0.40$ and $r^2 = 0.38$ (right) (the number of sample data points NSDP = 300). The increase of the distance between transects (compare the third row with fourth) and the decrease of the autocorrelation radius (compare the left column with the middle and right columns) both lead to lower correlation between the reconstructed field and that originally generated. Thus, the conclusion can be made that the adequacy of a reconstructed field to its original depends upon the D/R ratio.

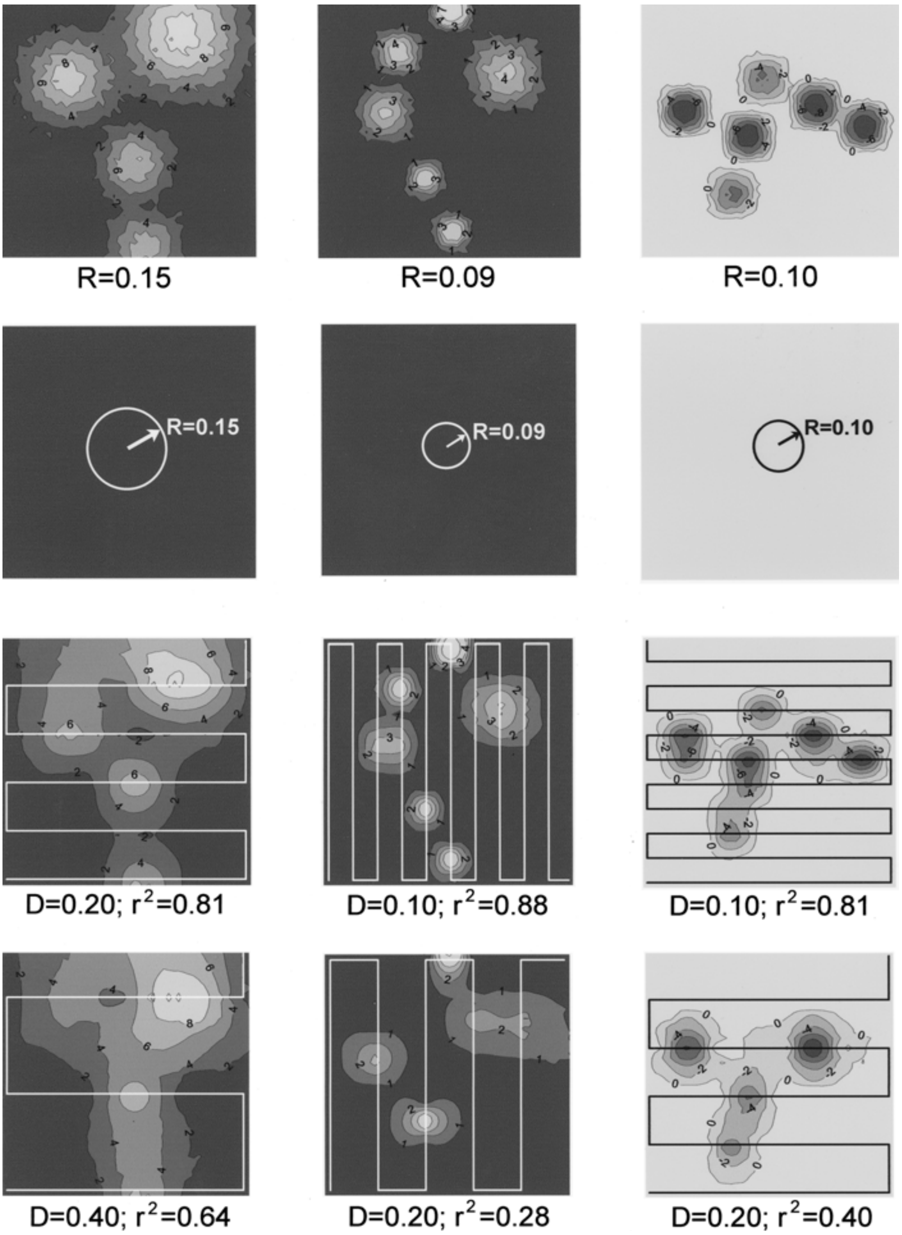


FIGURE 21.1 First row: Original patchy isotropic fields. Second row: Autocorrelation circles for the fields. Third and fourth rows: Paths of the simulated surveys and the reconstructed fields. Right column represents the gaps (the isolines are labeled by negative numbers).

Original anisotropic fields are generated as a set of anisotropic patches with similar spatial orientation (Figure 21.2, first row). Anisotropic fields are characterized by the anisotropy ratio (AR) and anisotropy angle (AA).[15] Autocorrelation radii in directions of abscissa and ordinate axes are equal to $R_X = 0.27$ and $R_Y = 0.13$ (AR = 4, AA = 0°, left), $R_X = 0.06$ and $R_Y = 0.23$ (AR = 6, AA = 90°, middle), or $R_X = 0.25$ and $R_Y = 0.06$ (AR = 6, AA = 0°, right) (second row). Simulations conducted indicate that the surveys in the direction of patch elongation permit us to reconstruct the original field more precisely than those in a perpendicular direction. For example, a survey carried out at D = 0.40 in the direction of patch elongation gives better results in terms of the original field adequacy ($r^2 = 0.92$, left; $r^2 = 0.76$, middle; or $r^2 = 0.79$, right) (NSDP = 200, third row) than even a survey conducted with the smaller distance between transects (D = 0.20, left; D = 0.22, middle; or D = 0.16, right) in a perpendicular direction ($r^2 = 0.66$, NSDP = 300, left; $r^2 = 0.16$, NSDP = 300, middle; or $r^2 = 0.26$, NSDP = 400, right) (fourth row). The conclusion can be made that the direction of the patch elongation is the optimal direction of an acoustic survey.

The distribution fields were rotated to a certain angle in order to examine the efficiency of surveys which, for some reason, cannot be conducted in the optimal direction. For each rotated original anisotropic fields (an example of rotation for 30° is given in Figure 21.3, first row), the autocorrelation radii in directions of abscissa and ordinate axes are equal to $R_X = 0.25$ and $R_Y = 0.18$ (AR = 4, AA = 30°, left), $R_X = 0.11$ and $R_Y = 0.20$ (AR = 6, AA = 120°, middle), or $R_X = 0.24$ and $R_Y = 0.16$ (AR = 6, AA = 150°, right). In the case of the left field, the survey carried out at the distance between transects D = 0.40 with the angle between the survey direction and that of patch elongation equaling 60° gives poorer results in terms of the original field adequacy ($r^2 = 0.69$, NSDP = 200, fourth row) than the survey with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.85$, NSDP = 200, third row). For the middle field, the survey carried out at D = 0.40 with the angle between the survey direction and that of patch elongation equaling 60° permits us to reconstruct the original field with lower adequacy ($r^2 = 0.25$, NSDP = 200, fourth row) than the survey with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.72$, NSDP = 200, third row). Finally, for the right field, the survey carried out at D = 0.40 with the angle between the survey direction and that of patch elongation equaling 60° gives worse results in terms of the original field adequacy ($r^2 = 0.66$, NSDP = 200, fourth row) than the survey with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.80$, NSDP = 200, third row). Thus, the results of experiments indicate that a decrease in the angle between the survey direction and that of patch elongation leads to the higher coefficient of determination between the reconstructed field and that originally generated. This is explained by the fact that D/R ratio in the direction of patch elongation is smaller than that in any other direction or, ultimately, by the relationship of the autocorrelation radii located in two reciprocally perpendicular directions (the autocorrelation radius situated at a smaller angle to the major axis of autocorrelation ellipse is always larger than that located at a larger angle to it).

Let us consider the same set of problems in application to the zigzag survey pattern. The examples of isotropic original fields, given in Figure 21.1, first row, are reproduced in Figure 21.4, first row. In the case of the left field, the surveys carried

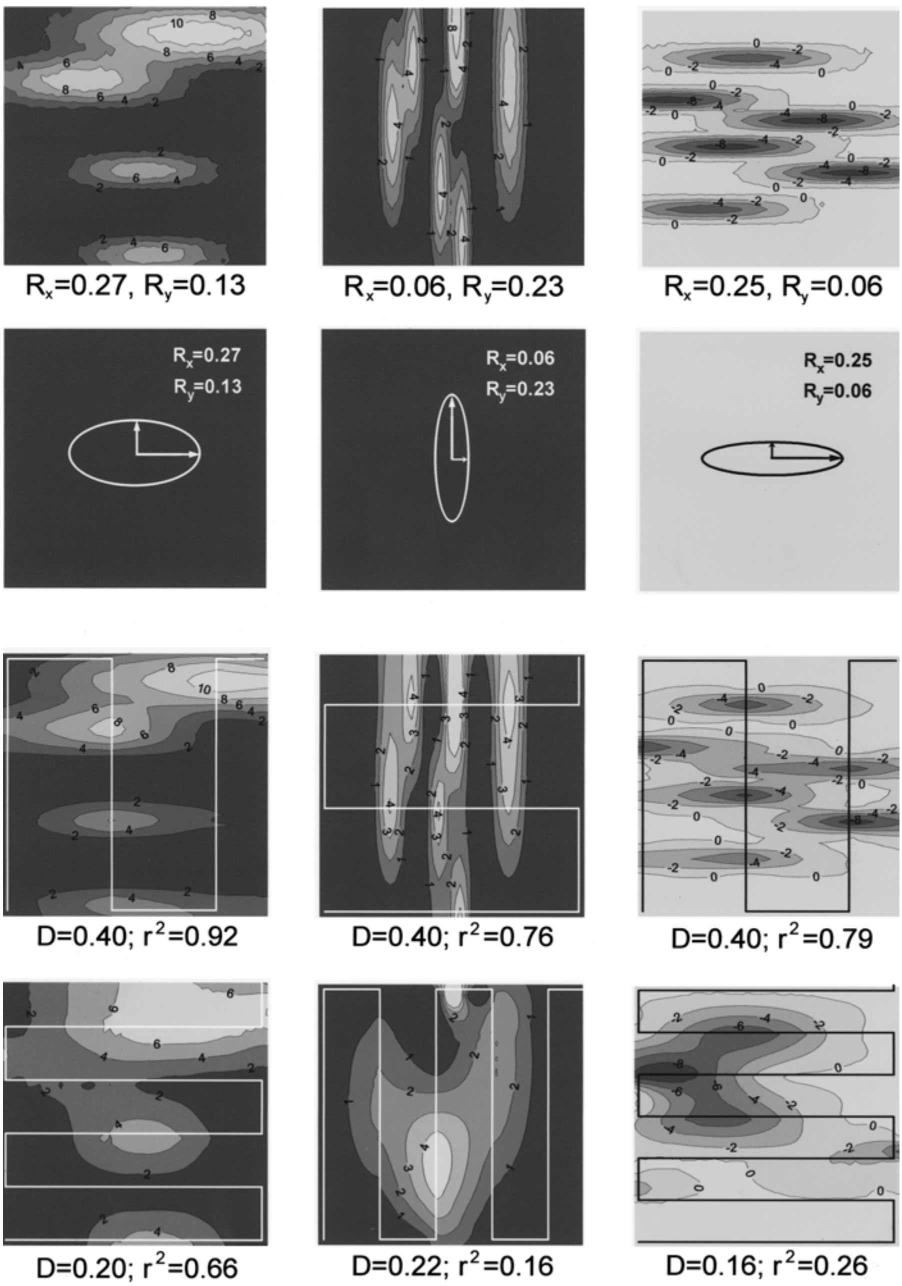


FIGURE 21.2 *First row*: Original patchy anisotropic fields. *Second row:* Autocorrelation ellipses for the fields. *Third* and *fourth rows:* Paths of the simulated surveys conducted in the directions of major or minor axes of autocorrelation ellipses and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

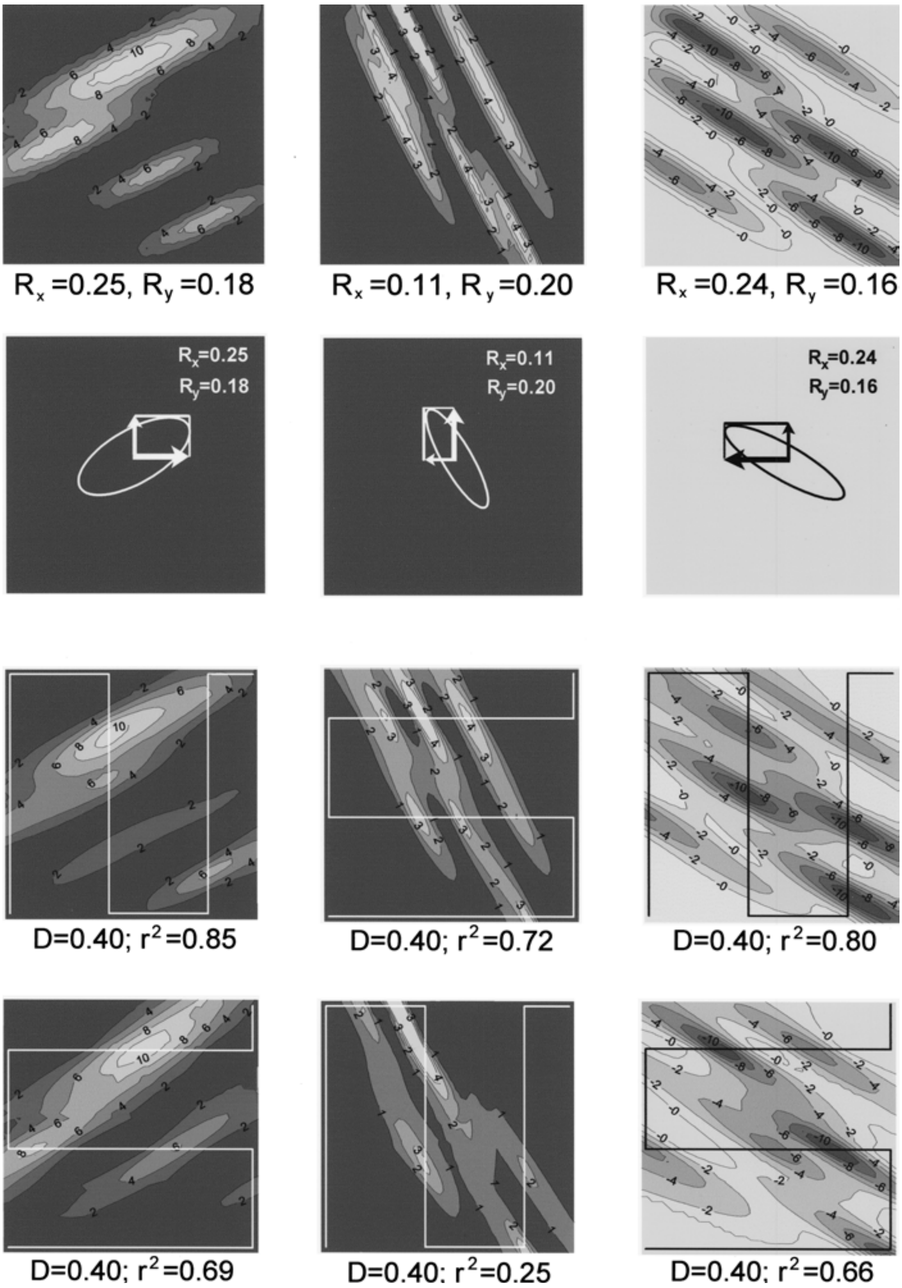


FIGURE 21.3 *First row:* Original patchy rotated anisotropic fields. *Second row:* Autocorrelation ellipses for the fields. *Third* and *fourth rows:* Paths of the simulated surveys conducted in the arbitrary directions and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

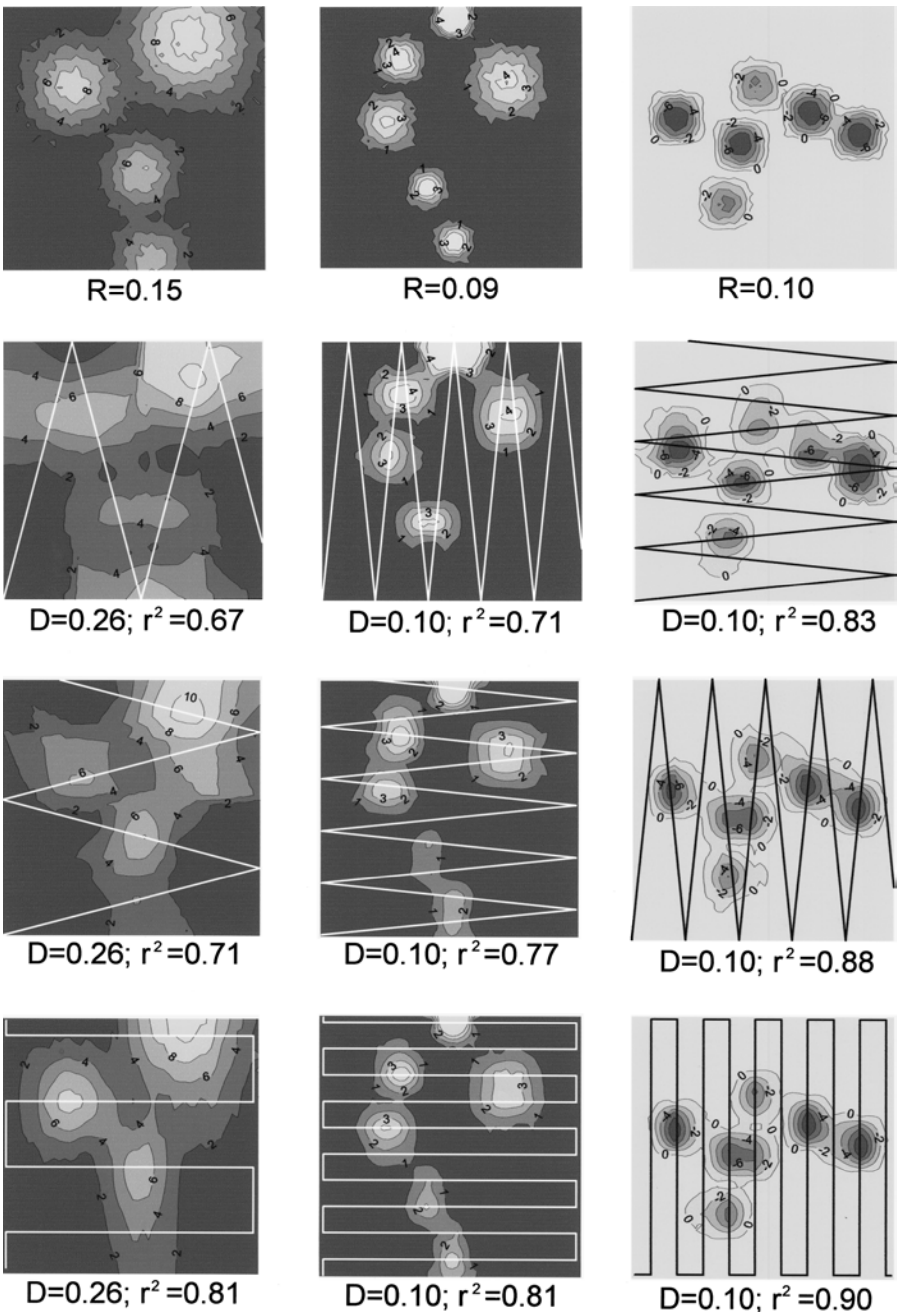


FIGURE 21.4 *First row:* Original patchy isotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

out at the distance between transects D = 0.26 in two perpendicular directions give similar results of $r^2 = 0.67$ (second row) and $r^2 = 0.71$ (third row) (NSDP = 198). For the middle field, the surveys conducted at D = 0.10 in two perpendicular directions give similar results of $r^2 = 0.71$ (second row) and $r^2 = 0.77$ (third row) (NSDP = 499). For the right field, the surveys carried out at D = 0.10 in two perpendicular directions give similar results of $r^2 = 0.83$ (second row) and $r^2 = 0.88$ (third row) (NSDP = 499). Thus, regarding zigzag transects, as well as in the case of parallel transects, the adequacy of reconstructing an isotropic field to that originally generated is actually independent on the direction of the survey performed. As shown by the simulations, the decrease of the ratio of the distance between transects to autocorrelation radius for the field (D/R) leads to the higher coefficient of determination between the reconstructed field and that originally generated, and vice versa.

To receive an objective comparison of the results of surveys by zigzag and parallel patterns throughout the paper, not only the transect spacing is taken equal but also their location (each parallel transect crosses the point where the middle of a corresponding zigzag transect is situated). A comparison shows that the zigzag patterns lead to poorer results than the parallel patterns do ($r^2 = 0.71$ and $r^2 = 0.81$, left; $r^2 = 0.77$ and $r^2 = 0.81$, middle; $r^2 = 0.88$ and $r^2 = 0.90$, right) (third and fourth rows).

The examples of original anisotropic fields, given in Figure 21.2, first row, are reproduced in Figure 21.5, first row. In the case of the left field, a survey carried out at the distance between transects D = 0.28 in the direction of patch elongation gives even better result in terms of the original field adequacy ($r^2 = 0.90$, NSDP = 180) (third row) than a survey conducted at D = 0.16 in a perpendicular direction ($r^2 = 0.64$) (second row). For the middle field, a survey carried out at D = 0.34 in the direction of patch elongation gives even better result in terms of the original field adequacy ($r^2 = 0.86$, NSDP = 153) (third row) than a survey conducted at D = 0.10 in a perpendicular direction ($r^2 = 0.67$) (second row). For the right field, a survey carried out at D = 0.48 in the direction of patch elongation permits us to reconstruct the original field much better ($r^2 = 0.72$, NSDP = 114) (third row) than a survey conducted at D = 0.12 in a perpendicular direction ($r^2 = 0.45$) (second row). Thus, the simulations conducted indicate that the surveys in the direction of patch elongation permit us to reconstruct the original field more precisely than those in a perpendicular direction. A comparison of the surveys by zigzag transects (third rows) with those by parallel transects (fourth rows) shows that the former lead to poorer results than the latter do ($r^2 = 0.90$ and $r^2 = 0.96$, left; $r^2 = 0.86$ and $r^2 = 0.92$, middle; $r^2 = 0.72$ and $r^2 = 0.79$, right).

The examples of rotated original anisotropic fields, given in Figure 21.3, first row, are reproduced in Figure 21.6, first row. In the case of the left field, the survey carried out at the distance between transects D = 0.28 with the angle between the survey direction and that of patch elongation equaling 60° gives poorer results in terms of the original field adequacy ($r^2 = 0.64$, second row) than the survey conducted at D = 0.34 with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.88$, NSDP = 154, third row). For the middle field, the survey carried out at D = 0.30 with the angle between the survey direction and that of

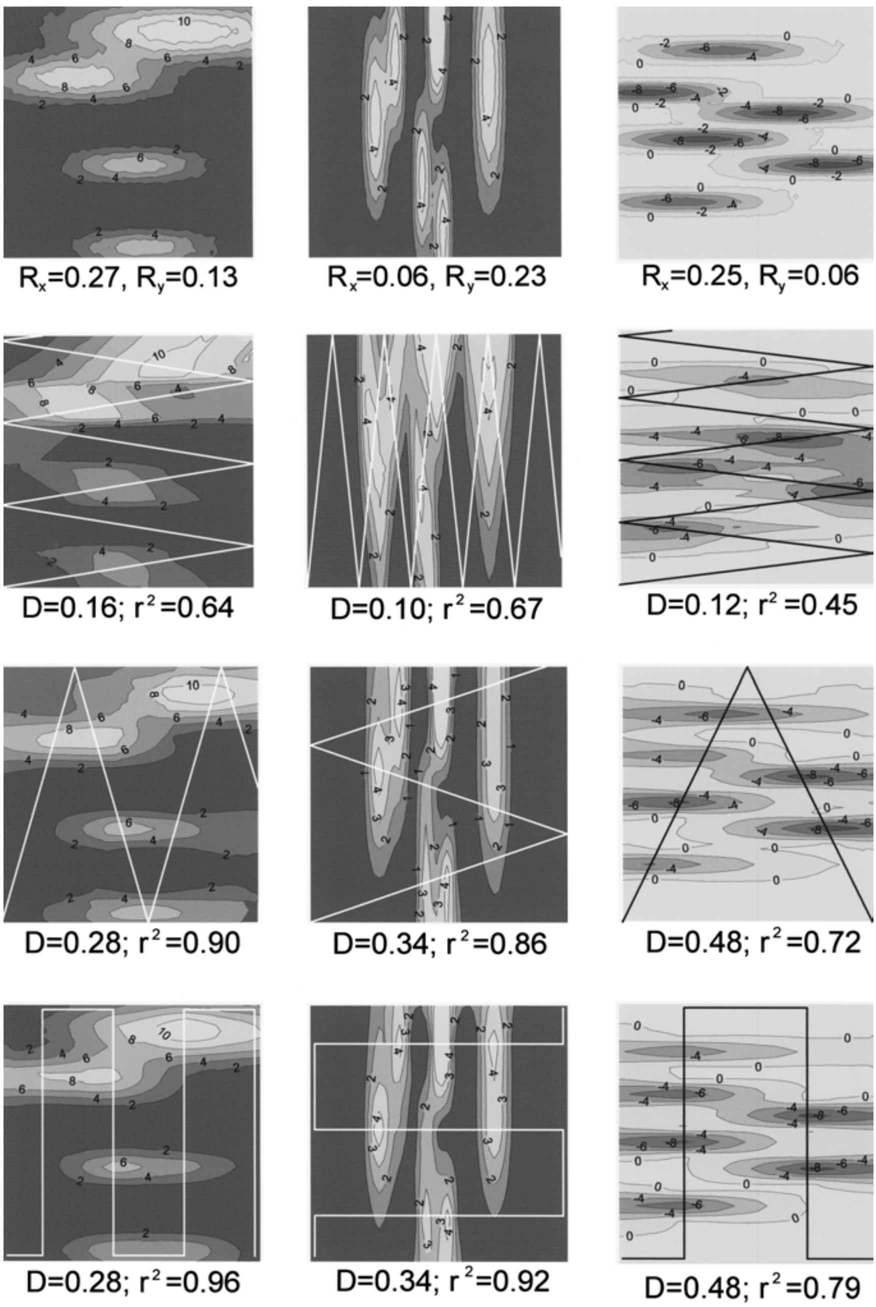


FIGURE 21.5 *First row:* Original patchy anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the directions of major or minor axes of autocorrelation ellipses and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

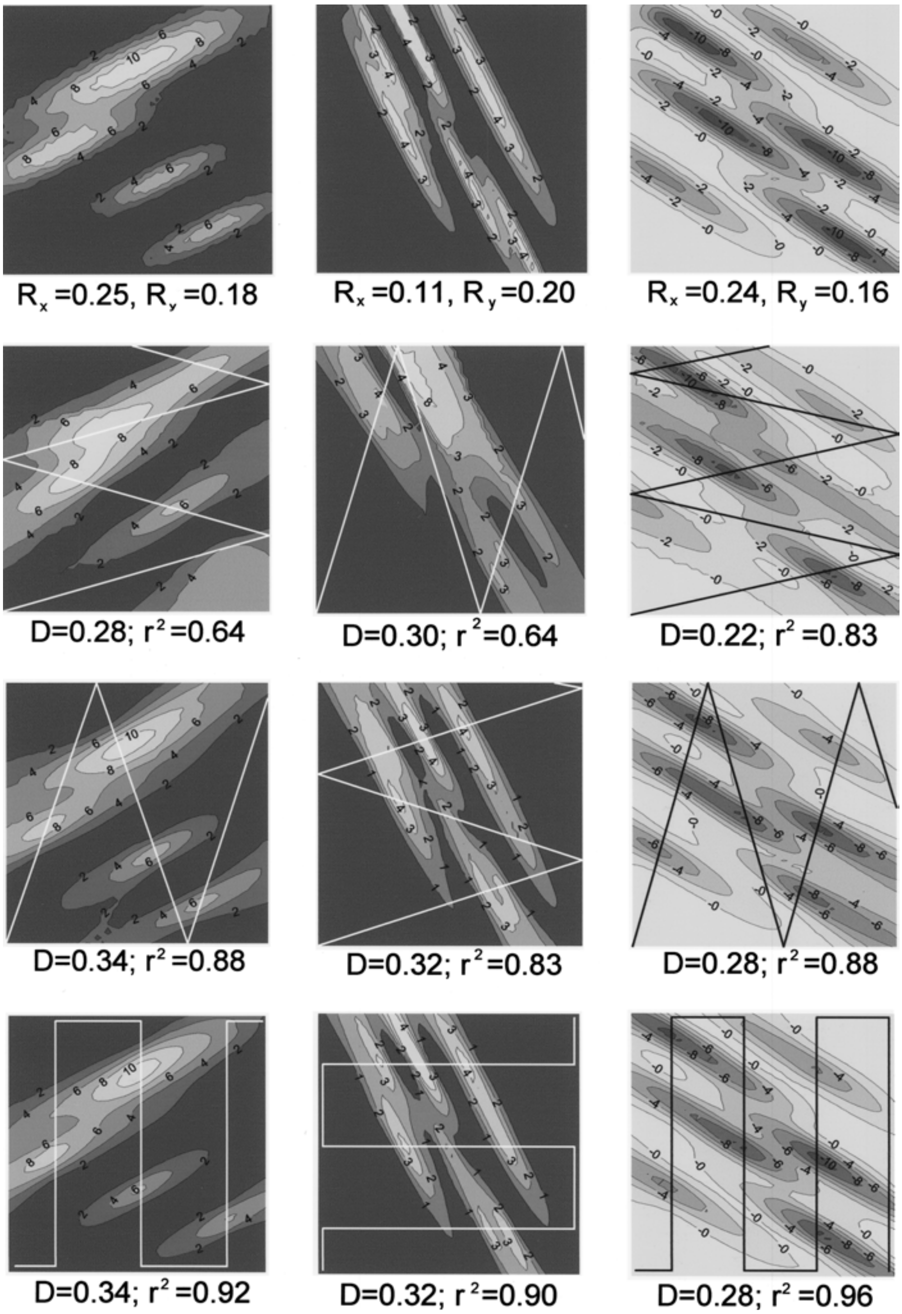


FIGURE 21.6 *First row:* Original patchy rotated anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the arbitrary directions and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

patch elongation equaling 60° permits us to reconstruct the original field with lower adequacy ($r^2 = 0.64$, second row) than the survey conducted at D = 0.32 with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.83$, NSDP = 160, third row). Finally, for the right field, the survey carried out at D = 0.22 with the angle between the survey direction and that of patch elongation equaling 60° gives worse results in terms of the original field adequacy ($r^2 = 0.83$, second row) than the survey conducted at D = 0.28 with the angle between the survey direction and that of patch elongation equaling 30° ($r^2 = 0.88$, NSDP = 179, third row). Thus, the results of experiments indicate that a decrease in the angle between the survey direction and that of patch elongation leads to the higher coefficient of determination between the reconstructed field and that originally generated. The explanation is similar to that given in application to the parallel survey pattern. A comparison of the surveys by zigzag transects (third rows) with those by parallel transects (fourth rows) shows that the former lead to poorer results than the latter do ($r^2 = 0.88$ and $r^2 = 0.92$, left; $r^2 = 0.83$ and $r^2 = 0.90$, middle; $r^2 = 0.88$ and $r^2 = 0.96$, right).

The results of simulated surveys with various ratios of the distance between transects to autocorrelation radius for the field (D/R) are considered in Figure 21.7. The existence of a distinct relationship fitting the generalized dataset (r^2 vs. D/R) confirms the possibility of using the autocorrelation radius as a parameter of a field when choosing the distance between transects. In this figure, all the range of possible values for coefficient of determination is conditionally divided into an area of high values (>0.70) and an area of low ones (<0.70). With the D/R ratio increasing to the critical value, the coefficient of determination remains in the first area, while with further increase, this coefficient falls into the second one. The critical value of the D/R ratio depends on the survey pattern: D/R = 1.5 to 2.0, regarding the parallel pattern; D/R = 1.0 to1.5, in respect to the zigzag one. For this reason, we suggest choosing the distance between transects from the condition that the D/R ratio is equal to a critical value. This distance ensures the match not less than $r^2 > 0.70$ between the field reconstructed on the basis of the data of the survey designed and that really existing in the water body, although unknown.

As in Figure 21.7, the regression curve regarding zigzag transects is located everywhere below that for parallel transects; the former allow less adequate reconstruction of a field than the latter do. A result of a specific mathematical experiment regarding zigzag transects may be above that for parallel transects. This indicates that there is a great probability that the conclusion is correct.

Fixed sampling effort. In all of the following examples, the overall survey paths regarding the parallel and zigzag patterns are equal to each other. To demonstrate the best results regarding the parallel or zigzag pattern, each example given below is selected from the entirety of surveys with the same overall path but various starting points.

The examples of original isotropic fields, given in Figure 21.1, first row, are reproduced in Figure 21.8, first row. The larger the patches, the smaller the required overall survey path (compare the surveys presented on the left column with those given on the middle and right ones). A comparison of the surveys by parallel transects (second row) with those by zigzag transects (third row) shows that the former lead to poorer results than the latter do: $r^2 = 0.66$ and $r^2 = 0.72$ (NSDP = 200, left column),

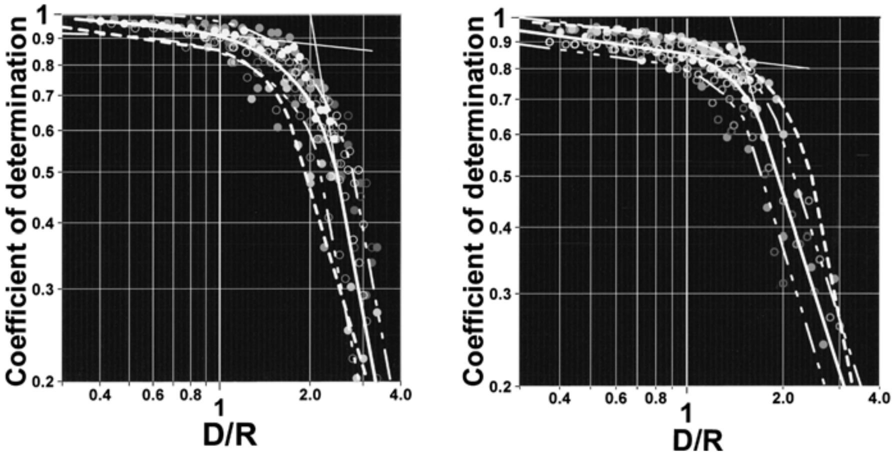


FIGURE 21.7 Coefficient of determination between the reconstructed field and that originally generated as a function of the D/R ratio with regard to various survey patterns (*left*: parallel; *right:* zigzag). Filled symbols correspond to surveys of patchy distribution fields; empty symbols: gappy fields. Colors correspond to the results of simulated surveys of the fields shown in Figure 21.1 to Figure 21.6. Dark blue: Figure 21.1 and Figure 21.4, *left*; red: Figure 21.1 and Figure 21.4, *middle* and *right*; crimson: Figure 21.2 and Figure 21.5, *left*; light blue: Figure 21.2 and Figure 21.5, *middle* and *right*; yellow: Figure 21.3 and Figure 21.6, *left*; green: Figure 21.3 and Figure 21.6, *middle* and *right* (surveys corresponding to the empty dark blue, crimson, and yellow circles are not shown in Figure 21.1 to Figure 21.6). The coefficient of determination regarding isotropic fields is obtained by averaging the results of the surveys carried out in the directions of abscissa and ordinate axes. The solid line represents the nonlinear regression ensuring the minimal sum of squared residuals (the ends of the straight lines indicate the bend of the nonlinear regression). The dashed lines are the borders of the confidence interval with the probability of p = 0.99. The thick dashed line corresponds to the regression for the survey pattern shown in the opposite position of the figure (*left:* zigzag; *right:* parallel). (See color insert following page 490.)

$r^2 = 0.72$ and $r^2 = 0.81$ (NSDP = 400, middle column), or $r^2 = 0.71$ and $r^2 = 0.74$ (NSDP = 400, right column).

The examples of original anisotropic fields, given in Figure 21.2, first row, are reproduced in Figure 21.9, first row. The direction of patch elongation for the left and right fields is the abscissa, while for the middle one it is the ordinate; as shown earlier, these are the optimal directions for acoustic survey. A comparison of the surveys by parallel transects (second row) with those by zigzag transects (third row) shows that the former lead to poorer results than the latter do: $r^2 = 0.91$ and $r^2 = 0.92$ (left column), $r^2 = 0.89$ and $r^2 = 0.92$ (middle column), or $r^2 = 0.80$ and $r^2 = 0.82$ (right column) (NSDP = 200).

The examples of rotated original anisotropic fields, given in Figure 21.3, first row, are reproduced in Figure 21.10, first row. A comparison of the surveys by parallel transects (second row) with those by zigzag transects (third row) shows that the former lead to poorer results than the latter do: $r^2 = 0.86$ and $r^2 = 0.89$

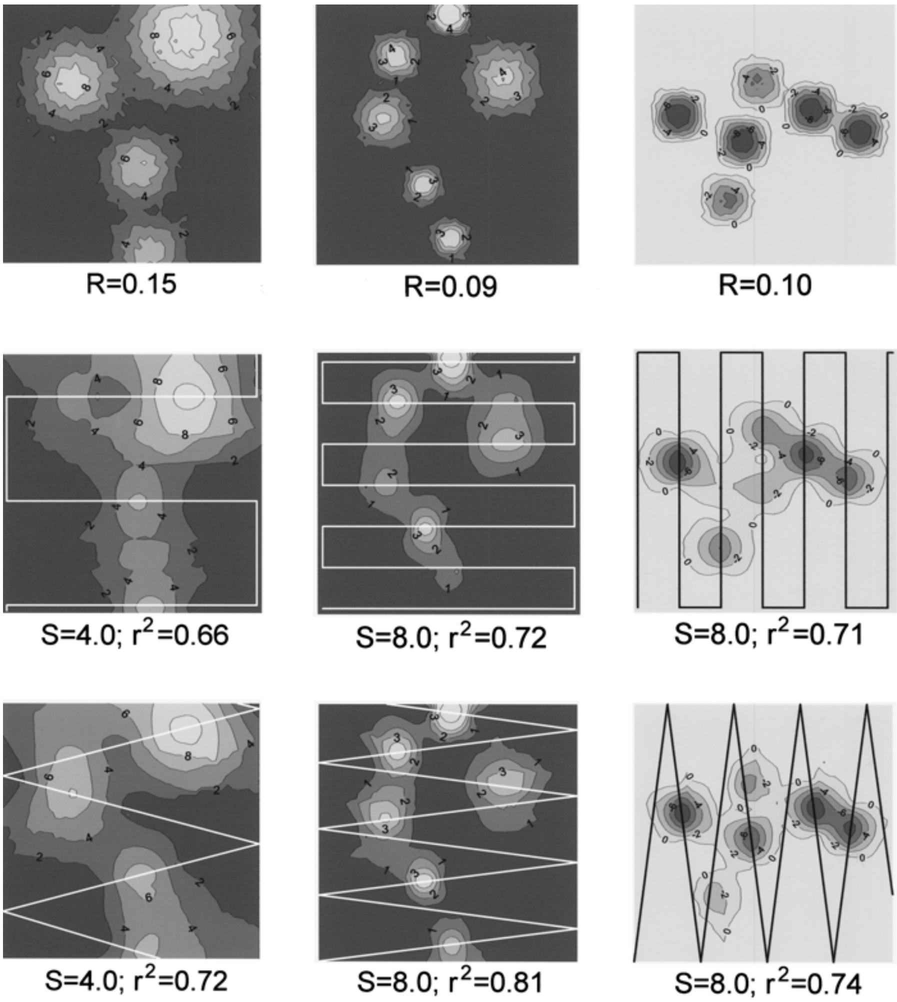


FIGURE 21.8 *First row:* Original patchy isotropic fields. *Second* and *third rows:* Paths of simulated surveys by parallel or zigzag transects and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

(left column), $r^2 = 0.78$ and $r^2 = 0.87$ (middle column), or $r^2 = 0.70$ and $r^2 = 0.73$ (right column) (NSDP = 200).

As the path of the survey crosses the field in various directions, the average value of the autocorrelation radius should be determined. Because further along it is intended to use the ratio of S/R_{av} to average the ratio and to represent the results in logarithmic scale, the use of the geometric mean is preferable.[14] Regarding the zigzag pattern, crossing of a field by a transect, oriented in an arbitrary direction, results in appearance of the two components of the autocorrelation radius: in the direction perpendicular to a survey and in the direction of a survey (the latter

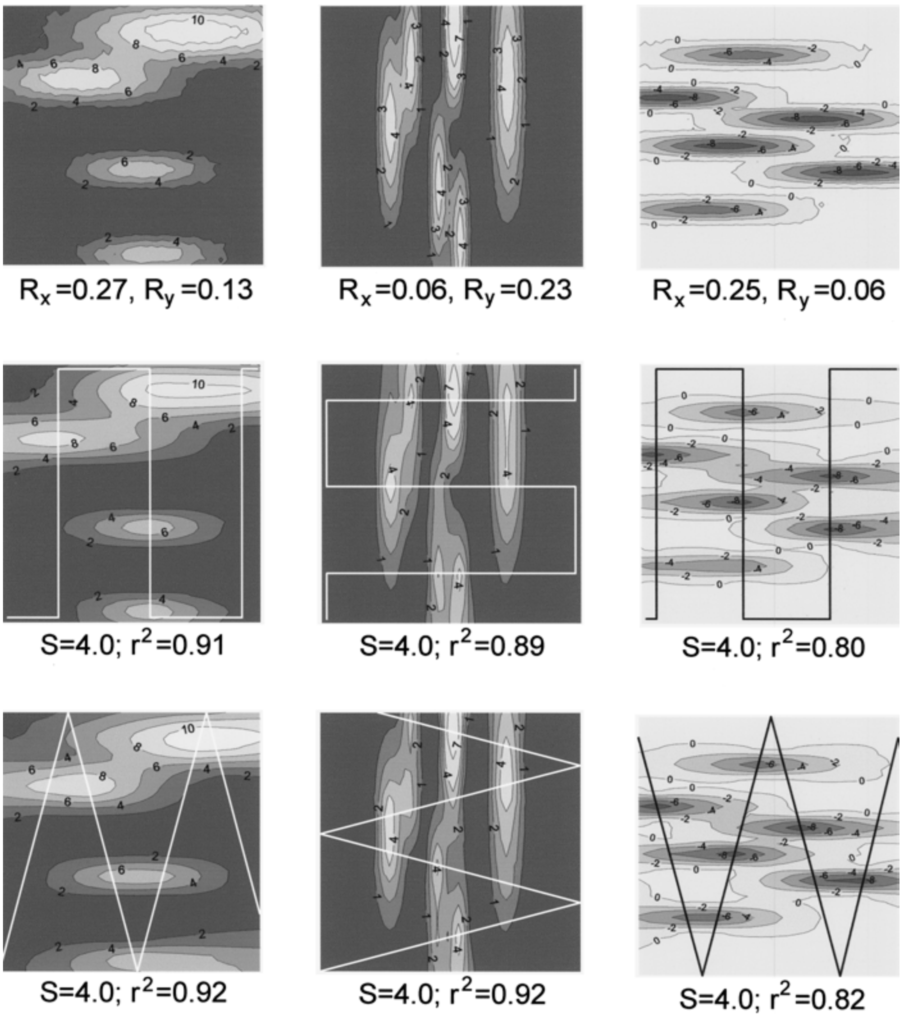


FIGURE 21.9 *First row:* Original patchy anisotropic fields. *Second* and *third rows:* Paths of simulated surveys by parallel or zigzag transects conducted in the directions of major axes of autocorrelation ellipses and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

represents, in fact, a part of the component of the autocorrelation radius). The sum of these partial components, appearing as a result of all the transects, is equal to the full component of the autocorrelation radius.[21] Regarding the parallel pattern, the only difference consists of the fact that the projection of the autocorrelation radius coincides with the radius itself. Therefore, for both patterns, the average autocorrelation radius can be calculated from the following formula:

$$R_{av} = \sqrt[n+1]{(R_p)^n R},$$

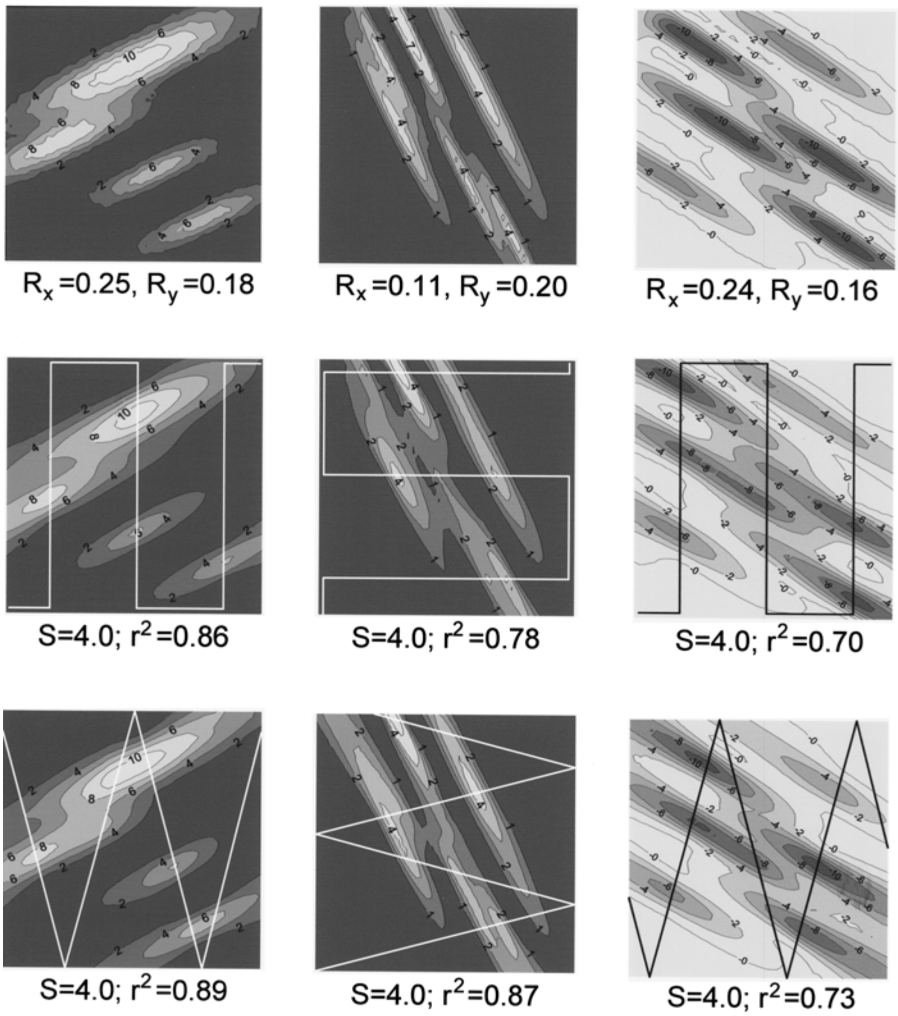


FIGURE 21.10 *First row:* Original patchy rotated anisotropic fields. *Second* and *third rows:* Paths of simulated surveys by parallel or zigzag transects conducted in arbitrary directions and the reconstructed fields. *Right column* represents the gaps (the isolines are labeled by negative numbers).

where n is the number of transects, R and R_p are the autocorrelation radii in the direction of a survey and in the perpendicular direction, respectively.

The results of simulated surveys with various ratios of the overall survey path to the autocorrelation radius for the field (S/R_{av}) are considered below (Figure 21.11). The existence of a distinct relationship fitting the generalized dataset (r^2 vs. S/R_{av}) confirms the possibility of using this ratio in determining the efficiency of acoustic surveys. The results of the surveys conducted with the same overall paths but various starting points are presented in Figure 21.11 as columns. As seen in the figure, the

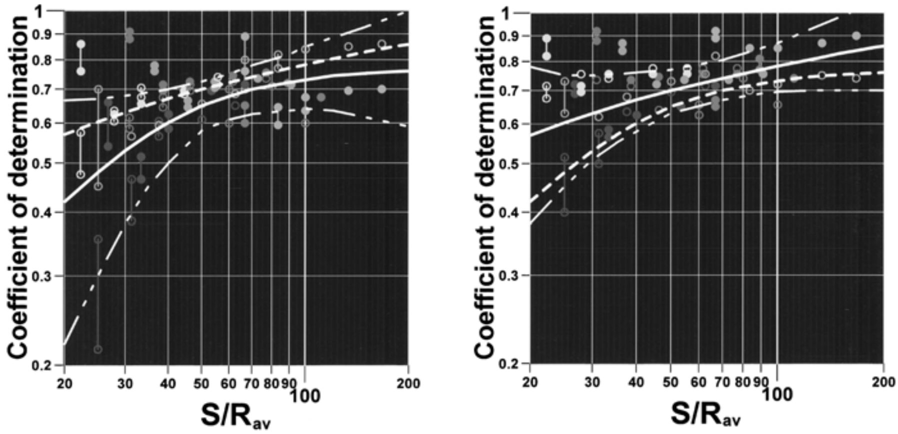


FIGURE 21.11 Coefficient of determination between the reconstructed field and that originally generated as a function of the S/R$_{av}$ ratio with regard to various survey patterns (*left:* parallel; *right:* zigzag). Filled symbols correspond to surveys of patchy distribution fields; empty symbols: gappy fields. Colors correspond to the results of simulated surveys of the fields shown in Figure 21.8 to Figure 21.10. Dark blue: Figure 21.8, *left*; red: Figure 21.8, *middle* and *right*; crimson: Figure 21.9, *left*; light blue: Figure 21.9, *middle* and *right*; yellow: Figure 21.10, *left*; green: Figure 21.10, *middle* and *right* (surveys corresponding to the empty dark blue, crimson, and yellow circles are not shown in Figure 21.8 to Figure 21.10). The coefficient of determination regarding isotropic fields is obtained by averaging the results of the surveys carried out in the directions of abscissa and ordinate axes. The solid line represents the nonlinear regression ensuring the minimal sum of squared residuals. The dashed lines are the borders of the confidence interval with the probability of p = 0.99. The thick dashed line corresponds to the regression for the survey pattern shown in the opposite position of the figure (*left:* zigzag; *right:* parallel). (See color insert following page 490.)

regression curve regarding parallel transects is located everywhere below that for zigzag transects. Therefore, the results obtained permit us to conclude that the former allow, as a rule, less adequate reconstruction of a field than the latter do. The result of a specific mathematical experiment regarding parallel transects may be above that for zigzag transects. This indicates that there is a great probability that the conclusion is correct.

Choice of the unit of sampling distance. Let us consider the choice of this unit in application to the parallel survey pattern. The examples of original isotropic fields, given in Figure 21.1 (first row, left, and middle) are reproduced in Figure 21.12 (first row, left, and middle); for the right field, the autocorrelation radius R = 0.16. The following distances between transects are taken as examples: D = 0.22 (left), D = 0.14 (middle), or D = 0.24 (right). In the case of the left field, the increase of the unit of sampling distance from d = 1/50 to d = 1/7 results in lowering the coefficient of determination from $r^2 = 0.83$ (NSDP = 300, second row) to $r^2 = 0.74$ (NSDP = 40, third row). For the middle field, the increase of the unit of sampling distance from d = 1/50 to d = 1/12 results in lowering the coefficient of determination from $r^2 = 0.77$ (NSDP = 450, second row) to $r^2 = 0.74$ (NSDP = 104, third row). For the

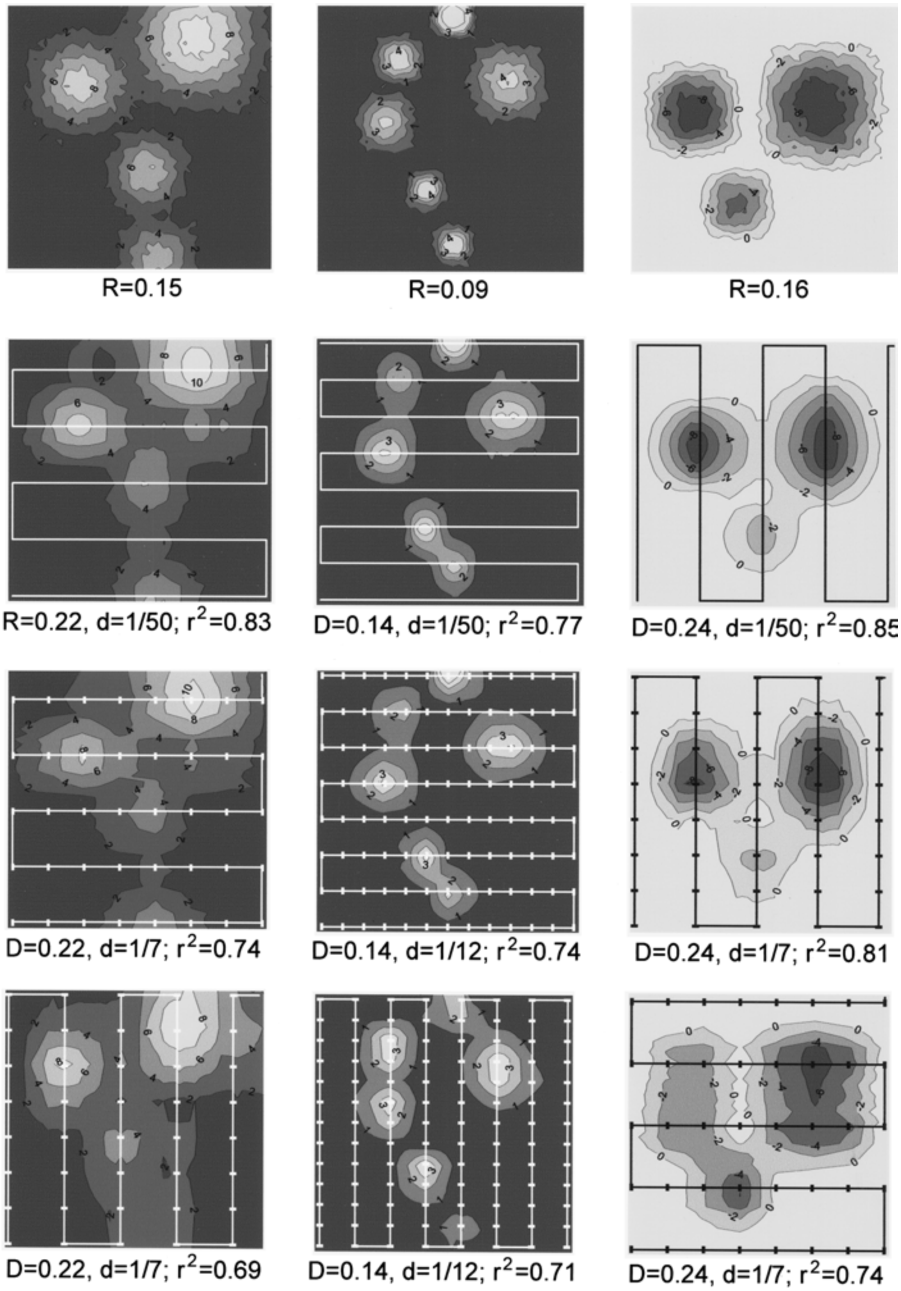


FIGURE 21.12 *First row:* Original patchy isotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys and the reconstructed fields (*third* and *fourth rows* show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

right field, the increase of the unit of sampling distance from d = 1/50 to d = 1/7 results in lowering the coefficient of determination from $r^2 = 0.85$ (NSDP = 300, second row) to $r^2 = 0.81$ (NSDP = 40, third row). Thus, the experiments conducted confirm the fact that, in the case of isotropic fields, with a constant ratio of the distance between transects to the autocorrelation radius in the survey direction (D/R), the adequacy of reconstruction of a field depends on the ratio of the sampling distance unit to the autocorrelation radius in the direction perpendicular to the survey (d/R_p).

In the case of the left field, the surveys carried out at D = 0.22 and d = 1/7 in the two perpendicular directions give similar results of $r^2 = 0.74$ (third row) and $r^2 = 0.69$ (fourth row); for the middle field, the surveys conducted at D = 0.14 and d = 1/12 in the two perpendicular directions give similar results of $r^2 = 0.74$ (third row) and $r^2 = 0.71$ (fourth row); for the right field, the surveys carried out at D = 0.24 and d = 1/7 in the two perpendicular directions also give similar results of $r^2 = 0.81$ (third row) and $r^2 = 0.74$ (fourth row). Thus, for isotropic fields and a given distance between transects and unit of sampling distance, the adequacy of reconstructing a distribution field is practically independent of the direction of the survey performed.

The examples of original anisotropic fields, given in Figure 21.2 (first row, left, and middle) are reproduced in Figure 21.13 (first row, left, and middle); for the right field, the autocorrelation radii $R_x = 0.13$ and $R_Y = 0.28$. As shown earlier, the optimal direction for the survey of the left field is the abscissa, while for the middle and right ones it is the ordinate. The following distances between transects are taken as examples: D = 0.40 (left), D = 0.32 (middle), or D = 0.40 (right). In the case of the left field, the increase of the unit of sampling distance from d = 1/50 to d = 1/8 results in lowering the coefficient of determination from $r^2 = 0.92$ (NSDP = 200, second row) to $r^2 = 0.86$ (NSDP = 27, third row). For the middle field, the increase of the unit of sampling distance from d = 1/50 to d = 1/8 results in lowering the coefficient of determination from $r^2 = 0.83$ (NSDP = 250, second row) to $r^2 = 0.66$ (NSDP = 36, third row). For the right field, the increase of the unit of sampling distance from d = 1/50 to d = 1/6 results in lowering the coefficient of determination from $r^2 = 0.90$ (NSDP = 200, second row) to $r^2 = 0.86$ (NSDP = 21, third row). Thus, the mathematical experiments confirm that in the case of anisotropic fields similar to that of isotropic ones, with a constant ratio of the distance between transects to the autocorrelation radius in the survey direction (D/R), the adequacy of reconstructing a field depends on the ratio of the sampling distance unit to the autocorrelation radius in the direction perpendicular to the survey (d/R_p).

In the case of the left field, a survey carried out at D = 0.40 and d = 1/8 in the direction of patch elongation gives almost the same results in terms of the original field adequacy ($r^2 = 0.86$, NSDP = 27, third row) as a survey at D = 0.12 and d = 1/4 conducted in a perpendicular direction ($r^2 = 0.90$, NSDP = 45, fourth row). For the middle field, a survey carried out at D = 0.32 and d = 1/8 in the direction of patch elongation permits us to reconstruct the original field with almost the same adequacy ($r^2 = 0.66$, NSDP = 36, third row) as a survey conducted at D = 0.10 and d = 1/3 in a perpendicular direction ($r^2 = 0.69$, NSDP = 40, fourth row). Finally, for the right field, a survey carried out at D = 0.40 and d = 1/6 in the direction of patch elongation gives even better results ($r^2 = 0.86$, NSDP = 21, third row) than a survey conducted

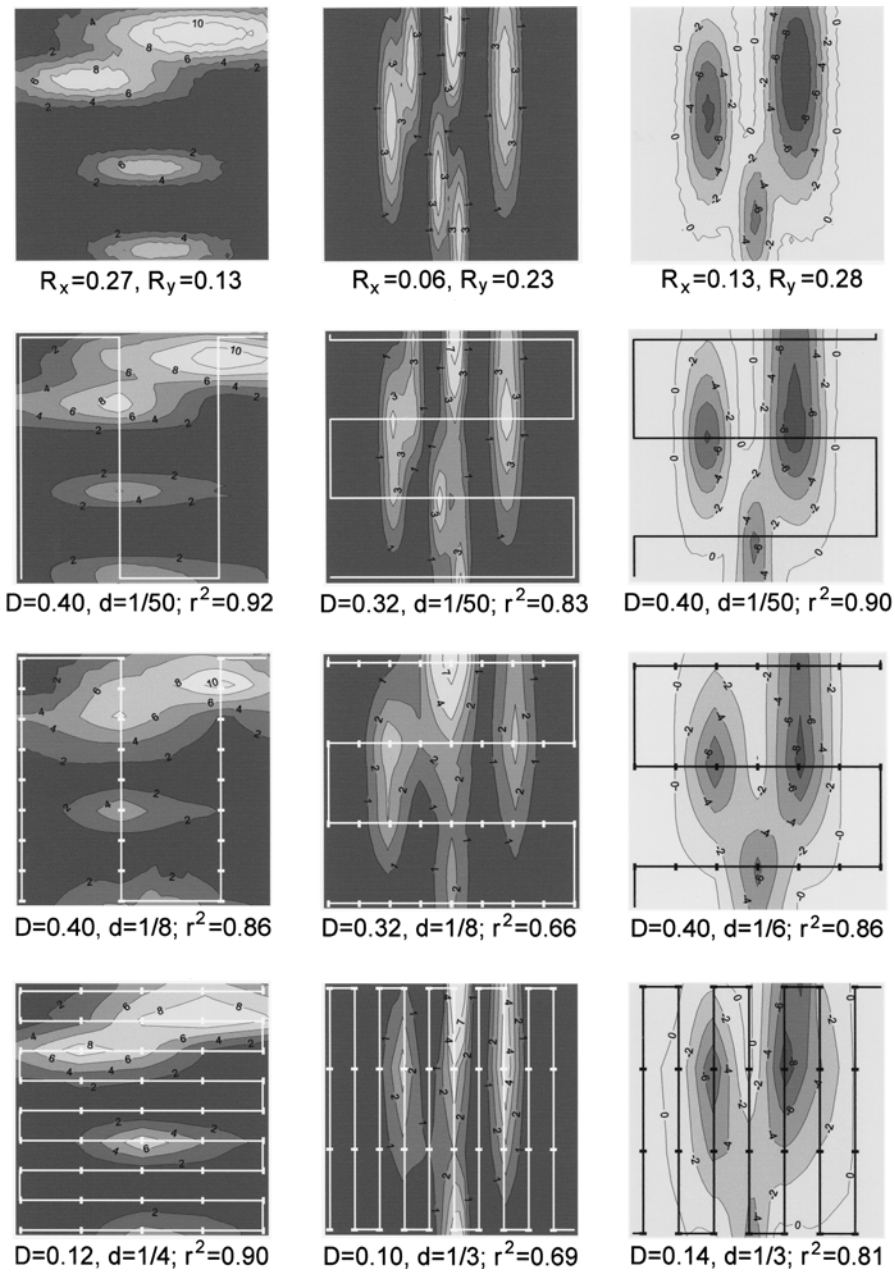


FIGURE 21.13 *First row:* Original patchy anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the directions of the major or minor axes of autocorrelation ellipses and the reconstructed fields (*third* and *fourth rows* show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

at D = 0.14 and d = 1/3 in a perpendicular direction (r^2 = 0.81, NSDP = 28, fourth row). Thus, the surveys in the direction of patch elongation with a larger distance between transects and a smaller unit of sampling distance permit us to reconstruct the original field with about the same adequacy as those in a perpendicular direction with a smaller distance between transects and a larger unit of sampling distance.

The examples of rotated original anisotropic fields, given in Figure 21.3 (first row, left, and middle) are reproduced in Figure 21.14 (first row, left, and middle); for the right field, the autocorrelation radii R_x = 0.18 and R_Y = 0.27. The increase in the angle between the survey direction and that of patch elongation makes us set a smaller distance between transects so that we may achieve the same adequacy of reconstructing an original patchy distribution field. At the same time, a larger unit of sampling distance can be set. For example, in the case of the left field, the survey carried out at D = 0.40 and d = 1/7 with the angle between the survey direction and that of patch elongation equaling 30° gives even better results in terms of the original field adequacy (r^2 = 0.77, NSDP = 24, third row) than the survey conducted at D = 0.28 and d = 1/6 with the angle between the survey direction and that of patch elongation equaling 60° (r^2 = 0.74, NSDP = 28, fourth row). For the middle field, the survey carried out at D = 0.30 and d = 1/12 with the angle between the survey direction and that of patch elongation equaling 30° permits us to reconstruct the original field with the same adequacy (r^2 = 0.61, NSDP = 52, third row) as the survey conducted at D = 0.14 and d = 1/7 with the angle between the survey direction and that of patch elongation equaling 60° (r^2 = 0.61, NSDP = 64, fourth row). Finally, for the right field, the survey carried out at D = 0.24 and d = 1/7 with the angle between the survey direction and that of patch elongation equaling 30° gives even better results in terms of the original field adequacy (r^2 = 0.88, NSDP = 40, third row) than the survey conducted at D = 0.16 and d = 1/5 with the angle between the survey direction and that of patch elongation equaling 60° (r^2 = 0.85, NSDP = 42, fourth row). This is explained, ultimately, by the relationship of the autocorrelation radii located in two reciprocally perpendicular directions.

The choice of the unit of sampling distance in application to the zigzag survey pattern is considered below. Let us note that with account of the fact that the unit of sampling distance (d) is given in proportion to a transect length, the values of d taken in the direction perpendicular to the survey are equal to those taken along the transect (Figure 21.15).

The examples of original isotropic fields, given in Figure 21.12, first row, are reproduced in Figure 21.16, first row. The following distances between transects are taken as examples: D = 0.24 (left), D = 0.10 (middle) or D = 0.24 (right). In the case of the left field, the increase of the unit of sampling distance from d = 1/50 to d = 1/6 results in lowering the coefficient of determination from r^2 = 0.71 (NSDP = 196, second row) to r^2 = 0.67 (NSDP = 23, third row). For the middle field, the increase of the unit of sampling distance from d = 1/50 to d = 1/8 results in lowering the coefficient of determination from r^2 = 0.79 (NSDP = 499, second row) to r^2 = 0.67 (NSDP = 88, third row). For the right field, the increase of the unit of sampling distance from d = 1/50 to d = 1/5 results in lowering the coefficient of determination from r^2 = 0.72 (NSDP = 199, second row) to r^2 = 0.64 (NSDP = 19, third row). Thus, the experiments conducted confirm the fact that regarding the zigzag pattern,

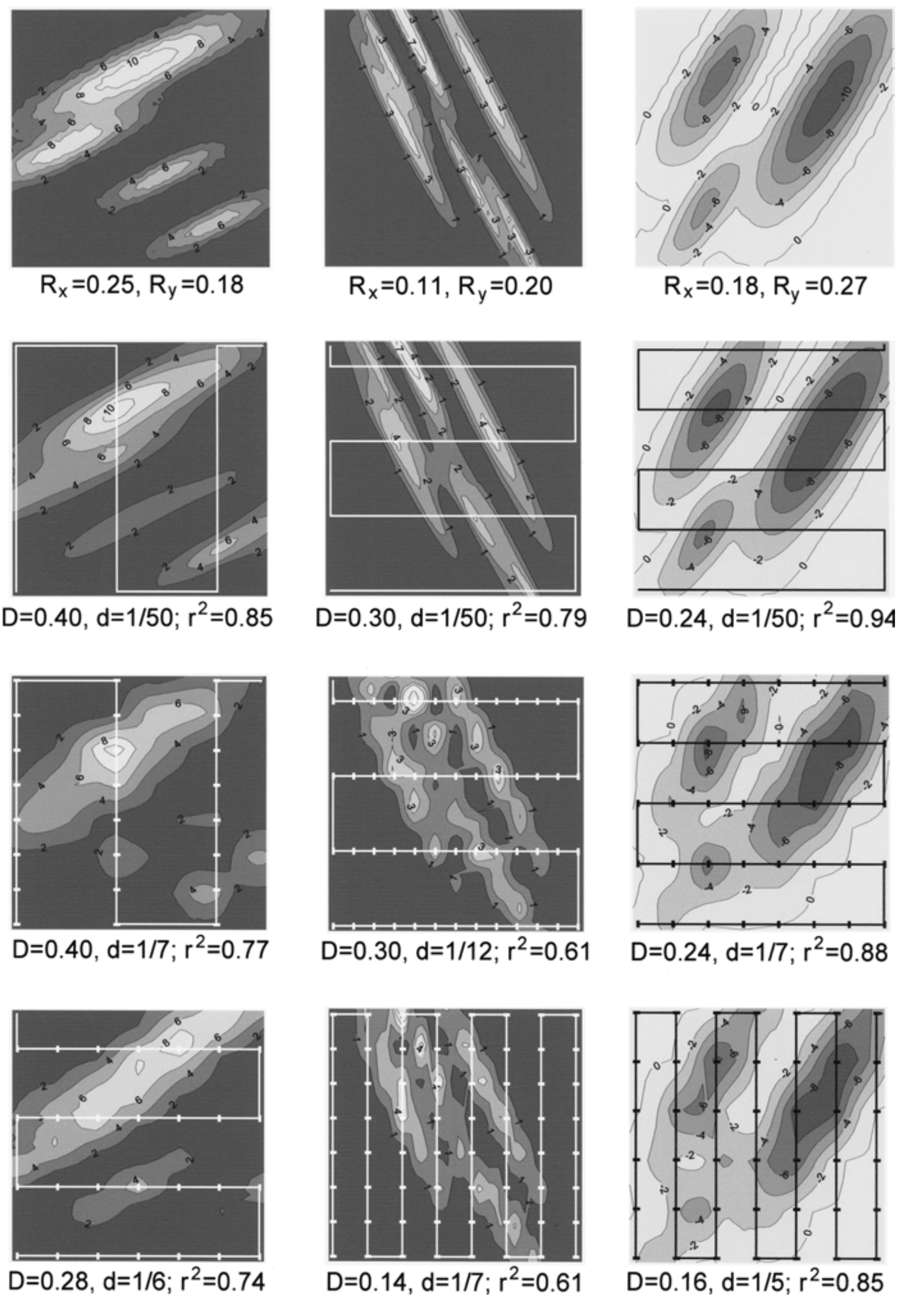


FIGURE 21.14 *First row:* Original patchy rotated anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the arbitrary directions and the reconstructed fields (*third* and *fourth rows* show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

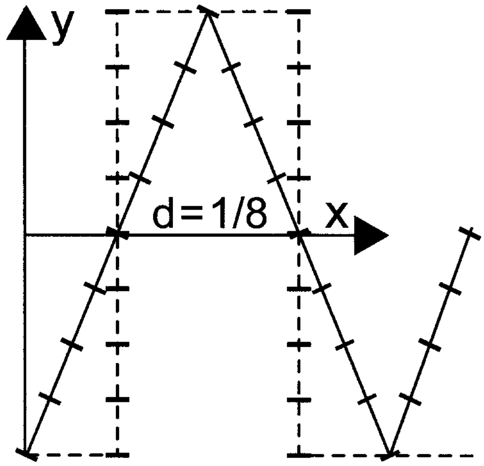


FIGURE 21.15 Map of acoustic surveys by various pattern of transects (zigzag: solid line; parallel: dashed line), along which a unit of sampling distance (d) is set. Axes correspond to the direction of the survey (x) and the perpendicular direction (y).

similar to the parallel one in the case of isotropic fields, with a constant ratio of the distance between transects to the autocorrelation radius in the survey direction (D/R), the adequacy of reconstruction of a field to its original depends on the ratio of the sampling distance unit to the autocorrelation radius in the direction perpendicular to the survey (d/R_p). A comparison shows that, with fixed transect spacing and a given number of sampling points on each full transect, the zigzag pattern leads to poorer results than the parallel pattern does ($r^2 = 0.67$ and $r^2 = 0.71$, left; $r^2 = 0.67$ and $r^2 = 0.71$, middle; $r^2 = 0.64$ and $r^2 = 0.74$, right) (third and fourth row).

The examples of original anisotropic fields, given in Figure 21.13, first row, are reproduced in Figure 21.17, first row. The following distances between transects are taken as example: D = 0.28 (left), D = 0.34 (middle) or D = 0.40 (right). In the case of the left field, the increase of the unit of sampling distance from d = 1/50 to d = 1/7 results in lowering the coefficient of determination from $r^2 = 0.94$ (NSDP = 180, second row) to $r^2 = 0.76$ (NSDP = 24, third row). For the middle field, the increase of the unit of sampling distance from d = 1/50 to d = 1/8 results in lowering the coefficient of determination from $r^2 = 0.90$ (NSDP = 152, second row) to $r^2 = 0.55$ (NSDP = 24, third row). For the right field, the increase of the unit of sampling distance from d = 1/50 to d = 1/8 results in lowering the coefficient of determination from $r^2 = 0.86$ (NSDP = 129, second row) to $r^2 = 0.81$ (NSDP = 20, third row). Thus, the mathematical experiments confirm that in the case of anisotropic fields, similar to that of isotropic ones, with a constant ratio of the distance between transects to the autocorrelation radius in the survey direction (D/R), the adequacy of a reconstructed field to its original depends on the ratio of the sampling distance unit to the autocorrelation radius in the direction perpendicular to the survey (d/R_p). A comparison of the surveys by zigzag transects (third row) with those by parallel transects (fourth row) shows that, with fixed transect spacing and a given number of sampling points on each full transect,

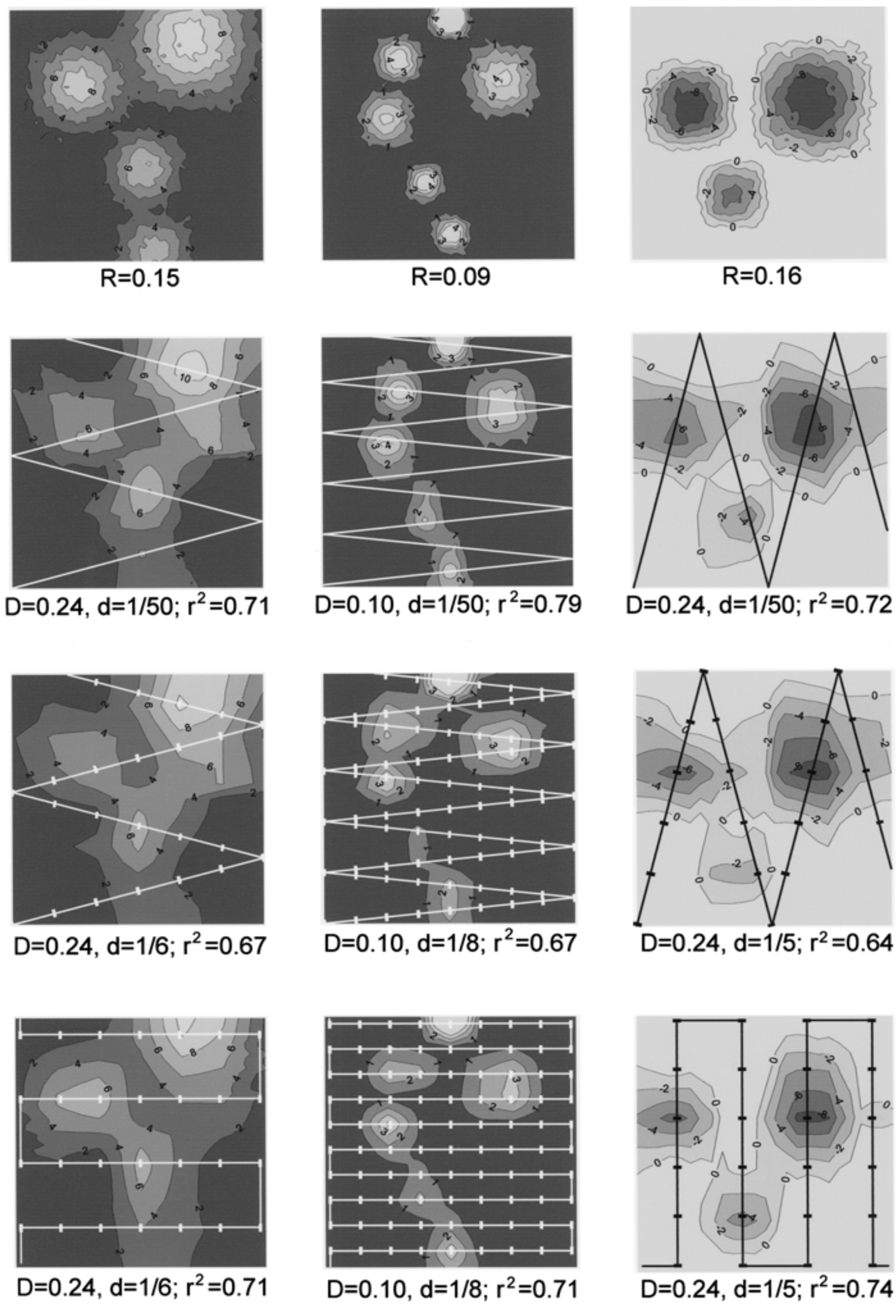


FIGURE 21.16 *First row:* Original patchy isotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys and the reconstructed fields (*third* and *fourth rows* show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

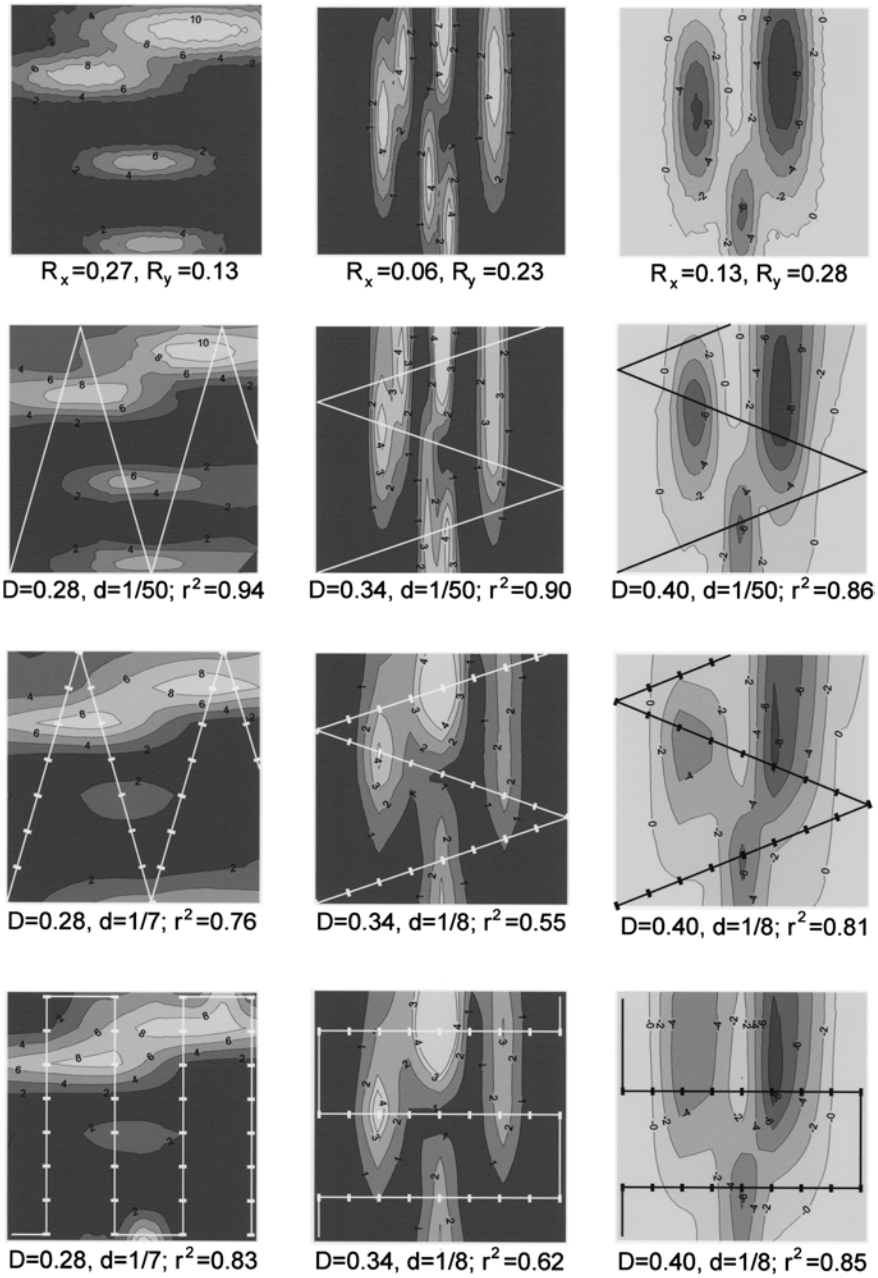


FIGURE 21.17 *First row:* Original patchy anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the directions of major axes of autocorrelation ellipses and the reconstructed fields (*third* and *fourth r*ows show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

the former lead to the poorer results than the latter do ($r^2 = 0.76$ and $r^2 = 0.83$, left; $r^2 = 0.55$ and $r^2 = 0.62$, middle; $r^2 = 0.81$ and $r^2 = 0.85$, right).

The examples of rotated original anisotropic fields, given in Figure 21.14, first row, are reproduced in Figure 21.18, first row. The following distances between transects are taken as example: D = 0.34 (left), D = 0.30 (middle) or D = 0.36 (right). In the case of the left field, the increase of the unit of sampling distance from d = 1/50 to d = 1/10 results in lowering the coefficient of determination from $r^2 = 0.88$ (NSDP = 156, second row) to $r^2 = 0.81$ (NSDP = 30, third row). For the middle field, the increase of the unit of sampling distance from d = 1/50 to d = 1/10 results in lowering the coefficient of determination from $r^2 = 0.85$ (NSDP = 161, second row) to $r^2 = 0.59$ (NSDP = 31, third row). For the right field, the increase of the unit of sampling distance from d = 1/50 to d = 1/7 results in lowering the coefficient of determination from $r^2 = 0.90$ (NSDP = 144, second row) to $r^2 = 0.81$ (NSDP = 20, third row). Thus, the mathematical experiments confirm that in the case of rotated anisotropic fields, with a constant ratio of the distance between transects to the autocorrelation radius in the survey direction (D/R), the adequacy of a reconstructed field to its original depends on the ratio of the sampling distance unit to the autocorrelation radius in the direction perpendicular to the survey (d/R_p). A comparison of the surveys by zigzag transects (third row) with those by parallel transects (fourth row) shows that, with fixed transect spacing and a given number of sampling points on each full transect, the former lead to poorer results than the latter do ($r^2 = 0.81$ and $r^2 = 0.85$, left; $r^2 = 0.59$ and $r^2 = 0.61$, middle; $r^2 = 0.81$ and $r^2 = 0.85$, right).

The results of the simulated surveys with various ratios of sampling distance units to autocorrelation radius (d/R_p) are considered below (Figure 21.19). The existence of a distinct relationship fitting the generalized dataset (r^2 vs. d/R_p) confirms the possibility of using the autocorrelation radius in the direction perpendicular to the survey as a parameter of a field in choosing the unit of sampling distance. In this figure, all the range of possible values for coefficient of determination is conditionally divided into an area of high values (>0.70) and an area of low ones (<0.70). With the d/R_p ratio increasing to the critical value, i.e., to 1.0 to 1.5, the coefficient of determination remains in the first area, while with further increase, this coefficient falls into the second one. For this reason, we suggest choosing the unit of sampling distance from the condition $d/R_p = 1.0$ to 1.5. This unit ensures the match not less than $r^2 > 0.70$ between the field reconstructed on the basis of the data of the survey designed and that really existing in the water body, although unknown.

As seen in Figure 21.19, the lines representing the nonlinear regression for both patterns practically coincide with each other. However, as mentioned earlier, the values of the unit of sampling distance are given in proportion to the transect length. Consequently, the absolute values of the sampling distance unit regarding the zigzag transects will be, in general, larger than those given in application to parallel transects. The extent of the increase depends on the inclination of a transect from the direction perpendicular to the survey.

Dynamic fields. Let us suppose that all of the patches move in the same direction, uniformly and rectilinearly; simulated surveys by parallel or zigzag transects are undertaken in the opposite/same directions relative to that of the patch movement; the speed of the patch movement is equal to a specific part of that of the survey

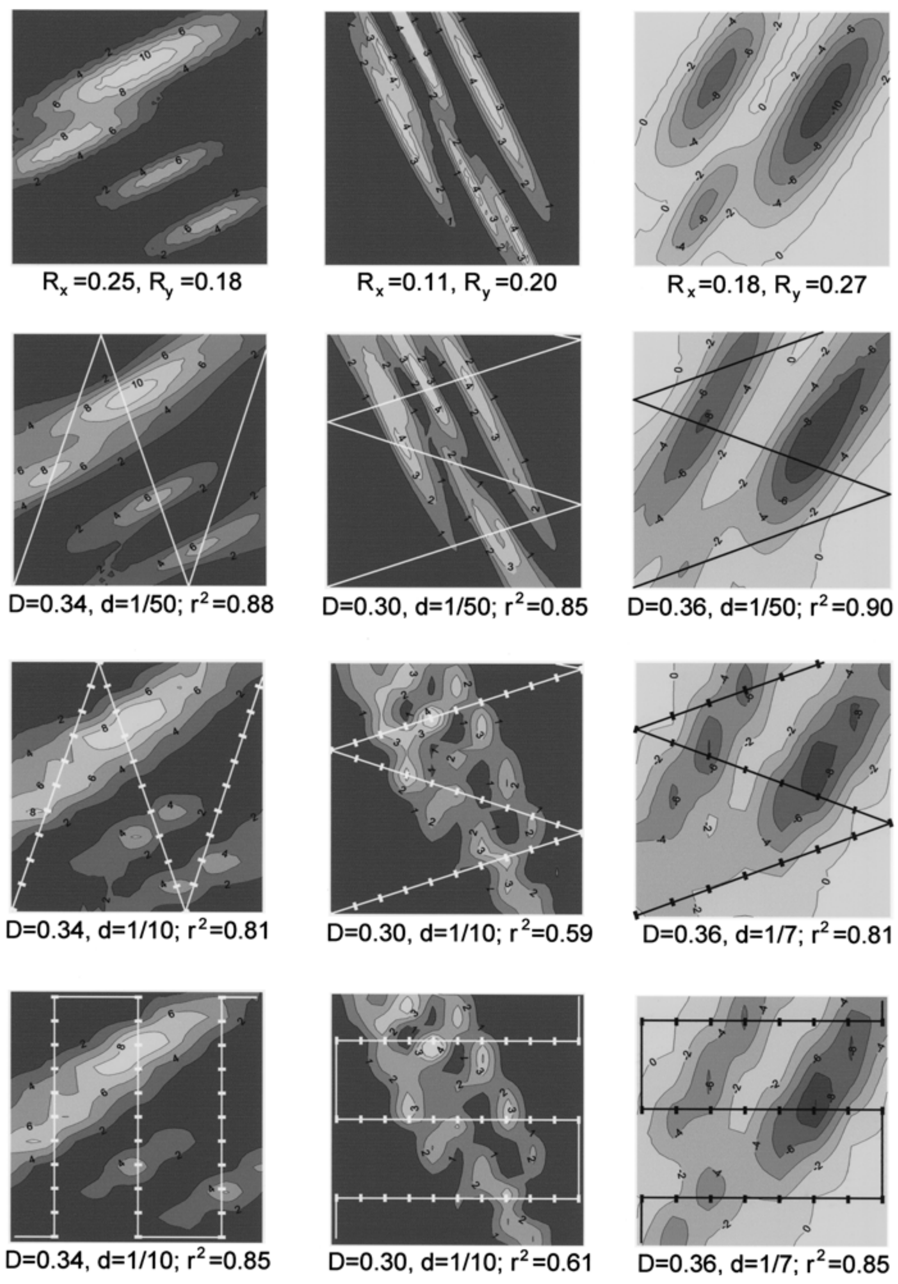


FIGURE 21.18 *First row:* Original patchy rotated anisotropic fields. *Second* through *fourth rows:* Paths of the simulated surveys conducted in the arbitrary directions and the reconstructed fields (*third* and *fourth rows* show the unit of sampling distance set along the transects). *Right column* represents the gaps (the isolines are labeled by negative numbers).

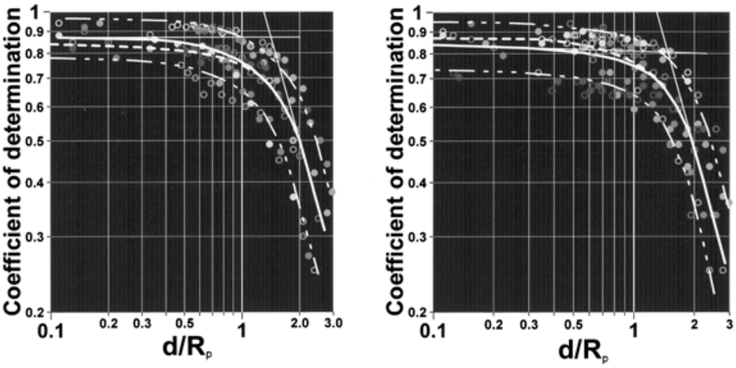


FIGURE 21.19 Coefficient of determination between the reconstructed field and that originally generated as a function of the d/R_p ratio with regard to various survey patterns (*left:* parallel; *right:* zigzag). Filled symbols correspond to surveys of patchy distribution fields; empty symbols: gappy fields. Colors correspond to the results of simulated surveys of the fields shown in Figure 21.12 to Figure 21.14 and Figure 21.16 to Figure 21.18. Dark blue: Figure 21.12 and Figure 21.16, *left* and *right*; red: Figure 21.12 and Figure 21.16, *middle*; crimson: Figure 21.13 and Figure 21.17, *left* and *right*; light blue: Figure 21.13 and Figure 21.17, *middle*; yellow: Figure 21.14 and Figure 21.18, *left* and *right*; green: Figure 21.14 and Figure 21.18, *middle* (surveys corresponding to the empty red, light blue, and green circles are not shown in Figure 21.12 to Figure 21.14 and Figure 21.16 to Figure 21.18). The coefficient of determination regarding isotropic fields is obtained by averaging the results of the surveys carried out in the directions of abscissa and ordinate axes. The solid line represents the nonlinear regression ensuring the minimal sum of squared residuals (the ends of the straight lines indicate the bend of the nonlinear regression). The dashed lines are the borders of the confidence interval with the probability of $p = 0.99$. The thick dashed line corresponds to the regression for the survey pattern shown in the opposite position of the figure (*left:* zigzag; *right:* parallel). (See color insert following page 490.)

(0.7, 0.5, or 0.3). Survey design is examined with regard to the reconstruction of a field in a stationary coordinate system (survey of a region); i.e., the field is reconstructed disregarding the information on the patch movement and compared with the average original field.

In Figure 21.20a and b, some specific conditions are, in particular, considered (see the caption of the figures). Both surveys are shown simultaneously (left column); to simplify the figures, the unit of sampling distance is not shown. The situation that could be described by the Doppler effect occurs when the survey and the patches move in the opposite directions (middle and right columns, second and third rows); the Doppler effect is also observed when the survey and the patches move in the same direction (middle and right columns, fourth and fifth rows).

In the case when the transect spacing is fixed (Figure 21.20a), the comparison of coefficients of determination obtained both in cases of survey movement in the opposite or same directions relative to that of the patch movement shows that the parallel pattern leads to better results than the zigzag pattern does ($r^2 = 0.67$ and $r^2 = 0.64$, middle column, second and third rows; $r^2 = 0.61$ and $r^2 = 0.58$; middle

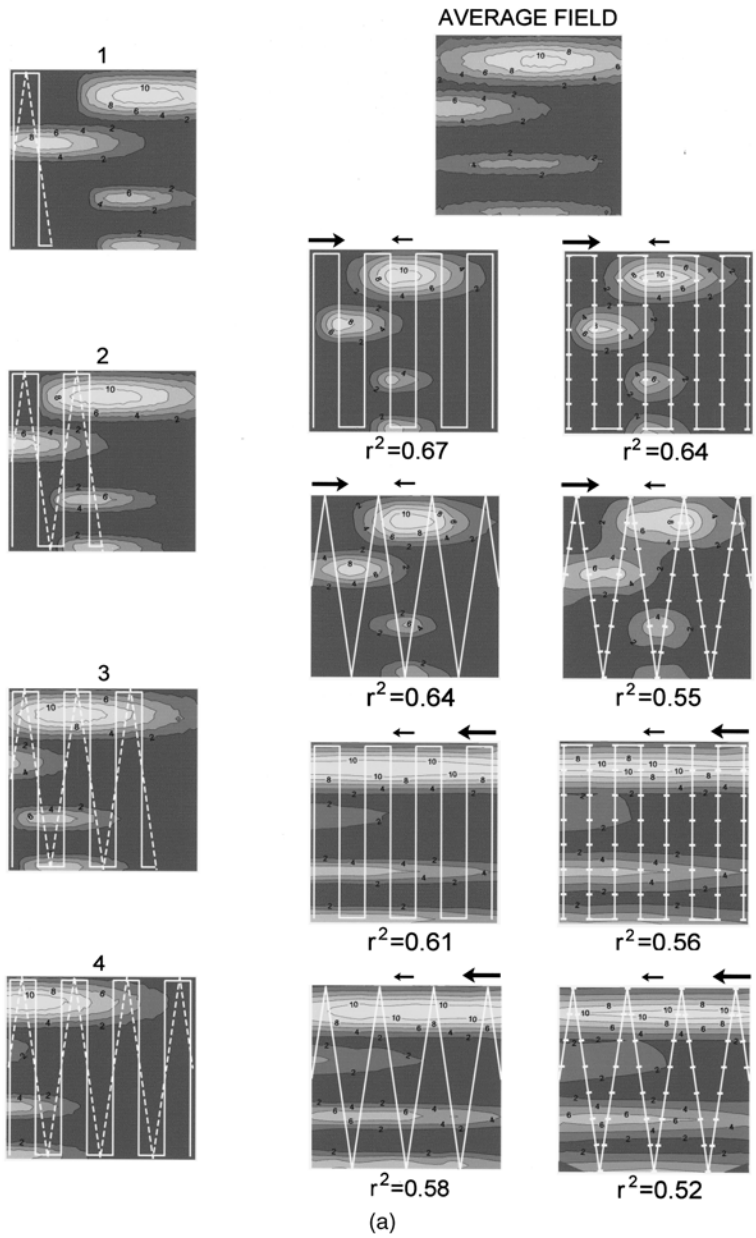
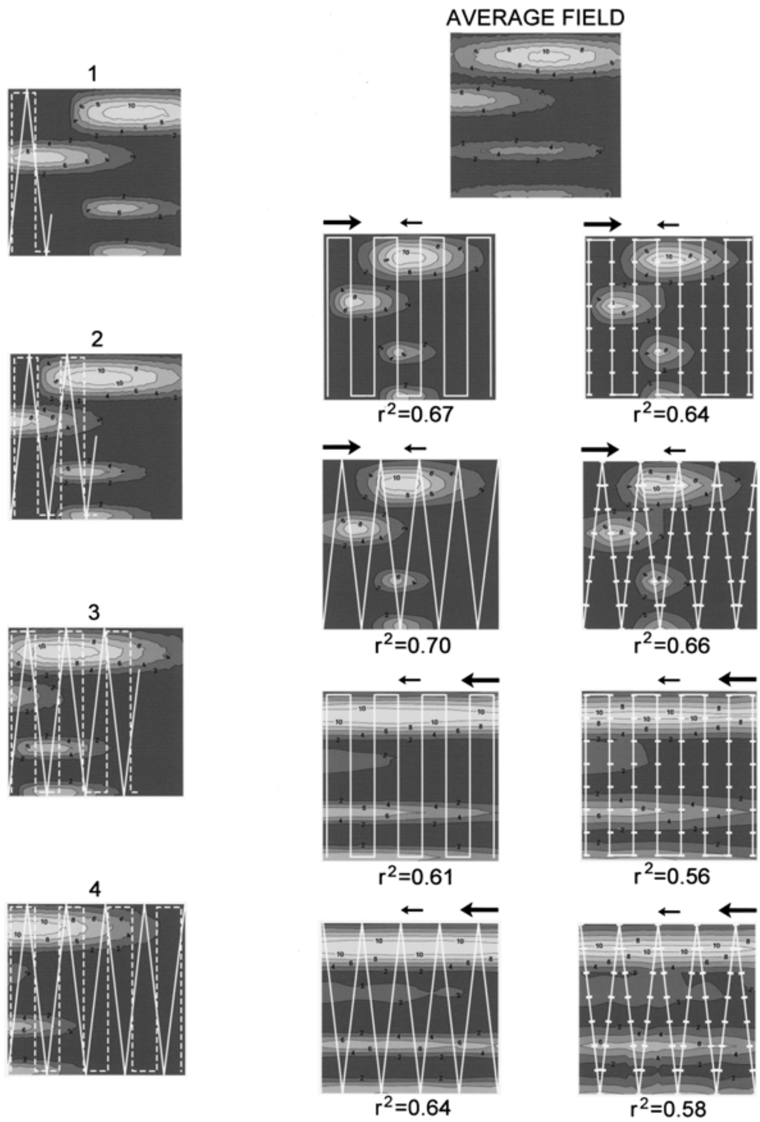


FIGURE 21.20 A moving field (initially presented in Figure 21.2, *first row*, *left*, as an immovable one) and the path of simulated survey corresponding to one of the following cases: (a) the transect spacing is fixed (the distance between transects D = 0.14); (b) the fixed sampling effort is allocated to the survey (the overall survey paths for both patterns S = 9.0). The patches move from right to left uniformly and rectilinearly; the survey is carried out in the opposite direction; the speed of the patch movement is 0.7 of that of the survey; the unit of



(b)

FIGURE 21.20 (CONTINUED) sampling distance d = 1/7. *Left column:* sequence of positions of the field and the surveys at the same moments of time: (a) the solid line indicates the parallel pattern; dashed line, the zigzag; (b) the solid line indicates the zigzag pattern; the dashed line, the parallel one. The rest of the figure, *top position:* the average original field. *Middle column:* the fields reconstructed from the results of the surveys by parallel or zigzag transects in the opposite (*second* and *third rows*) or same (*fourth* and *fifth rows*) directions. *Right column:* the same conditions but the unit of sampling distance is set along transects. Smaller arrows indicate the direction of the patch movement; larger ones, the directions of the surveys.

column, fourth and fifth rows). Similar results are obtained in the case where the unit of sampling distance is set along transects: (r^2 = 0.64 and r^2 = 0.55, right column, second and third rows; r^2 = 0.56 and r^2 = 0.52; right column, fourth and fifth rows).

In the case when the given sampling effort is allocated to a survey (Figure 21.20b), the comparison of the results shows that the parallel pattern leads to poorer results than the zigzag pattern does (r^2 = 0.67 and r^2 = 0.70, middle column, second, and third rows; r^2 = 0.61 and r^2 = 0.64; middle column, fourth, and fifth rows). Similar results are obtained in the case where the unit of sampling distance is set along transects: (r^2 = 0.64 and r^2 = 0.66, right column, second, and third rows; r^2 = 0.56 and r^2 = 0.58; right column, fourth, and fifth rows).

As shown by simulations, if the dimension of patches in the direction of movement exceeds that of a surveyed area, a survey in the opposite direction gives best results. In contrast, if the dimension of moving patches is smaller than that of a surveyed area, it is reasonable to carry out a survey in the same direction. The test of the criterion for choosing a survey direction confirms its applicability for the case when a unit of sampling distance is set along transects. In both cases, with lower speed of the patch movement, this regularity remains valid but becomes less distinct.

21.4 CONCLUSIONS

Survey design and the algorithm of data analysis for distribution fields consisting of gaps are exactly the same as they are for patchy distribution fields. A survey in the direction of patch elongation is optimal. A patchy field can be reconstructed properly ($r^2 > 0.70$) if $D/R < 1.5$ to 2.0 (regarding the parallel pattern) or $D/R < 1.0$ to 1.5 (in respect to the zigzag one) and $d/R_p < 1.0$ to 1.5. To some extent, it is possible to compensate the change in the adequacy of the reconstruction of a patchy distribution field by an increase of the distance between transects and a corresponding decrease of the sampling distance unit and vice versa. It is practically expedient to carry out a survey with a larger distance between transects and smaller unit of sampling distance. The absolute values of the sampling distance unit regarding the zigzag transects are, in general, larger than those given in application to parallel transects.

If the transect spacing is fixed, the parallel pattern allows (with great probability) more adequate reconstruction of an original distribution field (in cases of both immovable and movable fields) than the zigzag pattern does. In contrast, if a fixed sampling effort is allocated to a survey, the parallel pattern allows (with great probability) less adequate reconstruction of an original distribution field (in cases of both immovable and movable fields) than the zigzag pattern does.

If the dimension of patches in the direction of movement exceeds that of a surveyed area, a survey in the opposite direction gives best results. In contrast, if the dimension of moving patches is smaller than that of a surveyed area, it is reasonable to carry out a survey in the same direction.

ACKNOWLEDGMENTS

A major part of the material of the present chapter was initially published in References 18 to 21 and is reprinted with kind permission from Elsevier Science and Kluwer Academic Publishers. I am grateful to the Israeli Ministry of Absorption, the Israeli Ministry of Science and Technology, the Rich Foundation, the United States Agency for International Development, and the United States–Israel Binational Science Foundation, whose financial support made this study possible.

REFERENCES

1. Steele, J.H., Patchiness, in *The Ecology of the Seas*, Cushing, D.H. and Walsh, J.J., Eds., Blackwell Scientific Publications, London, 98–115, 1976.
2. Yudanov, K.I., *Interpretation of Echograms of Hydroacoustic Fish-Finding Instruments*, Translated from Russian, Israel Program for Scientific Translations, Jerusalem, 1971, 120 pp.
3. Gallager, S.M. et al., High-resolution observations of plankton spatial distributions correlated with hydrography in the Great South Channel, Georges Bank, *Deep-Sea Res*. II, 43(7,8): 1627–1663, 1996.
4. Greene, C.H., Wiebe, P.H., and Zamon, J.E., Acoustic visualization of patch dynamics in oceanic ecosystems, *Oceanography*, 7(1): 4–12, 1994.
5. Thompson, S.K, *Sampling*, John Wiley & Sons, 1992, 335 pp.
6. MacLennan, D.N. and Simmonds, E.J., *Fisheries Acoustics*, Chapman & Hall, London, 1992, 325 pp.
7. Yudanov, K.I., Kalikhman, I.L., and Tesler, W.D., *Manual of Acoustic Surveys*, Moscow, VNIRO (in Russian), 1984, 124 pp.
8. Kalikhman, I.L. et al., Choosing distance between acoustic survey tracks, in *Scientific Committee for the Conservation of Antarctic Marine Living Resources*, *Selected Scientific Papers*, V/BG/24, 1986, pp. 152–164.
9. Kalikhman, I.L., Correcting distance between acoustic survey tracks, in *Progress in Fisheries Acoustics*: *An Underwater Acoustic Group Conference Held at MAFF Fisheries Laboratory*, Lowestoft, England, March 21/22, 1989, Proceedings of the Institute of Acoustics, 11(3): 212–215, 1989.
10. Rose, G.A. and Leggett, W., The importance of scale to predator–prey spatial correlations: an example of Atlantic fishes, *Ecology*, 71(1): 33–43, 1990.
11. Kizner, Z.I., Tesler, W.D., and Zaripov, B.R., Methodological Recommendations on Mathematical Simulation of Commercial Concentrations and Treatment of Survey Data with Computers, Moscow, VNIRO (in Russian), 1983, 44 pp.
12. Zaripov, B.R., Kizner, Z.I., and Tesler, W.D., The analysis of methods used in the treatment of data obtained in echo surveys by means of a mathematical model of moving fish aggregations, *Fishery Acoustic Problems, Scientific Proceedings of VNIRO*, 25–36 (in Russian), 1983.
13. Williamson, N.T., Effect of serial correlation on precision of fish abundance estimates derived from quantitative echo sounder survey, *BIOMASS, Acoustic Krill Estimation Working Party News*, Vol. 2, 1980, 16 pp.
14. Croxton, F.E., Cowden, D.J., and Klein, S., *Applied General Statistics*, Sir Isaac Pitman & Sons, London, 1968, 754 pp.
15. Cressie, N.A.C., *Statistics for Spatial Data*, John Wiley & Sons, New York, 1991, 900 pp.

16. Gandin, L.S., *Objective Analysis of Meteorological Fields*, Translated from Russian, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.
17. Stull, R.B., *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1988, 666 pp.
18. Kalikhman, I. and Ostrovsky, I., Patchy distribution fields: survey design and adequacy of reconstruction, *ICES J. Mar. Sci*., 54: 809–818, 1997.
19. Kalikhman, I., Patchy distribution fields: sampling distance unit and reconstruction adequacy, *ICES J. Mar. Sci*., 58: 1184–1194, 2001.
20. Kalikhman, I., Patchy distribution fields: a zigzag survey design and reconstruction adequacy, *Environ. Monit. Assess*., 76: 275–289, 2002.
21. Kalikhman, I., Patchy distribution fields: sampling distance unit of a zigzag survey and reconstruction adequacy, *Environ. Monit. Assess.*, 80(1): 1–16, 2002.

22 Monitoring, Assessment, and Environmental Policy*

J.J. Messer

CONTENTS

22.1 THE USES OF MONITORING AND ASSESSMENT IN ENVIRONMENTAL POLICY-MAKING

Perhaps the earliest example of environmental monitoring, assessment, and policy-making was John Snow's careful collection of data (monitoring) on the incidence of cholera in 19th century London. Analysis of the data led him to identify a contaminated well on Broad Street as the most likely source of infection (assessment). The authorities subsequently closed the well by removing the pump handle (policy-making).[1] We have come a long way since then. The other chapters in this volume describe advances in monitoring technology and the implementation of major monitoring networks. Assessment has become increasingly complex and formalized. Policy-making has moved well beyond decisions by local authorities and is now national, multinational, and international in scope. In this overview, we examine

* This chapter has been approved for publication by the United States Environmental Protection Agency, but the opinions are those of the author, and do not necessarily reflect the official views or policies of the Agency. The author is grateful to C. Riordan, K. Thornton, and S.T. Rao for their helpful comments on a draft of this chapter.

1-56670-641-6/04/$0.00+$1.50

some examples of the roles that monitoring has played in assessment and environmental policy since Snow's earlier, simpler time.

Monitoring, in the context of this chapter, is the systematic collection of data for the purpose of checking on the environment, as opposed to collection of field data primarily to support a scientific study. *Assessment* is the process of analyzing and evaluating the resulting monitoring data, together with other scientific evidence, to support policy-making. Although science and assessment are inextricably linked, science is the discovery of knowledge through research, whereas assessment involves analyzing the quality of scientific understanding and bounding the uncertainties, so that decision-makers can act with an appropriate interpretation of the benefits, costs, and risks of alternative policies.[2] *Policy* is any course of action intended to guide decisions about whether and how to protect or restore the environment. Examples include treaties, legislation, executive orders, administrative rulemaking, execution of regulations, and even stimulation of private sector actions by government. Policy-making may occur within a given decision-making structure, or it may involve deliberations about the decision-making structure itself.[3]

Environmental policy-making can involve:

1. Identifying and analyzing environmental problems
2. Formulating policy and setting goals and priorities
3. Executing policy and managing programs
4. Evaluating policy and program performance

Monitoring and assessment can contribute to policy-making through one or more of these activities. This overview examines some of the significant national and international environmental policy issues of the past four decades to try to determine whether monitoring data either did or did not play a key role in each of the four policy-making activities. The international issues include acid precipitation, stratospheric ozone depletion, and global climate change; the national issues include criteria air pollutants and water quality management in the U.S. We conclude with a look at the use of monitoring and assessment to support environmental "report cards" as potential policy tools.

While it is tempting to try to make a case for the relative importance of monitoring vs. other sciences (e.g., modeling or toxicology) in any particular policy decision, the various science underpinnings are usually too inextricably linked to support any such conclusion. Consequently, when we conclude that monitoring played *a* key role, that is not to say that it played *the* key role. Even more important, in the words of the late Congressman George E. Brown, a long-time participant in environmental policy-making in the U.S. House of Representatives, "Political expediency will always play a greater role in policy-making than will analytical thinking, scientific or otherwise."[4] Congressman Brown's remarks remind us that the role that monitoring plays in policy-making may never be known precisely, and perhaps least of all by scientists seeking objective truth based on falsifiable data.

In order to keep this overview to a manageable length, we have limited its scope. The choice of examples is influenced by the author's experience with air and water policy in the U.S., and under-represents national examples from other countries and other environmental policy areas (e.g., natural resource management). Monitoring of pollutant emissions and of enforcement actions are not considered, even though they can and have played important roles in policy-making. Discussions of the examples rely heavily on overview papers which provide a wealth of information on other factors that influenced the corresponding policy decisions, and they should be consulted for primary literature sources not explicitly referenced in the text.

22.2 ACID PRECIPITATION

The early history of the acid precipitation problem is described by Likens[5] and Cowling,[6] and the more recent history by Clark et al.[7] and Sundqvist et al.[8] Acid precipitation was identified as a potential problem in 19th century England by Smith, and its harmful effects were explored in the 1950s by Gorham. However, identification of acid precipitation as a major issue by the public did not become widespread until newspapers in Sweden began extensive coverage of an analysis of monitoring data by Oden in the late 1960s. Monitoring data from a network put into place in western Europe in the early 1950s showed that acid precipitation had increased substantially in both intensity and geographic extent, and that the increases could be correlated with monitoring data on emissions from stationary sources, some of which were hundreds of miles distant. This public attention led the Swedish government to present a case study at the 1972 U.N. Stockholm Conference on the Human Environment, that led to the establishment of major government research programs in Europe and North America to more thoroughly analyze the problem.

In the U.S., monitoring at an experimental watershed at Hubbard Brook in New York and at a few stations run by individual investigators around the country demonstrated that acid precipitation was already a widespread phenomenon in eastern North America in 1976, and that changes in the ratios of sulfuric and nitric acids could be correlated with monitoring data from emissions sources. Monitoring of lakes and streams in the U.S. and Canada also showed large numbers of acidified surface waters and declines in certain fish species. The National Acid Deposition Network (NADP) was established in the late 1970s (see Chapter 27 this book) to collect data nationwide. In 1979 international agreements were in put in place between the U.S., Canada, and the European Community to analyze these problems and to seek potential solutions, and in 1980 Congress established what was to become the decade-long, $530 million National Acid Precipitation Program (NAPAP). NAPAP culminated in an Integrated Assessment in 1991 that analyzed the impact of alternative emissions reduction targets on the future acidity of surface waters in the eastern U.S., and the costs of the various alternatives.[6] The results of the NAPAP integrated assessment, which relied heavily on data on emissions, patterns of the acidity in precipitation from the NADP network, and statistically reliable patterns of the chemistry of lakes, streams, and soils established by systematic data collection by the National Surface Water Survey, showed that a 50% sulfate emissions reduction would significantly reduce the acidity of lakes and streams in sensitive parts of the U.S. and Canada.[9] Although there was concern about the potential effects of acid rain on forests in Europe and the U.S., many of the data were based on experimental studies, and even when surveys demonstrated higher than expected mortality of spruce-fir and sugar maple, there were enough confounding factors that the NAPAP Integrated Assessment did not conclude that there was clear monitoring evidence linking the declines to acid deposition.

It is certainly clear that analysis of monitoring data on both precipitation and surface waters played a key role in identifying the acid precipitation problem and bringing it to international attention. There has been much debate, however, about the role that the analysis of data in NAPAP (and NADP) ultimately played in formulating policy and setting goals and priorities in the acid rain provisions of the Clean Air Act Amendments (CAAA) of 1990.[10] The ultimate reduction target of 50% of sulfate emissions supported by the NAPAP assessment was exactly what was envisioned in a 1981 National Academy of Sciences Report that had set the stage for the acid rain policy debate in the U.S.[11]

Environmental monitoring appears to have played a modest role in executing and managing acid precipitation programs in the U.S. and Europe, even though the approaches are very different. In the U.S., the CAAA adopted a cap-and-trade approach that set a cap on sulfur emissions from certain categories of sources nationwide, and allowed sources to trade emissions credits to insure that the cap would be attained by reducing emissions at the facilities where the reductions were the most economical, irrespective of the effects on downwind ecosystems. The Act placed a premium on accurate emissions monitoring to manage the program, rather than environmental monitoring (unlike the provisions of the Act to manage criteria pollutants, which is discussed later in this overview). In Europe, the 1994 Sulfur Dioxide Protocol to the 1979 Convention on Long-Range Transport of Air Pollutants (LRTAP) employed the concept of “critical loads” to manage emissions. A critical load is a quantitative estimate of the atmospheric load of the pollutant below which significant harm is not expected to sensitive elements of the environment according to present knowledge. Instead of reducing emissions irrespective of downwind harm, critical loads of sulfate are established for downwind ecosystems, and a model is used to determine the required emission reductions upwind. A review of the most recent manual for determining critical loads reveals that the approaches rely more heavily on modeling and the results of field and laboratory experiments than on monitoring data.[12]

Acid precipitation (and some of its effects) continue to be monitored systematically in North America and Europe (see Chapter 27 this book), and monitoring certainly has been widely used in evaluating policy and program performance. Whether such evaluation has resulted in mid-course corrections in the control of acid precipitation is harder to assess. The first phase of controls on sulfur and nitrogen emissions under the CAA in the U.S. was not completed until 1995. At that time, the U.S. Environmental Protection Agency (EPA) prepared a report required by Congress to determine whether the CAA adequately protect sensitive areas of the U.S.[13] Limited monitoring data suggested that the capacity of forests to absorb most of the reactive nitrogen in precipitation (and thus to prevent acidification of soils and runoff) in the U.S. and Europe was being exceeded. Monitoring data from the acid-sensitive lakes of the Adirondack region, however, showed no discernable pattern over 12 years. Biogeochemical models suggested that further reductions in nitrogen emissions would be required, but due to uncertainties in the models, the

EPA concluded that no new recommendations were warranted at the time. The science advisory board that reviewed the report concluded that without direct monitoring evidence, it is unlikely that incremental regulations would be based on modeling evidence alone.

Do the monitoring data suggest additional reductions are needed? Phase 2 of the sulfate emissions reductions in the U.S. began in 1996. Monitoring of precipitation in the U.S. and Canada has since shown widespread reductions in sulfate since 1995, and monitoring of lakes in the U.S. has revealed recovery of one third of acid lakes to nonacidic status, but there have been no significant trends in nitrate in precipitation.[14–15] In 1998, the EPA enacted new controls on nitrogen oxide emissions that will substantially reduce nitrate in precipitation, but the primary purpose was to control regional ozone. In 2002, "Clear Skies" legislation was proposed that would further reduce acid precipitation but, again, as a beneficial side effect of reductions in ozone and particulate matter, which are thought to have substantial health effects.[16] Monitoring therefore has demonstrated that emissions reductions have had beneficial effects, but apparently has not led directly to modifications in controls.

22.3 STRATOSPHERIC OZONE DEPLETION

The scientific and policy history of the stratospheric ozone depletion problem has been described by Albritton,[17] Morrisette,[18] Abbat and Molina,[19] Rowlands,[20] and Bendick.[21] In 1974, Molina and Rowland hypothesized that chlorofluorcarbons (CFCs) could catalyze the reduction of ozone in the stratosphere, leading to an increase in ultraviolet B (UV-B) radiation at the Earth's surface. During the early 1970s, CFCs used as aerosol propellants constituted over 50% of total CFC consumption in the U.S. Widespread press coverage raised public concerns that the increase in UV-B radiation would lead to an increase in skin cancers. A consumer boycott followed in the U.S., and in 1977 the United Nations Environmental Program developed a nonbinding international effort to conduct research and monitoring on the ozone depletion problem. Nonessential uses of CFCs in aerosol containers were banned in Canada, the U.S., and a few European countries in 1978. This particular use of CFCs was reduced in the U.S. by approximately 95%, cutting total U.S. consumption of CFCs by nearly half. In the years following the aerosol ban, CFC use increased significantly in the refrigeration, foam, and solvent-using electronics industries and by 1985, CFC use in the U.S. had surpassed pre-1974 levels and represented 29% of global CFC usage.

Under the Clean Air Act of 1977, the EPA published a notice of advanced rulemaking in 1980 to further restrict the manufacture and use of aerosols, but did not act on it.[22] On the international front, the 1985 Vienna Convention for the Protection of the Ozone Layer continued to pursue a nonbinding approach to international controls on CFCs, and plans were made for the next international meeting in Montreal 2 years later. Then, in late 1985, an ozone "hole" was reported over Antarctica. Based on monitoring data from a ground network of Dobson monitors and confirmed by satellite measurements, the hole was actually found to be a 35% depletion in stratospheric ozone over Antarctica that had been occurring since 1957, showing a trend of accelerating loss since the mid-1970s. Two years later, in 1987, the Montreal Protocol was signed, committing 23 nations to a reduction of 50% of the most damaging Class I products (four CFCs and three halons, which contain bromine instead of chlorine) by 1998.

How big a role did monitoring data play in the 1978 ban and the Montreal Protocol? Daniel Albritton, who co-chaired the UNEP's scientific assessments of stratospheric ozone and was a key player at the meeting, recalled that as the Montreal Protocol was being negotiated, the ozone hole had not been explicitly linked to CFCs; it was not predicted by the atmospheric models being used at the time, and therefore was not explicitly considered.[17] With respect to monitoring data from a ground-based network of Dobson spectrophotometers begun during the International Geophysical Year in 1958 and from a solar backscatter ultraviolet (SBUV) instrument launched into orbit in 1978, he wrote that, "the observations, although they suggested a decrease whose rough magnitude was similar to that predicted, were not considered entirely believable. Theory, on the other hand, could justify some strong predictions." The theoretical models, however, predicted that without substantial reductions of CFC uses, there was a substantial risk of UV-B increases at high latitudes, and also of global warming. Richard Benedick, the chief U.S. negotiator on the Montreal Protocol, later said that the most extraordinary aspect of the treaty was that it imposed substantial costs against unproven dangers, "that rested on scientific theories rather than on firm data."[21] That situation was about to change.

Shortly after the Montreal Protocol was signed in 1987, an international team of scientists linked the ozone hole over Antarctica to CFCs. They later presented monitoring data from the Total Ozone Mapping Spectrometer (a satellite instrument) which showed that between 1979 and 1987 stratospheric ozone had decreased worldwide by 0.4% per year, a larger decline than anticipated at the time the Protocol was signed. Considering these findings, the London Convention in 1990 phased out production and importation of all CFCs by 2000, and the U.S. EPA proposed a conforming regulation in 1991.[22] Subsequent monitoring revealed that in January of 1992, stratospheric ozone had dipped as much as 20% in the Northern Hemisphere and as much as 45% for a few days over Russia. Later that year, the Copenhagen Convention in 1992 moved up the date for phase-out of the Class I compounds to January 1, 1996, and added two other compounds to the list. By 1993, 107 nations had become parties to the protocol.

Based on the remarks of Albritton and Benedick, key players at the Montreal meeting, we must conclude that monitoring data took a back seat to theory in identifying the problem of stratospheric ozone depletion. It is fairly clear, however, that the monitoring data became more important in focusing attention on the immediacy of the problem as targets and deadlines for phase-out of the Class I substances were tightened in the 1990 and 1992 amendments. Courtney Riordan, EPA's research director in the area, recalls Robert Watson of the White House Office of Science and Technology Policy briefing administration executives and Congress with a dramatic visualization of the ozone hole.[23] Also important in analyzing the problem, but not receiving as much attention in the overviews, were monitoring data from a worldwide network begun in 1978 that showed increasing levels of CFCs and other halocarbons in the background atmosphere over the period 1978 to 1992.[24] Morrisette argues that the ozone hole was a tangible, measurable impact that galvanized public opinion, and thus influenced the outcome in Montreal,[18] and coverage of the issue in the international popular press peaked in the late 1980s and early 1990s,[7] thus likely increasing public pressure to further tighten the deadlines in London and Copenhagen.

We can safely say that monitoring played a key role not only in analyzing the stratospheric ozone deletion problem, but in formulating national and international CFC policy and setting priorities, especially with respect to extent and rate of phase-outs. The compounds to be included in the phase-outs depend primarily on their Ozone Depletion Potentials (ODPs), which determine whether they are included in Class I or Class II (the latter have much lower ODPs and are being used as substitutes for Class I compounds until they, too, are phased out). Determination of ODPs is primarily based on theory and laboratory experiments, but an alternative empirical approach does utilize field observations and monitoring data.[25] The rate and extent of phase-outs appear from the overview papers to be primarily a function of the perceived need to do as much as possible, and as quickly as possible; the technological and economic feasibility, and the modeling results that forecast that the planned targets will significantly reduce the concentrations of ozone-depleting substances in the stratosphere.

The role of monitoring in executing policy and programs with respect to CFCs is apparently modest. Compliance with the Protocol and U.S. regulations relies on manufacturing, recycling, destruction, and import data, rather than on achieving an ambient standard. Monitoring plays a key role, however, in evaluating CFC policy and program performance. Monitoring of stratospheric ozone, background levels of CFCs and other halocarbons in the atmosphere, and more recently UV-B levels at the Earth's surface are analyzed and reassessed by international teams of scientists on a regular basis. The most recent report showed that ozone levels over the Antarctic have continued to decline and now are approximately 50% of the 1957 levels, and that levels continue to decline, but at a smaller rate, over many cities in midlatitudes.[26] It also showed that ozone-depleting compounds in the stratosphere peaked in 1994, and declined by about 5% through mid-2000 (expressed as the equivalent effective stratospheric chlorine), and that based on monitoring of hydrogen chloride and chlorine nitrate total column absorbance measurements, chlorine stopped increasing between 1997 and 1998 and has remained fairly constant since. The report concludes that monitored concentrations of ozone-depleting gases in the atmosphere are in line with expectations from the fully modified and adjusted Montreal Protocol, and that accelerating the rate or extent of phase-outs at this time would provide only modest improvements in the rate of ozone recovery to pre-1980 levels in the stratosphere. Therefore, no significant modifications of policy appear to be indicated by the monitoring data.

22.4 GLOBAL CLIMATE CHANGE

The scientific and policy history of the global climate issue has been described by many authors, but this discussion relies heavily on Hecht and Tirpak,[27] Keeling,[28] Pielke[29], Leaf,[30] and Clark et al.[7] The ideas that burning coal could increase carbon dioxide (CO_2) in the atmosphere enough to raise the global temperature significantly was first raised in 1896, but it was not until 1957 that Revelle and Seuss published a paper that drew scientists' attention to the "large-scale geophysical experiment" that humans were conducting with fossil fuel combustion. The global climate change debate since then has been driven primarily by modeling (again, accompanied by great scientific debate) changes in climate as a result of changing concentrations of carbon dioxide (CO_2), and other "greenhouse" gases, and by the effects of such changes and the economic costs of decreasing greenhouse gas emissions or mitigating the effects.

In the midst of the scientific debate about theory, however, monitoring played an important role in demonstrating that concentrations of greenhouse gases and global temperatures have continued to increase over the course of the scientific and policy debate, thus keeping the pressure on scientists and policy-makers alike. In the early 1950s, the scientific literature suggested that CO_2 concentrations were highly variable with latitude and time of day, which would make the determination of global trends very difficult. A young scientist named Charles Keeling made measurements of CO_2 on samples collected several times a day at a number of locations from Canada to South America and found, on the contrary, that daytime concentrations were all close to 310 ppm. His interest led to the beginning of a long-term data series of CO_2 measurements at the peak of Mauna Loa in Hawaii, the Antarctic, and two other locations, as part of the International Geophysical Year (IGY) in 1957. Keeling published his first results in 1960, showing a seasonal cycle in the northern hemisphere associated with plant growth that diminishes toward the equator, and a possible global year-to-year increase. At the time, the year-by-year rise appeared consistent with the amount of CO_2 from industrial activity. This later turned out to overestimate the importance of industrial emissions because the measurements were taken during an El Niño event, which tends to increase the levels in the atmosphere.

In 1969, Keeling reported that the CO_2 data from Mauna Loa now showed a definite upward trend of approximately 6 ppm over 10 years, with a strong seasonal cycle, and speculated that these data empirically supported Revelle's concerns about the potential for global warming. He published a subsequent paper in 1972 that showed that the trend line was increasing in slope (shown later to be the result of the periodic ENSO) and was consistent with some very simple compartmental models that showed that approximately half of the CO_2 was being absorbed into the ocean. Several international meetings were held on the global warming issue in the 1970s, culminating in the first World Climate Conference organized by the World Meteorological Organization (WMO) in Geneva in 1979. Although the analysis was published later, the late 1970s included a period of extremely unusual winters,[31] and there was concern at the time that the Earth may actually be heading into a period of glaciation rather than warming.

Research expenditures increased, and in 1985 the U.S. Department of Energy published a report concluding that some effects of global warming may already have been evident; CO_2 concentrations in the atmosphere (including Keeling's Mauna Loa data) continued to increase, as did Northern hemisphere land temperatures, sea surface temperatures, and sea level.[32] The report concluded that if emissions of greenhouse gases continued as expected, monitoring would either confirm the results predicted by the models or show that they "require extensive reconsideration." The United Nations Environment Programme (UNEP) published a conference report in 1986 that noted that monitoring showed that other important greenhouse gases (methane, CFCs and tropospheric ozone) were also increasing globally, further accelerating the risk of climate change. UNEP pressured the U.S. to support a convention to control greenhouse gases and to contribute to an international assessment effort, and the Climate Protection Act of 1987 was signed into law, requiring the U.S. EPA and DOE to develop policy options for dealing with the problem. In 1988, NASA scientist James Hansen testified to a Senate committee that he was 99% certain that global warming was underway. Hecht and Tirpack noted that monitoring data showed that at the time of Hansen's testimony, the world was suffering through one of the warmest years thus far on record, and that CO_2 concentrations had risen to 350 ppm.[27]

The International Governmental Panel on Climate Change (IPCC), jointly created by WMO and UNEP to bring together a global network of over 2,000 scientists to inventory current scientific knowledge of the climate system, the effects of climate change, and possible response strategies, published its first report in 1990. Two years later, the Framework Convention on Climate Change was signed, committing the signatories to emissions reductions, but without required targets or timetables. The next IPCC scientific assessment of 1995 led with monitoring data that showed that greenhouse gas concentrations, mean surface temperatures, and sea levels had continued to rise.[33] Two years after that report, in 1997, more than 180 countries ratified the Kyoto Protocol, which did contain targets and timetables. The targets were not driven by attempts to achieve particular greenhouse gas levels but by what were economically feasible for the signatories, and compliance has not been sufficient to expect a significant change in either CO_2 concentrations or temperatures. We thus conclude that monitoring has not played a key role in global climate policy-making beyond its key role in identification and analysis of the problem that forced international action.

22.5 CRITERIA AIR POLLUTANTS IN THE U.S.

Title I of the Clean Air Act Amendments of 1990 provide one of the few examples we have of the importance of monitoring in executing policy and managing programs. Criteria pollutants are explicitly identified in the CAAA, and include lead, ozone, carbon monoxide, sulfur dioxide, nitrogen dioxide (NOX), and particulate matter (PM).[34] Primary National Ambient Air Quality Standards (NAAQS) are required to be set at an ambient level determined by EPA to be protective of human health regardless of costs or current ambient levels. Each state develops a State Implementation Plan (SIP) that allows it to determine how to best achieve the NAAQS by controlling the various sources within the state. States are required to monitor the concentrations of these pollutants according to strict protocols at more than 18,000 sites in the NAMS/SLAMS network (see Chapter 27 this book). If the concentrations of any of these pollutants exceed their respective standards in a particular area, then the area is determined to be in nonattainment, and the SIP must be adjusted to achieve attainment in a reasonable period of time.

An area in nonattainment may even lose highway construction grants from the federal government. Thus, there is a direct feedback between monitoring and controls on pollutant sources.

Although monitoring does not play a direct role in establishing targets for the NAAQS themselves, it can play a role in setting emissions reductions targets to achieve the NAAQS, as well as setting priorities for which NAAQS get the greatest attention. NAAQS are required under the CAAA to be reviewed every 5 years, but the EPA frequently falls behind schedule. In the early 1990s, a series of statistical analysis or mortality data and PM data from the NAMS/SLAMS network indicated that nonattainment of the current NAAQS for PM_{10} (the fraction of particulate matter smaller than 10 μm) may account for more than 60,000 deaths in the U.S. annually.[35] These results received considerable press coverage, and a public interest group sued the EPA to promptly review the NAAQS. At the time, the available toxicological data did not seem to explain or support this finding, but eventually the EPA did revise the PM NAAQS, including the addition of a new NAAQS for an even smaller particulate fraction ($PM_{2.5}$). Again, that was largely identified though epidemiological studies.[36]

The criteria pollutants also offer examples of the role of monitoring in evaluating policy and program performance. Each year, the EPA publishes a report on the status and trends in air quality. The most recent report concluded that, despite progress, approximately 133 million people live in counties where monitored air in 2001 exceeded at least one of the NAAQS, usually because of ozone and particulate matter.[37] Consequently, the EPA was proposing rules to reduce emissions from certain road, nonroad mobile, and stationary combustion sources. Moreover, the EPA also had submitted to Congress Clear Skies legislation that, if enacted, would mandate reductions of particle- and ozone-forming compounds from power generators by 70% from current levels through a nationwide cap and trade program. The EPA was not only more aggressively pursuing rulemakings under the existing statute, but seeking new statutory authority based on monitoring data that showed the current rules were not achieving the NAAQS. Furthermore, emissions monitoring data for both sulfur dioxide and NOX showed that cap-and-trade approaches were not leading to large regional shifts in emissions, which strengthened the EPA's commitment to a cap-and-trade approach for the three pollutants covered by the proposed new legislation.[38]

Although monitoring does not lead directly to the establishment of NAAQS, two examples show how it can nonetheless be important in other air quality regulations. The case for lead (a criteria pollutant) is nicely discussed in a volume edited by Ratcliff.[39] Careful monitoring of atmospheric background lead in the atmosphere at Mauna Loa in the 1970s showed that most atmospheric lead must come from tetraethyl lead in gasoline, as opposed to some global geochemical source. Monitoring of human blood lead levels also showed a close correlation with ambient levels of lead in the atmosphere in the U.S. over the 1970s. These observations likely played a key role in the decision to ban lead in gasoline in the U.S. in 1986. Likewise, a decision by the EPA in 2000 to regulate mercury emissions from power plants[40] relies on a risk assessment that hinges on an extensive program monitoring mercury in fish tissue in the 1970s and 1980s.[41]

22.6 WATER QUALITY IN THE U.S.

In an influential paper in 1971, Wolman concluded that increasing pressures on U.S. rivers may have been outstripping then-current investments in water quality management, but that water quality monitoring was inadequate to determine if things were getting better or worse, or to determine the level and types of expenditures needed.[42] The next year, the Federal Water Pollution Control Act Amendments of 1972 were drafted so that water quality monitoring would drive the nation's entire approach to water quality management, including assessment, standard setting, planning, discharge permitting, construction funding, and accountability. Savage provides an excellent analysis.[43] Under the Act, each state would survey its waters and assign designated uses (e.g., fishing, body contact, water supply) to each one. Water quality standards to achieve each designated use would be set and reevaluated periodically. Point-source discharges (e.g., wastewater treatment plants) would receive permits that would ensure that water quality standards were not exceeded in the receiving water body at the point of discharge. If water quality monitoring revealed that a water body was still not achieving standards, the state then must determine the total maximum daily load (TMDL) that would be allowable for each pollutant that was causing the violation of the standard, and develop a watershed management plan to insure that the TMDLs were not exceeded. The plan could include management of nonpoint sources pollution that were otherwise not required to have a discharge permit. Every 2 years, the states would report to the EPA on the extent and causes of nonattainment of water quality standards, and federal funding for construction of wastewater treatment plants and watershed management would be tied to the extent of waters not achieving standards in each state. Monitoring also would be used to reassess both the designated uses for each water body, and the adequacy of the standards to protect each designated use. The system thus would address all four policy purposes for monitoring.

It is likely that all of the states use their water quality monitoring programs to identify and analyze problems, to formulate policy and set goals and priorities, to execute and manage their water quality control programs, and to evaluate their performance, but these decisions are not conveniently documented in the literature. Adequate funding to support the level of monitoring required to make the system work as planned was never made available, and the system still suffers from the consequences at the national policy level. In 2000, the U.S. General Accounting Office reported that the States had "little of the information needed to assess the quality of their waters and, in particular, to those that are impaired — a particularly serious problem, given resources needed to address such impairments."[44] Although this statement recalls that of Wolman three decades earlier, the EPA and the states have made progress in the last few years in improving monitoring designs (see Chapter 22 this book) and in providing additional funding to execute them.[43]

In addition to the water quality monitoring by the states, the U.S. government also conducts or sponsors its own water quality monitoring programs. We have already seen now monitoring of lakes sponsored by EPA has revealed that considerable recovery from acidic status has occurred as a result of SO_2 emissions reductions, but that additional controls will be needed to further reduce the number of acidic lakes.[37] The monitoring of mercury in fish noted previously was conducted by the U.S. Fish and Wildlife Service.[45] The U.S. Geological Survey NASQAN program (see Chapter 27 this book) has been particularly instrumental in tracking long term patterns and trends in the export of certain chemicals from watersheds and, for example, has identified tremendous increases in the export of nitrogen from the Mississippi River watershed over the past four decades, which is contributing to a very large "dead zone" of anoxic water in the Gulf of Mexico.[46] An Action Plan describing a national strategy to reduce the frequency, duration, size, and degree of oxygen depletion of the hypoxic zone of the northern Gulf of Mexico was submitted as a Report to Congress on January 18, 2001, but there are as yet no enforceable targets or deadlines.[47] Again, most such monitoring data sets seem to play the largest policy role in identifying and analyzing problems.

22.7 ENVIRONMENTAL "REPORT CARDS"

If monitoring and assessment have contributed significantly to policy decisions, why is the public not more aware of the importance of monitoring data? There has been growing international attention to providing "report cards" to the public on trends in the condition of the environment,[48] but the idea is certainly not a new one. The National Environmental Policy Act of 1969 requires the President to report annually to Congress on the state and condition of the environment, on current and foreseeable trends, on the adequacy of available resources, on the progress of programs aimed at protecting the environment, and on a program for remedying any deficiencies in these programs, including recommendations for any new legislation.[49] Being the foremost item on the list, it would therefore seem that Congress intended that environmental monitoring data would play a key role in guiding national environmental policy. A review of the annual reports, however, reveals substantially more data on polluting activities (e.g., water withdrawals, pollutant emissions, vehicle miles traveled) and administrative programs (e.g., permits written, expenditures on control or clean-up), than on trends in the condition of the environment itself. The 21st Annual Report of 1990, for example, shows little environmental monitoring data other than those described in the examples in this overview.[50] In the same year, the Administrator of the EPA wrote that he thought that the EPA did an exemplary job of protecting public health and the quality of the environment, but concluded, "Now the bad news. I cannot prove it."[51]

Only slight progress has been made since 1990. As of 2003, the European Environmental Agency (EEA) maintains data on 92 environmental indicators, of which only 16 involved environmental monitoring,[52] an eclectic mix that includes human exposure to traffic noise and ozone, pollutants in rivers, and fragmentation in grasslands, and fisheries stocks. On the global scale, the UNEP developed its Global Environment Outlook (GEO) project in response to the environmental reporting requirements of Agenda 21 and to a UNEP Governing Council decision in May 1995 requesting the production of a comprehensive global state of the environment report. In order to prepare the second report in 2000, the GEO Data Working Group identified approximately 90 variables associated with data sets from 202 countries.[53] Analysis by the group identified so many inconsistencies and other quality problems with these data that in the final report, the 90 variables were reduced to 15, none of which actually represented measures of the state of the environment. The most recent GEO report, Global Environmental Outlook 3, includes a Web-based data portal with access to over 400 environmental data sets, but only lists a few state variables, including the extent of degradation of agricultural land, extent of deforestation, fish catch, and the number of endangered species of invertebrates.[54]

In the U.S., the Government Performance and Results Act of 1993 (GPRA) now requires all federal agencies to develop annual performance plans, to establish performance goals, and to express those goals in an objective, quantifiable, and measurable form.[55] The act has provided considerable impetus for the EPA and other agencies with environmental missions to identify indicators of program performance to correspond to their GPRA goals. Some of the indicators are associated with monitoring data on the actual state of the environment, but the numbers so far have been modest. The Heinz Center, working with a group of public and private sector partners, developed a report entitled The State of the Nation's Ecosystems.[56] The report includes 103 indicators, of which only 33 were judged by the authors to have adequate data for national reporting. Even more recently, The U.S. Environmental Protection Agency Draft Report on the Environment included 112 indicators of the state of the environment (exclusive of public health indicators), of which only 21 were deemed adequate for national reporting, based on the availability and representativeness of national environmental monitoring data.[57] The majority of the 21 indicators relied on data from the sources identified in the examples in this overview, but a number of new ones involved land cover and land use indicators derived from globally consistent satellite data. This situation is expected to improve in the next report. The Forest Health Monitoring and Forest Inventory and Analysis programs of the U.S.D.A. Forest Service are on the verge of providing nationwide data on ecological condition of forests across the U.S. (see Chapter 30 this book), and the EPA Environmental Monitoring and Assessment Program is doing the same for estuaries nationwide (see Chapter 29 this book) and conducting regional pilot projects on streams and rivers.

It is not clear exactly how or whether environmental report cards like those discussed above (and others published by nongovernmental environmental organizations) have affected policy-making to the extent seen for acid rain, stratospheric ozone depletion, or global climate change. A primary intention of all of these projects is to provide information about environmental trends to the attentive public (including lawmakers), who will in turn put pressure on environmental agencies to take any necessary corrective actions. Of course, it is seldom quite that simple. For example, Healy and Ascher explain how an increase in monitoring mandated by Congress to support public participation in forest management decisions resulted in assessment becoming such a complex task that the result was to "shift power away from nonexpert actors, undermine rights arguments, polarize debates over appropriate resource use, and delay timely decision-making."[58] Congressman Brown also questioned whether the truly objective data needed to make policy even exist, in the sense that experts from opposing sides in environmental debates always seem to claim that the data support their position. He concluded that the most promising roles for environmental data were in identifying problems and evaluating outcomes in order to provide mid-course corrections in policy.[4] On balance, it still seems better to have data than not, but unless and until more monitoring data become available, all these points will remain moot.

22.8 SUMMARY AND CONCLUSIONS

We have seen in the examples in this chapter that monitoring and assessment can play a key role in any or all of the four areas of policy-making identified at the beginning of this overview. The most frequent role in the examples was identifying and analyzing environmental problems, followed by evaluating policy and program performance. Key roles in formulating policy and setting goals and priorities and executing policy and managing programs were less frequent, but the examples for criteria air pollutants and water quality management in the U.S. show how important these roles can be for integrated policy approaches that are built around a central core of monitoring and assessment. These results are probably reasonably representative of air and water-related policy-making based on ambient environmental monitoring. If monitoring of pollutant emissions and pollution control actions had been included, the importance of monitoring in setting goals and executing and managing programs would increase substantially. Examples from other countries or other environmental policy areas (e.g., natural resource management) may also have revealed different patterns. As we look to the future, growth of interest in environmental report cards, accountability legislation like the Government Performance and Results Act of 1993, and shifts between receptor-oriented regulatory strategies such as "critical loads" and emissions-oriented strategies such as "cap-and-trade" also could shift the future balance of the roles monitoring and assessment play in environmental policy-making.

In any case, this overview should make clear why monitoring and assessment can and should be so important to environmental policy-making. Recognition of this importance is the key to both designing and maintaining regional and global monitoring networks. Too often, monitoring is seen as an expensive and less-worthy drain on funding that could otherwise be spent on research or pollution control. The examples in this overview should demonstrate otherwise: that monitoring, designed with a view toward explicitly supporting one or more of the four types of policy decisions, can lead to better and more efficient environmental policies and programs. This fact makes the contributions in this book on environmental monitoring all the more important.

REFERENCES

1. Summers, J., *Soho — A History of London's Most Colourful Neighborhood*, Bloomsbury, London, 1989, pp. 113–117.
2. Cowling, E.B., The performance and legacy of NAPAP, *Ecol. Appl.*, 2, 111, 1992.
3. Laswell, H.D., *A Pre-view of Policy Sciences*, Elsevier, New York, 1972.
4. Brown, G.E., Science's real role in policy-making, *C&EN*, 101, May 31, 1993.
5. Likens, G.E., Acid precipitation, *C&EN*, 84, Nov 22, 1976.

6. Cowling, E.B., Acid precipitation in historical perspective, *Environ. Sci. Technol.*, 18, 110, 1982.
7. Clark, W. et al., Acid rain, ozone depletion, and climate change: An historical overview, in *Learning to Manage Global Environmental Risks*, *Vol. 1: A Comparative History of Social Responses to Climate Change*, *Ozone Depletion and Acid Rain*, Clark, W. et al., Eds., MIT Press, Boston, MA, 2001, chap. 2.
8. Sundkvist, G., Letell, M., and Lidskog, R., Science and policy in air pollutant abatement strategies, *Environ. Sci. Policy*, 5, 147, 2002.
9. NAPAP, *1990 Integrated Assessment Report*, U.S. National Acid Precipitation Assessment Program, Washington, D.C., 1991.
10. Winstanley, D. et al., Acid rain and policy making, *Environ. Sci. Policy*, 1, 51, 1998.
11. National Research Council, *Acid Deposition — Atmospheric Processes in Eastern North America*, National Academy Press, Washington, D.C., 1983.
12. *Manual on Methodologies and Criteria for Mapping Critical Levels/Loads and Geographical Areas where they are Exceeded,* International Cooperative Programme on Modelling and Mapping of Critical Loads and Levels and their Air Pollution Effects, Risks and Trends, http://www.oekodata.com/icpmapping/html/manual.html, 1996.
13. Renner, R., "Scientific uncertainty" scuttles new acid rain standard, *Environ. Sci. Technol*., 29, 464A, 1995.
14. International Joint Commission, *Air Quality Agreement Progress Report 2002*, Washington, D.C., and Ontario, Canada, 2002.
15. U.S. EPA, Response Of Surface Water Chemistry To The Clean Air Act Amendments Of 1990, EPA/620/R-02/004, Office of Research and Development, Research Triangle Park, NC, 2002.
16. U.S. EPA, The Clear Skies Initiative, Section B; Human Health and Environmental Benefits, Office of Air and Radiation, www.epa.gov/clearskies, 2002.
17. Albritton, D., Stratospheric ozone depletion: Global processes, in *Ozone Depletion*, *Greenhouse Gases*, *and Climate Change*, National Academy of Sciences, Washington, D.C., 1989, chap. 3.
18. Morrisette, P.M., The evolution of policy responses to stratospheric ozone depletion, *Nat. Resour. J.*, 29, 793, 1989.
19. Abbatt, J.P.D., and Molina, M.J. Status of stratospheric ozone depletion, *Annu. Rev. Energy Environ*., 18, 1, 1993.
20. Rowlands, I.H., The fourth meeting of the parties to the Montreal Protocol: Report and reflection, *Environment*, 35, 25, 1993.
21. Benedick, R., *Ozone Diplomacy: New Directions In Safeguarding The Planet*, Harvard University Press, Cambridge, MA, 1991.
22. Protection of Stratospheric Ozone, *Fed. Reg.*, 56, 4396, 1991.
23. Personal communication, C. Riordan, Formerly Director of Office of Ecological Processes and Effects, EPA, ORD, April 25, 2003.
24. Prinn, R., A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE, *J. Geophys. Res.*, 105, 17,751, 2000.
25. Wuebbles, D., Weighting functions for ozone depletion and greenhouse gas effects on climate, *Annu. Rev. Energy Environ*., 20, 45, 1995.
26. Scientific Panel of the Montreal Protocol on Substances that Deplete the Ozone Layer, *Scientific Assessment of Ozone Depletion*, United Nations Environment Programme, Nairobi, August, 2002.
27. Hecht, A. and Tirpak, D., Framework agreement on climate change: a scientific and policy history, *Climatic Change*, 29, 371, 1995.

28. Keeling, C.D., Rewards and penalties of monitoring the Earth, *Annu. Rev. Energy Environ.*, 23, 25, 1998.
29. Pielke, R.A., Jr., Policy history of the U.S. Global Change Research Program: Part I. Administrative development, *Glob. Environ. Chang.*, 19, 9, 2000.
30. Leaf, D., Managing global atmospheric change, *Hum. Ecol. Risk. Assess.*, 7, 1211, 2001.
31. Kerr, R., Wild string of winters confirmed, *Science*, 227, 506, 1985.
32. Kerr, R., Greenhouse warming still coming, *Science*, 232, 573, 1986.
33. Intergovernmental Panel on Climate Change, *Climate Change 1995*, World Meteorological Organization and United Nations Environmental Programme, Nairobi, 1995.
34. U.S. EPA, *Plain English Guide to the Clean Air Act*, EPA-400-K-93-001, Office of Air and Radiation, Washington, D.C., 1993.
35. Kaiser, J., Showdown over clean air science, *Science*, 277, 466, 1997.
36. Kaiser, J., Evidence mounts that tiny particles can kill, *Science*, 289, 22, 2000.
37. U.S. EPA, *Latest Findings on National Air Quality: 2001 Status And Trends*, 454-F-00-002, Office of Air and Radiation, Washington, D.C., 2002.
38. Personal communication, R. Birnbaum, U.S. EPA Office of Atmospheric Programs, Washington, D.C., March 31, 2002.
39. Ratcliffe, J., Ed., *Lead in Man and the Environment*, Ellis Horwood, Chichester, England, 1981.
40. Regulatory Finding on the Emissions of Hazardous Air Pollutants from Electric Utility Steam Generating Units, *Fed. Reg.*, 65, 79825, 2000.
41. U.S. EPA, *Mercury Study Report to Congress. Volume VI: An Ecological Assessment for Anthropogenic Mercury Emissions in the United States*, EPA-452/R-97-008, Office of Air Quality Planning and Standards and Office of Research and Development, Washington, D.C., 1997.
42. Wolman, G., The nation's rivers, *Science*, 174, 905, 1971.
43. Savage, R., Sample problem, *Environ. Forum*, September/October 2002.
44. *Water Quality: Key EPA and State Decisions Limited by Inconsistent and Incomplete Data*, GAO/RCED-00-54, U.S. General Accounting Office, Washington, D.C., 2000.
45. Lowe, T. et al., National Contaminant Biomonitoring Program: Concentrations of seven elements in freshwater fish, *Arch. Environ. Contam. Toxicol.*, 14, 363, 1985.
46. Goolsby, D., Long term changes in concentrations and flux of nitrogen in the Mississippi River Basin, U.S.A, *Hydrological Processes*, 15, 1209, 2001.
47. Mississippi River/Gulf of Mexico Watershed Nutrient Task Force, *Action Plan for Reducing*, *Mitigating*, *and Controlling Hypoxia in the Northern Gulf of Mexico*, Washington, D.C., 2001.
48. Parker, J. and Hope, C., The state of the environment, *Environment*, 34, 19, 1992.
49. National Environmental Policy Act, 42 U.S.C. 4341.
50. Council on Environmental Quality, *Environmental Quality — 21st Annual Report*, Executive Office of the President, Washington, D.C., 1990.
51. Reilly, W., Measuring for environmental results, *EPA J.*, 15, 2, 1990.
52. http://themes.eea.eu.int/all indicators_box
53. Van Woerden, J., Ed., *Data issues of global environmental reporting: Experiences from GEO-2000*, UNEP/DEIA&EW/TR.99-3 and RIVM 402001013, 1999.
54. UNEP, *Global Environmental Outlook 3*, United Nations Environment Programme, Nairobi, 2003.
55. Government Performance and Results Act, PL 103-62, Congress of the United States, Washington, D.C.

56. The Heinz Center, *The State of the Nation's Ecosystems*, Cambridge University Press, New York, 2002.
57. EPA, *The U.S. Environmental Protection Agency Draft Report on the Environment*, Technical Document, EPA 600-R-03-050, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C., 2003.
58. Healy, R. and Ascher, W., Knowledge in the policy process: incorporating new environmental information in natural resources policy making, *Policy Sci.*, 28, 1, 1995.

23 Development of Watershed-Based Assessment Tools Using Monitoring Data

S.L. Osowski

CONTENTS

1-56670-641-6/04/$0.00+$1.50

23.1 INTRODUCTION

This chapter discusses the development of watershed-based environmental assessment tools and the use of monitoring data. The impetus for the use of such data, formulation of tools, and the uses of such data and tools comes from the environmental assessment process. For example, how does one assess indirect or cumulative effects? How does one prioritize areas for further investigation or for mitigation within the environmental assessment process? Cumulative impact assessments,[1–5] use of GIS technology,[6–8] watershed-based approaches,[9–10] and similar decision-making tools[11–13] have recently been the subject of journal articles and have been included in the agendas at environmental policy and scientific meetings. Monitoring data, GIS, and the formulation of holistic (i.e., watershed-based) tools seems to be the preferred method to accomplish these goals.

In this chapter, we will discuss the kinds of monitoring data that can be used, background on environmental assessment, cumulative impacts, watershed-based tools, and a tool developed by U.S. Environmental Protection Agency Region 6 to aid in environmental assessment preparation and review, as well as factors for potential developers to include when developing tools of their own.

23.2 BACKGROUND PRINCIPLES AND CONCEPTS

23.2.1 CUMULATIVE IMPACT ASSESSMENT

The word "cumulative" has been defined in several different ways, depending on context. Words that are similar or even overlapping with cumulative include "aggregate," "indirect," and "secondary" impacts. For example, within risk assessment, "aggregate" refers to the amount of one biologically available chemical from multiple exposure paths,[14] whereas "cumulative" refers to the accumulation of a toxin (or toxic effect) from multiple exposure routes and multiple contaminants (with a common toxicity).[14–15] Traditional risk assessment treats multiple exposures as independent events.[16]

Within the National Environmental Policy Act (NEPA), "cumulative" refers to past and present actions. These actions could identify a significant cumulative impact on the environment; however, there is little agreement as to how past and present actions should be considered in the assessment process and, commonly, past conditions are included as a definition of the existing or baseline conditions within the assessment process.[17] According to McCold and Saulsbury,[17] using a point in time when the environmental resource or condition was most abundant is a suitable baseline. Incorporating past and present conditions as part of the baseline negates their contribution towards cumulative effects.[17]

As NEPA practitioners have discovered, environmental assessments on single projects and the decisions arising from them do not mean that cumulative effects are assessed or determined to be insignificant. The traditional single media approach does not address complex environmental relationships.[18] Single projects with minimal impacts may accumulate over time and space and then may equal a significant impact[19] or as Kahn[20] termed it, the "tyranny of small decisions made singly." Cumulative impacts are not often fully addressed due to the complexity of these impacts, the lack of available data on their consequences, and the desire to limit the scope of environmental analysis. Unfortunately, cumulative impacts are rarely considered in decision-making processes because the methods available (e.g., statistical, models, etc.) are not practical in a regulatory arena.[21] With the development and use of GIS, investigators could identify large scale impacts[22] and impacts that were cumulative.[23] Mitigation opportunities are also affected by an inadequate cumulative impacts assessment.[17] Abbruzzese and Leibowitz[21] developed a framework for comparing landscape units by allowing consideration of cumulative impacts, especially in management decisions, since the goal was a general evaluation of a region as a whole. They used four indices in their evaluation: (1) a function index that measured the amount of a specific ecological attribute, (2) the value of the ecological attribute or function related to social goals, (3) the functional loss of the function or attribute (i.e., cumulative impacts on the function/attribute), and (4) the ability to replace the specific ecological attribute and its function (i.e., replacement potential).

23.2.2 Watershed-Based Assessments

The holistic nature of watershed level assessments incorporates cumulative impacts, in that multiple stressors (biological, socioeconomic, chemical, etc.) can be analyzed over a large spatial scale,[24] either one watershed or the aggregation of several. With the introduction and subsequent increase in the use of spatial analysis tools such as Geographic Information System (GIS), regionally scaled projects, planning, and processes such as those that use the ecoregion[18] or watershed[24–29] as a base unit have become more commonplace. Reasons for using the watershed as the base unit for landscape-level assessments include functionality, biophysical processes, naturally defined area vs. politically defined area, environmental impact assessment, holism, socioeconomic, and comparability/compatibility with other programs or areas.[24,27,29] These tools have also inspired scientists concerned about landscape level patterns and change and their effect on terrestrial and aquatic communities.[27,30] For example, Steiner et al.[27–28] stated that watersheds provide a framework in which to evaluate hydrological processes on wildlife habitat and land suitability for human development (residential, commercial, and industrial) in a way better than other methods or scales. Using a watershed approach with risk assessment can lead to the increased use of monitoring data.[24] Watershed-level assessments are more holistic than assessments performed locally or those based on political boundaries because of their ability to relate potentially unrelated factors[31] and for comparisons at other scales (e.g., several watersheds can be aggregated).[32]

The watershed approach has also been used to analyze environmental problems that do not fit well into traditional programs or assessment methods (e.g., nonpoint source water pollution, regional studies)[24,33] and those problems needing more holistic or comprehensive analysis (including decision making). Watershed-level assessments also lead to intergovernmental coordination on regulatory and management initiatives.[24,27]

23.2.3 Use of Monitoring Data

What kinds of monitoring data are useful in the environmental assessment process? Table 23.1 shows several datasets that can be used in environmental assessment.[34] Some datasets will be almost standard for nearly all projects, whereas others are more

TABLE 23.1
Sources of Monitoring Data Used in Watershed Environmental Assessment Tools

Environmental Feature	Source	Database	Description	Scale and Accuracy	Date
Air quality resources	TCEQ and EPA R6	Nonattainment	Ozone nonattainment and near-nonattainment areas	County level, 1:100,000	2002
Agricultural resources	USGS	NLCD	Agricultural land classification	30 meter resolution	1992
Aquatic resources					
Hydrologic data	U.S. EPA/USGS	NHD	U.S. hydrographic dataset	1:100,000	2000
Hydrologic data	TWDB	Reservoirs to be included in the 1996 Water Plan	Generally reservoirs w/authorized capacity of 5000+ acre-ft and authorized diversion of water for consumptive municipal or industrial use		1997
	GLO	Coastal Management Zone Boundary	Inland extent of areas subject to regulation under the TX Coastal Management Program		
	Bureau of Transportation Statistics	National Waterway Network	Shipping waterways in and around the U.S.	1:100,000	2001
	U.S. Bureau of the Census	TIGER	Hydrologic data	1:100,000	2000
Water quality	TCEQ	Designated Stream Segments of Concern	Impaired waters from 1999 303(d) list	1:63,360–1:250,000	1999
Wetlands	USGS	NLCD	Wetlands land classification	30 m resolution	1992
Terrestrial resources					
Soils	NRCS	STATSGO	State soils layer	1:250,000	1994
Soils	NRCS	SSURGO	County soils layer	1:24,000	Varied
Vegetation	Texas Tech University	GAP	Vegetation and species habitat	30 m	1998
Vegetation	TPWD	Vegetative Types of TX	TX Vegetation/Habitat	1:250,000	1982

Managed lands	Varied	Managed Lands	Parks, forest, wildlife refuges	Varied	Varied
Land use/land cover	USGS	NLCD	Wildlife habitat	30 m	1992
Threatened and endangered species/ sensitive habitats	TPWD	BCD	Quad/county level species lists	7.5′ quadrangle and county	1994
	USFWS	Potential T&E Habitat in S/SE Texas	Potential habitat in S/SE Texas	County level	2001/2002
	U.S. EPA	Potential Habitat Index	Model of highly sensitive habitat	30 m	1992/2002
	GLO	Priority Protection Habitat Areas (Upper and Lower Coast)	Areas along coast of sensitive coastal habitats or species	1:24,000	1995/1998
	GLO	Bird Rookeries	Bird rookeries along coast	1:24,000	
	TPWD	Ecological Stream Segments of Concern	Ecological significant river/stream segments	1:100,000	1995
Hazardous Waste and Brownfields	U.S. EPA	Envirofacts	EPA permitted facilities	Point data—varied accuracy	Varied
	U.S. EPA	Toxic Release Inventory	Toxic release sites	Point data—varied accuracy	2000
	U.S. EPA	Superfund Sites	Federal and state superfund sites	Point data—varied accuracy	2002
	TCEQ	Hazardous Waste Sites	Federal and state hazardous waste sites	Point data—varied accuracy	2002
	TCEQ	Radioactive Waste Sites	Radioactive waste sites	1:24,000	2000
	TCEQ	Landfills	Municipal solid waste landfills	Point data—varied accuracy	1996
	TXDOT	TXDOT Maintenance Facilities	TXDOT maintenance facilities	1:2,000,000??	2000

(*continued*)

TABLE 23.1 (Continued)
Sources of Monitoring Data Used in Watershed Environmental Assessment Tools

Environmental Feature	Source	Database	Description	Scale and Accuracy	Date
Historic, Archeological and Cultural Resources					
Managed Lands (4(f) potential)	Varied	Managed Lands	National parks, forest, and refuges; state parks and wildlife areas	Varied	Varied
Archeological	THC/TXDOT	Archeological Site Distribution in the I-69 Corridor	Density map derived from known distribution of sites	1:24,000	Varied
Archeological	THC	THC Atlas	Archeological data	1:24,000	Varied
Cultural	TNRIS and TIGER	Indian reservation boundaries	Indian reservation boundaries	1:24,000	2000 ?
Cultural	USGS	GNIS	Physical and geographical feature names	1:24,000	1981
Cultural	THC	Historic markers	Historic roadway signs	Point data—varied accuracy	2002
Cultural	THC	Historic national Register properties	Historic national register properties	Point data—varied accuracy	2002
Cultural	TXDOT	Historic off-system bridges	Historic off-system bridges	1:24,000	2001
Geology	BEG	Geologic data	BEG Geology of South Texas	1:250,000	
Topography	USGS	NED	Elevation data	30 m resolution	Varied
Groundwater/Aquifers	TNRIS/TWDB	Major/minor aquifers	Major and minor aquifers of TX	1:250,000	
	U.S. EPA Region 6	Sole Source aquifers	TX sole source aquifers	1:100,000	1996
Watersheds	USGS	8-digit Hydrologic Units	8-digit hydrologic units of the U.S.	1:250,000	1995
Floodplains	FEMA	Q3 Flood Data	100 year/500 year flood plains	1:24,000	Varied
Social/Economic/EJ	U.S. Bureau of the Census	PL94-171	Population and minority data	Block level	1990/2000
	U.S. Bureau of the Census	SF3A	Population, housing, income	Block group level	1990/2002

Miscellaneous	USGS/TOPP	DOQQ	DOQQ	1 m	Varied
	USGS	Digital Raster Graphic	7.5′ topographic maps	1:24,000	Varied
	U.S. Bureau of the Census	TIGER	Urbanized areas	1:100,000	2000
	TXDOT	County Boundaries	County boundaries	1:24,000	2000
	TWDB	Colonias[a]	Locations of colonias	Point data—varied accuracy	1996
	TXDOT	TXDOT District Boundaries	District boundaries	1:24,000	1994
		Aerial Photos	B&W aerial photos		2001
	NASA	Landsat	Satellite imagery	30 m resolution	1996
	TIGER	TIGER	State and federal congressional districts	1:100,000	2000
	U.S. Bureau of the Census	TIGER	Pipelines/utilities	1:100,000	2000
	U.S. Bureau of the Census	TIGER	Railroads	1:100,000	2000
	EPA Region 6	Schools	Schools—address matched using TEA listing	100 m	2002

Note: TCEQ = Texas Commission on Environmental Quality, EPA = Environmental Protection Agency, USGS = U.S. Geological Survey, NLCD = National Land Cover Data, NHD = National Hydrography Dataset, TWDB = Texas Water Development Board, GLO = General Land Office, TIGER, STATSGO = State Soil Geographic Database, SSURGO = Soil Survey Geographic Database, GAP, BCD = Biological and Conservation Data, TXDOT = Texas Department of Transportation, THC = Texas Historical Commission, TNRIS = Texas Natural Resources Information System, GNIS = Geographic Names Information System, BEG = Geology, NED = National Elevation Database, FEMA = Federal Emergency Management Agency, USGS/TOPP, DOQQ = Digital Orthophoto Quarter-Quad, NASA = National Aeronautical and Space Administration, TEA = Transportation Equalization Act.

[a] Colonias are unincorporated residential areas, typically near the U.S.–Mexico border or border towns, where municipal services are lacking (garbage disposal, sewage disposal, drinking water, plumbing to home). The lack of public services in these populated areas increases the chance of environmental contamination and resulting disease.

Source: Lueckenhoff, D.L. and J.E. Danielson, 2002, personal communication.

specific in their use. One caveat, though, is that all of the datasets that one would potentially use were collected for different purposes. For example, air monitoring data may have been collected under the Clean Air Act for permitting or enforcement, not for the environmental assessment process. Therefore, the use of this data serves as a screening-level evaluation to prioritize areas for further investigation. Data may come from different sources that have different levels of confidentiality or restrictions on their use or release to the public.

23.2.4 Use of GIS

GIS is used in the development of assessment and screening tools not only because of its spatial data visualization abilities (i.e., maps of different data layers, coverages, landscape level, etc.), but also because of its modeling and analysis functions, including landscape metrics (e.g., FRAGSTATS), and other calculations (e.g., population density and hydrological functions). Thus, GIS has become a vital research and assessment tool,[22,35–39] although Smit and Spaling[15] predicted that GIS would not be broadly used for cumulative impacts assessment. When used at the watershed or landscape level, GIS can identify and prioritize areas for protection of animal movement by evaluating different land management uses.[36]

Since complicated modeling and analysis tools are less likely to be used in regulatory processes, Leibowitz et al.[40] suggest six properties of GIS assessment tools. These properties include (1) simplicity (not needing expert modeling abilities), (2) use of available data (rather than experimentation), (3) analytical (not needing numerical simulation), (4) approximate (need matches level of effort), (5) measurable change, and (6) expandable (use in more sophisticated models).

The use of GIS and monitoring data has also made the assessment of secondary, indirect, or cumulative effects much easier than in times past. Past evaluations of these effects relied on a nearly completely subjective approach by the investigator, but with the proliferation of computer technology, including GIS, this is changing.

23.2.5 Decision Structures

Most tools use some sort of criteria or factors to evaluate the data layers used in the assessment.[28,41–43] These ranks or scores help to simplify the analysis,[24] normalize disparate data sets onto one nominal scale,[36,44] and provide an easily understandable format to communicate the results to various audiences. These “scores” are helpful in comparing NEPA alternatives or other aspects of projects since the score represents the relative value of one alternative to another.[21,28,44] It also identifies “red flags”[45] or issues that are inadequately addressed or are issues of concern within the environmental assessment process. These scoring systems may represent the difference between an ideal state of the environment and reality.[46] However, this simple type of data integration has been criticized.[47]

When building an assessment tool, one of the things to consider is whether to weight individual criteria[21,36] or to consider them all of equal weight. If weights are chosen, then the importance of the decision increases.[28]

23.3 U.S. EPA REGION 6 EXAMPLE: GIS SCREENING TOOL (GISST) DEVELOPMENT

23.3.1 BACKGROUND

The GISST is a watershed-based environmental assessment tool developed to provide a more systematic approach to considering cumulative impacts in making environmentally sound decisions. It is designed to better understand the potential significance of cumulative effects and to facilitate communication of technical and regulatory data with industry, the public, and other stakeholders. The GISST is not a training manual for impact assessment and users should be familiar with environmental impact assessment (EIA) in order to appropriately consider the vulnerabilities of and potential impacts on the affected environment. In addition to being used in an enforcement setting to correct violations, indices such as GISST can increase the quality of existing databases in a logical, objective framework to strengthen agency decision-making. EPA and others[48] are moving toward watershed or geographic approaches to assessment.[10,49] Most GIS tools are identification tools showing where certain features are on the landscape. GISST is a prioritization tool—that is, given several options, it shows which one has the least potential impact or is more vulnerable.

The GISST helps to focus the agency's assessment of potential impacts under NEPA and points to ways to monitor the effectiveness of project controls and holistic mitigation. As a screening tool, GISST aids industry or permitees, agencies, groups, and the public in comparing facilities and NEPA alternatives, and locating most of vulnerable areas. Such screening tools help establish better communication among stakeholders.[48]

The original drive for the development of GISST began as a way for Region 6 NEPA staff to more objectively evaluate the information submitted by applicants and the potential cumulative impacts of swine feedlots in Oklahoma and present this information to the decision-maker, the EPA Regional Administrator, to determine where concentrations of swine feedlots (CAFOs) might have constituted a potential significant adverse impact in a watershed.[50] Since that time GISST has been applied to a variety of programs and projects.

23.3.2 DEVELOPMENT

GISST is a system that uses GIS coverages and imposes a scoring structure on this data so that decisions can be made. GISST considers environmental vulnerabilities and potential impacts by using USGS watershed subunits called Hydrologic Unit Codes (HUCs).[51] However, the same criteria and principles have been applied to individual projects, usually large scale complex projects (see case study). The watershed subunit is created by merging watershed area data and state stream segment information to form the base analytical unit of the tool. Depending on the state and locality, anywhere from an 8-digit to 14-digit HUC can be used. Higher level HUCs represent a finer grain than lower numbered HUCs.[51] The mathematical algorithm has been employed in several EPA Region 6 applications for consistency and ease of use. The scoring structure consists of criteria, using 1 as low concern and 5 as high concern, based on available datasets and expert input. These individual criterion scores can be compared among the base units one is interested in (e.g., watersheds, facilities, NEPA alternatives). The 1 to 5 scale, which is also consistent with other regional programs, keeps the ranking system simple, with as small a number as possible to capture a sense of greater or lesser environmental concern.

The method that the GISST uses in terms of scoring and ranking could be considered as a multicriteria evaluation or MCE.[15,36,43] MCE can include standardization of criterion scores, multiplication by weighting factor, and/or addition of all criterion scores.[43] Criteria are evaluated using a mathematical formula; however, there are cases where a simple summation of the criteria scores provides a more appropriate assessment (e.g., cumulative impacts). The GISST equation has three parts, but can be modified, depending on project needs and data availability, specifically to a watershed or larger project area:

1. Environmental vulnerability, DV (average for all vulnerability criteria)
2. Environmental impact, D_I (average for all impact criteria)
3. Area (of the watershed, A_{WS}, and total known projects, A_I)

The unitless GISST algorithm is as follows: GISST = $[\Sigma (A_I/A_{WS})] \times D_V \times D_I$. The individual criteria selected, including the area criterion, are dependent on the needs of, and appropriateness to, specific projects. GISST is flexible in that portions of the equation can be used or not, as appropriate. For example, a user may only want to determine the relative environmental vulnerability of two riparian areas or watershed subunits, or a user may want to know the potential impacts to those areas in addition to the environmental vulnerability. The user selects the appropriate criteria to use (Table 23.2). Other projects may use the county or other polygon instead of the watershed subunit and A_{WS}. The EPA Region 6 developers stress that the individual criterion scores may be more important in communicating environmental concerns rather than final GISST scores.

The development of criteria forces decision-makers to determine the comparative risk or impact of five options. In principle, this is a very difficult process, and scores/criteria may cause disagreements or controversy. GISST also makes all stakeholders aware of what resources will be evaluated and the associated risk (score) that environmental assessors are willing to acknowledge. Screening models such as the GISST can lead to decisions to prioritize certain aspects of facility or project operations for environmental review. GISST is different from other GIS tools in several ways, the most important of which is the scoring structure. Most GIS tools are used as mapping tools in which the user gets a map and then must decide what constitutes greater or lesser environmental concerns. By already having a scoring structure in place in the form of criteria (1 to 5 scoring), GISST results are more objective and less subjective. Therefore, GISST becomes an effective communication tool and can assist in streamlining projects or program needs.

TABLE 23.2
Listing of Final and Provisional Criteria Developed and Used for GISST

Category of Criterion
Water Quality
Surface water use
Water quality (STORET Data)
Rainfall
Water releases
Surface water quantity
Distance to surface water
Groundwater probability
Groundwater quality
Unified watershed assessment
Average stream flow
Sole source aquifer
Major/minor aquifer
Aquifer/geology rating
Channelization
Individual well water
Septic tank and cesspool use
TRI reported water releases
Soil permeability
Surface water area
100-year floodplain (P)
500-year floodplain (P)
Water design flow data (P)
Ecological
Wildlife habitat
Wildlife habitat quality (land use data)
Landscape texture
Landscape aggregation
Patch area
Habitat fragmentation
Endangered and threatened species
Road density
Watershed area
Wetlands (P)
Protected habitats (P)
Agricultural lands (P)
Air Quality
Air quality
TRI reported air releases
Severity of ozone pollution (P)
Ozone nonattainment (P)
Pollution Prevention
Environmental assessment (P)
Pollution prevention (P)
Model energy code (P)
Energy-efficient office equipment (P)
Energy efficient appliances (P)
Lighting system upgrade (P)
Million Solar Roofs Initiative (P)
Federal Energy Management Program (P)
Socioeconomic
Colonias[a]
High school education
Educational achievement ranking
Economic
Minority
Age
Children
Older population
Pregnancy
Population change
Population density
Total population
Houses lacking complete plumbing
Telephone communications
Ability to speak English
Linguistic isolation
Foreign born
Cultural resources
Age of homes (P)
Employment (P)
Toxicity
Toxicity weighted TRI water releases
Toxicity weighted TRI air releases
Toxicity weighted RCRA-BRS data
Other industries, pollution sources
RCRA permitted units (P)
RCRA hazardous waste disposal (P)

(*continued*)

TABLE 23.2 (Continued)
Listing of Final and Provisional Criteria Developed and Used for GISST

Category of Criterion	Category of Criterion
CAFO	
Livestock population density	Transportation near CAFOs
Lagoon loading rate	Density of CAFOs
Lagoon treatment system liner	Proximity of CAFOs
Land application technology	Unregulated CAFO facilities (P)
Nitrogen budget	
Phosphorus budget	**General**
Lagoon storage capacity	Density of National Historical Places (P)
Well head protection	Proximity of National Historical Places (P)
Employment in CAFO industry	Density of managed lands (P)
Odor	Proximity of managed lands (P)

Note: (P) indicates provisional criteria. Provisional criteria are those that have not been used, do not have a database to support their use, or are in the process of being developed, peer reviewed, and finalized; the underlying data and GIS coverages are dynamic and therefore the criteria may change as data sources become available.

[a] Colonias are unincorporated residential areas, typically near the U.S.–Mexico border, where municipal services are lacking (garbage disposal, sewage disposal, drinking water, and plumbing to homes). The lack of public services in these populated areas increases the chance of environmental contamination and resulting disease.

23.3.2.1 Area Criterion

$[\Sigma(A_I/A_{WS})]$ is the ratio of the cumulative area affected to the total area of evaluated watershed subunit, expressed as a percentage.

23.3.2.2 Vulnerability Criteria

The degree of vulnerability, D_V, is the sum of individual criterion scores divided by the number of vulnerability factors used. The vulnerability criteria are intentionally unweighted, reflecting a decision by the tool development team that the number of criteria used reflects the nature and purpose of the project for which it is used. In effect, the number of criteria for a certain environmental resource weights that feature more than an environmental resource with only one criterion. For example, one might use four water-related criteria, but only one economic criterion; therefore the final GISST score would be weighted toward water issues. The application of a criterion is dependent on the availability of data for a particular geographic area. Consequently, a particular criterion may not be used until a viable dataset becomes available. Many of the criteria reflect the questionnaire categories in Canter and Kamath,[1] although research was not available at the time the GISST was developed.

23.3.2.3 Impact Criteria

D_I is the sum of individual impact criterion scores divided by the number of impact factors used. They reflect industry specific impacts and not all may be used for a specific project. Many are also dependent on data and information from the individual facility or entity being evaluated. Therefore, stakeholders must have a clear understanding of the tool and a willingness to participate by providing sensitive data concerning their operations.

23.3.2.4 Criteria Groups

Criteria, whether impact or vulnerability, can be placed into broad groups: water quality, ecological, air quality, socioeconomic, toxicity, CAFOs, pollution prevention, and enforcement/compliance (Table 23.2). The use of water quality criteria will give the user an overall sense of surface and groundwater quantity and quality. Several different data sources were used.[50] Depending on the project, the user may not want to use all of the water quality criteria available or there may be a gap in which the user should develop a new criterion to meet his/her needs. In general, ecological criteria provide the user with what conditions are like for nonhuman organisms in the project area (wildlife habitat, endangered species, fragmentation, etc.). They describe the watershed or landscape (large scale) and the project area (small scale). Several criteria in this section (and others as well) can be "flipped." For example, we have chosen large tracts of wildlife habitat as the most vulnerable condition. Conventional wisdom suggests that large unbroken tracts of habitat are better able to support large species (e.g., black bear), migratory species (e.g., bald eagle), and maintain the functioning of communities and ecosystems. Certain large migratory species may serve as "umbrella species" for smaller, less mobile species (e.g., amphibians, insects). However, the most vulnerable condition could be the very small remnant patches of a particular habitat type. Without proper connectivity, however, small remnants of habitat will probably not support certain species.

Socioeconomic criteria are important for a number of reasons, including the requirements to assess environmental justice, NEPA requirements, and to prepare an effective public involvement strategy. Many of the criteria are useful for this last purpose, especially if English is not the primary language or the literacy level of the community is not high. Socioeconomic criteria are important in that an individual's place of residence, diet, exposure to occupational hazards, ability to receive adequate health care (both preventive and post-injury) may be controlled by income and education. For environmental justice (Title VI complaints), only three criteria need be used (economic, minority, and total population). For NEPA assessments, several others can be used to determine whether the proposed project will have a beneficial or adverse effect on the local population. Other criteria go even further in helping EPA staff prepare an effective public education and involvement campaign.

Depending on the type of project the user is trying to assess with GISST, toxicity criteria may be very important. These criteria help to determine what pollution sources are in proximity to the proposed project and the amount of releases (air and water) from facilities. In assessing cumulative or aggregate health effects, these criteria become extremely important in the decision as to whether further field investigations are needed.

23.3.3 Uses

23.3.3.1 Case Study: Swine Concentrated Animal Feeding Operation (CAFO) New Source Determination

This case study shows how GISST assists in the complete NEPA process, including identification of baseline conditions and potential impacts, avoidance and mitigation of impacts, monitoring of mitigation commitments, and enforcement of Clean Water Act violations.

Region 6 performs NEPA review for New Source Determinations for National Pollutant Discharge Elimination System (NPDES) permits in states where these federal programs have not been authorized/delegated. Oklahoma does not have NPDES permit authority for CAFOs. At the time in 1997, many states had been embroiled in controversies related to large CAFOs. Supporters of CAFOS argued that their facilities were simply another agricultural activity, protected in many states by right-to-farm laws that supported local economies. Opponents of CAFOs argued that the facilities were under-regulated industrial operations that resulted in environmental and public health risks. As such, the public was often divided and EPA was looked upon as an objective third-party to fairly evaluate these controversial issues.

The environmental issue was that very large (4 million animals/year) swine CAFOs were becoming established in a one-to-two-county (watershed) area in Oklahoma and that there could be possible (1) leaching from lagoon and/or land application areas, (2) odor from the facility (lagoons and land application of swine waste), and (3) health concerns due to dead animal disposal. This leaching might cause nitrate contamination of groundwater which also serves as drinking water for some residents. Using GIS coverages and information in the applicant Environmental Information Document (EID), GISST showed that several criteria (e.g., nitrate–nitrite exceedances, probability of the water table within 6 ft of the surface, proximity of CAFOS to each other) scored high (5, on a 1 to 5 scale). The facility did well on the use of control technologies (e.g., lagoon liner, innovative sprayer technology). Figure 23.1 to Figure 23.3 show D_V, D_I, and final CRIA/GISST scores (note: the acronym was changed from CRIA to GISST after the pilot study), respectively, for five subwatersheds in Oklahoma. These scores determine which watersheds

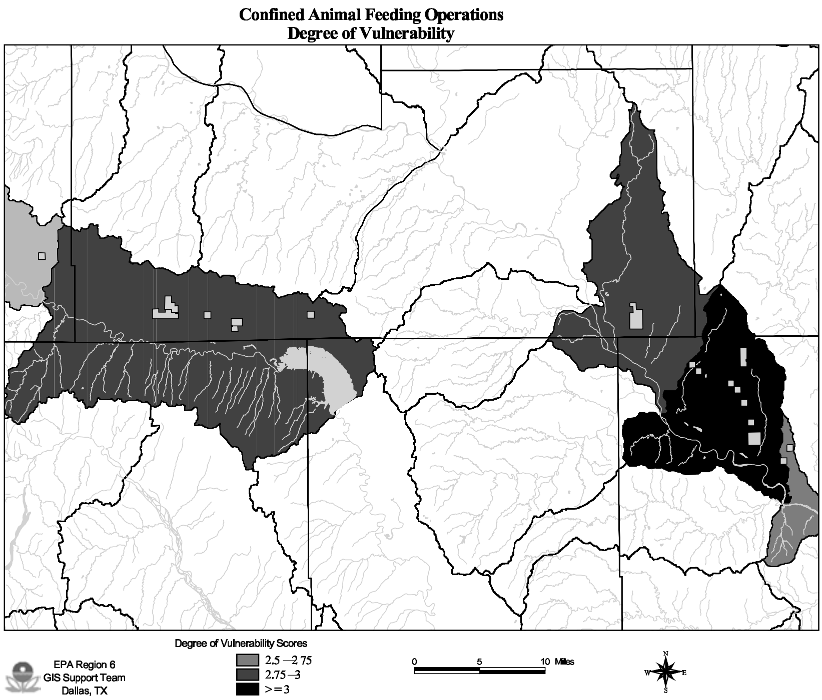


FIGURE 23.1 Degree of vulnerability for five subwatersheds in Oklahoma. Individual facilities are pictured in light gray. Darker colors indicate a higher degree of environmental vulnerability.

may be more vulnerable and/or have more impacts from the cumulative effects of multiple swine feedlots. Although not a measure of cumulative effects, D_I (Figure 23.4) can be calculated for individual facilities.

The Regional Administrator determined that the CAFO would not have their NPDES permit approved and a finding of no significant impact (FNSI) for the environmental assessment (EA) until a monitoring protocol and schedule could be agreed upon, given that the GISST had identified groundwater contamination as a potential significant impact. Monitoring (well) reports were submitted by the facility quarterly. At least one of these reports showed nitrate exceedances and possible groundwater contamination. This information was given to inspectors and enforcement officers who followed up with enforcement actions.

23.3.3.2 IH-69 NAFTA International Trade Corridor

IH-69 is a congressionally mandated interstate highway from Detroit, MI, to Brownsville and McAllen, TX. It runs 1600 mi, 1000 of which occur in Texas

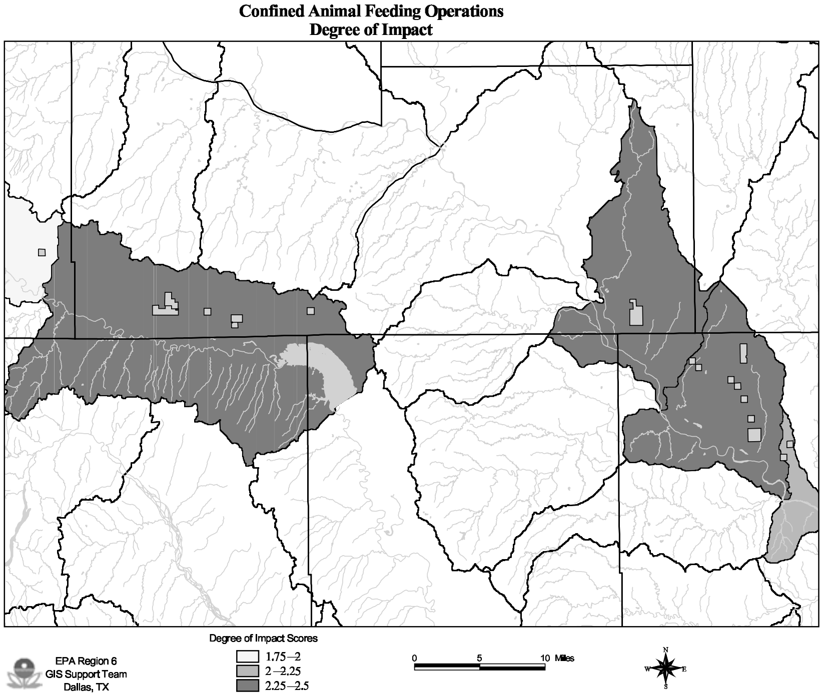


FIGURE 23.2 Degree of impact for five subwatersheds in Oklahoma. Individual facilities are pictured in light gray. Darker colors indicate a higher degree of impact from the cumulative effect of all pictured facilities in the subwatershed.

(Figure 23.5). Its stated goals are to facilitate trade between Mexico, Canada, and the U.S. and to encourage economic development and transportation access to rural communities along the route. Issues for this large and complex project include the potential for large scale environmental impacts due to road alignment and construction, and coordination among the many state and federal agencies involved. The environmental goal is avoidance of critical ecological and environmental areas during the alignment determination process. Using GISST in 1 km^2 throughout Segments of Independent Utility (SIU), GISST identified areas of high vulnerability for possible NEPA alternatives and mitigation opportunities (Figure 23.6 and Figure 23.7). Using 1 km^2 is a new application of the GISST methodology. Traditionally, criteria and the final GISST scores were calculated on a watershed subunit basis. However, it is more appropriate to use the SIU as the base unit for the I-69 project. The use of GISST is a way for EPA to become involved early in the NEPA process to make environmental concerns known as a way to streamline the process for such a large complex project.

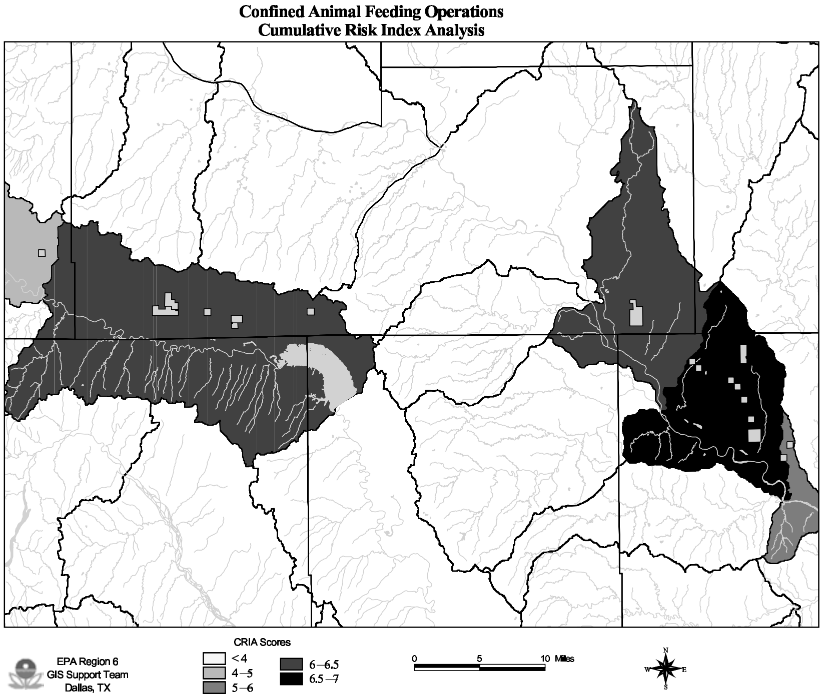


FIGURE 23.3 Final CRIA/GISST score for five subwatersheds in Oklahoma. Darker colors indicate vulnerable environmental conditions and potentially significant impacts from multiple swine feedlots (pictured in light gray).

23.4 LESSONS LEARNED

There are several benefits that users have noted since GISST became available:

- *Improved quality of review:* Comments can be compiled earlier, proactively, and are issue specific. Traditional NEPA comment letters are generic in that they refer to regulations and not to information contained in the NEPA document.
- *Early actions driven by technological capabilities:* EPA has been criticized for accepting information and analysis from applicants and contractors without verifying the information to some degree.
- *Wholesale approach:* GISST allows us to serve more customers by getting more information to more people relatively quickly.
- *Consistency:* GISST can develop into a region-wide capacity for high quality reviews and document preparation.
- *Institutional knowledge base:* As staff retires or moves to different jobs, knowledge of programs and regulations is lost. GISST criteria and scoring system capture this knowledge and enhance it through technology.

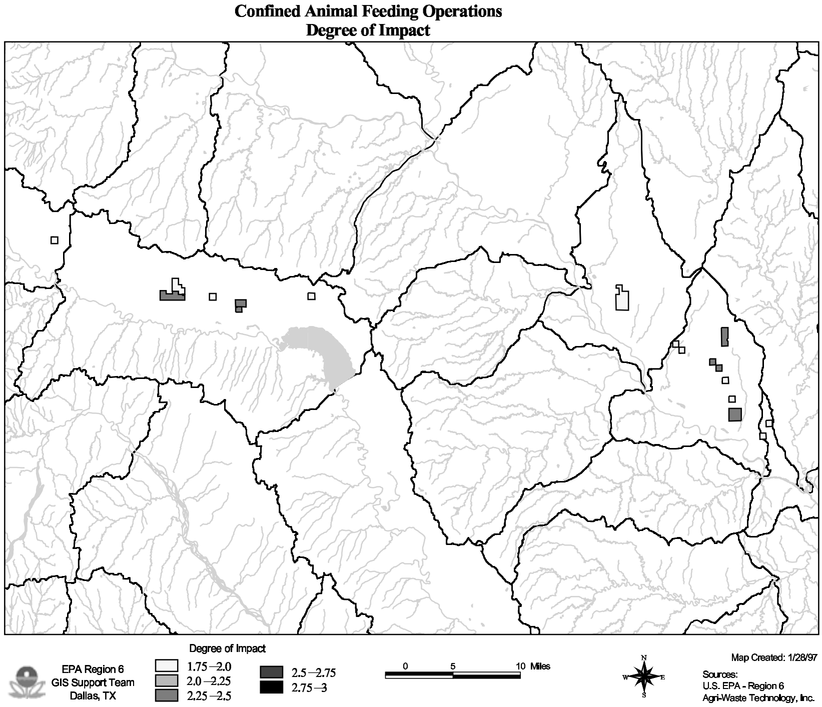


FIGURE 23.4 Degree of impact for each swine feedlot (CAFO) facility in five subwatersheds in Oklahoma. Darker colors indicate a higher degree of impact from the pictured facility.

- *Screening level:* GISST is not time- or labor-intensive, but designed to point out "red flags" to prioritize where additional resources might be used or additional information and analysis is needed.
- *Transparency:* GISST was developed in-house so users know how it works. One can compare this to purchased software packages that are "black boxes" where a user enters information, but has no idea how the "answer" is calculated. GISST users have more information on how each criterion is calculated and how it fits in with other criteria.
- *Flexible:* New criteria can be added/changed as needed.
- *Scaling:* GISST can be applied to local projects encompassing one facility or to regional projects such as interstate highways.

A few obstacles were noted as well. GISST may cause an information overload. For example, if a user had five NEPA alternatives and used 40 GISST criteria, the resulting matrix can be quite large and all of this information is accessible approximately 2 h after the GIS program is initiated. The EPA Region 6 developers stress looking for red flags—criterion scores of 4 or 5 that might indicate an environmental

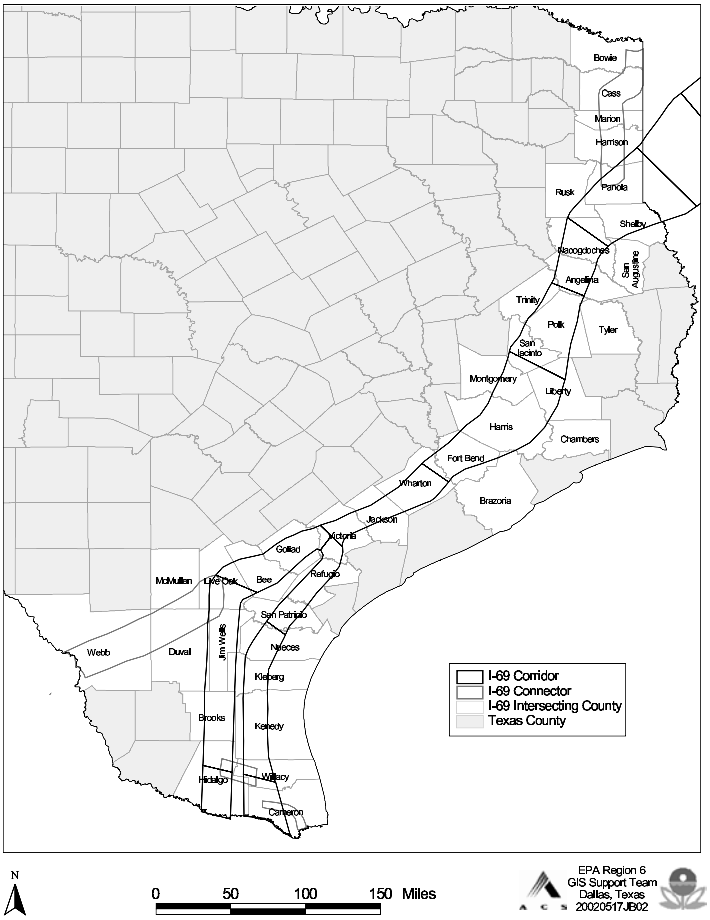


FIGURE 23.5 IH-69, NAFTA International Trade Corridor in Texas.

problem. Using GISST may increase workload because it is a wholesale approach. It takes approximately 2 h to get a wealth of information that previously was not available or only available after weeks of data collection. Therefore, staff may be able to review more documents, etc. in a shorter timeframe. The GISST is a screening-level tool only. It does not replace traditional risk assessment or field investigations. It can only point the user in the direction of where problems are likely to happen or where resources should be directed for additional studies.

23.5 CONCLUSION

In a climate of dwindling resources, the use of monitoring data and GIS to identify and prioritize areas of environmental vulnerability or impact will surely increase. Site-specific analysis has traditionally been the way environmental concerns were

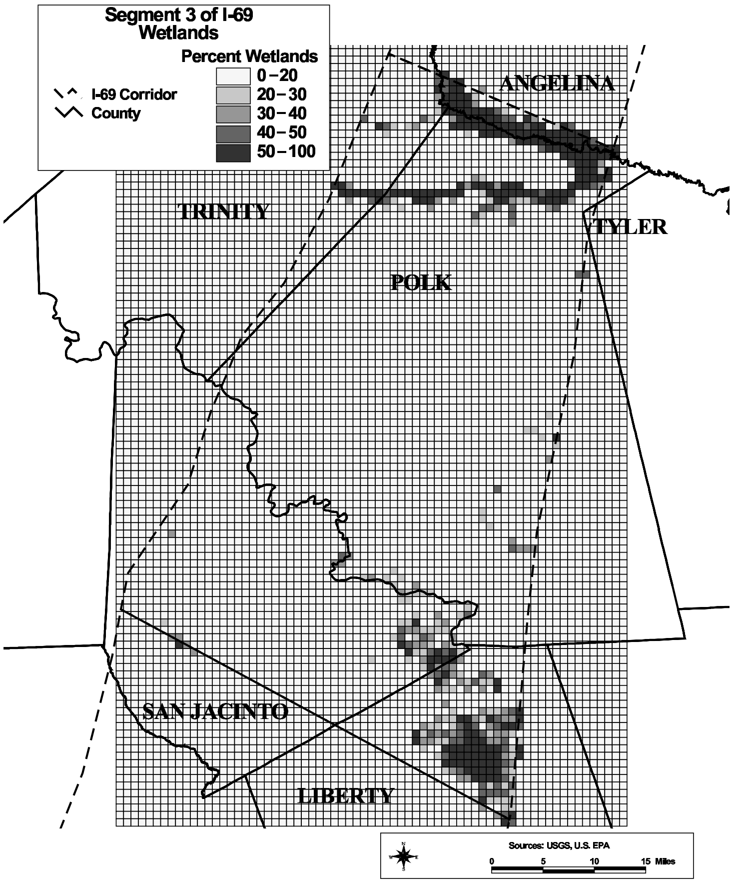


FIGURE 23.6 GISST Scores calculated for 1 km blocks within SIU 3 for the wetlands criterion. Darker colors indicate more vulnerable wetland areas (score of 5).

identified. However, the recognition that impacts may have broader or more significant effects at the watershed or landscape level, along with the development of sophisticated computer visualization and modeling tools, means that tools such as the GISST and those of others will become more prominent. Other organizations that have developed similar tools include The Nature Conservancy, other EPA offices, the U.S. Forest Service, and some state agencies. These types of tools are ever evolving as data become available.

ACKNOWLEDGMENTS

The author would like to thank Dominique Lueckenhoff, Dave Parrish, Gerald Carney, Joe Swick, Hector Pena, and Jeff Danielson, the co-developers, users, and proponents of the GISST tool; Jeff Danielson for GIS expertise; and Rob Lawrence

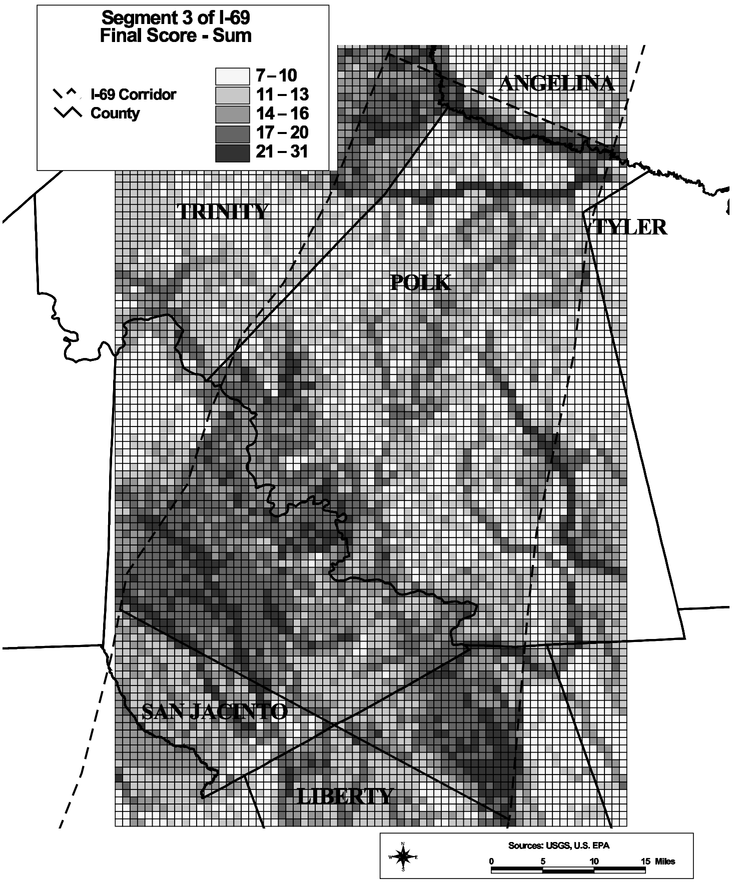


FIGURE 23.7 Summation of 20 GISST vulnerability criteria scores, calculated for 1 km blocks within IH69 SIU 3. Darker colors indicate a higher vulnerability (score of 5).

for administrative support. The views represented in this chapter are those of the author and do not constitute EPA policy.

REFERENCES

1. Canter, L. W. and J. Kamath. 1995. Questionnaire checklist for cumulative impacts. *Environ. Impact Assess. Rev.* 15: 311–339.
2. Rees, W. E. 1995. Cumulative environmental assessment and global change. *Environ. Impact Assess. Rev.* 15: 295–309.
3. Cox, L. H. and W. W. Piegorsch. 1996. Combining environmental information. I: Environmental monitoring, measurement, and assessment. *Environmetrics* 7: 299–208.

4. Piegorsch, W. W. and L. H. Cox. 1996. Combining environmental information. II: Environmental epidemiology and toxicology, *Environmetrics* 7: 309–324.
5. Burris, R. K. and L. W. Canter. 1997. Cumulative impacts are not properly addressed in environmental assessments. *Environ. Impact Assess. Rev.* 17: 5–18.
6. Peccol, E., C. A. Bird, and T. R. Brewer. 1996. GIS as a tool for assessing the influence of countryside designations and planning policies on landscape change. *J. Environ. Manage.* 47: 355–367.
7. Dale, H., A. W. King, L. K. Mann, R. A. Washington-Allen, and R. A. McCord. 1998. Assessing land-use impacts on natural resources. *Environ. Manage.* 22: 203–211.
8. Zhang, M., S. Geng, and S. L. Ustin. 1998. Quantifying the agricultural landscape and assessing spatio-temporal patterns of precipitation and groundwater use. *Landscape Ecol*. 13: 37–53.
9. Wang, X. and Z.-Y. Yin. 1997. Using GIS to assess the relationship between land use and water quality at a watershed level. *Environ. Int.* 23: 103–114.
10. Caruso, B. S. and R. C. Ward. 1998. Assessment of nonpoint source pollution from inactive mines using a watershed-based approach. *Environ. Manage.* 22: 225–243.
11. Howard, D. C. and R. G. H. Bunce. 1996. The countryside information system: a strategic-level decision support system. *Environ. Monit. Assess.* 39: 373–384.
12. Partidario, M. R. 1996. Strategic environmental assessment: key issues emerging from recent practice. *Environ. Impact Assess. Rev.* 16: 31–55.
13. Laskowski, S. L. and F. W. Kutz. 1998. Environmental data in decision making in EPA regional offices. *Environ. Monit. Assess.* 51: 15–21.
14. Moschandreas, D. J. and S. Karuchit. 2002. Scenario-model-parameter: a new method of cumulative risk uncertainty analysis. *Environ. Int.* 28: 247–261.
15. Smit, B. and H. Spaling. 1995. Methods for cumulative effects assessment. *Environ. Impact Assess. Rev.* 15: 81–106.
16. EPA. 1999. Guidance for performing aggregate exposure and risk assessments.
17. McCold, M. and J. W. Saulsbury. 1996. Including past and present impacts in cumulative impact assessments. *Environ. Manage.* 20: 767–776.
18. Mysz, A. T., C. G. Maurice, R. F. Beltran, K. A. Cipollini, J. P. Perrecone, K. M. Rodriguez, and M. L. White. 2000. A targeting approach for ecosystem protection. *Environ. Sci. Policy* 3: 347–35.
19. Theobald, D. M., J. R. Miller, and N. T. Hobbs. 1997. Estimating the cumulative effects of development on wildlife habitat. *Landscape Urban Plan.* 39: 25–36.
20. Kahn, A. E. 1966. The tyranny of small decisions: market failures, imperfections, and the limits of economics. *KYKLOS* 19: 23–45.
21. Abbruzzese, B. and S. G. Leibowitz. 1997. A synoptic approach for assessing cumulative impacts to wetlands. *Environ. Manage.* 21: 457–475.
22. O'Neill, R. V., K. H. Riitters, J. D. Wickham, and K. B. Jones. 1999. Landscape pattern metrics and regional assessment. *Ecosys. Health* 5: 225–233.
23. Odum, W. E. 1982. Environmental degradation and the tyranny of small decisions. *BioScience* 32: 728–729.
24. Serveiss, V. B. 2002. Applying ecological risk principles to watershed assessment and management. *Environ. Manage.* 29: 145–154.
25. Dickert, T. G. and A. E. Tuttle. 1985. Cumulative impact assessment in environmental planning: a coastal wetland watershed example. *Environ. Impact Assess. Rev.* 5: 37–64.
26. Espejel, I., D. W. Fischer, A. Hinojosa, C. Garcia, and C. Levya. 1999. Land use planning for the Guadalupe Valley, Baja California, Mexico. *Landscape Urban Plan.* 45: 219–232.

27. Steiner, F., J. Blair, L. McSherry, S. Guhathakurta, J. Marruffo, and M. Holm. 2000a. A watershed at a watershed: the potential for environmentally sensitive area protection in the upper San Pedro Drainage Basin (Mexico and USA). *Landscape Urban Plan.* 49: 129–148.
28. Steiner, F., L. McSherry, and J. Cohen. 2000b. Land suitability analysis for the upper Gila River watershed. *Landscape Urban Plan.* 50: 199–214.
29. Tinker, D. B., C. A. C. Resor, G. P. Beauvais, K. F. Kipfmueller, C. I. Fernandes, and W. L. Baker. 1998. Watershed analysis of forest fragmentation by clearcuts and roads in a Wyoming forest. *Landscape Ecol.* 13: 149–165.
30. Jones, K. B., A. C. Neale, M. S. Nash, R. D. Van Remortel, J. D. Wickham, K. H. Riitters, and R. V. O'Neill. 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States Mid-Atlantic Region. *Landscape Ecol.* 16: 301–312.
31. Miller, W., M. Collins, F. Steiner, and E. Cook. 1998. An approach for greenway suitability analysis. *Landscape Urban Plan.* 42: 91–105.
32. Montgomery, D. R., G. E. Grant, and K. Sullivan. 1995. Watershed analysis as a framework for implementing ecosystem management. *Water Resour. Bull.* 31: 369–385.
33. Boughton, D. A., E. R. Smith, and R. V. O'Neill. 1999. Regional vulnerability: a conceptual framework. *Ecosys. Health* 5: 312–322.
34. Lueckenhoff, D.L. and J.E. Danielson, 2002, personal communication.
35. Ji, W. and P. Leeberg. 2002. A GIS-based approach for assessing the regional conservation status of genetic diversity: an example from the southern Appalachians. *Environ. Manage.* 29: 531–544.
36. Clevenger, A. P., J. Wierzchowski, B. Chruszcz, and K. Gunson. 2002. GIS-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. *Conserv. Biol.* 16: 503–514.
37. Dale, V. H., R. V. O'Neill, F. Southworth, and P. Pedlowski, 1994. Modeling effects of land management in the Brazilian Amazonian settlement of Rondonia. *Conserv. Biol.* 8: 196–206.
38. Treweek, J. and N. Veitch. 1996. The potential application of GIS and remotely sensed data to the ecological assessment of proposed new road schemes. *Global Ecol. Biogeogr. Lett.* 5: 249–257.
39. Iverson, L. R., D. L. Szafoni, S. E. Baum, and E. A. Cook. 2001. A riparian wildlife habitat evaluation scheme developed using GIS. *Environ. Manage.* 28: 639–654.
40. Leibowitz, S. G., C. Loehle, B.-L. Li, and E. M. Preston. 2000. Modeling landscape functions and effects: a network approach. *Ecol. Model.* 132: 77–94.
41. Karydis, M. 1996. Quantitative assessment of eutrophication: a scoring system for characterizing water quality in coastal marine ecosystems. *Environ. Monit. Assess.* 41: 233–246.
42. Xiang, W.-N. 2001. Weighting-by-choosing: a weight elicitation method for map overlay. *Landscape Urban Plan.* 56: 61–73.
43. Store, R. and J. Kangas. 2001. Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modeling. *Landscape Urban Plan.* 55: 79–93.
44. Wickham, J. D., K. B. Jones, K. H. Riitters, R. V. O'Neill, R. D. Tankersley, E. R. Smith, A. C. Neale, and D. J. Chaloud. 1999. An integrated environmental assessment of the Mid-Atlantic Region. *Environ. Manage.* 24: 553–560.
45. Theobald, D. M., N. T. Hobbs, T. Bearly, J. A. Zack, T. Shenk, and W. E. Riebsame. 2000. Incorporating biological information in local land-use decision making: designing a system for conservation planning. *Landscape Ecol.* 15: 35–45.

46. Tran, L. T., C. G. Knight, R. V. O'Neill, E. R. Smith, K. H. Riitters, and J. Wickham. 2002. Fuzzy decision analysis for integrated environmental vulnerability assessment of the Mid-Atlantic Region. *Environ. Manage.* 29: 845–859.
47. Suter, G. W. 1993. A critique of ecosystem health concepts and indices. *Environ. Toxicol. Chem.* 12: 1533–1539.
48. Costanza, R. and M. Ruth. 1998. Using dynamic modeling to scope environmental problems and build consensus. *Environ. Manage.* 22: 183–195.
49. TNRCC. 1996. *The statewide watershed management approach for Texas-a guidance manual for TNRCCs Office of Water Resource Management.* August 29, draft.
50. Osowski, S. L., J. D. Swick, Jr., G. R. Carney, H. B. Pena, J. E. Danielson, and D. A. Parrish. 2001. A watershed-based cumulative risk impact analysis: environmental vulnerability and impact criteria. *Environ. Monit. Assess.* 66: 159–185.
51. Cederstrand, J. and A. Rea. 1995. *Watershed boundaries for Oklahoma*. OFR 95-727. U.S. Geological Survey, Oklahoma City, OK.

24 Bioindicators for Assessing Human and Ecological Health

J. Burger and M. Gochfeld

CONTENTS

1-56670-641-6/04/$0.00+$1.50

24.1 INTRODUCTION

The public, tribal nations, governmental agencies, scientists, health professionals, conservation agencies, and other stakeholders are increasingly interested in assessing the well-being of humans, other species, and their ecosystems. This interest has led to the establishment of local, state, and federal agencies, as well as conservation/preservation societies, to preserve, protect, and manage our environments and their ecological resources. While there are several paradigms to assess human and ecological health, human health risk assessment and ecological risk assessment have emerged as separate paradigms embodying the disciplines of toxicology and exposure assessment.[1,2] Other regulatory frameworks have been developed for ecological and human health risk assessment/management.[3,4] Human health risk assessment is often called environmental risk assessment, but the use of environment in this context is confusing.

There are many papers and books devoted to risk assessment, but usually they examine either human health assessment[5–7] or ecological risk methods.[8–14] Some volumes have included both human and ecological health risk assessment,[15] but they are usually treated in separate papers; integration is left to the reader. However, there have been some attempts to show the interconnections between human and ecosystem health (see Reference 16) and several thoughtful papers dealing with the interface between human and ecological health.

In this chapter we propose that assessing human and ecological health involves establishing biomonitoring plans that use indicators and biomarkers of exposure and effects. We suggest that risk assessors can develop bioindicators which are relevant to both human and ecological health. Since humans are only one of many receptors within ecosystems, they should be considered in the context of other species—particularly those they depend upon for food, clothing, shelter, tools, and weapons. Moreover, long-term monitoring programs require the interest and support of the general population, as well as government commitment, since public funds are needed to conduct these programs. Such interest is more easily gained if the bioindicators provide information about both human and ecosystem health.

24.1.1 Risk Assessment Paradigms

The National Research Council (NRC) codified the human health risk assessment paradigm in the early 1980s,[1] identifying four phases: hazard identification, dose-response analysis, exposure assessment, and risk characterization. This formalization was essential for human health risk assessment because it ensures consistency and provides the opportunity to compare across hazards, exposures, human populations, and sensitive subpopulations. Although human health risk assessment is complex and rife with uncertainties, the human population encompasses a single species, and risk assessment typically focuses on a limited range of morbidity or mortality endpoints.

Moreover, it is clear that individuals matter in human health risk assessment, particularly during the risk management and risk communication phase.

Ecological health is difficult to define, but we mean it to include both the health of individuals and species, as well as that of functioning ecosystems. Ecological risk assessment is inherently much more complex than human risk assessment because any ecosystem has dozens to hundreds of species (each with endpoints), as well as higher-order interactions that are critical to ecosystem functioning.[10,12,17–19] Ecological risk assessment is often more of an iterative process than is human health risk assessment because of its complexity.

Efforts to develop an integrated approach to ecological and human health risk assessment are desirable for efficacy, cost-effectiveness, and to garner public support.[18,20,21] However, the distinction between human health risks at a given point and the vast spatial context of ecological risks sometimes makes integrating human and ecological health risk assessment difficult.[19]

In 1993, the NRC suggested that the same paradigm should apply to both human and ecological health risk assessment, although the magnitude and importance of the basic steps might differ.[2] The refined risk assessment paradigm included problem identification, hazard identification, exposure assessment, exposure-response assessment, and risk characterization, while risk management was integrated as the endpoint or end-user of the risk assessment outcome. The NRC report also noted, as had others, that research, validation, and monitoring were critical aspects of the science that should impact all phases of risk assessment and risk management.[2] While some aspects of risk assessment such as extrapolation have often considered both humans and ecological receptors,[22] this approach has been rare for most issues, particularly for bioindicator development.

24.1.2 Other Risk Considerations

It soon became apparent that risk assessment was only one aspect of understanding and solving environmental problems— risk management and risk communication were also key elements in the process. The Presidential/Congressional Commission on risk assessment and risk management treated risk assessment as just one component within a larger risk management framework, which included problem definition, option selection, remediation, and evaluation as a context, with stakeholders involved at every step.[23] The process was iterative, requiring the revisiting of different aspects of the risk process. Monitoring is clearly an important part of both risk assessment and risk management to assess the outcomes, efficacy, and permancy of intervention or remediation and to provide early warning of any remaining problems.

24.2 MONITORING

Monitoring or surveillance are key to assessing the status or well-being for humans and other receptors within functioning ecosystems. Data can be obtained from many sources, involving many abiotic and biotic systems at a variety of spatial and temporal scales. Ideally, monitoring data are tailored to meet the needs of a particular question or situation and normally provide information on status and trends. Monitoring can provide early warning of any changes that could result in significant risk.

24.2.1 Types of Monitoring Data

Monitoring data may reflect abiotic systems (air, water, soil, sediment), biological processes (numbers of organisms, mortality rates, reproductive rates), biochemical markers (enzyme activity), or toxicological markers (blood lead, urinary metabolites). While biological processes have usually involved individuals or populations, recent attention has focused on ecosystem structure and function, such as species diversity, productivity, nutrient cycles, and food-web relationships. Similarly, there are larger-scale human processes (disease rates, migrations).

Sources of monitoring data may be low tech (field observations) or high tech (real-time data acquisition by satellites). Datasets may be sparse (one observation per year) or dense (updated several times a minute), and the scale of spatial resolution varies greatly as well.

24.2.2 Uncertainty for Both Ecological and Human Health Risk Assessments

Monitoring data can be qualitative but are usually quantitative in nature. Acquiring a set of monitoring data should allow extraction for spatial or temporal trends, for comparison among sites or regions, and for a recognition of the underlying distribution of values from which parameters (at least mean and variance) can be extracted. Much environmental data have an underlying log-normal distribution.

24.2.2.1 Sources of Variability

The quality of the input data will have at least four sources of variability: intrinsic variation, sampling errors, analytic errors, and random errors. Intrinsic variation refers to the variation within the biological system — which may be particularly important in the selection of indicators. This is real biological variability, often with a genetic or biochemical basis. It cannot be reduced.

Sampling errors occur during the sampling phase, either in collection of biological data (reproductive success, numbers) or of samples (e.g., tissues). The individuals sampled may not be representative of the population or phenomena. Improved design and random and/or stratified sampling to increase samples can reduce uncertainty.

Some databases have built-in quality assurance components which operate at the design and laboratory phase to minimize sampling and analytic errors (see EPA's data quality objectives[24]). Analytic errors range from mix-up of samples or mislabeled samples to errors introduced during storage, preparation, compositing, extraction, analysis, calculation, and reporting. Chain of custody procedures can reduce some of the errors, but built-in error-trapping algorithms should be used to reduce others, and quality control strategies, including use of reference laboratories, will reduce others. For some analyses many values may fall below the method's limit of quantification. When many or most values are in the nondetectable range, it is difficult to adequately characterize the distribution.

24.2.2.2 Quality Assurance

Monitoring data and the subsequent use of bioindicators or biomarkers should employ standard quality assurance and quality control (QA/QC) procedures to insure confidence in the overall quality of the data. These procedures involve methods of assuring the accuracy and validity of the data and vary depending upon whether data are field or laboratory-generated. Ultimately, however, the user must decide whether monitoring data are sufficient in type, quantity, or quality to support a human or ecological risk assessment, and managers and policy-makers must decide if it is useful for their particular needs. Then society must decide whether the monitoring data are relevant. Conversely, attention to data quality objectives will enhance the cost-effectiveness of sampling and focus on relevance and necessity.

24.2.2.3 Tools

A number of tools are available to aid in developing monitoring schemes, including sophisticated toxicity tests, remote sensing, GIS, and spatially explicit simulation models.[25–27] GIS tools are now available not only to show the spatial and temporal patterns of specific indicators, bioindicators, and biomarkers, but also to show relationships between different types of indicators. The tools, however, should not dictate the choice of indicators or biomarkers and, in most cases, suites of indicators will be required.[28]

24.3 BIOINDICATORS

Indicators such as bioindicators and biomarkers are the key components of biomonitoring schemes. Identification of indicators for both exposure and effects is also critical. Since it is not possible to monitor all species, interactions, and functions of ecosystems, the development of bioindicators and biomarkers is critical.[13,15,21,29] Most books devoted to ecological risk assessment provide methods for evaluation at different levels of ecological organization but do not provide a comprehensive plan for any one habitat or land type, although some have provided plans for regions.[30–32] Excellent methods are available,[14] but authors seldom commit to a specific plan.

Monitoring plans should take into account both the value and vulnerability of the ecosystems, as well as the relative susceptibility of these ecosystems.[33,34] Monitoring schemes will be most useful if (1) they include many species representing different trophic levels, (2) indicator selection is based on sound quantitative databases, and (3) caution is used in interpreting population trends, contaminant levels, and other parameters.[13,21,29]

There are some established large-scale monitoring plans, such as the Environmental Monitoring and Assessment Programs (EMAP) of the U.S. Environmental Protection Agency (EPA).[35] The EMAP program distinguishes three types of indicators: (1) response indicators that quantify conditions of the ecosystem, (2) exposure indicators that can be related to direct exposure, and (3) stress indicators that relate to the probable sources of pollution or degradation.[36]

The National Oceanic and Atmospheric Administration's Status and Trends Program provides information on many marine species. They use a few bioindicators to examine marine pollution, providing data on population dynamics and stability.[37]

In the Great Lakes, populations of colonial and fish-eating birds such as herring gulls (*Larus argentatus*) and their eggs are used as bioindicators of the water and environment.[38–40] Levels of PCBs and other contaminants in the Great Lakes were associated with chick abnormalities, parental neglect, and reproductive impairment in some fish-eating birds, leading to population declines.[39,41] Other seabirds and predatory birds have been used elsewhere,[42,43] allowing comparisons among geographical regions.

24.3.1 Characteristics of Indicators

24.3.1.1 Features Essential for Support

Several authors have argued that monitoring plans must be developed in such a way that they have long-term support, or they will not be conducted for a long enough time period to be useful.[31,44,45] This is especially true of long-term stewardship, where the needs for maintenance is projected for hundreds of years. Similarly, indicators must be selected to maximize their biological, methodological, and societal relevance (Table 24.1). To be biologically relevant, an indicator must exhibit changes in response to a stressor but not be so sensitive that changes occur when there is no cause for concern. The response should not be so sensitive that it indicates trivial or biological unimportant variations. The changes must be attributable to a particular stressor and important to the well-being of the organism.[14] Further, the changes being measured should reflect not only impairment to the species itself but to the populations and communities.

While biological relevance is the key feature of a bioindicator, it must also be methodologically relevant,[21] an aspect often ignored in indicator selection. A good indicator should be easy for scientists to measure, for managers to use in their resource management, and for regulators to employ in compliance mandates. Ease of measurement is a key characteristic and includes such aspects as clarity in objectives, ease of identification of important features, and ease of data gathering and analysis.

Societal relevance is also an important attribute of a useful indicator. Without such support it is unlikely that the indicator will be used over a wide enough spatial and temporal scale to provide meaningful information.[46] Society must be willing both to pay for the implementation of a biomonitoring plan (with specific indicators) and to act on the results. Thus, charismatic species such as bald eagles (*Haliaeetus leucocephalus*) or peregrine falcons (*Falco peregrinus*) are often used as indicators.[46] To many they symbolize a wild predator. In the 1950s, population failures of these and many other top-level predators served as indicators of chlorinated hydrocarbon pesticide contamination.

Society is usually interested in the well-being of humans and populations. In ecological systems, well-being often takes the form of being interested in population stability.

TABLE 24.1
Features of Bioindicators for Human and Ecological Health Assessment

Biological Relevance	Provides early warning Exhibits changes in response to stress Changes can be measured Intensity of changes relate to intensity of stressors Change occurs when effect is real Changes are biologically important and occur early enough to prevent catastrophic effects Change can be attributed to a cause
Methodological Relevance	Easy to use in the field Can be used by nonspecialists Easy to analyze and interpret data Measures what it is supposed to measure Useful to test management questions Can be used for hypothesis testing Can be conducted in reasonable time
Societal Relevance	Of interest to the public Easily understood by the public Methods transparent to the public Measures related to human health or ecological integrity Cost-effective Of interest to regulators and public policy makers

Note: These features apply for making decisions before and after remediation and restoration and to evaluate the efficacy and integrity of management, remediation, or other environmental actions.

Source: After Burger, J., *Strat. Environ. Manage*., 1, 351, 1999a[57]; Burger, J. and Gochfeld, M., *Environ. Monit. Assess*., 66, 23, 2001[21]; Carignan, V. and Villard, M.A., *Environ. Monit. Assess*., 78, 45, 2002.[29]

Population stability (assessment endpoint), however, is difficult to measure, and biologists usually measure a characteristic such as the number of individuals (measurement endpoint). The EPA has thus distinguished between assessment endpoints (the societal goal) from measurement endpoints (what can be measured[3,31]). Human and ecological well-being, however, also includes social/economic features, requiring indicators as well (Figure 24.1). To be most useful, ecological and human health indicators should be combined.

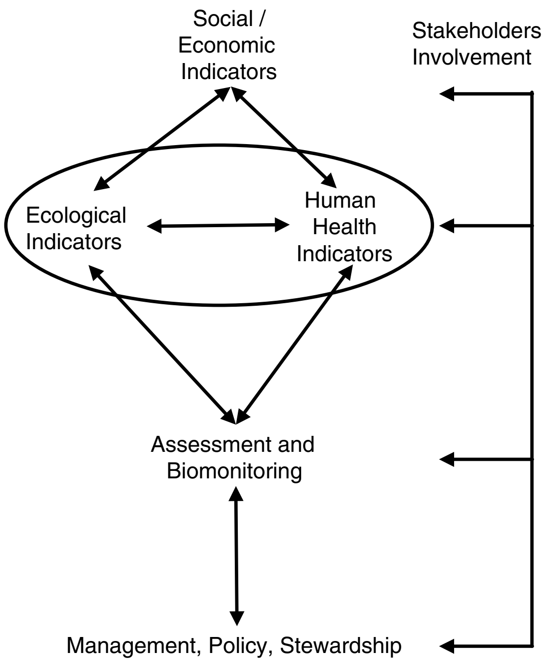


FIGURE 24.1 Schematic showing the relationship between social/economic and human/ecological indicators.

24.3.1.2 Selecting Indicators for Biological, Physical, and Chemical Stressors

The fields of toxicology and biology have largely developed bioindicators for their respective stressors without considering the implications of other stressors. That is, environmental protection agencies have dealt primarily with understanding how levels of contaminants in organisms and their effects have varied over time and space, often in relation to point-source pollution. Environmental conservation programs have examined changes in wildlife and habitats as a function of predator changes, invasive species, weather, and human disturbance. Although not the primary focus of this chapter, it is important to remember that biological, physical, and chemical stressors all affect humans and other receptors, and interpretation of data from monitoring schemes with indicators should consider all stressors.

24.3.1.3 Indicators and Levels of Biological Organization

For many years, and to some extent even today, indicators have been developed for either human or ecological risk assessment. This is certainly a valid approach for many specific problems, but it is less likely to garner public support for a long-enough period to be maximally useful to managers, regulators, and public policy makers. One reason that indicators are often developed separately for human or ecological health is that ecologists are fundamentally interested in each level of biological organization from the species to landscapes.[47] Usually humans are interested in individuals (health of one person) or of populations (public or community health).

We submit, however, that all levels of biological organization (and of indicators) are pertinent to both ecological and human health because all aspects tell us something that is essential for the ecosystems that must support humans and other species. All species live within ecosystems and ecosystems that must provide them with resources for survival and reproduction (Table 24.2).

24.3.2 Selecting Indicators for Human and Ecological Health Assessment

24.3.2.1 Efficacy

We argue that public support will be strongest for monitoring schemes when the indicators tell us something about both ecological as well as human health and well-being. Combining ecological and human health science to arrive at joint indicators provides more useful information to managers, regulators, and policy-makers as well as the public.[48] Since bioindicators normally are single species, it is important to remember that a wide range of other indicators are also necessary to assess community and ecosystem health, such as species richness, productivity, energy and nutrient flows, and biological integrity,[14,49–53] as well as environmental indicators (such as air or water quality). Each of these, however, also tells us something about human health and well-being since people live within, and derive all their resources from, ecosystems (Table 24.2).

Bioindicators can, and should, be developed which are useful for assessing human health, nonhuman health,[21,54] and the health of the ecosystem.[31,55,56] That is, if top-level carnivores are used as bioindicators of human and ecological health,[46] the information gathered as a bioindicator is also a measure of the individual health of the carnivore itself. Thus, with careful selection of bioindicators, managers and risk assessors can optimize how much information they gain for their investment.

24.3.2.2 Single Species as Indicators of Human and Ecological Health

While it is essential to select indicators that reflect all levels of biological organization, it is often the single species which are the most useful for monitoring programs. While caution should be applied in using single species as indicators,[32] they are easily monitored, their population sizes and reproductive rates can be measured, their tissue contaminants can be measured, and their overall health status can be ascertained.[21,57] Information about single species can usually provide information about both current status and future health.[58]

If selected carefully, single species can provide information about their own populations, species that consume them or are eaten by them, their ecosystems, as well as to humans who consume them.

TABLE 24.2
Usefulness of Indicators at Different Biological Levels of Organization to Human and Ecological Health

Ecological Level	Type of Indicator	Ecological Health	Human Health
Individual	Contaminant levels Lesions Disease Tumors Infertility	Used to evaluate health of individuals; for evaluation of risk to higher-level consumers	Used to evaluate health of individual people; for early warning of potential public health issues
Population	Reproductive rates Growth rates Movements Biomass	Used to evaluate health of populations of species, particularly endangered or threatened species	Used to evaluate public health; can be used to compare different regions, ethnicities, ages or other parameters (often referred to as community by public health officials)
Community	Foraging guilds Breeding guilds (groups of related species)	Measures health of species using the same niche, such as colonial birds nesting in a colony or foraging animals such as dolphins and tuna	Ecological indicators can be used for assessing recreational values (such as fishing/bird-watching), as well as consumption rates of fish and other resources
Ecosystem	Species diversity Decomposition rates Erosion rates Primary productivity Energy transfer	Measures changes in relative presence of species, how fast nutrients will become available, how fast nutrients in soil will no longer be available, how much photosynthesis is occurring	Used to assess the health and well-being of human economic endeavors, recreation, ecological services, and aesthetics
Landscape	Relative amounts of different habitats Patch size Corridors between habitat types	Measures dispersion of different habitat types, indicates relative species diversity values	Used for assessment of global services, crowding, diversity of ecosystems

24.3.3 Examples of Indicators of Human and Ecological Health

There are a wide range of species that can be used as indicators of ecological and human health. Below we provide five vertebrate examples to illustrate the ways such species are useful for understanding aspects of human health and of ecological health at several levels of biological organization.

All are fairly common, geographically widespread species which are sufficiently abundant so that using them as bioindicators does not directly affect their populations.

24.3.3.1 Largemouth Bass

Largemouth bass (*Micropterus salmoides*) are ideal as indicators because they are widespread, numerous, and are a popular sport fish.[59–61] Other intermediate-sized predatory fish are also ideal, including bowfin (*Amia calva*), pickerel (*Esox niger*), perch (*Perca flavescens*), and bluefish (*Pomatomus saltatrix*). Largemouth bass, as with a number of other sport fish, are particularly useful because of the risk they pose to higher-level carnivores, such as larger fish, predatory birds, and humans.

Understanding reproductive success, growth rates, survival, and population dynamics of bass in a region can lead to information that can serve as a baseline for establishing trends. Once established, temporal patterns in these variables can be a useful indicator of bass populations, community structure, and fish guilds. Since fishing is such a popular pastime, and bass are a preferred fish in many regions,[59,62,63] indicators of bass population stability lead directly to establishing creel and size limits for fishing.

Bass and other predatory fish are also useful as bioindicators of contamination (Figure 24.2). A wide range of higher-level carnivores eat bass and other intermediate-sized fish, allowing for continued bioaccumulation and magnification up the food chain, particularly for mercury. Mercury levels also increase with size and age of fish.[59,64–68]

Bass as a bioindicator of contaminant exposure have proven useful in South Carolina.[59–61,69] Mercury and selenium levels in the muscle tissue of 11 fish species from the Savannah River were compared along three stretches of the river: upstream, along, and downstream of the Department of Energy's Savannah River Site (SRS), a former nuclear material production facility. We tested the null hypothesis that there were no differences in mercury and selenium levels in fish tissue as a function of species, trophic level, and location along the river. There were significant interspecific differences in mercury levels, with bowfin (*Amia calva*) and bass having the highest levels. As expected, these differences generally reflected trophic levels, indicating the importance of sampling a species that is at a high trophic level if only one species can be examined. Mercury levels were positively correlated with body mass for 9 of the 11 species including bass. The mercury and selenium levels in fish tissue from the Savannah River are similar to, or lower than, those reported in many other studies, and in most cases pose little risk to the fish themselves or to other aquatic consumers, although levels in bowfin and bass are sufficiently high to pose a potential threat to high-level consumers such as herons, eagles, and humans.

It was our intention to use bass as a bioindicator of both ecological and human health. We conducted a risk assessment for people consuming these fish, based on site-specific data on consumption patterns.[70–72] This assessment indicated that even white fishermen at or below the median fish consumption level exceeded the Hazard Index for bass (and bowfin), while black people at the median consumption level exceeded the Hazard Index for 8 of the 11 fish, including bass. A scenario in which the quantities of fish were adjusted for preferences yielded similar results. These assessments

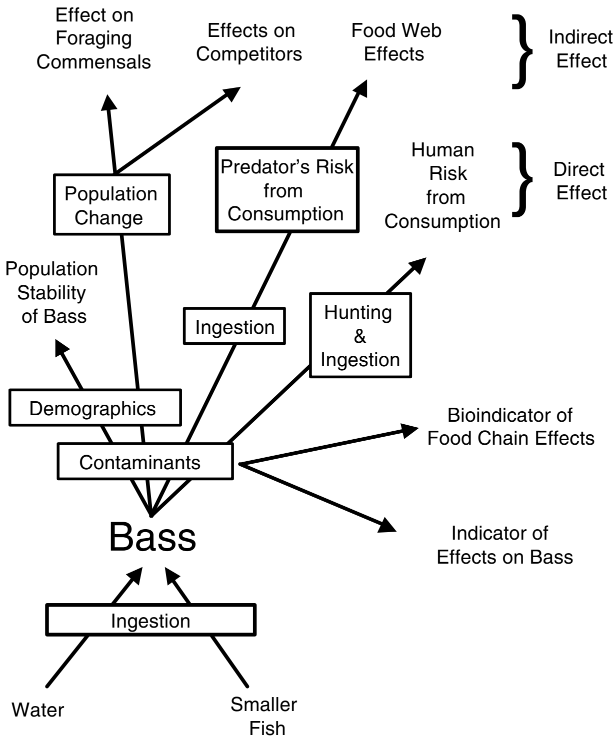


FIGURE 24.2 Schematic of bioindicator properties of bass and other predatory fish. Ingestion exposures of bass are shown on bottom of diagram. Measurement of population, reproduction, growth, human disturbance level, and contaminant levels can all serve as indicators.

suggested that over half of the black fishermen were exceeding the Hazard Index for most fish, and white fishermen who consumed fish at the average consumption level also exceeded the Hazard Index. The conversion for women, based on consumption rates and meal size, was 0.7. Caution must be used in evaluating the Hazard Indices because they incorporate safety factors. Over 35% of the bass sampled exceeded the 0.5 ppm level which most states and countries use in setting safe consumption levels.[60]

Radiocesium levels in the fish, including bass, were also examined.[61] The levels in the fish themselves were not high enough to cause a problem for the fish or to most higher level consumers, except for humans. The lifetime cancer risk for humans was calculated using the cancer slope factor of 3.2×10^{11}/pCi, and site-specific fish consumption from Savannah River fishermen. Using mean ^{137}Cs concentrations and median fish consumption for 70 years for black males, the groups with the highest consumption, the excess lifetime risk associated with the eight species of fish in the Savannah River ranged from 9×10^{7} to 1×10^{5}. The same calculation for fish from Steel Creek on the Department of Energy's SRS gave risk estimates from 1.4 to 8×10^{5} part on radiocesium.[61]

Bass thus meet the biological, methodological, and societal relevance for an indicator. Indeed, fish are particularly useful because they can indicate the well-being of their own populations, other guilds and communities, and overall ecosystem health, as well as being directly relevant to humans, including aspects of health, recreation, industry, and aesthetics.

24.3.3.2 Water Snakes

Water snakes (i.e., *Nerodia sipedon*) are very useful as bioindicators because they are very common aquatic organisms that are top-level predators. The snakes spend a great deal of time in the water, forage in the water, and return to the water for protection from predators.[73,74] They do not move great distances, usually remaining within about 250 m,[75] and thus represent local exposure. Water snakes living in a marshy habitat select dead cattail clumps for basking or the low branches of willow trees within the marsh.[76] They are generally active from April to October, even in the northern part of their range.[73]

They are useful indicators because they feed on fish which are eaten by many other predators, including human fishermen. Information on contaminant levels can be an indicator of exposure of the water snakes themselves and of the well-being of their populations. But contaminant levels in water snakes can also serve as an indicator of exposure of consumers that eat them and of organisms that are at the same trophic level (such as a number of birds and other mammals, and humans, Figure 24.3).

They and other species of snakes have been used as indicators of contaminants,[77,78] and have proven useful in examining contaminants near hazardous waste sites where runoff into surface water is problematic.[79] Because they are so common in such a large geographical region, they will be particularly useful in a comparative sense and to evaluate ecosystems over large regions. Contaminants are currently being examined in streams and rivers with water snakes in New Jersey, Tennessee, and South Carolina.[80,81]

While their use as a bioindicator for environmental pollution is clear, their accuracy as indicators for both ecological risk and human health risk is less clear. Water snakes can be used as an indicator of human disturbance and quality of ecosystems. Burger[82] examined the behavior of water snakes along a walking path adjacent to the Raritan Canal in New Jersey. She examined the hypothesis that disturbance to water snakes was directly related to human pedestrians. Nearly 40% of the variability in the distance to first respond for water snakes (N = 135) was accounted for by the distance the snake was from the path, the number of observers, and the number of people currently using the trail. As the number of pedestrians on the path increased, water snakes responded when people were farther from them. In this case, the snakes served as a bioindicator of human disturbance and density of human use. In addition to providing a measure of well-being for the water snakes and other organisms living in this aquatic environment, this indicator provides information on the quality of the outdoor experience for people using the trail. Increasingly, with more and more urbanization and concentration of people along coastal areas, we will need indicators of the qualities green places provide.

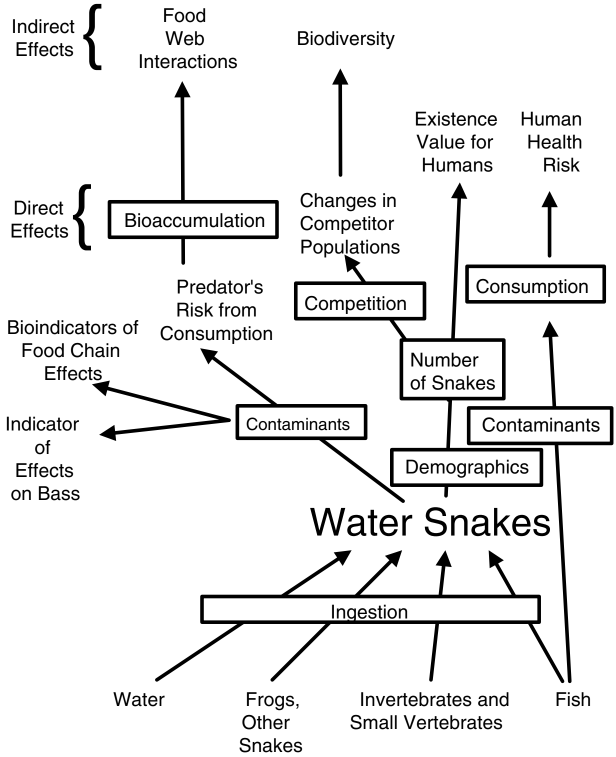


FIGURE 24.3 Schematic of bioindicator properties of water snakes. Ingestion exposures to water snakes are shown on bottom of diagram. Measurement of population, reproduction, growth, human disturbance level, and contaminant levels can all serve as indicators.

Water snakes meet the criteria of biological relevance because they are top-level predators, changes they experience can be measured, and these changes are indicative of stress. They are methodologically relevant because they are common and easy to monitor and catch, and their use can involve clear-cut objectives as well as allowing hypothesis testing. They also meet the criteria of societal relevance because their use is easily understood, it is scientifically defensible, and does not involve undue harm to the snake populations themselves.

24.3.3.3 Mourning Doves

Birds in general are useful bioindicators because they are widespread, often common, and the public is interested in their well-being. Moreover, because they are diurnal, they are very visible, and any harm to large numbers of birds in populated areas is immediately obvious.

Mourning doves are particularly useful bioindicators because they are eaten by a wide variety of organisms, including humans. They are very common throughout the U.S., nest in a large geographical area, occur in a wide variety of habitats, and

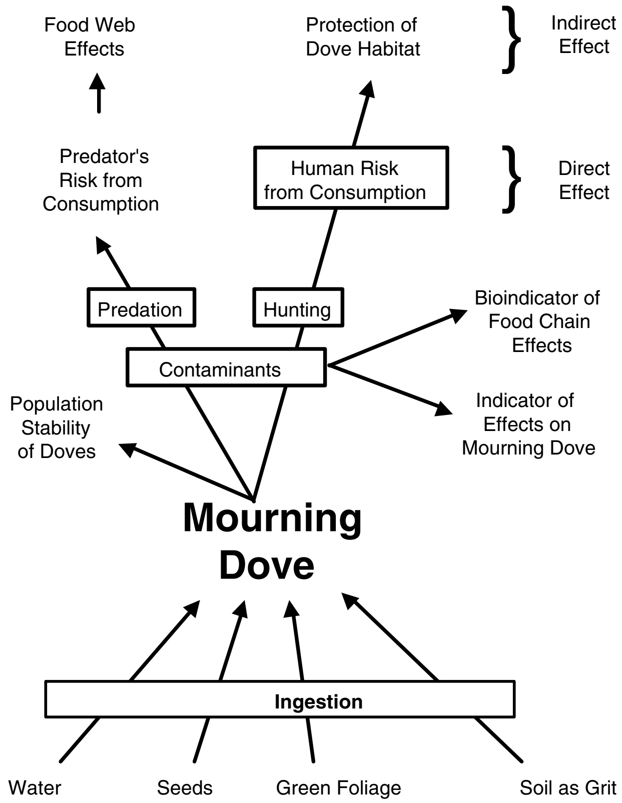


FIGURE 24.4 Schematic of bioindicator properties of mourning doves. Unlike many vertebrate indicators, doves eat soil directly as grit.

serve as prey for a diversity of predators.[83] Surprisingly, they are the most widely hunted of all migratory game birds, including ducks,[83] with nearly 70 million harvested annually.[84] Dove hunting is particularly popular in the southern U.S., where the season can extend for many months.[85]

Mourning doves are relatively low on the food chain because they eat primarily seeds. However, they pick up pollutants in the soil because they ingest soil as grit to aid in digestion.[83] They and children have the tendency to eat dirt and from many of the same playgrounds and open fields. Information on contaminant levels in dove tissues can be an indicator of exposure to toxins or the health of the dove populations, as well as of the potential well-being of organisms that consume them (such as hawks, predatory mammals, and human hunters, Figure 24.4).

We used data on heavy metals and radiocesium in doves from South Carolina to illustrate how mourning doves can serve as a bioindicator of both ecological and human health.[54,85,86] The work was conducted in South Carolina at and adjacent to the Department of Energy's SRS where radiocesium and metal contamination is of interest because of potential ecosystem and human health risks. Heavy metals and radiocesium were analyzed in different tissues because they are useful as indicators of different components of the ecosystem. Contaminant levels in muscle, however, were useful as bioindicators of exposure to other organisms that might eat them, including humans; liver levels were useful because they indicated potential damage to the doves themselves.

Adverse health effects to the doves can be assessed by comparing the toxic levels in their tissues with those known to cause ill effects in controlled laboratory studies, including closely related species.[87] The mean levels in the liver of mourning doves averaged 0.2 ppm in Par Pond (at the SRS), and 0.5 ppm from the town of Jackson (outside the Department of Energy site). Thus, it is clear that the lead levels in the mourning doves were below the level known to cause sublethal reproductive and behavioral effects in birds. Lead levels of 5 ppm or more in the liver are considered toxic.[88]

Dove muscle tissue was used to assess potential risk for organisms that consumed the doves because it is the tissue that is usually eaten by predators, scavengers, and humans. Thus, doves are useful as indicators of food-web effects, as well as of surface-soil contamination because they eat soil as grit. While mercury was not a concern because the levels in all tissues were nondetectable, lead was a concern because of the relatively high levels in the doves collected at Jackson, a public hunting field.[54]

Small children comprise the most sensitive subgroup consuming dove meat containing this lead. The potential risk to children was calculated using the EPA's Integrated Exposure Uptake/Biokinetic Model for lead.[89] We assumed that they consume the muscle from 4 doves/d for the 76-d hunting season (based on the dove hunting season at Jackson and consumption information from local dove hunters). Although most children would not eat so many doves, this is a reasonable exposure scenario for a family in which a parent was an avid dove hunter and which regularly consumed dove meat. We used the EPA model and the Centers for Disease Control "level of concern" (10 μg/dl) to calculate whether the consumption of dove meat would adversely impact children. The percent of children with blood lead exceeding 10 μg/dl would rise from 0.3% predicted with no dove consumption to 1.1% for 2- to 3-year olds and from 0.1% to 0.5% for 4- to 5-year-olds who ate dove meat readily when available. For the exposed subpopulation of children, the risk of having an elevated lead level, though not great, is increased 3- to 5-fold.

As might be expected because the SRS belongs to the Department of Energy, the risk to humans from radiocesium was higher for doves from Par Pond on the SRS, because the levels of radiocesium were an order of magnitude higher than at Jackson. While we conducted a standard risk assessment, we report here the number of doves that can be eaten without exceeding the one in a million increased cancer risk. We found that a hunter could eat only 152 doves from Par Pond without exceeding the one in a million increased cancer risk, while the same hunter could eat 3800 doves from Jackson. Clearly no individual eats 3800 doves during the hunting season. Using this method, easily understood by teenage dove hunters, it is possible to see that the risk is much higher from consuming Par Pond doves, compared to those obtained from the public hunting grounds. Adult males might regularly eat 10 or more at one time, based on local consumption rates. If so, then an adult male hunter could only eat doves from Par Pond for 15 d, well under the number of legal hunting days.

These data from South Carolina illustrate three main points: (1) contaminants data can be used to assess the direct risk to human and other consumers from consumption, (2) contaminants data can be used to assess risk to the doves themselves, and (3) risk information can (and should) be presented in a way that is easily and graphically understood.

Many other aspects of the biology of mourning doves can be used as bioindicators, including total numbers (population size). However, population size is of limited value since doves are mobile. Dove harvest (doves shot/season) may be a better indicator of populations, particularly on a nationwide scale. Given the enormous hunting effort, obtaining a sample of doves from hunters or trapping them is feasible, which makes obtaining a large sample for contaminant assessment possible.

Mourning doves meet the criteria of biological relevance (even though they are not top-level predators), methodological relevance (feasible to sample and analyze), and societal relevance (a familiar garden bird which is the most widely hunted gamebird in North America).

24.3.3.4 Herring Gulls/Other Seabirds

Seabirds are particularly useful as bioindicators of environmental pollution because they are very long-lived (some live as long as humans), are often top-level carnivores, and are often sufficiently common so that their populations are not impacted by their use.[90] Herring gulls, their feathers, and their eggs have been monitored in the Great Lakes,[38,40] along the Atlantic coast[91], and in Europe[92] as bioindicators of environmental contamination (Figure 24.5).

Feathers are a useful bioindicator because they reflect circulating levels at the time the feathers were formed, they usually reflect local exposure, metals are sequestered in feathers as an excretion method, and metal levels in feathers are correlated with internal tissue levels. The largest database for comparing heavy metal levels in birds is for feathers[93] because collecting feathers is a noninvasive procedure that can be applied to a wide range of species, including endangered species. For most species, plucked breast feathers grow back in 2 to 3 weeks.[94] The methodology for determining potentially harmful effects of given contaminant levels is to compare the levels of metals in the feathers of the species of interest to (1) levels in other species and (2) to levels known to cause harmful effects in laboratory studies.[90,95] This procedure can be applied to a wide range of contaminants (see Figure 24.6). There are limitations since the proportion of a dose that ends up in the feathers differs among contaminants and with the time course of exposure (acute vs. chronic).

In addition to using tissues and population levels as indicators of population, community and ecosystem health, herring gulls and other seabirds can be used as indicators of exposure of other higher trophic levels, such as humans, because seabirds often eat the same fish that people eat.[90]

24.3.3.5 Raccoons

Raccoons are useful bioindicators because they are common and widespread throughout the U.S., occupy a variety of habitats from rural to urban, are omnivores

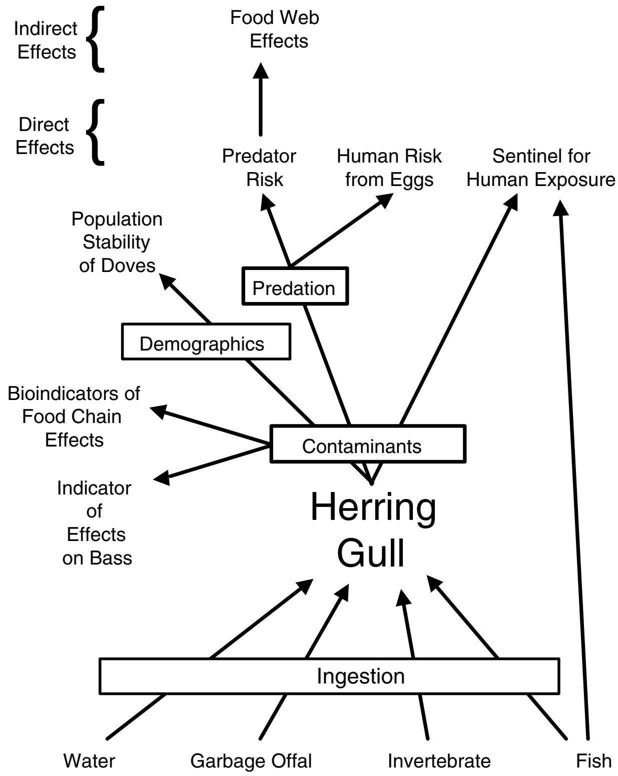


FIGURE 24.5 Schematic of herring gulls, showing not only species and ecosystem effects but also their value as sentinels.

and eat some organisms that are high on the trophic scale, are relatively sedentary, and are hunted and eaten in some parts of the U.S.[96] Raccoon hunting is a popular sport in the south, and raccoons are eaten by people or their pets.[97,98]

Raccoons obtain contaminants from their water and food which includes fruits, nuts, seeds, vegetable crops, invertebrates, small vertebrates (frogs, snakes), fish, and anything else they can find including garbage.[99] They in turn are eaten by larger predators, including humans. Their decaying carcasses are eaten by scavengers, and microbes return the rest of the contaminants and nutrients to the ecosystem (Figure 24.7). Understanding contaminant levels in their tissues can provide information on the health of raccoon populations themselves, on their predators (including humans), on other organisms that occupy the same trophic level, and on the rest of the ecosystem through the ultimate decomposition of their carcasses.

Methods to assess the risk from raccoons are similar to those employed for doves and water snakes: (1) levels in raccoon tissues are compared to levels known to cause lethality or sublethal effects in mammals, (2) levels in raccoon tissue are evaluated for potential food chain effects on their predators, (3) raccoon muscle tissue levels are assessed for potential effects to human consumers (and to companion pets), and (4)

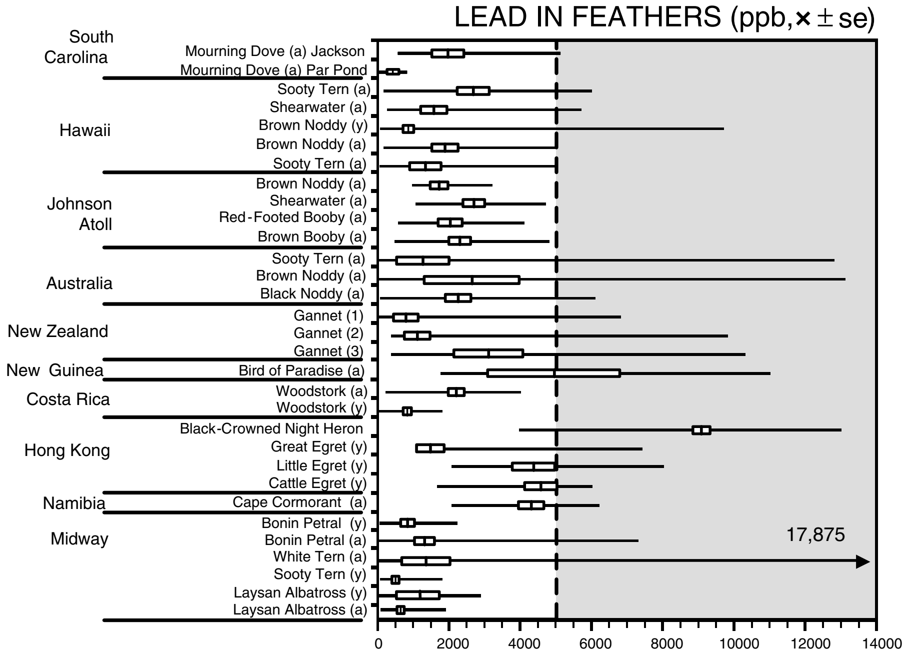


FIGURE 24.6 Lead levels in a range of species, showing both the mean values for many different studies (after Burger, J., *Rev. Environ. Toxicol.*, 5, 203, 1993)[92] and the adverse effects level from laboratory studies (after Burger, J. and Gochfeld M., *J. Toxicol. Environ. Health,* Part B, 3, 59, 2000).[90] Levels from new areas of interest can be compared to these values to ascertain whether there are ill effects.

tissue levels can provide insight into levels in other organisms of the same trophic level. Since raccoons are higher on the food chain than doves, they provide information on a different trophic level, an important aspect of any overall biomonitoring scheme.

The utility of using raccoons as indicators for their populations, other species and communities, ecosystems, and for humans was examined at the SRS in South Carolina. Raccoons collected off the site had significantly lower levels of mercury and selenium in both the liver and kidney than those on the SRS.[69] Further, some of the variation in contaminant body burdens was attributable to trophic feeding position, providing a way of obtaining information on several trophic levels.[100]

The mercury levels in the raccoons from hunting grounds were very low, indicating that people hunting there were not at risk, since human consumption is generally low.[101] However, a number of raccoons collected on SRS had levels above the U.S. Fish and Wildlife Service standard of 1.1 ppm, suggesting that they would pose a problem to human hunters (if there were hunting on site) and to other organisms that might eat the raccoons.[101]

Raccoons proved useful as indicators for humans and other receptors with respect to radiocesium as well.[100] There were significant differences among levels, with

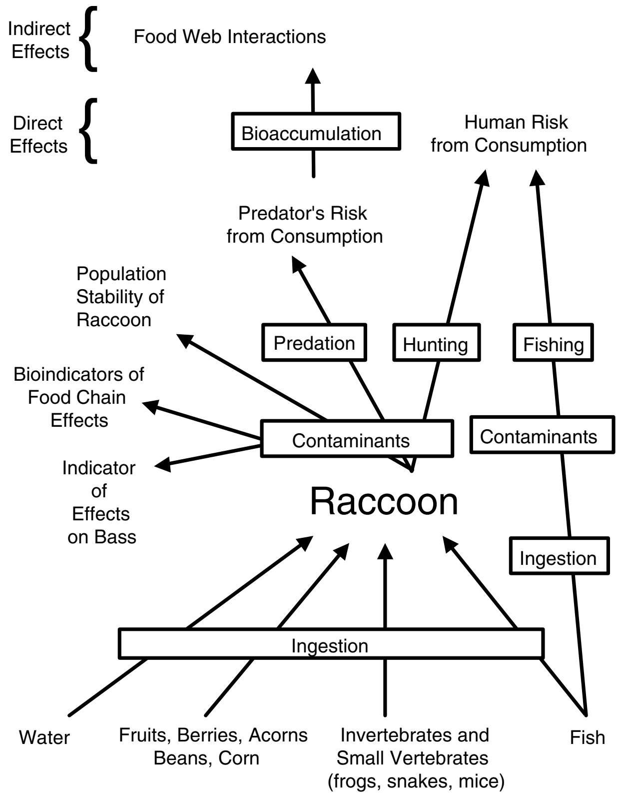


FIGURE 24.7 Schematic of bioindicator properties of raccoons, which have a highly varied diet and are very widespread in North America.

those from contaminated sites being higher than those from reference sites. Levels were sufficiently low that they provided no risk to themselves or other animals that consumed them. Only one of the raccoons exceeded the European Economic Community limit for radiocesium in edible muscle (0.6 Bq, EEC,[102]).

Raccoons meet the criteria of biological relevance (they eat a wide range of organisms and in turn are eaten by many organisms), methodological relevance (they are widespread, common, relatively easy to trap), and societal relevance (they are familiar and are hunted and consumed by a significant proportion of the population in the southeastern U.S.). Moreover, people find them charismatic and would respond negatively if large numbers turned up dead).

24.4 CONCLUSIONS

Developing biomonitoring plans is essential for a number of societal and public health concerns, including remediation, restoration, wildlife and habitat management, and general well-being of a diversity of receptors. Biomonitoring plans require a number of indicators and biomarkers which address different levels of biological organization from the individual to ecosystems and landscapes. While traditional risk assessors have either concentrated on ecological or human health endpoints, it is becoming increasingly clear that bioindicators that address both human health issues and ecological health issues will garner the most support from the public, and will be more likely to be continued.

Single species are most useful as indicators of human and ecological health because they can be used as indicators of (1) the health of the species itself, (2) the species that prey upon the species, (3) the food web that the species is part of, (4) the community, and (5) humans that depend upon the species (or its predators, prey, or competitors) for food, fiber, other resources, or for aesthetic or existence values.

A number of examples of bioindicators, ranging from fish and snakes to birds and mammals, were given to illustrate the varied ways that single species can be used as bioindicators of ecosystem and human health. Endpoints include biomarkers, sentinels, and indicators of behavioral disruption. While a number of community and ecosystem indicators are available (and extremely important to scientists understanding ecosystem structure and function), they are generally harder to use, more difficult to understand, and less appealing to the general public. Ecologists and other resource managers should consider developing suites of single-species indicators that can serve a dual role in providing information about ecosystems structure and function as well as human health.

ACKNOWLEDGMENTS

Many people contributed to various aspects of this research and we thank them. C. S. Boring, C. Jeitner, K. F. Gaines, R. A. Kennamer, C. Lord, M. McMahon, T. Benson, C. Safina, T. Shukla, S. Shukla, and I. L. Brisbin, Jr. worked on some of the original studies upon which this chapter is based. B. D. Goldstein, E. Faustman, J. Moore, and C. Powers provided valuable comments on the research or manuscript. The research reported herein was conducted under Rutgers protocol 97-017 and 86-016, and was funded by the Consortium for Risk Evaluation with Stakeholder Participation (CRESP) through the Department of Energy cooperative agreement (AI # DE-F-G-26-00NT-40938), by NIEHS (ESO 5022), and the Environmental and Occupational Health Sciences Institute. The views expressed in this chapter are solely the responsibility of the authors and do not represent those of the funding agency.

REFERENCES

1. National Research Council (NRC), *Risk Assessment in the Federal Government, Managing the Process*, National Academies Press, Washington, D.C., 1983.
2. National Research Council, *Issues in Risk Assessment*., National Academies Press, Washington, D.C., 1993.
3. Norton, S.B., Rodier, D.R., Gentile, J.H., van der Schalie, W.H., Wood, W.P., and Slimak, M.W., A framework for ecological risk assessment at the EPA, *Environ. Toxicol. Chem.*, 11, 1663, 1992.

4. Rand, G.M. and Zeeman, M.G., Ecological risk assessment: approaches within the regulatory framework, *Human Ecol. Risk Assess*., 4, 853, 1998.
5. Bailar, J.C., III, Needleman, J., Berney, B.L., and McGinnis, J.M., *Assessing Risks to Health*, Auburn House, Westport, CT, 1993.
6. National Research Council (NRC), *Building Consensus through Risk Assessment and Management of the Department of Energy's Environmental Remediation Program,* National Academies Press, Washington, D.C., 1994.
7. Mendelsohn, M.L., Peters, J.P., and Normandy, M.J., *Biomarkers and Occupational Health*, Joseph Henry Press, Washington, D.C., 1995.
8. Sheehan, P.J., Miller, D.R., Butler, G.C., and Bourdeau, P. (Eds.), *Effects of Pollutants at the Ecosystem Level*, John Wiley & Sons, Chichester, U.K., 1984.
9. National Research Council, *Ecological Knowledge and Environmental Problem Solving*, National Academies Press, Washington, D.C., 1986.
10. Bartell, S.M., Gardner, R.H., and O'Neill, R.V., *Ecological Risk Estimation*, Lewis Publishers, Boca Raton, FL, 1992.
11. Cairns, J., Jr., Niederlehner, B.R., and Orvos, D.R., *Predicting Ecosystem Risk*, Princeton Scientific Publishing, Princeton, NJ, 1992.
12. Suter, G.W., II (Ed.), *Ecological Risk Assessment*, Lewis Publishers, Boca Raton, FL, 1993.
13. Peakall, D., *Animal Biomarkers as Pollution Indicators,* Chapman & Hall, London, U.K., 1992.
14. Linthurst, R.A., Bourdeau, P., and Tardiff, R.G., *Methods to Assess the Effects of Chemicals on Ecosystems*, John Wiley & Sons, Chichester, U.K., 1995.
15. Piotrowski, J.K., Individual exposure and biological monitoring, in *Methods for Estimating Risk of Chemical Injury: Human and Non-human Biota and Ecosystems,* Vouk, V.B., Butler, G.C., Hoel, D.G., and Peakall, D.B., Eds., John Wiley & Sons, Chichester, U.K., 1985, pp. 123–135.
16. DiGiulio, R.T. and Monosson, E., *Interconnections between Human and Ecosystem Health*, Chapman & Hall, London, U.K., 1996.
17. Forman, R.T.T. and Godron, M., *Landscape Ecology*, John Wiley & Sons, New York, 1986.
18. Burger, J. and Gochfeld, M., Ecological and human health risk assessment: a comparison, in *Interconnections between Human and Ecosystem Health,* DiGuilio, R.T. and Monosson, E., Eds., Chapman & Hall, London, U.K., 1996, pp 127–148.
19. Suter, G.W., II, Integration of human health and ecological risk assessment, *Environ. Health Perspect*., 105, 1282, 1997.
20. Harvey, T., Mahaffrey, K.R., Vasquez, S., and Dourson, M., Holistic risk assessment: an emerging process for environmental decisions, *Regul. Toxicol. Pharmacol*., 22, 110, 1995.
21. Burger, J. and Gochfeld, M., On developing bioindicators for human and ecological health, *Environ. Monit. Assess*., 66, 23, 2001.
22. Munns, W.R., Jr. and MacPhail, R., Extrapolation in human health and ecological risk assessments: proceedings of a symposium, *Human Ecol. Risk Assess*., 8, 1, 2002.
23. President's Commission, Presidential/Congressional Commission on Risk Assessment and Risk Management, U.S. Government Printing Office, Washington, D.C., 1997.
24. Environmental Protection Agency, *Guidance for the Data Quality Objective Process*, EPA/600/R-96/055, Washington, D.C., 2000, http://www.epa.gov/quality/qa_docs.html.
25. Cairns, J., Jr. and Niederlehner, B.R., Genesis and future needs, in *Predicting Ecosystem Risk*, Cairns, J., Jr., Niederlehner, B.R., and Orvos, D.R., Eds., Princeton Scientific Publishing, Princeton, NJ, 1992, pp. 327–344.

26. Cairns, J., Jr. and Niederlehner, B.R., Developing a field of landscape ecotoxicology, *Ecol. Appl*., 6, 780, 1996.
27. Aspinall, R. and Pearson, D., Integrated geographical assessment of environmental contamination in watch catchments: linking landscape ecology, environmental modeling and GIS, *J. Environ. Manage.*, 59, 299, 2000.
28. Harwell, M.A. and Kelly, J.R., Indicators of ecosystem recovery, *Environ. Manage.,* 14, 527, 1990.
29. Carignan, V. and Villard, M.A., Selecting indicator species to monitor ecological integrity: a review, *Environ. Monit. Assess.*, 78, 45, 2002.
30. Hunsaker, C., Carpenter, D., and Messer, J., Ecological indicators for regional monitoring, *Bull. Ecol. Soc. Am.*, 71, 165, 1990.
31. Suter, G.W., II, Endpoints for regional ecological risk assessment, *Environ. Manage.,* 14, 9, 1990.
32. Cairns, J., Jr., The genesis of biomonitoring in aquatic ecosystems, *Environ. Profe*., 12, 169, 1990.
33. Burger, J., Method for and approaches to evaluating susceptibility of ecological systems to hazardous chemicals, *Environ. Health Perspect.*, 105, 843, 1997.
34. Burger, J. and Gochfeld, M., On developing bioindicators for human and ecological health, *Environ. Monit. Assess*., 66, 23, 2001.
35. Summers, K., Robertson, A., and Johnston, J., Monitoring the Condition of Estuarine Shallow Water Habitats, Marine and Estuarine Shallow Water Science and Management Conference, 42 pp., 1995.
36. Messer, J.J., Linthurst, R.A., and Overton, W.S., An EPA program for monitoring ecological status and trends, *Environ. Monit. Assess*., 17, 67, 1991.
37. O'Connor, T.P. and Ehler, C.N., Results from the NOAA National Status and Trends Program on distribution and effects of chemical contamination in the coastal and estuarine United States, *Environ. Monit. Assess.,* 17, 33, 1991.
38. Fox, G.A., Gilman, A.P., Peakall, D.B., and Anderka, F.W., Behavioral abnormalities of nesting Lake Ontario herring gulls, *J. Wildl. Manage*., 42, 477, 1978.
39. Fox, G.A., Gilbertson, M., Gilman, A.P., and Kubiak, T.J., A rationale for the use of colonial fish-eating birds to monitor the presence of developmental toxicants in Great Lakes fish, *J. Great Lakes Res*., 17, 151, 1991.
40. Peakall, D.B. and Fox, G.A., Toxicological investigations of pollutant-related effects in Great Lakes gulls, *Environ. Health Perspect*., 71, 187, 1987.
41. Gilbertson, M., Kubiak, T., Ludwig, J., and Fox, G., Great Lakes embryo mortality, edema and deformities syndrome (CLEMEDS) in colonial fish-eating birds: similarity to chick edema disease, *J. Toxicol. Environ. Health*, 33, 455, 1991.
42. Newton, I., Long-term monitoring of organochlorine and mercury residues in some predatory birds in Britain, *Acta 20 Congr. Int. Congr*., 20, 2487, 1991.
43. Thompson, D.R., Furness, R.W., and Monteiro, L.R., Seabirds as biomonitors of mercury inputs to epipelagic and mesopelagic marine food chains, *Sci. Total Environ*., 213, 299, 1998.
44. Stout, B.B., The good, the bad and the ugly of monitoring programs: defining questions and establishing objectives, *Environ. Monit. Assess*., 26, 91, 1993.
45. O'Connor, J.S. and Dewling, R.T., Indices of marine degradation: their utility, *Environ. Manage*., 10, 335, 1986.
46. Fox, G. (Ed.), *Bioindicators as a Measure of Success for Virtual Elimination of Persistent Toxic Substances,* International Joint Commission, Hull, Quebec, Canada, 1994.
47. Holl, K.D. and Cairns, J., Jr., Landscape indicators in ecotoxicology, in *Handbook of Ecotoxicology*, Hoffman, D.J., Rattner, B.A., Burton, G.A., Jr., and Cairns, J., Jr., Eds., CRC Press, Boca Raton, FL, 1995, pp. 185–197.

48. Meffe, G.K. and Viederman, S., Combining science and policy in conservation biology, *Wild. Soc. Bull*., 23, 327, 1995.
49. Karr, J.R., Yant, P.R., and Fausch, K.D., Spatial and temporal variability in the Index of Biotic Integrity in three Midwestern streams, *Trans. Am. Fish. Sci*., 116, 1, 1987.
50. Karr, J.R., Measuring biological integrity, in *Principles of Conservation and Biology,* 2nd ed., Meffe, G.K. and Carroll, C.R., Eds., Sinauer Assoc., Sunderland, MA, 1997.
51. Karr, J.R. and Chu, E.W., Biological monitoring: essential foundation for ecological risk assessment, *Human Ecol. Risk Assess*., 3, 993, 1997.
52. Karr, J.E. and Chu, E.W., *Restoring Life in Running Waters: Better Biological Monitoring*, Island Press, Washington, D.C., 1999.
53. Kimberling, D.N., Karr, J.R., and Fore, L.S., Measuring human disturbance using terrestrial invertebrates in the shrub-steppe of eastern Washington (USA), *Ecol. Indicators*, 1, 63, 2001.
54. Burger, J., Kennamer, R.A., Brisbin, I.L., and Gochfeld, M., Metal levels in mourning doves from South Carolina: Potential hazards to doves and hunters, *Environ. Res*., 75, 173, 1997.
55. Slocombe, D.S., Environmental monitoring for protected areas: review and prospect, *Environ. Monit. Assess*., 21, 49, 1992.
56. Wilson, J.G., The role of bioindicators in estuarine management, *Estuaries*, 17, 94, 1994.
57. Burger, J., Environmental monitoring on Department of Energy lands: the need for a holistic plan, *Strat. Environ. Manage*., 1, 351, 1999a.
58. Davis, G.E., Design elements of monitoring programs: the necessary ingredients for success, *Environ. Monit. Assess*., 26, 99, 1993.
59. Burger, J., Gaines, K.F., Boring, S., Stephens, W.L., Jr., Snodgrass, J., and Gochfeld, M., Mercury and selenium in fish from the Savannah River: species, trophic level, and locational differences, *Environ. Res*., 87, 108, 2001.
60. Burger, J., Gaines, K.F., and Gochfeld, M., Ethnic differences in risk from mercury among Savannah River fishermen, *Risk Anal*., 21, 533, 2001.
61. Burger, J., Gaines, K.F., Stephens, W.L., Jr., Boring, C.S., Brisbin, I.L., Jr., Snodgrass, J., Peles, J., Bryan, L., Smith, M.H., and Gochfeld, M., Radiocesium in fish from the Savannah River and Steel Creek: potential food chain exposure to the public, *Risk Anal*., 21, 545, 2001.
62. Fleming, L.E., Watkins, S., Kaderman, R., Levin, B., Ayyar, D.R., Bizzio, M., Stephens, D., and Bean, J.A., Mercury exposure in humans through food consumption from the Everglades of Florida, *Water Air Soil Pollut*., 80, 41, 1995.
63. Campbell, K.R., Dickey, R.J., Sexton, R., and Burger, J., Fishing along the Clinch River arm of Watts Bar Reservoir adjacent to the Oak Ridge Reservation, Tennessee: behavior, knowledge, and risk perception, *Sci. Total Environ*., 288, 145, 2002.
64. Phillips, G.R., Lenhart, T.E., and Gregory, R.W., Relations between trophic position and mercury accumulation among fishes from the Tongue River Reservoir, Montana, *Environ. Res*., 22, 73, 1980.
65. Lange, T.R., Royals, H.E., and Connor, L.L., Mercury accumulation in largemouth bass (*Micropterus salmoides*) in a Florida lake, *Arch. Environ. Contam. Toxicol*., 27, 466, 1994.
66. Denton, G.R.W. and Burdon-Jones, C., Trace metals in fish from the Great Barrier Reef, *Mar. Poll. Bull*., 17, 210, 1996.
67. Bidone, E.D., Castilhos, Z.C., Santos, T.J.S., Souza, T.M.C., and Lacerda, L.D., Fish contamination and human exposure to mercury in Tartarugalzinho River, Northern Amazon, Brazil: A screening approach, *Water Air Soil Pollut*., 97, 9, 1997.

68. Peterson, S.A., Herlihy, A.T., Hughes, R.M., Motter, K.L., and Robbins, J.M., Level and extent of mercury contamination in Oregon, USA, lotic fish, *Environ. Toxicol. Chem*., 21, 2157, 2002.
69. Burger, J., Gaines, K.F., Lord, C., Shukla, S., and Gochfeld, M., Metal levels in raccoon tissues: differences on and off the Department of Energy's Savannah River Site in South Carolina, *Environ. Monit. Assess*., 74, 67, 2002.
70. Burger, J., Recreation and risk: potential exposure, *J. Toxicol. Environ. Health*, 52, 269, 1997.
71. Burger, J., Risk, a comparison of on-site hunters, sportsmen, and the general public about recreational rates and future land use preferences for the Savannah River Site, *J. Environ. Plan. Manage*., 43, 221, 2000.
72. Burger, J., Stephens, W., Boring, C.S., Kuklinski, M., Gibbons, J.W., and Gochfeld, M., Factors in exposure assessment: ethnic and socioeconomic differences in fishing and consumption of fish caught along the Savannah River, *Risk Anal*., 19, 427, 1999.
73. King, R.B., Population ecology of the Lake Erie water snake (*Nerodia sipedon insularum), Copeia*, 757, 1986.
74. King, R.B., Microgeographic, historical, and size-correlated variation water snake diet composition, *J. Herpetol*., 27, 90, 1993.
75. Mills, M.S., Hudson, C.J., and Berna, H.J., Spatial ecology and movements of the brown water snake (*Nerodia taxispilota*), *J. Herpetol.,* 51, 412, 1995.
76. Weatherhead, P.J. and Robertson, I.C., Thermal constraints on swimming performance and escape response of northern water snakes (*Nerodia sipedon*), *Can. J. Zool*., 70, 94, 1992.
77. Bauerle, B., Spencer, D.L., and Wheeler, W., The use of snakes as a pollution indicator species, *Copeia*, 1875, 366, 1975.
78. Burger, J., Trace element levels in pine snake hatchlings: tissue and temporal differences, *Arch. Environ. Contam. Toxicol*., 22, 208, 1992.
79. Fontenot, L.W., Nobelt, G.P., Akins, J.M., Stephens, M.D., and Cobb, G.P., Bioaccumulation of polychlorinated biphenyls in ranid frogs and northern water snakes from a hazardous waste site and a contaminated watershed, *Chemosphere*, 40, 803, 2000.
80. Burger, J., Jeitner, C., Jensen, H., Fitzgerald, M., Carlucci, S., Shukla, S., Ramos, R., and Gochfeld, M., Habitat use in basking Northern water (*Nerodia sipedon*) and Eastern garter (*Thamnophis sirtalis*) snakes in New Jersey, *Copeia*, submitted ms.
81. Murray, S., Gaines, K., and Burger, J., Using water snakes as bioindicators of radionuclide contamination, unpublished ms., CRESP, 2003.
82. Burger, J., The behavioral response of basking Northern water (*Nerodia sipedon*) and Eastern garter (*Thamnophis sirtalis*) snakes to pedestrians in a New Jersey park, *Urban Ecosyst*., in press.
83. Mirarchi, R.E. and Baskett, T.S., Mourning Dove (*Zenaida macroura*), in *Birds of North America,* Poole, A. and Gill, F., Eds., American Ornithologists' Union, Philadelphia, PA. 1994.
84. Baskett, B.W. and Sayre, M.W., Characteristics and importance, in *Ecology and Management of the Mourning Dove,* Baskett, T.S., Sayre, M.W., Tomlinson, R.E., and Mirarchi, R.E., Eds., Stackpole Books, Harrisburg, PA, 1993, pp. 1–6.
85. Kennamer, R.A., Brisbin, I.L., Jr., McCreedy, C.D., and Burger, J., Radiocesium in mourning doves: effects of a contaminated reservoir drawdown and risk to human consumers, *J. Wildl. Manage*., 62, 497, 1998.
86. Burger, J., Kennamer, R.A., Brisbin, I.L., and Gochfeld, M., A risk assessment for consuming doves, *Risk Anal*., 18, 563, 1998.

87. Burger, J. and Gochfeld, M., Lead and neurobehavioral development in gulls: a model for understanding effects in the laboratory and the field, *NeuroToxicology*, 18, 279, 1997.
88. Ohlendorf, H., Marine birds and trace elements in the temperate North Pacific, in *The Status, Ecology, and Conservation of Marine Birds of the North Pacific*, Canadian Wildlife Service Special Publication, Ottawa, 1993.
89. Environmental Protection Agency, Uptake/biokinetic model for lead (Version 0.99d), U.S. Environmental Protection Agency, Washington, D.C., 1994.
90. Burger, J. and Gochfeld, M., Effects of chemicals and pollution on seabirds, in *Biology of Marine Birds,* Schreiber, E.A. and Burger, J., Eds., CRC Press, Boca Raton, 2001, pp. 485–525.
91. Burger, J., Heavy metals and selenium in herring gull (*Larus argentatus*) nesting in colonies from eastern Long Island to Virginia, *Environ. Monit. Assess*., 48, 285, 1997.
92. Becker, P.H., Thyen, S., Mickstein, S., Sommer, U., and Schmieder, K.R., Monitoring pollutants in coastal bird eggs in the Wadden Sea, *Wadden Sea Ecosyst. Rep*., 8, 59, 1998.
93. Burger, J., Metals in avian feathers: bioindicators of environmental pollution, *Rev. Environ. Toxicol*., 5, 203, 1993.
94. Burger, J., Nisbet, I.C.T., and Gochfeld, M., Metal levels in regrown feathers: assessment of contamination on the wintering and breeding grounds in the same individuals, *J. Toxicol. Environ. Health*, 37, 363, 1992.
95. Burger, J. and Gochfeld M., Effects of lead on birds (Laridae): a review of laboratory and field studies, *J. Toxicol. Environ. Health,* Part B, 3, 59, 2000.
96. Burger, J., Sanchez, J., Gibbons, J.W., Benson, T., Ondrof, J., Ramos, R., McMahon, M.J., Gaines, K., Lord, C., Fulmer, M., and Gochfeld, M., Attitudes and perceptions about ecological resources and hazards of people living around the Savannah River Site, *Environ. Monit. Assess.*, 57, 195, 1999.
97. South Carolina Department of Natural Resources, 1995–6 commercial fur harvest summary, *Furbearer Res. Bull.,* (Fall) 1996, 1, 7.
98. South Carolina Department of Natural Resources, The impact of sport raccoon hunting on deer movement and deer hunting success, *Furbearer Res. Bull*., (Fall) 1996, 2–4.
99. Burger, J., *Animals in Towns and Cities,* Kendall-Hunt Publishing, Dubuque, IA, 1999.
100. Gaines, K.F., Romanek, C.S., Boring, S., Lord, C.G., Gochfeld, M., and Burger, J., Using raccoons as an indicator species for metal accumulation across trophic levels: a stable isotope approach, *J. Wildl. Manage*., 66, 811, 2002.
101. Lord, C.G., Gaines, K.F., Boring, C.S., Brisbin, I.L., Gochfeld, M., and Burger, J., Raccoon (*Procyon lotor*) as a bioindicator of mercury contamination at the U.S. Department of Energy's Savannah River Site, *Arch. Environ. Contam. Toxicol*., 43, 356, 2002.
102. European Economic Community, Derived Reference Levels as a Basis for the Control of Foodstuffs Following a Nuclear Accident: A Recommendation from the Group of Experts Set Up under Article 31 of the Euratom Treaty, European Economic Community Regulation 170/86, Brussels, Belgium, 1986.

25 Biological Indicators in Environmental Monitoring Programs: Can We Increase Their Effectiveness?

V. Carignan and M.-A. Villard

CONTENTS

25.1 INTRODUCTION

Human activities have gradually altered the natural environment of North America since the colonization of the land. At the time, natural disturbance regimes created a dynamic mosaic of successional stages throughout the landscape (the shifting mosaic hypothesis[1]) to which species had to adapt. Contemporary land use, on the

1-56670-641-6/04/$0.00+$1.50

other hand, reflects an entirely different situation in which human actions are the dominant structuring elements in most landscapes and natural disturbance regimes often have much less influence than they had prior to human settlement.[2] Changes in the rate and extent of disturbances brought about by human activities affect ecological integrity (*sensu* Karr and Dudley[3]) to the point that many species which were adapted to historical disturbance regimes are now becoming threatened or endangered.

In the hope of curbing a potential biodiversity crisis, many agencies are allocating considerable resources to the monitoring of environmental change and its effects on the native flora and fauna. For these agencies and many other organizations, biological indicators possess an undeniable appeal as they provide a time- and cost-efficient alternative to assess the impacts of environmental disturbances on the resources of concern. However, the actual sensitivity of various indicators to environmental change has yet to be demonstrated and their uncritical use fosters the risk of underestimating the complexity of natural systems.[4]

Despite the abundant criticism on the use of biological indicators,[5–9] natural resources managers and researchers are likely to continue using them until better approaches are proposed. Consequently, it is crucial that their conceptual and operational limitations be clearly identified and accounted for, so as to guide their use in environmental monitoring. Therefore, this chapter aims to review the basic steps in the development of a management or monitoring program incorporating the use of biological indicators. A particular emphasis will be placed on the selection of an appropriate set of biological indicators.

25.2 DEVELOPING A COMPREHENSIVE ENVIRONMENTAL MONITORING PROGRAM

The development of an environmental monitoring program essentially follows a series of steps which progressively increase the knowledge of the condition of the ecosystem as well as of the means to reduce the stress on specific components. These steps are identified below and detailed further in the following sections:

1. Biodiversity assessment: How do the current and the pristine state of the ecosystem compare? Is there evidence for ecosystem degradation? If so, which ecosystem components have been affected/degraded by environmental changes?
2. Formulation of precise, goal-oriented, management objectives: What is the desired state of the ecosystem?
3. Selection of relevant biological indicators: What species, structures, or processes can provide surrogate measures of the state of the ecosystem?
4. Selection of parameters to measure the status of the selected biological indicators (e.g., abundance, biomass, reproductive success)
5. Implementation of conservation actions to mitigate disturbances. What management actions can be taken to bring the ecosystem to the desired state?

25.2.1 Biodiversity Assessment

When assessing the state of biodiversity in a region, one must keep in mind that ecosystems are dynamic and, consequently, ecological integrity exists as many possible combinations of structural and compositional variables. This implies that ecosystems do not exhibit a unique undisturbed state, i.e., climax that can be maintained indefinitely. Rather, they exhibit a suite of conditions over all space and time, and the processes that generate these dynamics should be maintained.[10]

The assessment of the state of biodiversity for a given region requires us to determine the current state of biodiversity relative to agreed-upon reference conditions. Defining the reference conditions for regional ecosystems (i.e., ranges in ecosystem parameters that would be observed in the absence of anthropogenic effects) is an essential step in environmental monitoring programs because the results will serve as a benchmark to assess current and future conditions, and this may also help to formulate management goals. Ray[11] defined a reference condition as "the biodiversity resulting from the interactions between the biota, the physical environment, and the natural disturbance regime in the absence of the impact of modern technological society." However, there are substantial difficulties in establishing appropriate reference conditions since one must not only have adequate data on those conditions but also information on their range of variation.[12,13] Unfortunately, such data are often lacking[14] and, if they can be found, it is usually only for short periods, making it difficult to determine whether current dynamics actually fall within the natural range of variation (e.g., vegetation succession). Consequently, natural ranges of variation are virtually unknown for many ecosystems[15] and, in many studies, perceived optimal conditions often serve as a substitute. The most important criterion for the use of optimal present-day conditions as a reference would be that the site has been held in a state of minimal human impact for sufficient time to justify the assumption that its current state does represent natural or, at the very least, sustainable conditions.[16]

At this step, it is important to have some knowledge of the biological indicators likely to be used in the monitoring program so that the data necessary to interpret trends in these indicators are collected. For example, if a bird species assemblage associated with mature forests is used as an indicator of forest stand condition in the landscape, reference conditions on the size distribution, composition, and structure of stands may be the only data that need to be collected. In subsequent steps, managers might investigate why differences arose between current and reference stand conditions (e.g., alterations to disturbance regimes) and recommend appropriate management actions.

25.2.2 Formulation of Management Objectives

Before implementing a monitoring program, managers must have clear objectives about the desired state of biodiversity.[17,18] For example, one might perceive the presence of a certain number of breeding pairs of bald eagles (an indicator species for the state of prey stocks such as fish) as desirable. More generally, however, the main objective of ecosystem managers should be to maintain or restore the natural state and dynamics of the ecosystem, which may include[19,20]:

1. The maintenance or restoration of viable populations of all native species in natural patterns of abundance and distribution
2. The maintenance of key geomorphological, hydrological, ecological, biological and evolutionary processes within normal ranges of variation
3. The encouragement of land uses that are compatible with the maintenance of ecological integrity and discouragement of those that are not

In the real world, however, the socioeconomic and political context often influences the degree to which such objectives can be reached.[21,22] For example, the reintroduction of large predators in some protected areas cannot be done without considering the potential interactions between these predators and livestock present in nearby ranches.

25.2.3 Selection of Relevant Indicators

Brooks et al.[23] defined indicators as "measures, variables, or indices that represent or mimic either the structure or function of ecological processes and systems across a disturbance gradient." Indicators can reflect biological, chemical, and physical aspects of the ecosystem, and have been used or proposed to characterize ecosystem status, track, or predict change, and influence management actions.[24] They can also be used to diagnose the cause of an environmental problem[25] or to quantify the magnitude of stress on ecosystems.[26]

Indicators were originally used in studies describing species–habitat associations[27] as well as in crop production (e.g., indicators of soil fertility[28]). More recently, they have been proposed (1) as surrogates for the measurement of water, air, or soil quality to verify the compliance of industries to particular antipollution laws,[29] (2) for the assessment of habitat quality,[30,31] and (3) to detect the effects of management activities on certain species.[32,33] Additionally, indicators have frequently been incorporated into policies and regulations[34,35] and used to monitor the degree of ecological integrity in aquatic[36,37] and terrestrial[38] ecosystems.

Because managers cannot possibly measure all potentially relevant indicators in an ecosystem, the choice of what to measure is critical. In general, indicators must capture the complexity of the ecosystem yet remain simple enough to be monitored relatively easily over the long term. A set of indicators should possess some or all of the following qualities (expanded from Reference 9):

1. Provide early warning signs, i.e., indicate an impending change in key characteristics of the ecosystem.[39]
2. Provide continuous assessment over a wide range and intensity of stresses.[40] This allows the detection of numerous impacts to the resource of concern and also means that an indicator will not bottom out or level off at certain thresholds.[39,41]
3. Have a high specificity in response. This may be critical to establish causal relationships and, hence, appropriate management decisions.[5,40]

4. Be cost-effective to measure, i.e., amenable to simple protocols applicable even by nonspecialists.[42,43]
5. Be easily communicated to nonscientists and decision makers. The Common Language Indicator Method developed by Schiller et al.[44] is particularly interesting in this respect. These authors found that nonscientists better understand information on contamination of forest plants by air pollution than specific information on individual measures (e.g., foliar chemistry and lichen chemistry).

There are three broad categories of indicators: biological (e.g., species, populations, communities), structural (e.g., stand structure and patch configuration in the landscape), and process-based (e.g., frequency and intensity of fires or flooding events). In this chapter, however, we restrict our discussion to biological indicators because they tend to be the primary tool on which we rely to make management recommendations.

25.2.3.1 Biological Indicators

A rich terminology has been developed to describe the various roles played by different types of biological indicators.[45–48] Indicators may (1) act as surrogates for larger functional groups of species, (2) reflect key environmental variables, or (3) provide early warning signs of an anticipated stressor (e.g., forest birds as indicators of the progression of maple dieback in Quebec,[49] or plants[50] and soil properties[51] as indicators of trampling effects). The capacity of biological indicators to fulfill such roles has, however, received much criticism and warrants further discussion.

25.2.3.1.1 Criticisms about the Use of Biological Indicators

Species-based approaches have been criticized on the grounds that they do not provide whole-landscape solutions to conservation problems, that they cannot be applied at a rate sufficient to address the urgency of the threats, and that they consume a disproportionate amount of conservation funding.[52–54] Furthermore, Schiller et al.[44] argued that "because the act of selecting and measuring indicators involves a human cognitive and cultural action of observing the environment in a particular way under certain premises and preferences, indicator information implicitly reflects the values of those who develop and select them." These flaws have been confirmed recently by Andelman and Fagan.[6] They found that biological indicator schemes did not perform substantially better than randomly selected sets of a comparable number of species, thus refuting the claim that umbrella, flagship, and other types of biodiversity indicator schemes had any special utility as conservation surrogates for the protection of regional biota. These results are not surprising because, even from purely theoretical considerations, the indicator species approach to maintaining populations of all vertebrate species cannot be expected to work well. First, paleoecological evidence is inconsistent with the notion of persistent associations among species at any scale[55]; "species may simply live in the same places because they coincidentally share a need for a similar range of physical conditions, rather than because of complex, coevolved, interactions."[56] Second, because no two species occupy the same niche, it is unlikely that there will be a complete overlap in the broad distribution of indicator species and the full suite of taxa for which they are supposed to be indicative.[5,57–60] Indeed, Prendergast et al.[61] found that only 12% of hotspots in bird species richness coincided with those of butterflies. Similar results (10 to 11%) were obtained by Lawton et al.[62] using birds, butterflies, flying beetles, canopy beetles, canopy ants, leaf-litter ants, termites, or soil nematodes. As suggested by Pärt and Söderström,[63] differences in the occurrence of different taxa in a particular region could be related to the environmental characteristics to which each taxon responds (e.g., birds may react more strongly to landscape context than do plants).

A majority of studies report a lack of spatial coincidence in diversity hotspots (birds, ants, and plants[64]; butterflies and plants[17]; birds and plants[63,65]; butterflies and moths[66]). Positive correlations in the species richness of taxa occupying the same area have been found in butterflies and plants[67]; tiger beetles, butterflies, and birds[68]; and birds and butterflies.[69] Considering the differences in the ecological requirements of the species or taxa examined, these results are not too surprising. One might expect that it would be more useful to select indicator species from groups of species with similar resource use (i.e., a guild indicator species) than to expect a given taxon to indicate several different groups. Unfortunately, even in such cases there is little assurance that habitat suitability or population status of one species will parallel those of other species in the guild.[5,7,70,71] Although species in a guild exploit the same type of resources, they do not necessarily respond the same way to other habitat characteristics.[72,73] Furthermore, the life history of many species is often partly or completely unknown, and this adds to the uncertainty of species reaction to environmental changes and to the difficulty of extrapolating from one species to another. Thus, the patterns in response to ecosystem change exhibited by different species within the same guild may not be readily predictable, even among groups of closely related taxa (forest birds[57,74,75]; arboreal marsupials[76]). Declines in populations of one member of a guild could therefore be hidden by a general increase in the populations of others.[59] Finally, Jaksic[77] showed, using an assemblage of raptor species as an example, that both the composition and number of guilds may change through time following resource depletion (e.g., fewer guilds when prey diversity is low). Correspondingly, the observable guild structure of communities or assemblages may not reflect organizing forces such as competition; rather, it may simply represent a group of species responding opportunistically to changing resource levels.[77] Therefore, guilds may merely be a tool helping managers to determine which habitat factors are important in management decisions by providing insight into general changes in resource availability or other structuring elements and processes that may affect specific guilds; they may not have further predictive value.[74]

Other difficulties are associated with the indicator's ability to detect responses to disturbance or to show sensitivity to specific disturbance types. First, reaction time depends on the assemblages targeted for study, taxa with short generation times reacting more quickly than those with longer generations.[78] However, smaller organisms may also adapt more rapidly to changes,[79] making them less sensitive and, thus, less useful as indicators. Second, species may be affected by factors unrelated to the integrity of the focal ecosystem and exhibit population fluctuations that are not seen in sympatric species (e.g., disease, parasites, competition, predation, conditions in

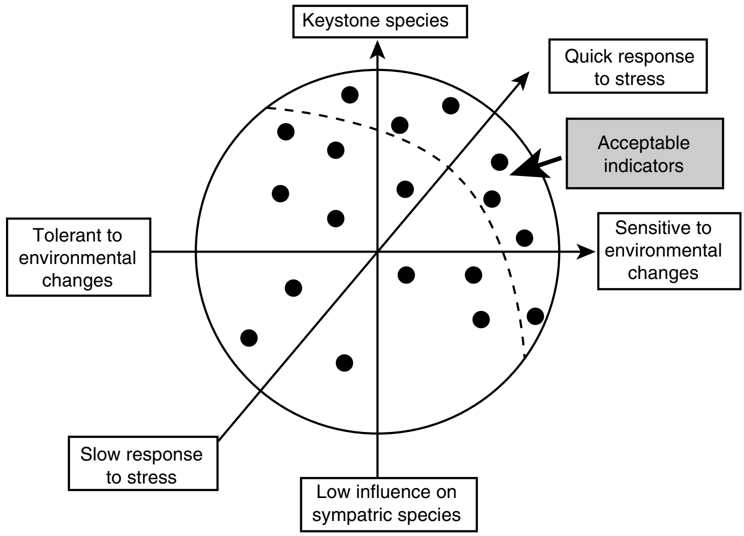


FIGURE 25.1 Schematic representation of the decision process involved in the selection of biological indicators, shown here as dots.

other areas for migratory species, and stochastic variations[80]). For these reasons, it can be inappropriate to consider the occurrence and abundance of indicator species as an indication of integrity without concurrent knowledge on the state of other elements within the ecosystem.[81]

25.2.3.1.2 Minimizing the Disadvantages of Biological Indicators

In the preceding section, we reviewed many of the flaws attributed to biological indicators both at the conceptual and operational level. However, we believe that these flaws do not discredit the use of biological indicators but rather that they emphasize the importance of exercising caution when selecting indicators for monitoring purposes. To assist managers and researchers in the selection of appropriate and representative sets of biological indicators, we suggest using three criteria (Figure 25.1):

1. The species should ideally have a strong influence on sympatric species.
2. The species should have been shown to be sensitive to environmental changes. This criterion will tend to favor the selection of ecological specialists and, therefore, species that may provide early-warning signs of disturbances. By definition, these species tend to occupy less frequent habitat types and, thus, smaller habitat patches.
3. The species should quickly respond to a given stress. This allows us to apply management actions without delay to mitigate the sources of disturbance. This criterion will tend to favor the selection of smaller organisms with shorter generation times (e.g., invertebrates), which may benefit from more local conservation actions (e.g., soil rehabilitation).

The set of biological indicators selected according to these three criteria should be sensitive to disturbances taking place over different spatial and temporal scales. Therefore, it should provide a useful mean to monitor the evolution of the state of the ecosystem along every step of the management program.

25.2.3.2 Pros and Cons of Different Taxa as Biological Indicators

Many taxa have been examined as potential indicators of biodiversity (see Reference 9). Invertebrates in general have been shown to be sensitive and accurate indicators of ecosystem integrity, presumably because environmental factors (moisture gradient, soil density, and altitude[82]) play a greater role in shaping species assemblages than biological relationships such as competition, predation and parasitism.[83] However, Davies and Margules[84] warned against generalizing the reactions of one invertebrate taxon to others since there are still considerable gaps in taxonomic knowledge and because, from what is known, they show markedly different responses to habitat alterations. Karr[36] also argued that invertebrates may not be the best indicators because they require a high degree of taxonomic expertise, and they are difficult and time-consuming to sample, sort, and identify. In addition to these problems, invertebrates seem to mainly react to environmental changes over fine spatial scales and, hence, may be inadequate indicators for organisms reacting to changes over larger scales. On the other hand, larger organisms may, in the same way, represent poor umbrellas for species mainly reacting to fine-scale disturbances. The low correspondence among indicators reacting to changes over different spatial or temporal scales reflects differences in their rates of population increase, generation times, and habitat specificity.[85] Consequently, both small and large organisms are, by themselves, inadequate indicators. Environmental monitoring programs should thus consider them together or in conjunction with other taxa.

Birds may offer a compromise and provide a good indication of the status of certain components of ecosystems since they have been shown to respond to environmental changes over several spatial scales.[86–88] Bird species occupying higher trophic levels (carnivores, piscivores, etc.) may also prove to be good biodiversity indicators since they are closely associated with the state of the food web on which they rely (e.g., great white heron and fish supply[30]). However, using birds as indicators carries certain disadvantages, mainly because they are highly mobile. Thus, they may be less reliable indicators of local conditions because populations can be affected by habitat changes elsewhere within their home range, in the surrounding landscape, or in other parts of their range.[86]

Thus, each taxon has its advantages and limitations and using only one or a few indicator taxa to monitor ecological integrity could provide a distorted picture.[5] Consequently, many authors[89–91] advocate the use of a greater taxonomical variety of biological indicators. However, as pointed out by Simberloff,[60] one must be careful not to consider too many indicator species as this would defeat the original purpose, i.e., reduce the amount of data that need to be collected to monitor ecological integrity.

25.2.3.3 Choosing the Appropriate Parameters to Monitor Biological Indicators

Once managers have selected potential biological indicators, they have to identify appropriate parameters to monitor their response to environmental change. Parameters such as density, abundance, or species richness are often used in environmental monitoring programs. However, many authors suggest that these metrics are, by themselves, inadequate predictors of population persistence[92] because abundance at a given site does not necessarily reflect biotic and abiotic characteristics.[93,94] Furthermore, abundance varies as a function of numerous factors, many of which may operate entirely independent of habitat conditions at a particular site.[80,95,96] Thus, natural fluctuations in abundance can be difficult to distinguish from those associated with human activities.[97]

With regard to summary statistics such as species richness which combines presence/absence of species with distinct life histories, Conroy and Noon[92] concluded that they are "unlikely to be useful, may be misleading and, at a minimum, are highly scale-dependent." Diversity indices overlook many important variables and thus oversimplify exceedingly complex systems.[36] They may also mask important changes among assemblages, such as the gain of exotic species.[98,99]

Reproductive success may be a better index for predicting the persistence of species than mere presence or abundance because (1) secondary population parameters (e.g., abundance) may show time lags in their response to habitat alterations[100] whereas primary parameters such as reproductive success respond immediately and (2) primary parameters are more representative of variations in resource availability or interspecific interactions than secondary ones. However, reproductive success is notoriously time-consuming to quantify in the field, at least directly, which is why alternative methods have been proposed for monitoring purposes at least in the case of songbirds.[95,101–104] However, these methods either require further validation or are not very cost-effective over large spatial scales.

25.2.4 Study Design Considerations

A monitoring program should include a clear definition of the experimental units and sample populations to ensure sufficient replication to allow statistical testing[29] and the consideration of how the data will be analyzed so as to optimize statistical power.[34,105] It is critical to consider the relative risks of committing type I and type II errors when designing a monitoring program. The key problem managers face in detecting significant trends is that the sources of noise are quite difficult to separate from deterministic changes.[87,106] Even when they are far removed from human activities, ecosystems show a high degree of variability over different temporal and spatial scales in their species composition, structure, and function. Population size, for example, tends to be very noisy even when there is no net long-term trend.[107] Furthermore, time lags in population response to habitat degradation suggest that by the time a decline is detected, it may be too late to take necessary management actions.[108] In this context, it may be appropriate to relax the alpha level to 0.10 or even to 0.20 (instead of the usual 0.05) since it is generally preferable to spend extra efforts investigating a few false reports of change than to have waited for a definitive result of change, at which time it may be too late to react, and fewer management options may exist.[106,109]

25.3 CONCLUSION

The vast body of literature concerning biological indicators that has been published in the last decade or so features a debate between proponents and opponents of their use in environmental monitoring programs. It is in our opinion that much of the criticism concerning the potential limitations and constraints of biological indicators does not preclude their use, but rather points out the need for more stringent selection criteria and a more cautious interpretation of their response to environmental change. Managers and researchers now recognize the importance of (1) selecting a wider variety of biological indicators based on a solid quantitative approach using data from the focal region and (2) incorporating them within a comprehensive monitoring program that pays attention to the interpretation of their response in the face of a myriad of potential causal and confounding factors. Furthermore, a consensus has emerged on the need to monitor biological indicators over multiple spatial scales.[25,47,110,111] Although we have not included indicators at higher levels of organization (e.g., landscape structure indices, ecosystem processes) in this chapter, we consider them to be complementary to biological indicators. A reduction in the proportion of forest cover in the landscape could partly explain population declines observed in an indicator species.

Management actions based on the interpretation of monitoring data represent the final step in an environmental monitoring program. Biological indicators themselves can then be used to determine the success of such actions. Researchers and managers will have to work together on a continuous basis to ensure that such actions are taken at the right time, that these actions are based on the best possible information available, that their outcome is carefully monitored, and that appropriate corrections are made if necessary. This is the basis of active adaptive management, and we hope that our institutions will allow this process to take place on a much larger scale than it does currently. The future of our ecosystems depends on this continuous learning process.

REFERENCES

1. Bormann, F.H. and Likens, G.E., *Pattern and Process in a Forested Ecosystem*, Springer-Verlag, New York, 1979.
2. Fuller, J.L. et al., Impact of human activity on regional forest composition and dynamics in central New England, *Ecosystems***,** 1, 76–95, 1988.
3. Karr, J.R. and Dudley, D.R., Ecological perspective on water quality goals, *Environ. Manage.*, 5, 55–68, 1981.
4. De Leo, G.A. and Levin, S., The multifaceted aspects of ecosystem integrity, *Conserv. Ecol.* (online), 1, 1–22 (http://www.consecol.org/vol1/iss1/art3), 1997.
5. Landres, P.B., Verner, J., and Thomas, J.W., Ecological use of vertebrate indicator species: a critique, *Conserv. Biol.* 2, 316–328, 1988.

6. Andelman, S.J. and Fagan, W.F., Umbrellas and flagships: efficient conservation surrogates or expensive mistakes? *Proc. Natl. Acad. Sci. U.S.A.*, 97, 5954–5959, 2000.
7. Lindenmayer, D.B., Margules, C.R., and Botkin, D.B., Indicators of biodiversity for ecologically sustainable forest management, *Conserv. Biol.*, 14, 941–950, 2000.
8. Lindenmayer, D.B. et al., The focal-species approach and landscape restoration: a critique, *Conserv. Biol.*, 16, 338–345, 2002.
9. Carignan, V. and Villard, M.-A., Selecting indicator species to monitor ecological integrity: a review, *Environ. Monit. Assess.*, 78, 45–61, 2002.
10. Christensen, N.L. et al., The report of the Ecological Society of America committee on the scientific basis for ecosystem management, *Ecol. Appl.*, 6, 665–691, 1996.
11. Ray, G.C., Ecological diversity in coastal zones and oceans, in *Biodiversity*, Wilson, E.O., Ed., National Academy Press, Washington, D.C., 1988, pp. 36–50.
12. Kelly, J.R. and Harwell, M.A., Indicators of ecosystem recovery, *Environ. Manage.*, 14, 527–545, 1990.
13. Landres, P.B., Morgan, P., and Swanson, F.J., Overview of the use of natural variability concepts in managing ecological systems, *Ecol. Appl.*, 9, 1179–1188, 1999.
14. Miller, J.A., Biosciences and ecological integrity, *BioScience*, 41, 206–210, 1991.
15. Davis, G.E., Design elements of monitoring programs: the necessary ingredients for success, *Environ. Monit. Assess.*, 26, 99–105, 1993.
16. Andreasen, J.K. et al., Considerations for the development of a terrestrial index of ecological integrity, *Ecol. Indicators*, 1, 21–36, 2001.
17. Kremen, C., Assessing the indicator properties of species assemblages for natural areas monitoring, *Ecol. Appl.*, 2, 203–217, 1992.
18. MacDonald, L.H., Developing a monitoring project, *J. Soil Water Conserv.*, 49, 221–227, 1994.
19. Noss, R.F., The wildlands project: land conservation strategy, *Wild Earth* (special issue), 10–25, 1992.
20. Grumbine, R.E., What is ecosystem management? *Conserv. Biol.*, 8, 27–38, 1994.
21. Regier, H.A., The notion of natural and cultural integrity, in *Ecological Integrity and the Management of Ecosystems*, Woodley, S., Kay, J., and Francis, G., Eds., St. Lucie Press, Delray Beach, FL, 1993, pp. 3–18.
22. Brunner, R.D. and Clark, T.W., A practice-based approach to ecosystem management, *Conserv. Biol.*, 11, 48–58, 1997.
23. Brooks, R.P. et al., Towards a regional index of biological integrity: the example of forested riparian ecosystems, *Environ. Monit. Assess.*, 51, 131–143, 1998.
24. Kurtz, J.C., Jackson, L.E., and Fisher, W.S., Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development, *Ecol. Indicators*, 1, 49–60, 2001.
25. Dale, V.H. and Beyeler, S.C., Challenges in the development and use of indicators, *Ecol. Indicators*, 1, 3–10, 2001.
26. Hunsaker, C.T. and Carpenter, D.E., *Environmental Monitoring and Assessment Program: Ecological Indicators*, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, NC, 1990.
27. Hall, H.M. and Grinnell, J., Life-zone indicators in California, *Proc. California Acad. Sci.*, 9, 37–67, 1919.
28. Shantz, H.L., Natural Vegetation as an Indicator of the Capabilities of Land for Crop Production in the Great Plains Area, USDA Bulletin Bureau of Plant Industry 210, 1911, 91 pp.
29. MacDonald, L.H. and Smart, A., Beyond the guidelines: Practical lessons for monitoring, *Environ. Monit. Assess.*, 26, 203–218, 1993.

30. Powell, G.V.N. and Powell, A.H., Reproduction by great white herons *Ardea herodias* in Florida Bay as an indicator of habitat quality, *Biol. Conserv.*, 36, 101–113, 1986.
31. Canterbury, G.E. et al., Bird communities and habitat as ecological indicators of forest condition in regional monitoring, *Conserv. Biol.*, 14, 544–558, 2000.
32. Roberts, T.H. and O'Neil, L.J., Species selection for habitat assessments, *T. North Am. Wildl. Nat. Res. Conf.*, 50, 352–362, 1985.
33. Suter, G.W., II, Endpoints for regional ecological risk assessment, *Environ. Monit. Assess.*, 14, 9–23, 1990.
34. McKenney, D.W. et al., Workshop results, in *Towards a Set of Biodiversity Indicators for Canadian Forests: Proceedings of a Forest Biodiversity Indicators Workshop*, McKenney, D.W., Sims, R.A., Soulé, M.E., Mackey, B.G., and Campbell, K.L., Eds., Natural Resources Canada, Sault Ste.-Marie, Ontario, 1994, pp. 1–22.
35. Parks Canada Agency, Protecting Ecological Integrity with Canada's National Parks. Vol. II: Setting a New Direction for Canada's National Parks. Report of the Panel on the Ecological Integrity of Canada's National Parks, Minister of Public Works and Government Services, Ottawa, Canada, 2000.
36. Karr, J.R., Assessment of biotic integrity using fish communities, *Fisheries*, 6, 21–27, 1981.
37. Harig, A.L. and Bain, M.B., Defining and restoring biological integrity in wilderness lakes, *Ecol. Appl.*, 8, 71–87, 1998.
38. Bradford, D.F. et al., Bird species assemblages as indicators of biological integrity in Great Basin rangeland, *Environ. Monit. Assess.*, 49, 1–22, 1998.
39. Noss, R.F., Indicators for monitoring biodiversity: A hierarchical approach, *Conserv. Biol.*, 4, 355–264, 1990a.
40. O'Connell, T.J., Jackson, L.E., and Brooks, R.P., A bird community index of biotic integrity for the Mid-Atlantic highlands, *Environ. Monit. Assess.*, 51, 145–156, 1998.
41. Woodley, S., Monitoring and measuring ecosystem integrity in Canadian National Parks, in *Ecological Integrity and the Management of Ecosystems*, Woodley, S., Kay, J., and Francis, G., Eds., St. Lucie Press, Delray Beach, FL, 1993, pp. 155–176.
42. Davis, G.E., Design of a long-term ecological monitoring program for Channel Islands National Park, CA, *Nat. Area. J.*, 9, 80–89, 1989.
43. di Castri, F., Vernhes, J.R., and Younès, T., Inventoring and monitoring biodiversity: a proposal for an international network, *Biol. Int.*, 27, 1–27, 1992.
44. Schiller, A. et al., Communicating ecological indicators to decision makers and the public, *Conserv. Ecol.* (online), 1–17, http://www.consecol.org/journal/vol5/iss1/art19, 2001.
45. Lambeck, R.J., Focal species: a multi-species umbrella for nature conservation, *Conserv. Biol.*, 11, 849–856, 1997.
46. Caro, T.M. and O'Doherty, G., On the use of surrogate species in conservation biology, *Conserv. Biol.*, 13, 805–814, 1999.
47. Noss, R.F., Assessing and monitoring forest biodiversity: a suggested framework and indicators, *For. Ecol. Manage.*, 115, 135–146, 1999.
48. Armstrong, D., Focal and surrogate species: getting the language right, *Conserv. Biol.*, 16, 285–286, 2002.
49. DesGranges, J.-L., Mauffette, Y., and Gagnon, G., Sugar maple forest decline and implications for forest insects and birds, *T. North Am. Wildl. Nat. Res. Conf.*, 52, 677–689, 1987.
50. Liddle, M.J., A selective review of the ecological effects of human trampling on natural ecosytems, *Biol. Conserv.*, 7, 17–36, 1975.
51. Brown, J.H., Jr., Kalisz, S.P., and Wright, W.R., Effects of recreational use on forested sites, *Environ. Geol.*, 1, 425–431, 1977.

52. Franklin, J.F., Preserving biodiversity: species, ecosystems, or landscapes? *Ecol. Appl.*, 3, 202–205, 1993.
53. Hobbs, G.J., Jr., Ecological integrity, new Western myth: a critique of the Long's Peak report, *Environ. Law*, 24, 157–169, 1994.
54. Walker, B., Conserving biological diversity through ecosystem resilience, *Conserv. Biol.*, 9, 747–752, 1995.
55. Hunter, M.L., Jacobson, G.J., Jr., and Webb, T., III, Paleoecology and the coarse-filter approach to maintaining biological diversity, *Conserv. Biol.*, 2, 375–385, 1988.
56. Gleason, H.A., The individualistic concept of the plant association, *Bull. Torrey Bot. Club*, 53, 7–26, 1926.
57. Mannan, R.W., Morrison, M.L., and Meslow, E.C., The use of guilds in forest bird management, *Wildl. Soc. Bull.*, 12, 426–430, 1984.
58. Niemi, G.J. et al., A critical analysis on the use of indicator species in management, *J. Wild. Manage.,* 61, 1240–1252, 1997.
59. Hutto, R.L., Using landbirds as an indicator species group, in *Avian Conservation: Research and Management*, Marzluff, J.M. and Sallabanks, R., Eds., Island Press, Washington, D.C., 1998, pp. 75–91.
60. Simberloff, D., Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? *Biol. Conserv.*, 83, 247–257, 1998.
61. Prendergast, J.R. et al., Rare species, the coincidence of diversity hotspots and conservation strategies, *Nature*, 365, 335–337, 1993.
62. Lawton, J.H. et al., Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest, *Nature*, 391, 72–76, 1998.
63. Pärt, T. and Söderström, B., Conservation value of semi-natural pastures in Sweden: contrasting botanical and avian measures, *Conserv. Biol.*, 13, 755–765, 1999a.
64. Järvinen, O. and Vaïsänen, R.A., Quantitative biogeography of Finnish land birds as compared with regionality in other taxa, *Ann. Zool. Fenn.*, 17, 67–85, 1980.
65. Pärt, T. and Söderström, B., The effects of management regimes and location in landscape on the conservation of farmland birds in semi-natural pastures, *Biol. Conserv.*, 90, 113–123, 1999b.
66. Ricketts, T.H., Daily, G.C., and Ehrlich, P.R., Does butterfly diversity predict moth diversity? Testing a popular indicator taxon at local scales, *Biol. Conserv.*, 103, 361–370, 2002.
67. Thomas, C.D. and Mallorie, H.C., Rarity, species richness, and conservation: butterflies of the Atlas Mountains in Morocco, *Biol. Conserv.*, 33, 95–117, 1985.
68. Pearson, D.L. and Cassola, F., World-wide species richness patterns of tiger beetles (Coleoptera: Cicindelidae): indicator taxon for biodiversity and conservation studies, *Conserv. Biol.*, 6, 376–391, 1992.
69. Blair, R.B., Birds and butterflies along an urban gradient: Surrogate taxa for assessing biodiversity? *Ecol. Appl.*, 9, 164–170, 1999.
70. Block, W.M., Brennan, L.A., and Gutiérrez, R.J., Evaluation of guild-indicator species for use in resource management, *Environ. Manage.*, 11, 265–269, 1987.
71. Hutto, R.L., Reel, S., and Landres, P.B., A critical evaluation of the species approach to biological conservation, *Endangered Species Update*, 4, 1–4, 1987.
72. Landres, P.B., Use of the guild concept in environmental impact assessment, *Environ. Manage.*, 7, 393–398, 1983.
73. Martin, T.E., Avian life history evolution in relation to nest sites, nest predation and food, *Ecol. Monogr.*, 65, 101–127, 1995.
74. Szaro, R.C., Guild management: an evaluation of avian guilds as a predictive tool, *Environ. Manage.*, 10, 681–688, 1986.

75. Thiollay, J.-M., Influence of selective logging on bird species diversity in a Guianan rain forest, *Conserv. Biol.*, 6, 47–63, 1992.
76. Lindenmayer, D.B., Cunningham, R.B., and McCarthy, M.A., Landscape analysis of the occurrence of arboreal marsupials in the montane ash forests of the central highlands of Victoria, southeastern Australia, *Biol. Conserv.*, 89, 83–92, 1999.
77. Jaksic, F.M., Abuse and misuse of the term "guild" in ecological studies, *Oikos*, 37, 397–400, 1981.
78. Niemëla, J., Langor, D., and Spence, J.R., Effects of clear-cut harvesting on boreal ground beetles assemblages (Coleoptera: Carabidae) in Western Canada, *Conserv. Biol.*, 7, 551–561, 1993.
79. Peters, R.H., *The Ecological Implications of Body Size,* Cambridge University Press, New York, 1983.
80. Steele, B.B., Bayn, R.L., Jr., and Grant, C.V., Environmental monitoring using population of birds and small mammals: analyses of sampling effort, *Biol. Conserv.*, 30, 157–172, 1984.
81. Wicklum, D. and Davies, R.W., Ecosystem health and integrity? *Can. J. Bot.*, 73, 997–1000, 1995.
82. Luff, M.L., Eyre, M.D., and Rushton, S.P., Classification and prediction of grassland habitats using ground beetles (Coleoptera: Carabidae), *J. Environ. Manage.*, 35, 301–315, 1992.
83. Schoener, T.W., Patterns in terrestrial vertebrates versus arthropod communities: do systematic differences in regularity exist? in *Ecology*, Diamond, J. and Case, T.J., Eds., Harper and Row, New York, 1986, pp. 556–586.
84. Davies, K.F. and Margules, C.R., Effects of habitat fragmentation on carabid beetles: experimental evidence, *J. Anim. Ecol.*, 67, 460–471, 1998.
85. Murphy, D.D., Freas, K.E., and Weiss, S.B., An environmental-metapopulation approach to population viability analysis for a threatened invertebrate, *Conserv. Biol.*, 4, 41–51, 1990.
86. Temple, S.A. and Wiens, J.A., Bird populations and environmental changes: can birds be bio-indicators? *Am. Birds*, 43, 260–270, 1989.
87. Villard, M.-A., Merriam, G., and Maurer, B.A., Dynamics in subdivised populations of neotropical migratory birds in a temperate fragmented forest, *Ecology*, 76, 27–40, 1995.
88. Drapeau, P. et al., Landscape-scale disturbances and changes in bird communities of eastern boreal mixed-wood forest, *Ecol. Monogr.*, 70, 423–444, 2000.
89. Cairns, J., Jr., The myth of the most sensitive species, *BioScience*, 36, 670–672, 1986.
90. Griffith, J.A., Connecting ecological monitoring and ecological indicators: a review of the literature, *J. Environ. Syst.*, 26, 325–363, 1997.
91. Whitford, W.G. et al., Vegetation, soil, and animal indicators of rangeland health, *Environ. Monit. Assess.*, 51, 179–200, 1997.
92. Conroy, M.J. and Noon, B.R., Mapping of species richness for conservation of biological diversity: Conceptual and methodological issues, *Ecol. Appl.*, 6, 763–773, 1996.
93. Van Horne, B., Density as a misleading indicator of habitat quality, *J. Wildl. Manage.*, 47, 893–901, 1983.
94. Vickery, P.D., Hunter, M.L., Jr., and Wells, J.V., Is density an indicator of breeding success? *Auk*, 109, 706–710, 1992.
95. Lancia, R.A. et al., Validating habitat quality assessment: An example, *T. N. Am. Wildl. Nat. Res. Conf.*, 47, 96–110, 1982.
96. Askins, R.A. and Philbrick, M.J., Effect of changes in regional forest abundance on the decline and recovery of a forest bird community, *Wilson Bull.*, 99, 7–21, 1987.

97. Karr, J.R. and Chu, E., *Restoring Life in Running Waters: Better Biological Monitoring*, Island Press, Washington, D.C., 1999.
98. Noss, R.F., Can we maintain biological and ecological integrity? *Conserv. Biol.*, 4, 241–243, 1990b.
99. Lindenmayer, D.B., Future directions for biodiversity conservation in managed forests: indicator species, impact studies and monitoring programs, *For. Ecol. Manage.*, 115, 277–287, 1999.
100. Wiens, J.A., Rotenberry, J.T., and Van Horne, B., A lesson in the limitations of field experiments: shrubsteppe birds and habitat alteration, *Ecology*, 67, 365–376, 1986.
101. DeSante, D.F., Monitoring avian productivity and survivorship (MAPS): a sharp, rather than blunt, tool for monitoring and assessing landbird populations, in *Populations*, McCullough, D.R. and Barrett, R.H., Eds., Elsevier Applied Science, London, U.K., 1992, pp. 511–521.
102. Buford, E.W., Capen, D.E., and Williams, B.K., Distance sampling to estimate fledgling brood density of forest birds, *Can. Field-Nat.*, 110, 642–648, 1996.
103. Martin, T.E. et al., *BBIRD (Breeding biology research and monitoring database) field protocol*, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MO, 1997.
104. Gunn, J.S. et al., Playbacks of mobbing calls of black-capped chickadees as a method to estimate reproductive activity of forest birds, *J. Field Ornithol.*, 71, 472–483, 2000.
105. Steedman, R. and Haider, W., Applying notions of ecological integrity, in *Ecological Integrity and the Management of Ecosystems*, Woodley, S., Kay, J., and Francis, G., Eds., St. Lucie Press, Delray Beach, FL, 1993, pp. 47–60.
106. Gibbs, J.P., Droege, S., and Eagle, P., Monitoring populations of plants and animals, *BioScience*, 48, 935–940, 1998.
107. Palmer, M.W., Potential biases in site and species selection for ecological-monitoring, *Environ. Monit. Assess.*, 26, 277–282, 1993.
108. Doak, D.F., Source-sink models and the problem of habitat degradation: general models and applications to the Yellowstone grizzly, *Conserv. Biol.*, 9, 1370–1379, 1995.
109. Kurzejeski, E.W. et al., Experimental evaluation of forest management: the Missouri Ozark forest ecosystem project, *T. North Am. Wildl. Nat. Res. Conf.*, 58, 599–609, 1993.
110. White, P.S. and Bratton, S.P., After preservation: philosophical and practical problems of change, *Biol. Conserv.*, 18, 241–255, 1980.
111. Woodley, S. and Theberge, J., Monitoring for ecosystem integrity in Canadian National Parks, in *Science and the Management of Protected Areas*, Willison, J.H.M., Bondrup-Nielsen, S., Drysdale, C., Herman, T.B., Munro, N.W.P., and Pollock, T.L., Eds., Elsevier, New York, 1992, pp. 369–377.

26 Judging Survey Quality in Biomonitoring

H.Th. Wolterbeek and T.G. Verburg

CONTENTS

26.1 INTRODUCTION

Biomonitoring, in a general sense, may be defined as the use of bioorganisms to obtain quantitative information on certain aspects of the biosphere (see Puckett,[1] Garty,[2] Markert et al.,[3,4] or Wittig[5] for clear overviews of what is meant by the terms *monitors*, *indicators*, or *collectors*). The relevant information in biomonitoring may be deduced from changes in the behavior of the organism (Herzig et al.,[6] Impact: occurrence of the species, ecological performance, morphology) and may also be obtained from assessment of relevant substances in the monitor tissues. In metal air pollution biomonitoring, information is mostly based on the determination of the

1-56670-641-6/04/$0.00+$1.50

metal content of the monitor. The present paper discusses quality aspects of (larger-scaled) biomonitoring surveys, thereby essentially using insights and literature data on trace element air pollution. Although the scale of the survey may have various dimensions (e.g., time, distance, amount of samples), and although many quality considerations apply for any scale, most of the presently used data originate from geography-related scales. The latest Europe-wide moss monitoring survey may illustrate how big these scales may become; it included the participation of some 30 European countries.[7]

26.2 SOME BASICS OF BIOMONITORING

26.2.1 Dose–Response Relationships

To avoid any lengthy discussion on the terminology of what is meant by monitors or indicators in the context of metal air pollution, the present chapter generally handles the term biomonitoring as "the use of the biomonitor organism to get information of elemental deposition and/or atmospheric levels, thereby including impact information because the dose–response relationships should be quantified as far as is possible." The latter means that, although the information on impact also serves its own additional purposes,[4] if we regard the elemental levels in the biomonitor as a response to ambient elemental levels (air deposition = dose), and if we restrict ourselves to the context of the dose–response relationship, impact on the biomonitor physiology should be seen as relevant because it may cause changes in the nature of this dose–response relationship (see Garty[2] for a review on the impact on lichen physiology of metals such as Pb, Fe, Cu, Zn, Cd, Ni, Cr, and Hg). Moreover, natural and anthropogenic causes for variabilities in ambient macro- and microconditions such as acidity (SO_2), temperature, humidity, light, altitude, or ambient elemental nutritional occurrences may cause the biomonitor to exhibit variable behavior (see Seaward et al.[8] for altitude effects on lichen responses). Part of this variance may show as local variance[9] but it may be clear that this variable behavior becomes a problem when it seriously affects the biomonitor in its accumulative responses.

26.2.2 Goals of the Survey

In the presently adopted context, the goal of the survey may be the assessment of (large-scaled) geographical patterns (levels, deposition) in trace element air pollution, or in multielement approaches; it may even go a little bit further in the sense that profiles of emission sources may be determined on basis of the correlations between the abundances of the elements.

For nearly all used biomonitor species, it has been demonstrated that their element contents indeed reflect atmospheric trace element levels or deposition.[10–12] However, they are also influenced by (local) soil dusts,[10,13,14] and may be affected by their substrates.[15,16] Regardless of the statistical methods used in determining correlations between element occurrences, the survey's goals prescribe that the survey's quality should be assessed in terms of its geographical resolution, keeping in mind that local soil dusts or substrates may affect initial data.

In the literature, biomonitoring species for trace element air pollution are often selected on the basis of criteria such as *specificity* (which means that accumulation is considered to occur from the atmosphere only, see Rühling[13]), *accumulation ratios*,[1,10] or a well-defined *representation of a sampling site*,[9] the latter in terms of relatively small variabilities in local metal contents. The question may be raised here as to what extent these biomonitor properties give any valuable information about the quality of the survey (high accumulation or small local variances are not goals of the survey). The concept of local variance, however, is interesting in the context of the monitor's requirements and may be used in an idea of measurable quality.

26.3 INTRODUCING MEASURABLE ASPECTS OF SURVEY QUALITY

26.3.1 Vitality and Dose–Response

As indicated above, the geographical comparability of the biomonitor's responses should be seen as dependent on the geography-related variability of its behavior; the latter may vary with time and ambient overall conditions.

Seasonal effects on plant elemental concentrations have been described by Markert and Weckert,[17] Ernst,[18] and Markert.[19] In general terms, these effects may be ascribed to both elemental leaching and increased availability by rainfall[20] and to seasonally varying degrees of dilution by mass increments, the latter due to seasonal variations in growth rates.[19] At this point, metal phytotoxic effects on plant physiology should be regarded. Growth is often used as a striking marker for strong physiological disorder[21] but effects on metal accumulation already occur when this visible growth symptom is less pronounced or even absent. One of the most direct effects on the cellular level is the alteration of the plasma membrane permeability, which may cause excess leakage of ions[22] and may have effects on metal accumulation characteristics.[23]

Strongly acidic precipitation, largely associated with atmospheric SO_2,[24,25] may yield lower moss metal concentrations[26]; toxic action in plants is indicated, especially for SO_2 and NO_2.[27–29]

Ambient SO_2 has been considered in initial studies with bark,[30,31] and estimations were based on measurements of bark S and bark acidity. In a later bark study, Wolterbeek et al.[32] examined relationships between sulfate, ammonia, nitrate, acidity, and trace metals. Bark sulfate, ammonia, and nitrate were interpreted as not significantly affecting bark metal retention, but Ca and Hg were affected by acidity. For bark, the Ca loading in particular may determine the buffering capacity with respect to incoming acidic precipitation[33]; further neutralization may be brought about by alkalizing effects from atmospheric NH_3.[29]

Based on moss, lichen, and bark data, Wolterbeek and Bode[34] proposed to supplement the trace metal analysis in biomonitors with the determination of pH, NH_4, NO_3, and SO_4. Furthermore, parallel to comparisons in metal contents, biomonitors may be taken into comparative determinations of vitality. These may include the assessment of membrane leakage, stress-ethylene production, the rate of photosynthesis, spectral reflectance, chlorophyll content, etc.[23,27,35,36] One of the first biomonitoring studies ever to combine metal contents with assessment of vitality was an Argentinian lichen survey by Jason et al.[37]: Judgement of geographical differences in metal contents was performed only after concluding that no significant differences could be observed in selected signs of lichen vitality.

26.3.2 Time

As indicated above, seasonal effects on the biomonitor's elemental content may be ascribed to the combined effects from accumulation, release, rainfall, growth, etc. The total illustrates the dynamic behavior of the biomonitor: it continuously accumulates and releases elements of interest. For mosses, once-accumulated elements are generally presumed to be retained for an infinite length of time and all dynamics are thus pressed into element-specific retention efficiencies.[13] These efficiencies, however, may very well comprise the effects of combined uptake and release.

As pointed out by Reis et al.,[38,39] both uptake and release processes show that the biomonitor should be regarded as reflecting a certain period of atmospheric element levels or deposition, the rate characteristics of both accumulation and release implying that this reflected period should be considered as element specific.

The time dedicated to survey field work and the sampling of biomonitor material should be planned as short, relative to the extent of the reflection period. Reis[20] gives rough data for a number of elements and for a Portugal lichen survey, which indicate that surveys should be performed in terms of weeks rather than months, the specifics, of course, depending on the elements of interest. In practice, survey aspects such as time and geography may become governed by other aspects such as available personnel, handling means/capacity, or costs,[34] but the above suggests that especially the time-drifting of a survey should receive ample attention; in extreme cases the survey should not be carried out.

26.3.3 Local and Survey Variances

Generally, and although the sampling site may be regarded as the basic unit of the survey, it is mostly left out in any discussion of the goals, results, or implications of the survey. In reality, the sampling site is simply selected as a spot of (geographical) dimensions which is small relative to the dimensions of the survey. Implicitly, it is assumed that the sampling site is essentially homogeneous with respect to the investigated variation in survey parameters. The determination of the local variance implies that all aspects of the survey are taken into account: the selection of the biomonitor species, the definition of the sampling site, sampling, sample handling, elemental analysis, etc. Wolterbeek et al.[9] compared instrumental variances (element analyses) with local variances and concluded that analytical uncertainties generally do not contribute to a significant extent to local variances (see Figure 26.1). They concluded that in larger-scaled surveys any attempt to improve analytical precision may be regarded as meaningless. This very point implies that the selection of a biomonitor on the basis of its accumulation factor (which may be seen as resulting in better analytical precision) has limited value in terms of survey quality.

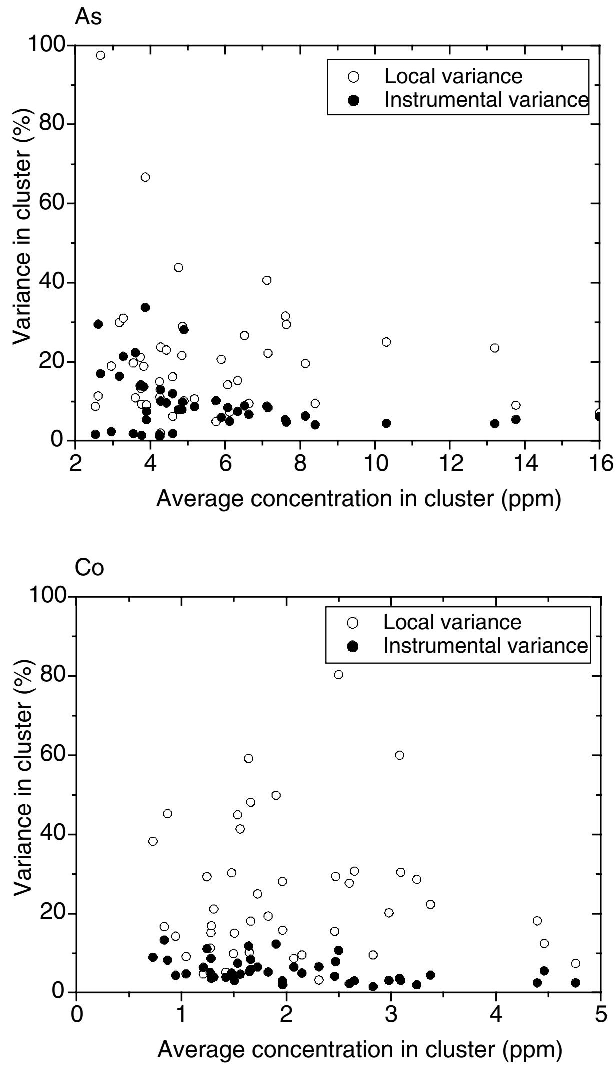


FIGURE 26.1 Instrumental variances (repeatability in elemental determinations, closed symbols) and local variances (variances in element concentrations in biomonitor tissues within a sampling site, open symbols) illustrated by results for arsenic (As) and cobalt (Co) in lichens. The data are calculated from survey results by Sloof and Wolterbeek (1991). The clusters indicate results from fivefold local sampling throughout the total survey area (The Netherlands, 30,000 km^2). (Wolterbeek, H.Th. and Verburg, T.G., unpublished results.)

It should also be noted that the local variance as such does not contain any direct information about the survey's quality; thus, it cannot be seen as decisive criterion for biomonitor selection. However, as with the sampling site, the local variance may be regarded as a basic unit of the survey; the local variance should be small relative to the total variance of the survey.

26.4 THE CONCEPT OF THE SIGNAL-TO-NOISE RATIO

Wolterbeek et al.[9] introduced the concept of the survey signal-to-noise ratio as a measurable aspect of survey quality. In this concept, both the local variance and the survey variance are taken into consideration, with the local variance seen as the survey noise and the survey variance regarded as the survey signal. Both signals are combined into the signal-to-noise ratio by which the survey is given a measurable expression of quality, in the sense that the geographical resolution of the survey is presented. The approach means that all survey aspects such as biomonitor species, accumulation factor, season, sampling site, or local variance become part of the signal-to-noise properties of the survey.

In the literature on biomonitoring, generally all information is about the signal; hardly any survey report contains information on the noise. Since the concept says that the noise is as essential as the survey signal, the inevitable conclusion is that hardly any survey report contains information on quality. Wolterbeek et al.[9] give signal-to-noise data for surveys with mosses, lichens, and tree bark performed in The Netherlands, Slovenia, the Czech Republic, and the Chernobyl region in Ukraine—(element-specific) data range from 1 to 6. Their report also discusses possible approaches by factor-analytical techniques aimed at data cleanup (removal of soil effects), or source profile isolation (dedicated surveys: new datasets on specific sources only), and the consequences for the resulting newly derived signal-to-noise ratios. Ratios down to 0.1 indicated the virtual absence of reliable information on nonselected elements in dedicated surveys, whereas ratios up to 13 illustrated quality improvement for selected source-specific elements.

It should be noted here, however, that whereas the survey variances are directly implicated by the survey's results, the local variances fully depend on the researcher's interpretation of the sampling site, both in terms of the site dimensions and the approaches in local multifold sampling. Therefore, they deserve further attention in future studies.

26.5 VARIANCES

26.5.1 LOCAL SAMPLING SITES

In practice, surveys are often set up by a preset sampling grid of a density tuned by available personnel, the available time, the analytical capacity, or costs. In some cases *a priori* knowledge of the survey area directs subarea grid densities for regions of special interest. Grid densities may be higher in urban (industrial) than in rural (remote) subareas of the survey.[13,40] Grid densities may also be ruled by circumstances: during field work it may become clear that the intended biomonitors cannot always be found. The consequence is that grid densities vary, or that more than a single species is to be used. Of course the risk to be forced to use more than a single species is becoming larger with any increase in survey area or variability in environmental conditions. The implication is that interspecies calibrations should be carried out, and that additional variances are introduced into the survey. Uncertainties associated with interspecies comparisons are reported as up to 50% or thereabouts.[41]

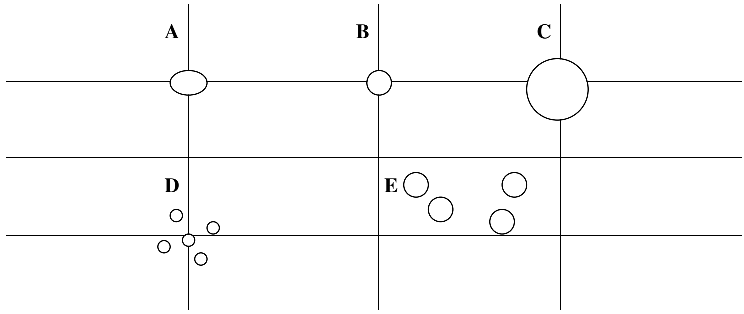


FIGURE 26.2 Sampling grids. The lines represent a survey sampling grid. The symbols indicate possible sampling strategies and selected sampling site dimensions: (A) sites at grid crossings, with site dimensions in line with the grid dimensions; (B), (C) sampling sites at grid crossings, (B) and (C) differing in size; (D) multiple small-sized sites at grid crossings, the sampling site dimensions depending on the desired multitude of subsamples; (E) multiple sites (regular or irregular distributions) covering grid cell areas. Note that (A) through (D) are grid node sampling (they could also be set as grid center sampling), and that (E) is grid cell sampling. Wolterbeek, H.Th. and Verburg, T.G (unpublished results).

The sampling grid dictates the coordinates of the sites to be visited for sampling, but hardly any survey protocol comprises (scientific) arguments to back up the dimensions or desired characteristics of the actual sampling spots (Figure 26.2). In reports, spots may be defined as "at least 300 m away from the nearest road"[7] or "open spots in the forest,"[13] or imply an actual sampling area of about "50 × 50 m,"[19] but further sampling protocols are mostly dedicated to the prescription of sampling locations and sampling, sample handling, and sample storage procedures. Of interest in this context are the sampling flow charts given by Markert,[42] suggesting that sampling may involve errors up to 1000%, which implies that sampling is by far the most error-sensitive aspect of environmental surveys.

The total shows that, in surveys, two main difficulties occur—how to define a sampling site (see Figure 26.2), and how to sample in order to really represent the selected site. For combining data from individual sites into a geographically continuous representation (mapping), the reader is referred to Cressie.[43,44] The following paragraphs of the present paper are devoted to site representaton and its consequences. Interpolation approaches, however, will be addressed in Section 26.6.1, in suggested methods to use data from surrounding sites in judging local-site information.

26.5.2 Local Variances

In discussing outcomes from local sites, especially if these data are to be used in judging survey quality,[9] the results may be regarded both in the context of their *accuracy* and *precision* (see Figure 26.3). In many surveys, the information on precision is *a priori* lost by pooling of the subsamples before eventual analysis. In addition, local sampling mostly comprises taking samples at random in small numbers before pooling, and the question may be raised as to what extent this approach yields a reliable representation of the site. Presuming some level of homogeneity, the

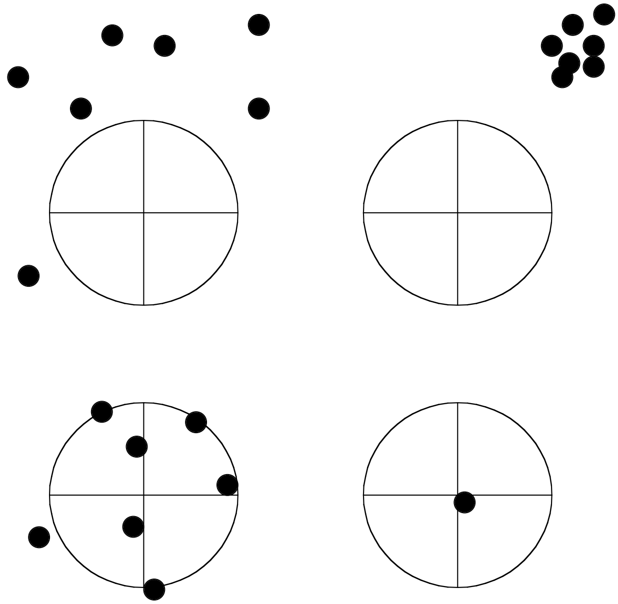


FIGURE 26.3 Accuracy (on target) and precision (repeatability). Upper left: Poor accuracy, poor precision. Upper right: Poor accuracy, good precision. Lower left: Good accuracy, poor precision. Lower right: Good accuracy, good precision. (Wolterbeek, H.Th. and Verburg, T.G., unpublished results.)

observed values may be part of a local normal distribution, but the small numbers of subsamples taken may implicate that sites are not (always) rightly represented, both in observed averaged levels (accuracy) and variances (precision). (See Table 26.1 for some illustrative values.)

26.5.2.1 Local Pooling and Homogenization

In pooling approaches, samples are mixed and milled before further processing. The implicit assumption is that mixing invariably results in a homogenized total sample. Wolterbeek and Verburg[45] studied the mixing approach by taking 32 tree bark subsamples from a local Dutch site, and performed metal determinations directly in the initial samples and also in 32 subsamples taken from the pooled bulk after thorough mixing. Their data are given in Table 26.2 for several elements. The results indicate that the average outcomes (based on n = 32 samples) may be trusted but that mixing does not always result in homogenization. This means that average local levels and uncertainties may only be obtained by analysis of a multitude of initially taken local samples. Survey costs and planning, however, mostly permit only small-numbered subsampling in the field and even smaller numbers of (sub)samples taken into eventual elemental analysis.[12]

TABLE 26.1
Local Sampling Sites

	Data from Local Sampling Sites (Means ± SD, in mg/kg)					
	Tree Bark (n = 32)		Moss (n = 25)		Soil (n = 25)	
Element	Initial	Bootstrapped	Initial	Bootstrapped	Initial	Bootstrapped
As	0.29 ± 0.06	0.29 ± 0.056	0.38 ± 0.12	0.38 ± 0.02	6.7 ± 1.9	6.7 ± 0.4
Ba	527 ± 1028	414 ± 619	—	—	92 ± 23	91 ± 1
Br	13 ± 3	13 ± 3	8.4 ± 1.5	8.0 ± 1.5	30 ± 6	30 ± 2
Cr	13 ± 7	13 ± 7	—	—	15 ± 4	15 ± 2
Cs	0.06 ± 0.02	0.06 ± 0.02	0.24 ± 0.06	0.24 ± 0.02	0.8 ± 0.2	0.8 ± 0.03
Cu	32 ± 5	32 ± 5	—	—	—	—
Fe	850 ± 220	850 ± 210	565 ± 87	566 ± 18	3678 ± 558	3716 ± 109
K	1240 ± 244	1230 ± 277	5173 ± 242	5220 ± 233	3060 ± 1277	3017 ± 234
Mg	666 ± 91	665 ± 93	1533 ± 242	1557 ± 44	—	—
Zn	60 ± 11	60 ± 10	73 ± 7	74 ± 1	77 ± 30	78 ± 7

Notes: Multiple samples were taken from tree bark, moss, and soil. Comparisons between means and variances (in mg/kg dry weight) for initial samples, and means and variances derived after bootstrapping. Bootstrapping methods were used to estimate the mean and variance of the concentration populations (Hall, 1986; Efron and Tibshirani, 1986; Diciccio and Romano, 1988). T-testing indicated the absence of any significant (P = 0.05) differences between the initial and the bootstrapped (normal) populations. Note the differences between biomaterials and soils. — = no analyses performed.

Source: Tree bark data taken from Wolterbeek, H.Th. and Verburg, T.G., Judging survey quality: local variances. *Environ. Monit. Assess.*, 73, 7–16, 2002. With permission.

26.5.2.2 Fivefold Subsampling and the Local Population

In recent biomonitoring surveys on trace element air pollution (apart from the majority of sites for which subsamples are pooled before further processing) in a limited number of sites, fivefold subsampling was performed with all subsamples processed separately in further procedures.[9,14] The results were used to judge both local averages and variances.

In studying this approach, Wolterbeek and Verburg[45] made use of multifold data on local sampling (n = 32, 25, and 25 for tree bark, moss, and soil respectively, see Table 26.1). Local populations of element concentrations were estimated from the available full data by bootstrapping methods (see Hall,[46] Efron and Tibshirani[47] or Diciccio and Romano[48] for details on bootstrapping), and repeated randomized n = 5 trials were taken out of the total number of local samples. Table 26.3 gives results obtained after 500 trials, and presents the mean and maximal increment factors F by which the trial local variance should be increased to ensure full compatibility with the actual local elemental concentration population. The applied student's t-testing implies a very strict verification of the local trial outcomes, although survey quality uses local variance only.[9] The Table 26.3 test compares both means and

TABLE 26.2
Element Concentrations (mg/kg ± SD) in Tree Bark (Delft, The Netherlands, n = 32) in Initial 32 Samples in a Bootstrap-Derived Population and in 32 Subsamples after Mixing

Element	Bark, Initial	Bark, Bootstrap	Bark, Mixed
As	0.29 ± 0.06	0.29 ± 0.05	0.30 ± 0.02
Ba	527 ± 1028	414 ± 619	433 ± 156
Br	13.3 ± 2.6	13.2 ± 2.6	12.6 ± 0.3
Cr	13.1 ± 7.3	13.2 ± 6.5	13.9 ± 0.8
Cs	0.06 ± 0.02	0.06 ± 0.02	0.06 ± 0.01
Cu	32 ± 5	32 ± 5	31 ± 5
Fe	850 ± 220	850 ± 210	860 ± 23
K	1240 ± 244	1230 ± 277	1260 ± 40
Mg	666 ± 91	665 ± 93	675 ± 88
Zn	60 ± 11	60 ± 10	61 ± 1

Note: Bootstrapping methods were used to estimate the mean and variance of the concentrations population (Hall, 1986; Efron and Tibshirani, 1986; Diciccio and Romano, 1988). The strong reduction in SD for Ba after bootstrapping suggests a nonnormal initial distribution of Ba values (see the initially very high relative Ba variance). After mixing, strongly reduced SDs are expected for all elements, but note the unchanged SD data for especially Cu and Mg. Also considering the absence of significant effects on SD from analytical routines (illustrated by Figure 26.1), the data indicate that mixing does not always result in homogenization.

Source: Data from Wolterbeek, H.Th. and Verburg, T.G., Judging survey quality: local variances. *Environ. Monit. Assess.*, 73, 7–16, 2002. With permission.

variances of the trial with the concentration population characteristics. Table 26.3 also presents the number of cases in which $F \neq 1$. The data indicate that, statistically speaking, n = 5 trials give reasonable results (agreement with local populations for all selected elements in >90% of the trials). However, the remaining up to 10% of all selected cases suggest that occurring deviations may have severe consequences. (Note the maximal F values and the 33 arsenic cases for which $F \neq 1$ [thus, in only about 7% of the total 500 cases] cause the mean F to shift from unit value to 1.34.)

26.5.2.3 The Local Site and the Survey

Table 26.4 gives elemental concentrations for surveys with tree bark, moss, and soil. These survey-level data may be compared directly to the (local) data shown in Table 26.1. Of course, the means vary; a selected site for multiple sampling does not

TABLE 26.3
Repeated (N = 500) Randomized Subsampling (n = 5) in Local Populations and T-Tests on Trial with Population

	F (Bark)			F (Moss)			F (Soil)		
Element	**Mean**	**Max**	**E**	**Mean**	**Max**	**E**	**Mean**	**Max**	**E**
As	1.34	32	33	1.20	11	42	1.13	9	25
Br	1.04	6	11	1.32	16	46	1.07	7	14
Cr	1.12	13	30	—	—	—	1.18	9	43
Cs	1.05	7	12	1.08	6	24	1.16	13	33
Cu	1.16	30	24	—	—	—	—	—	—
Fe	1.15	13	21	1.07	10	10	1.04	5	17
K	1.06	13	8	1.14	8	33	1.12	9	22
Mg	1.11	12	12	1.17	8	38	—	—	—
Zn	1.36	26	41	1.17	29	18	1.03	7	6

Note: Outcomes are expressed in increment factors F (means and maxima) in the local variance necessary to maintain full compatibility between trial outcome and population (based on t-test outcomes). E = number of trials with F ≠ 1. Initial local sampling was 32 for bark and 25, both for moss and soil.

Source: Data from Wolterbeek, H.Th. and Verburg, T.G., Judging survey quality: local variances. *Environ. Monit. Assess.*, 73, 7–16, 2002. With permission.

necessarily have averaged survey levels. Of interest are the survey variances, not only in comparing survey results between bark, moss, and soil, but also comparing the local with the survey tree bark data (Table 26.1 and Table 26.4). Note that the surveys show a higher variance which is strongly reduced by bootstrapping.

To see whether the survey means and variances could be estimated by using a limited number of observations, Wolterbeek and Verburg[45] performed repeated (N = 500) randomized subsampling (N = 5) in survey populations: Table 26.5 gives survey data for moss and soil in complete parallel to the local data presented in Table 26.3. The data indicate that the agreement between trial and survey is highly variable, probably due to (expected) skewed distributions for several elements. It should be noted here that the survey data in Table 26.5 also comprise all local problems discussed so far (Table 26.3). It may also be clear that the survey quality Q, defined as the ratio between survey and local variance, will suffer from both local and survey problems. This is illustrated in Table 26.6, where Q is given both on the initial suvey data, the bootstrapped (local) data, and on randomly selected 20 and 40% of the available local and survey data. The later data were used to estimate possibilities for judging survey Q on limited subsets: the incidence of >50% deviations of Q_3 from actual survey Q_1 (N, see Table 26.6) indicate that it is hardly possibly to get reliable information on survey Q by using subsurvey information.

TABLE 26.4
Survey Data

Survey Data: Element Concentrations (mg/kg dry wt., means ± SD) in Biomonitors

	Moss		Soil		Tree Bark	
Element	**Initial**	**Bootstrapped**	**Initial**	**Bootstrapped**	**Initial**	**Bootstrapped**
As	0.47 ± 0.17	0.44 ± 0.04	5.1 ± 2.8	5.7 ± 0.4	1.3 ± 0.6	1.3 ± 0.2
Ba	26 ± 8	21 ± 2	104 ± 36	106 ± 4	137 ± 98	116 ± 36
Br	6.7 ± 2.9	6.9 ± 1.0	22 ± 9	17 ± 1	24 ± 14	22 ± 6
Cr	5.1 ± 3.8	3.4 ± 0.4	20 ± 10	23 ± 2	14 ± 6	13 ± 1
Cs	0.23 ± 0.11	0.25 ± 0.04	0.68 ± 0.28	0.73 ± 0.03	0.31±0.17	0.34 ± 0.08
Cu	—	—	—	—	39 ± 20	30 ± 3
Fe	805 ± 412	637 ± 41	3716 ± 1576	4038 ± 329	2213 ± 980	2039 ± 345
K	5553 ± 1346	4993 ± 186	3971 ± 2075	3938 ± 222	1498 ± 451	1745 ± 207
Mg	1301 ± 256	1189 ± 88	—	—	—	—
Zn	64 ± 30	54 ± 8	69 ± 71	57 ± 7	220 ± 111	216 ± 29

Notes: Samples were taken from moss, soil, and tree bark. Comparisons between means and variances (in mg/kg dry weight) for initial samples, and means and variances derived after bootstrapping (100 bootstraps). Bootstrapping methods were used to estimate the mean and variance of the concentrations population (Hall, 1986; Efron and Tibshirani, 1986; Diciccio and Romano, 1988). — = no analyses performed.

TABLE 26.5
Repeated (N = 500) Randomized Subsampling (n = 5 sites) in Survey Populations and T-Tests on Trial with Population

	F (Moss)			F (Soil)		
Element	**Mean**	**Max**	**E**	**Mean**	**Max**	**E**
As	1.00	3	1	1.46	22	33
Br	1.00	1	0	1.82	16	130
Cr	1.34	7	88	1.00	1	0
Cs	1.00	1	0	1.77	56	77
Fe	1.38	10	101	1.00	1	0
K	1.56	17	70	1.34	20	50
Mg	1.00	1	0	—	—	—
Mn	1.44	9	135	1.02	5	4
Zn	1.01	3	3	1.20	6	66

Notes: Outcomes are expressed in increment factors F (means and maxima) in the survey variance necessary to maintain full compatibility between trial outcome and population (based on t-test outcomes). E = number of trials with $F \neq 1$. Initial number of sampling sites was 54 both for moss and soil.

Source: Data from Wolterbeek, H.Th. and Verburg, T.G., Judging survey quality: local variances. *Environ. Monit. Assess.*, 73, 7–16, 2002. With permission.

TABLE 26.6
Signal-to-Noise Ratios (Q = SV/LV) for Moss and Soil Surveys (54 Sites)

	Moss Survey					Soil Survey				
			Q_3					Q_3		
Element	Q_1	Q_2	a/b	Av ± SD	N	Q_1	Q_2	a/b	Av ± SD	N
As	1.4	2.8	a	1.6 ± 1.2	136	1.5	1.6	a	1.9 ± 1.4	111
			b	1.5 ± 0.5	78			b	1.6 ± 0.4	34
Br	1.9	4.9	a	2.2 ± 1.3	165	1.3	1.2	a	1.6 ± 1.0	128
			b	2.0 ± 0.7	76			b	1.4 ± 0.4	56
Cr	—	—	a	—	—	2.9	6.3	a	5.2 ± 5.3	266
			b	—	—			b	3.9 ± 2.6	190
Cs	1.8	4.7	a	2.3 ± 1.7	175	1.5	1.3	a	1.9 ± 1.1	155
			b	2.1 ± 0.9	117			b	1.6 ± 0.6	57
Fe	4.7	3.3	a	6.4 ± 6.8	233	2.8	4.4	a	3.8 ± 2.9	152
			b	5.2 ± 3.2	179			b	3.2 ± 1.1	63
K	1.3	1.2	a	2.0 ± 1.6	227	1.6	1.4	a	2.1 ± 1.4	163
			b	1.5 ± 0.7	127			b	1.8 ± 0.8	102
Mg	1.1	2.9	a	1.4 ± 1.2	118	—	—	a	—	—
			b	1.2 ± 0.3	36			b	—	—
Mn	4.9	3.1	a	6.2 ± 6.4	250	4.2	5.9	a	5.5 ± 5.2	217
			b	5.3 ± 2.8	147			b	4.8 ± 2.2	136
Zn	4.2	8.3	a	4.9 ± 3.4	105	4.6	11.	a	8.4 ± 13.8	247
			b	4.4 ± 1.3	35		6	b	5.9 ± 3.4	150

Notes: Q_1 = Q on actual survey; Q_2 = Q on bootstrapped local sites; Q_3 = Q based on a fraction of the survey's sampling sites (A = 20% survey and fivefold local; B = 40% survey and tenfold local); Q_3 determined in 500 trials; av = average; N = number of cases (out of 500 trials) that Q_3 is outside range of Q_1 ± 50%; SV = survey variance; LV = local variance.

Source: Data taken from Wolterbeek, H.Th. and Verburg, T.G., *Environ. Monit. Assess.*, 73, 7-16, With permission.

26.6 JUDGING LOCAL DATA BY USING NEARBY SITES

26.6.1 Interpolation

Apparently, generated local data suffer from uncertainties due to actual local-site variabilities combined with very small numbers of samples taken. Here, the question is what possibilities exist to use data from "nearby" sites to judge the quality of local data. To do this, first an interpolation approach should be selected, to be used for recalculations of local data. As a first exercise, Wolterbeek and Verburg (unpublished) generated Poisson data simply arranged along an X-axis. They also used original Na data from a Dutch moss dataset (Wolterbeek and Verburg, unpublished), and recalculated these data both by inversed distance weighing ($1/r^2$ and $1/r^3$, IDW2 and IDW3) and by kriging interpolation (no nugget effects, no drifts,

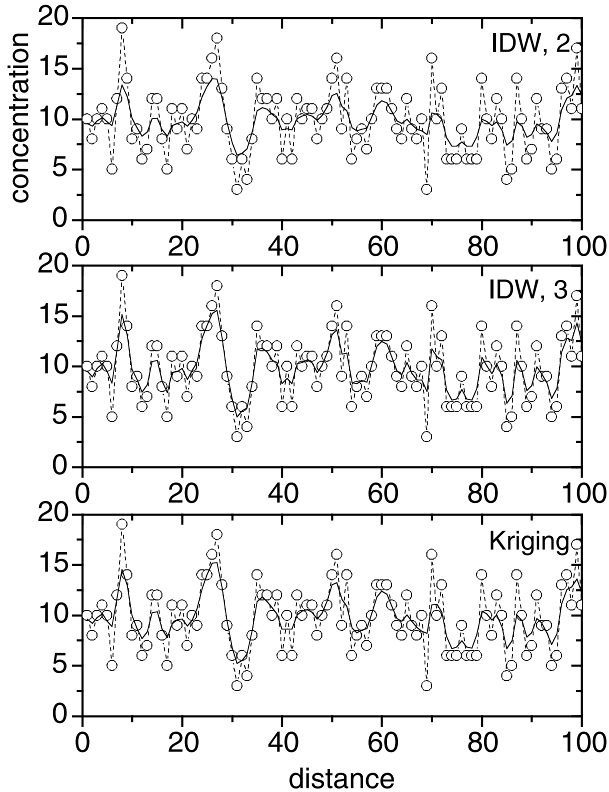


FIGURE 26.4 Calculation of element concentrations by interpolation approaches. The dots present randomly generated data following a Poisson distribution (n = 100) presented along a distance X-axis. The lines show the representation of the concentrations by *inversed distance weighing (1/r2 = IDW2, 1/r3 = IDW3)*, and by *kriging (linear variogram, no nuggets, no drifts)*. Data specifics are as follows. Original data: mean ± SD = 9.82 ± 3.37; IDW2: sum of squares = 457, mean ± SD = 9.82 ± 1.61; IDW3: sum of squares = 254, mean ± SD = 9.82 ± 1.61; kriging: sum of squares = 342, mean ± SD = 9.82 ± 2.04.

linear variogram; see also Cressie.[43,44]) The results are given in Figure 26.4 and Figure 26.5. Figure 26.4 shows that all interpolations yield original mean values, but it also shows that the kriging method best follows the original variance within the data. Figure 26.5 gives a more visual representation of the interpolation performances, which shows that kriging best preserves the original variance, especially visible for the Na moss dataset. Therefore, in further approaches, Wolterbeek and Verburg (unpublished) used the kriging method throughout.

26.6.2 Using Nearby Sites to Estimate Local Variance

Wolterbeek and Verburg (unpublished) used a theoretical "doughnut" approach to judge local variance based on the survey variance obtained from nearby sites. The doughnut approach is represented in Figure 26.6. Kriging interpolation was applied

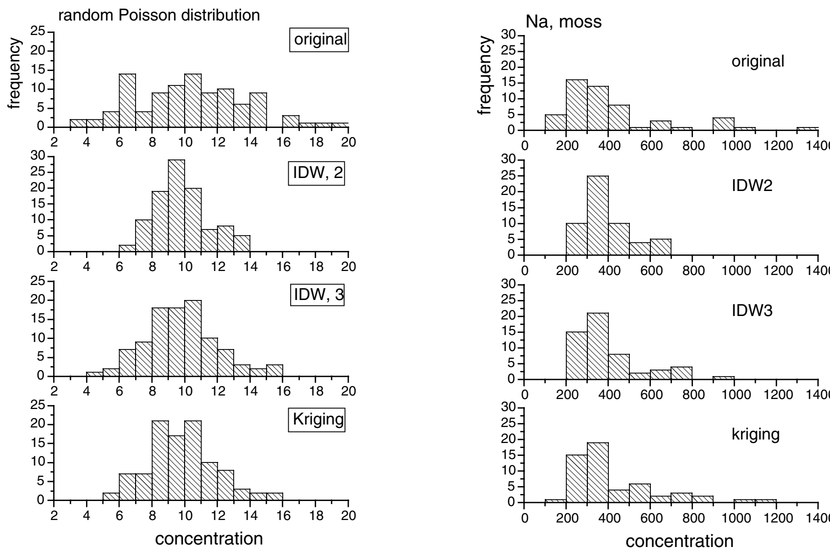
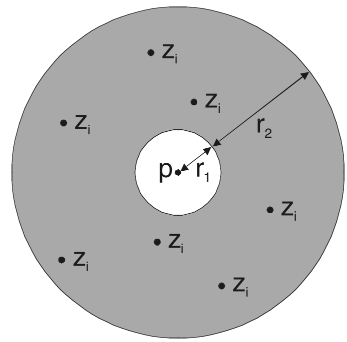


FIGURE 26.5 Fractional concentration distributions in randomly generated Poisson data (left, and see Figure 26.4) and in Na concentrations of a moss data set (right). The moss data comprise a 1995 survey throughout The Netherlands (unpublished results). Note especially the kriging performance in representing the Na concentrations in moss: original data: mean ± SD = 423 ± 256; IDW2: sum of squares (SS), mean ± SD = 2.1×10^6, 392 ± 115; IDW3 SS: mean ± SD = 2.0×10^6, 398 ± 163; kriging SS: mean ± SD = 1.7×10^6, 421 ± 216.

to estimate the value of sampling point *p* for every set of values *Zi,* randomly distributed throughout the gray-colored doughnut area. *Z* values were estimated within the given limits of variance in *Z* values. In the example (Figure 26.6), *Z* was taken as 100 ± 60, and both values and positions of *Z* were randomly and repeatedly varied (n = 25). The number of *Z* sites was taken as 4, 6, 8, or 10, to indicate variations in sampling site densities. The graph shows the variance of *p* in relation with the averaged distance to *p* of sites *Z*, and indicates that for a given survey variance the sampling site density can be used as a parameter to *a priori* judge the local variance in every local site *p*.

26.7 JUDGING SURVEY QUALITY BY RECALCULATING SURVEY DATA

In Section 26.4 and Section 26.5 the survey quality Q, expressed as the ratio between survey and local variance SV/LV, was derived from the original survey data based on the local variance estimated from multifold sampling and analyses in a limited number of sampling sites. The problems discussed in Section 26.5 imply that an alternative approach which essentially follows the nearby-sites approximations (see Section 26.6) may deserve further attention. Generalizing the thoughts on nearby



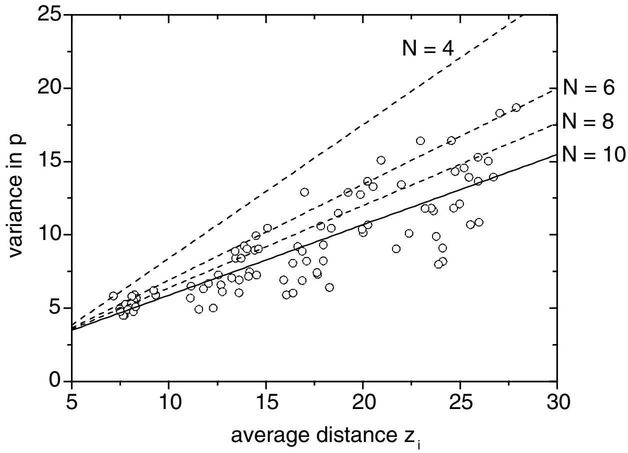


FIGURE 26.6 The "doughnut" surrounding sites approach. Above: The unknown element concentration of site p is estimated with help of N surrounding sites Z. The mean distance to p of sites Z can be estimated under conditions of variable r_1 and r_2. Under: Kriging was used to estimate the concentration and variance in p. The graph shows the variance in p in relation with the number N of sites Z and of the mean distance to p of Z. In all calculations, Z-sites were randomly chosen in the gray doughnut area. The dots shown in the figure represent the N = 10 situation, the lines represent N = 4, 6, 8, 10 sites. In calculations, r_1 was set to a constant value of 5, and r_2 was taken as 10, 18, 25, and 35 respectively (see the four dot clusters in the graph). For *Zi* values and positions and for each N situation, randomized values were generated (n = 25 trials) with normalized *ZI* mean ± SD = 100 ± 60. Note that the normalized Z mean value also results in averaged p = 100, and that the derived variances in p depend on the adopted variance in Z. The graph indicates that for a given survey variance, the sampling site density can be used as a parameter to *a priori* set the desired local variance in every sampling point p. (Wolterbeek and Verburg, unpublished.)

TABLE 26.7
Survey Quality Data for Na in a 1995 National Moss Survey in The Netherlands

Parameter	Original Data	Kriging Approach
Survey mean ± SD	423 ± 256	421 ± 216
Survey SD (%)	60	51
Local variance (%)	14 ± 7[a]	6 ± 3[b]
SV/LV	4.3	8.5

Notes: The data have been expressed as the ratio SV/LV between the survey variance SV and local variance LV. Outcomes are given based on the original Na concentration data and on the recalculated survey data by kriging approaches.

[a] Based on outcomes from multifold (mostly fivefold) subsampling and Na determinations in nine randomly selected sampling sites. The original-data approach to SV/LV was based on the survey variance and the determination of the (averaged) local variance by fivefold sampling in a limited number of sampling sites.
[b] Based on kriging-derived determinations of Na local variances in all 54 sampling sites. The kriging approach to SV/LV was based on the use of the full initial survey in recalculations of local values and local variances.

Source: Wolterbeek and Verburg, unpublished.

sites leads to a set-up in which the full survey is the starting point: the survey serves as a doughnut (Figure 26.6), and for each recalculation of an individual site, the specific site is taken out of the survey, and reestimated by kriging interpolation of the rest of the survey. Here, it should be noted, of course, that the distance weighing within kriging (directed by the characteristics of the variograms[43,44]) indicates that survey data contribute in distance-depending extents. The Na data from the moss survey may again serve as an example. Figure 26.7 shows results of the kriging approach, both in recalculated individual Na data and in Na geographical maps. The interpolation resulted in a new Na concentration for each individual sampling site, with more "smoothed" differences between nearby sites but without losing appreciable Na survey variance (Figure 26.5). Table 26.7 gives survey quality Q for both cases. The survey approach resulted in increased Q values, entirely based on decreases in the local variance of recalculated sites.

As can be seen from the upper graph in Figure 26.7, the interpolaton may result in relatively strong changes in Na values in specific sites. The data given in Table 26.3 and especially Table 26.5 indicate that local variances should be increased to make local data from the fivefold sampling compatible with local populations. Of interest here are the increments in local variances necessary to make the initial data compatible to the kriging-derived data. The two-tailed Z-testing of initial local data vs. kriging-derived local data indicate that the local variances from initial data should be raised from 14% (Table 26.7) to about 28% to ensure compatibility in >90% of

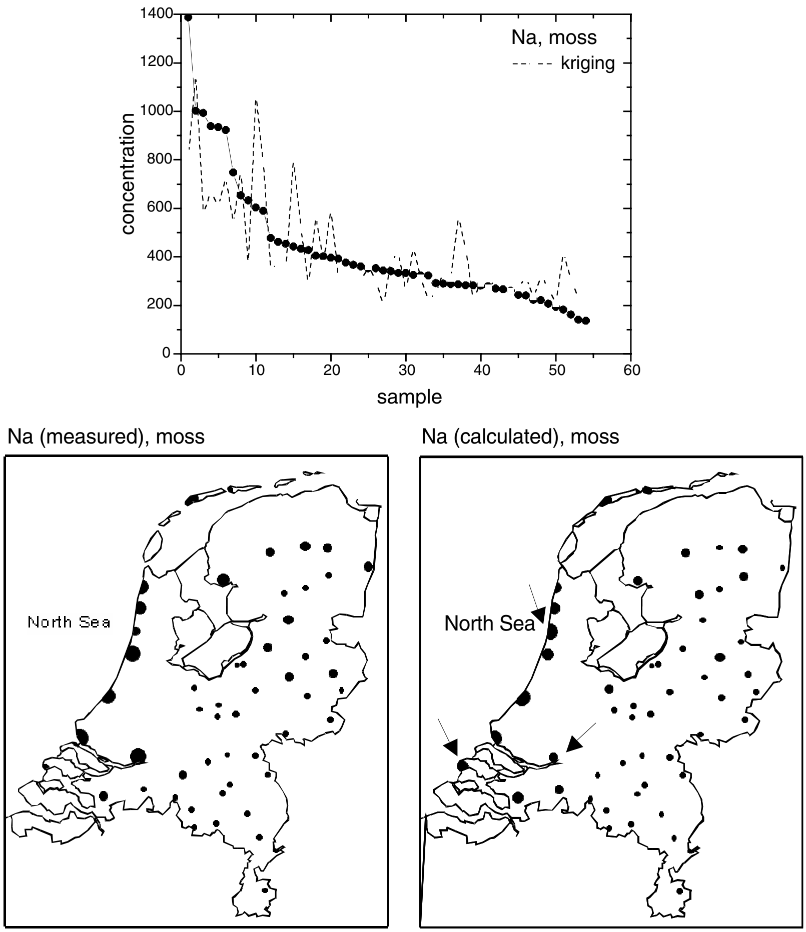


FIGURE 26.7 Reproduction of Na concentrations in moss samples (n = 54). Upper graph: The Na data (solid circles) are ordered in decreasing concentration, from the left side to the right side of the graph, the results by kriging are shown by the dotted line. Each of the individual Na concentrations was taken out of data set and recalculated by kriging, thereby using the rest of the Na data. It should be realized here that although each recalculated value is principally derived from the remaining n = 53 data, the practical distance weighing makes further-away points less important; in the given Na example only 2 to 6 neighboring sites participated to a relevant extent (>10%) in the determinations of the values. See Figure 26.5 for further information on the kriging performance on the Na data. (Wolterbeek and Verburg, unpublished). Lower maps: Moss results for Na (n = 54). The solid circle areas are proportional to the Na concentrations. Left: Initial data on Na concentrations in moss, throughout The Netherlands (Moss survey 1995, Wolterbeek and Verburg, unpublished results). Right: Na concentrations in moss, recalculated by kriging. Note the three arrows in the right-hand map, indicating prominent changes in initial Na concentrations. For both approaches in the moss Na surveys, signal-to-noise (SV/LV) ratios can be calculated. The direct approach (left map) in SV/LV was based on the survey variance and the determinaton of the average local variance by fivefold sampling in a limited number of sites, the kriging approach (right map) to SV/LV was based on the use of the initial survey in recalculations of local values and local variances. It should be noted that in the latter approach, survey variance is subject to slight shifts (see Figure 26.4 and Figure 26.5); for SV/LV data see Table 26.7.

all sites (Wolterbeek and Verburg, unpublished). This increment factor is in reasonable correspondence with the values indicated in Table 26.3 and Table 26.5 which were derived from the uncertainties arising from the limited fivefold local subsampling**.** In all, the survey approach (Section 26.7) may be a full alternative for the local-site approach (Section 26.5) in judging survey quality.

REFERENCES

1. Puckett, K.J., Bryophytes and lichens as monitors of metal deposition. *Bibliotheca Lichenologica* 30, 231, 1988.
2. Garty, J., Lichens as biomonitors for heavy metal pollution. In: *Plants as Biomonitors*: *Indicators for Heavy Metals in the Terrestrial Environment*, Markert, B., Ed., VCH Publishing, Weinheim, D, 1993, pp. 193–264, ISBN 3-527-30001-5.
3. Markert, B. et al., The use of bioindicators for monitoring the heavy-metal status of the environment. *J. Radioanal. Nucl. Chem.*, 240, 425, 1999.
4. Markert, B., Oehlmann, J., and Roth, M., Biomonitoring of heavy metals: definitions, possibilities, and limitations. In: *Proc. Int. Workshop on Biomonitoring of Atmospheric Pollution (with Emphasis on Trace Elements)*, *BioMAP*, Lisbon, 2000, p. 129, IAEA TECDOC 1152, IAEA, Vienna, Austria.
5. Wittig, R., general aspects of biomonitoring heavy metals by plants. In: *Plants As Biomonitors*, Markert, B., Ed., VCH Verlagsgesellschaft, Weinheim, Germany, 1993, pp. 3–27.
6. Herzig, R. et al., Lichens as biological indicators of air pollution in Switzerland: Passive biomonitoring as a part of an integrated measuring system for monitoring air pollution. In: *Element Concentration Cadasters in Ecosystems*, H. Lieth, B. Markert, Eds., CH Verlagsgesellschaft, Weinheim, Germany, 1990, pp. 317–332.
7. Buse A. et al., Heavy Metals in European Mosses: 2000/2001 Survey. UNECP ICP Vegetation, CEH Bangor, Bangor, 45 pp., 2003.
8. Seaward, M.R.D. et al., Recent levels of radionuclides in lichens from southwest Poland with particular reference to ^{134}Cs and ^{137}Cs. *J. Environ. Radioactiv.*, 7, 123, 1988.
9. Wolterbeek, H.Th., Bode, P., and Verburg, T.G., Assessing the quality of biomonitoring via signal-to-noise ratio analysis. *Sci. Tot. Environ.*, 180, 107, 1996.
10. Sloof, J.E., Environmental Lichenology: Biomonitoring Trace-Element Air Pollution. Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1993.
11. Rühling, A. and Tyler, G., Sorption and retention of heavy metals in the woodland moss Hylocomium splendens (Hedw.). Br. et Sch. *Oikos*, 21, 92, 1968.
12. Wolterbeek, H.Th. and Bode, P., Strategies in sampling and sample handling in the context of large-scaled plant biomonitoring surveys of trace-element air pollution. *Sci. Tot. Environ.*, 176, 33, 1995.
13. Rühling, A. (Ed.), Atmospheric Heavy Metal Deposition in Europe — Estimations Based on Moss Analysis. NORD 1994:9, Nordic Council of Ministers, AKA-PRINT, A/S, Arhus, 1994.
14. Kuik, P., Sloof, J.E., and Wolterbeek, H.Th., Application of Monte Carlo assisted factor analysis to large sets of environmental pollution data. *Atmos. Environ.*, 27A, 1975, 1993.
15. De Bruin, M. and Hackenitz, E., Trace element concentrations in epiphytic lichens and bark substrate, *Environ. Pollut.*, 11, 153, 1986.

16. Sloof, J.E. and Wolterbeek, H.Th., Substrate influence on epiphytic lichens. *Environ. Monit. Assess.*, 25, 225, 1993.
17. Markert, B. and Weckert, V., Fluctuations of element concentrations during the growing season of *Poytrichum formosum* (Hedw.). *Water Air Soil Pollut.*, 43, 177, 1989.
18. Ernst, W.H.O., Element allocation and (re)translocation in plants and its impact on representative sampling. In: *Element Concentration Cadasters in Ecosystems: Methods of Assessment and Evaluation,* Lieth, H. and Markert, B., Eds., VCH Publishing, New York, 1990, pp. 17–40, ISBN 0-89573-962-3.
19. Markert, B., Instrumental analysis of plants. In: *Plants as Biomonitors: Indicators for Heavy Metals in the Terrestrial Environment*, Markert, B., Ed., VCH Publishing, New York, 1993, pp. 65–104, ISBN 1-56081-272-9.
20. Reis, M.A., Biomonitoring and Assessment of Atmospheric Trace Elements in Portugal. Methods, Response Modelling and Nuclear Analytical Techniques. Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2001.
21. Van Gronsveld, J. and Clijsters, H., Toxic effects of metals. In: *Plants and the Chemical Elements. Biochemistry, Uptake, Tolerance, and Toxicity,* Farago, M., Ed., VCH Publishing, New York, 1994, pp. 149–178, ISBN 1-56081-135-8.
22. De Vos, R., Copper-induced oxidative stress and free radical damage in roots of copper tolerant and sensitive silene cucubalus. Ph.D. thesis, Free Univ., Amsterdam, The Netherlands, 120 pp., 1991.
23. Garty, J., Kloog, N., and Cohen, Y., Integrity of lichen cell membranes in relation to concentration of airborne elements. *Arch. Environ. Con. Toxicol.*, 34, 136, 1998.
24. Brown, D.H. and Beckett, R.P., Differential sensitivity of lichens to heavy metals. *Ann. Bot.*, 52, 51, 1983.
25. Seaward, M.R.D., The use and abuse of heavy metal bioassays of lichens for environmental monitoring. In: *Proc. 3rd Int. Conf. Bioindicators, Deteriorisations Regionis*, Liblice, Czechoslovakia, J. Spaleny, Ed., Academia, Praha, 1980, 375.
26. Gjengedal, E. and Steinnes, E., Uptake of metal ions in moss from artificial precipitation. *Environ. Monit. Assess.*, 14, 77, 1990.
27. Garty, J., Karary, Y., and Harel, J., The impact of air pollution on the integrity of cell membranes and chlorophyll in the lichen *Ramalina duriaei* (De Not.) Bagl. Transplanted to industrial sites in Israel. *Arch. Environ. Contam.* Toxicol., 24, 455, 1993.
28. Balaguer, L. and Manrique, E., Interaction between sulfur dioxide and nitrate in some lichens. *Environ. Exp. Bot.*, 31, 223, 1991.
29. De Bakker, A.J. and Van Dobben, H.F., Effecten van ammoniakemissie op epiphytische korstmossen; een correlatief onderzoek in de Peel. Rapport Rijksinstituut voor Natuurbeheer 88/35, 48 pp., Leersum, The Netherlands, 1988.
30. Stäxang, B., Acidification of bark of some deciduous trees, *Oikos,* 20, 224, 1969.
31. Härtel, O. and Grill, D., Die Leitfaehigkeit von Fichtenborken-Extrakten als empfindlicher Indikator fuer Luftverunreinigungen. *Eur. J. For. Pathol.*, 2, 205, 1979.
32. Wolterbeek, H.Th. et al., Relations between sulphate, ammonia, nitrate, acidity and trace element concentrations in tree bark in The Netherlands. *Environ. Monit. Assess.*, 40, 185, 1996.
33. Farmer, A.M., Bates, J.W., and Bell, J.N.B., Seasonal variations in acidic pollutant inputs and their effects on the chemistry of stemflow, bark and epiphytic tissues in three oak woodlands in N.W. Britain. *New Phytol.*, 118, 441, 191.
34. Wolterbeek, H.Th. and Bode, P., Strategies in sampling and sample handling in the context of large-scale plant biomonitoring surveys of trace element air pollution. *Sci. Tot. Environ.*, 176, 33, 1995.

35. Gonzalez, C.M. and Pignata, M.L., Effect of pollutants emitted by different urban-industrial sources on the chemical response of the transplanted *Ramalina ecklonii* (Spreng.) Mey. & Flot. *Toxicol. Environ. Chem.*, 69, 61, 1999.
36. Pignata, M.L. et al., Relationship between foliar chemical parameters measured in *Melia azedarach* L. and environmental conditions in urban areas. *Sci. Tot. Environ.*, 243, 85, 1999.
37. Jasan, R.C. et al., On the use of the lichen *Ramalina celastri* (Spreng.) Krog. & Swinsc. as an indicator of atmospheric pollution in the province of Córdoba, Argentina, considering both lichen physiological parameters and element concentrations. *J. Radioanal. Nucl. Chem.* (in press).
38. Reis, M.A. et al., Lichens (*Parmelia sulcata*) time response model to environmental elemental availability. *Sci. Tot. Environ.*, 232, 105, 1999.
39. Reis, M.A. et al., Calibration of lichen transplants considering faint memory effects. *Environ. Pollut.*, 120, 87, 2002.
40. Reis, M.A. et al., Main atmospheric heavy metal sources in Portugal by biomonitor analysis. *Nucl. Instrum. Methods in Phys. Res.* B 109/110, 493, 1996.
41. Wolterbeek, H.Th. et al., Moss interspecies comparisons in trace element concentrations. *Environ. Monit. Assess.*, 35, 263, 1995.
42. Markert, B., *Instrumental Element and Multi-Element Analysis of Plant Samples. Methods and Applications*. John Wiley & Sons, Chichester, U.K., 1996, ISBN 0-471-95865-4.
43. Cressie, N.A.C., The origins of kriging. *Math. Geol.*, 22, 239, 1990.
44. Cressie, N.A.C., *Statistics For Spatial Data.* John Wiley & Sons, New York, 900 pp., 1991.
45. Wolterbeek, H.Th. and Verburg, T.G., Judging survey quality: local variances. *Environ. Monit. Assess.*, 73, 7, 2002.
46. Hall, P., On the bootstrap and confidence intervals. *Ann. Stat.*, 14, 1431, 1986.
47. Efron, B. and Tibshirani, R., Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.*, 1, 54, 1986.
48. Diciccio, T.J. and Romano, J.P., A review of bootstrap confidence intervals. *J. R. Stat. Soc.*, B50(3), 338, 1988.

27 Major Monitoring Networks: A Foundation to Preserve, Protect, and Restore

M.P. Bradley and F.W. Kutz

CONTENTS

1-56670-641-6/04/$0.00+$1.50

27.1 INTRODUCTION

Ideally, major human and environmental monitoring networks should provide the scientific information needed for policy and management decision-making processes. It is widely recognized that reliable, comparable, and useable measurement results are a key component of effective monitoring and successful sustainable development policies. Monitoring should drive the planning process and provide the necessary data to evaluate the results of programs that were created, and then provide feedback to show what remains to be done.

Previous sections of this book have focused on the conceptual basis of monitoring systems, media specific monitoring, statistical design and sampling, and assessment, indicators, and policy. This final section introduces major monitoring networks.

The objectives of this introductory chapter are to (1) lay out the challenges associated with the development of a major monitoring network; (2) discuss some conceptual models for environmental and human health monitoring programs; (3) present some examples of current major environmental monitoring networks; (4) present information on human health monitoring; (5) articulate additional factors critical to effective monitoring networks; and (6) discuss potential future developments in monitoring networks.

Subsequent chapters in this section will provide detailed descriptions of four additional major monitoring networks:

- U.S. Clean Air Status and Trends Network (CASTNet)
- South African River Health Program (RHP)
- U.S. Environmental Monitoring and Assessment Program (EMAP)
- U.S. Forest Health Monitoring Program (FHM)

A fifth chapter will describe the U.S. Regional Vulnerability and Assessment Program (ReVA), which is designed to focus on integrating and synthesizing information on the spatial patterns of multiple exposures to allow a comparison and prioritization of risks.

27.2 CHALLENGES

While simple in concept, the task of developing a comprehensive program for environmental monitoring is extremely complex. Understanding the condition of the environment is difficult because the environment has many interacting components (e.g., soil, water, air, plants, and animals, including humans) that are affected by a variety of physical and biological conditions.

In addition, it is a challenge to design statistically defensible, effective, and efficient programs that will accomplish their goals with the least amount of money, time, and effort. Logistical limitations impose inherent tradeoffs among the number of variables that can be measured, the frequency at which they can be measured, and the number of sites involved.

Many government agencies have mandates that include monitoring, and these agencies spend considerable resources collecting environmental data. Regulatory requirements impose additional costs on the private sector and added to that complexity are the various universities, nongovernmental organizations (NGOs), and citizen–science groups who also engage in long-term environmental monitoring. Often the data are collected only for limited geographic areas or are not collected in a statistically representative manner. The lack of standardization of data or monitoring methods makes correlation across databases or across spatial scales problematic.

An added level of complexity comes with trying to link environmental monitoring and assessment with public health surveillance and assessment. Public health surveillance has largely focused on acute infectious diseases, chronic diseases, injuries, risk factors, and health practices, and the environmental agenda is separate from that of the traditional public health community. Information presently exists on chemical exposure and health status but not together. The removal of environmental health authority from public health agencies has led to fragmented responsibility, lack of coordination, and inadequate attention to the health dimensions of environmental problems (IOM, 1988). The regulatory infrastructure is driven by media-specific, source-specific, and probably molecule-specific approaches that shape everything from the way state and local health departments and environmental agencies do business to what research gets funded at universities (Burke, 1997). Public health systems lack even the most basic information about chronic disease and potential contributions of environmental factors.

Although written for coastal waters and estuaries, conclusions of the report *Managing Troubled Waters* from the National Research Council (1990) hold true for environmental monitoring as a whole. These conclusions include that monitoring would become even more useful under a comprehensive program documenting status and trends, and that a national survey would best combine intensive regional observations and cause–effects studies with a sparser national network of observations.

27.3 CONCEPTUAL MODELS

A comprehensive monitoring program must integrate across all facets of the environment (from the driving variables to the responding systems and across temporal and spatial scales) and must have the commitment to developing long-term databases (from decades to centuries distant) (U.S. EPA 2003a). Various conceptual models or frameworks have been developed to describe the comprehensive scope required of a major monitoring network. Some deal primarily with the spatial scales of monitoring activities, while others deal with the metrics and indicators monitored (Kutz et al. 1992a). Both are critical to the quality of the monitoring network.

27.3.1 The CENR Framework

The Committee on Environment and Natural Resources (CENR) was established by former U.S. President Clinton in recognition that the traditional single-agency, single-discipline way of solving problems was no longer adequate. The CENR recognized a high priority need to integrate and coordinate environmental monitoring and research networks and programs across agencies of the federal government

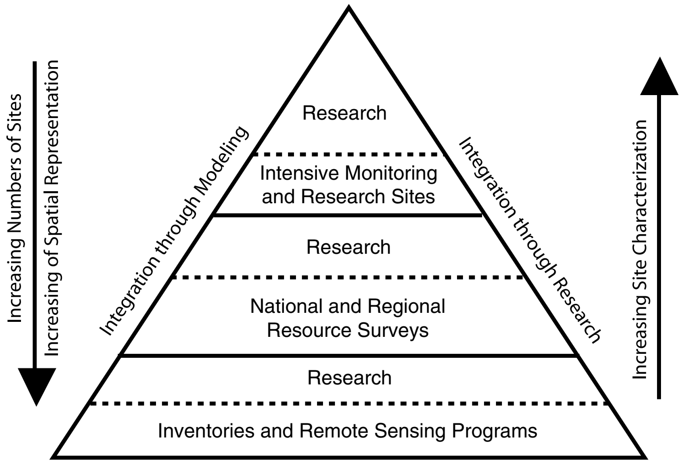


FIGURE 27.1 CENR Monitoring Framework, 1996.

(National Science and Technology Council, 1998). As an initial step, an interagency team of scientists and program managers produced a hierarchical design for integration of environmental monitoring activities.

The CENR Framework (Figure 27.1) links existing intensive ecological research and monitoring stations, regional surveys, remote sensing programs, and fixed-site monitoring networks in order to track complex environmental issues at a range of spatial and temporal scales. Each type of monitoring program yields unique or specific information, such as activities that (1) characterize specific properties of large regions by simultaneous and spatially intensive measurements sampling the entire region, (2) characterize specific properties of large regions by sampling a subset of the region, and (3) focus on the properties and processes of specific locations (National Science and Technology Council, 1997a).

A fundamental premise underlying this framework is that no single sampling design can effectively provide all of the information needed to evaluate environmental conditions and guide policy decisions. Ultimately, measurements at all three levels must be performed in a coordinated fashion, allowing an improved understanding of ecosystems and an improved ability to manage those systems for integrity and sustainability.

The first level includes inventories and remote sensing programs. These are extremely valuable for understanding the distribution and variations in land use, vegetative cover, ocean currents, and other surface properties of Earth, as well as for providing early warning of dangerous weather conditions. However, these programs are capable of focusing only on a small subset of the variables that are important for evaluating environmental condition and generally require extensive ground-based sampling to interpret the satellite images and to quantify their uncertainty.

A resource inventory is a complete description of the resource in question. Inventories typically involve documenting the number of physical features of a resource, such as the number and size of wetlands in a given area. Inventories do not involve regular, repeated sampling of the resource in question.

Remote sensing provides data or measurements collected as a series of contiguous and simultaneous measures across a large area. Remotely sensed surveys provide the capability to monitor a given area for changes in spectral (color) and spatial characteristics and can be conducted over a variety of spatial and temporal scales appropriate for specific applications or issues. The most common source of such data is from sensors mounted on fixed-wing aircraft or satellites.

The second level (survey programs) is designed, through sampling, to describe quantitatively and statistically characteristics of the entire population of a specific resource (e.g., lakes, birds, trees) in a given study region without completing a census of the resource.

Fixed-site monitoring networks are collections of permanent stations that make frequent measurements of a few specific environmental properties at a resolution that is sufficient to determine regional condition. The station density is generally not sufficiently high to accurately predict conditions at points between stations but these networks serve important functions such as regionalizing information on environmental parameters and forecasting environmental hazards.

Probability survey designs use a statistical sampling approach (similar to opinion polls) that is cost-effective and scientifically defensible. The strength of probability sampling is the ability to estimate the extent of condition, along with a statement about the uncertainty surrounding that estimate. Changes and trends in environmental condition can be determined if surveys are continued through time.

The third level (intensive monitoring and research sites) involves measuring all the major potential causes of environmental change at the same locations where environmental responses of concern to society are also measured. This form of monitoring permits the understanding of specific environmental processes sufficiently to allow prediction across space and time. Research at these sites on basic environmental processes related to hydrology, geology, biogeochemistry, atmospheric chemistry, population dynamics, and ecosystem dynamics has produced many major advances in environmental science (National Science and Technology Council 1997a).

The most significant aspect of this framework is that all three types of monitoring should be conducted in a coordinated fashion to provide for an integrated environmental monitoring and research capability.

Although U.S. agencies have been collaborating (National Oceanic and Atmospheric Administration (NOAA), U.S. Geological Survey (USGS), Environmental Protection Agency (EPA), U.S. Fish and Wildlife Service (U.S. F&WS)), no attempt has been made to utilize the CENR Framework in setting national priorities and agency budgets.

At a finer scale, the Delaware River Basin Collaborative Environmental Monitoring and Research Initiative (CEMRI) is a prototype environmental monitoring strategy that will link air quality, hydrological and forestry information across the landscape of the Delaware River Basin. CEMRI revolves around five issue-based studies and two integrating activities which in combination will create a long-term, multicomponent, and multiscale ecosystem monitoring strategy for the complex landscapes of the Delaware River Basin. CEMRI partners include USGS, U.S. Forest Service (USFS), and U.S. National Park Service (NPS).

In addition, several monitoring programs have embraced the principles of the Framework as evidenced by having the three tiers of monitoring. Two of these programs, the Environmental Monitoring and Assessment Program (EMAP) and the Forest Health Monitoring Program (FHM) are described in detail in subsequent chapters. The NPS Inventory and Monitoring Program (I&M) is being implemented in approximately 270 natural-resource parks over 4 years. Parks have been organized into 32 networks, linked by geography and shared natural-resource characteristics (Williams, 2001). Each network designs an integrated program to monitor physical and biological resources such as air and water quality, endangered species, exotic species, and other flora and fauna.

The CENR Framework does not explicitly address human health or socioeconomic endpoints; however, the three-tier approach could be used with any endpoints.

27.3.2 Framework for Environmental Public Health Tracking

The environment plays an important role in human development and health. Researchers have linked exposures to some environmental hazards with specific diseases; for example, exposure to asbestos has been linked to lung cancer. Other associations between environmental exposures and health effects are suspected but need further research; for example, the link between exposure to disinfectant by-products and bladder cancer. Currently, few systems exist at the state or national level to track many of the exposures and health effects that may be related to environmental hazards. Tracking systems that do exist are usually not compatible with each other, and data linkage is extremely difficult.

The U.S. Centers for Disease Control and Prevention (CDC) is developing a framework (CDC, 2003a) for the Environmental Public Health Tracking Network (Figure 27.2) that will integrate data about environmental hazards and exposures with data about diseases that are possibly linked to the environment. This system will allow federal, state, and local agencies, and others to do the following:

- Monitor and distribute information about environmental hazards and disease trends
- Advance research on possible linkages between environmental hazards and disease
- Develop, implement, and evaluate regulatory and public health actions to prevent or control environment-related diseases

In FY 2002, CDC received $17.5 million from the U.S. Congress to begin the first health-tracking pilot projects in targeted areas around the country. Grants were awarded to 17 states, 3 local health departments, and 3 schools of public health to begin development of a national, environmental public health tracking network and to develop environmental health capacity in state and local health departments.

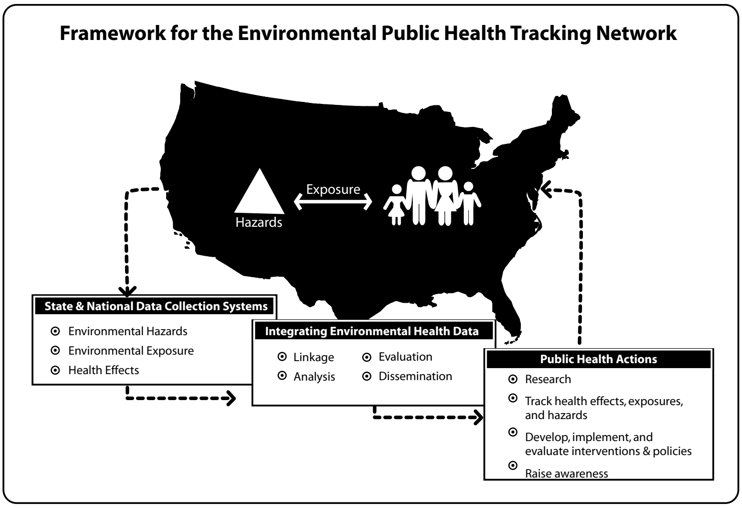


FIGURE 27.2 CDC's Framework for the Environmental Public Health Tracking Network. (http://www.cdc.gov/nceh/tracking/EPHT_diagram.htm)

The focus of this effort is on chronic disease and other noninfectious health effects that may be related to environmental exposures (Wiken 1986).

The Public Health Information Network will enable consistent use, security, and exchange of response, health, and disease tracking data between public health partners. This is supported though the Public Health Information Network's five key components: detection and monitoring, data analysis, knowledge management, alerting, and response (CDC 2003a).

27.3.3 Indicator Frameworks

Much of the early work on environmental monitoring focused on the condition or state of the environment but did not provide information about cause, associations, or management actions. The Pressure–State–Response (PSR) Framework and Driving Forces–Pressures–State–Exposure–Effects–Action (DPSEEA) Framework provide two more comprehensive perspectives.

27.3.4 Pressure–State–Response (PSR) Framework

The Organization for Economic Cooperation and Development (OECD) developed this framework (Figure 27.3) to link pressures on the environment as a result of human activities, with changes in the state (condition) of the environment (land, air, water, etc.). Society then responds to these changes by instituting environmental and economic programs and policies, which feed back to reduce or mitigate the pressures or repair the natural resource (OECD 1993).

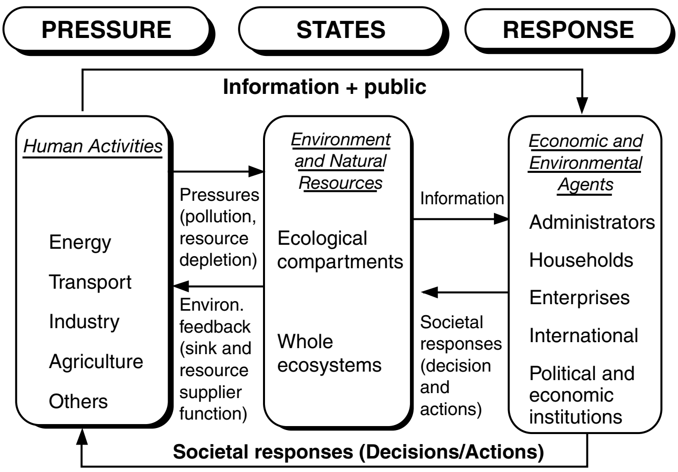


FIGURE 27.3 Pressure–state–response (PSR) framework.

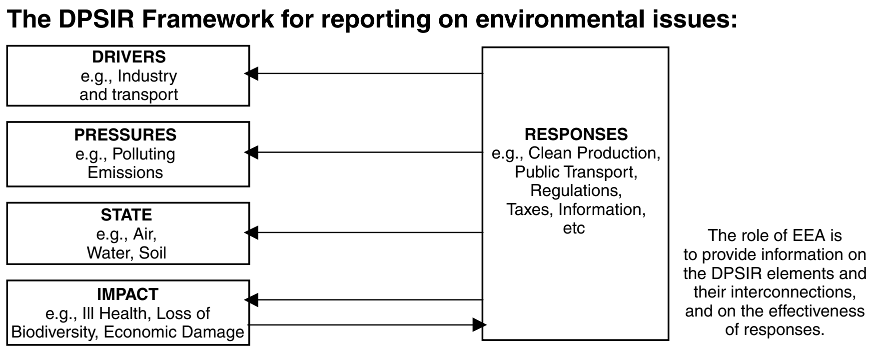


FIGURE 27.4 The DSPIR Framework (From European Environmental Agency, 1999, Conceptual Framework: How We Reason.)

The PSR framework is now widely used but is continuing to evolve. One of the main challenges has been differentiating between pressures and the state of the environment.

A development of PSR has been the Driving Force–Pressure–State–Impact–Response Framework (DPSIR) adopted by the European Environment Agency (EEA). The framework (Figure 27.4) assumes cause–effect relationships between interacting components of social, economic, and environmental systems, which are

- Driving forces of environmental change (e.g., industry and transport)
- Pressures on the environment (e.g., discharges of wastewater or polluting emissions)
- State of the environment (e.g., water quality in rivers and lakes)

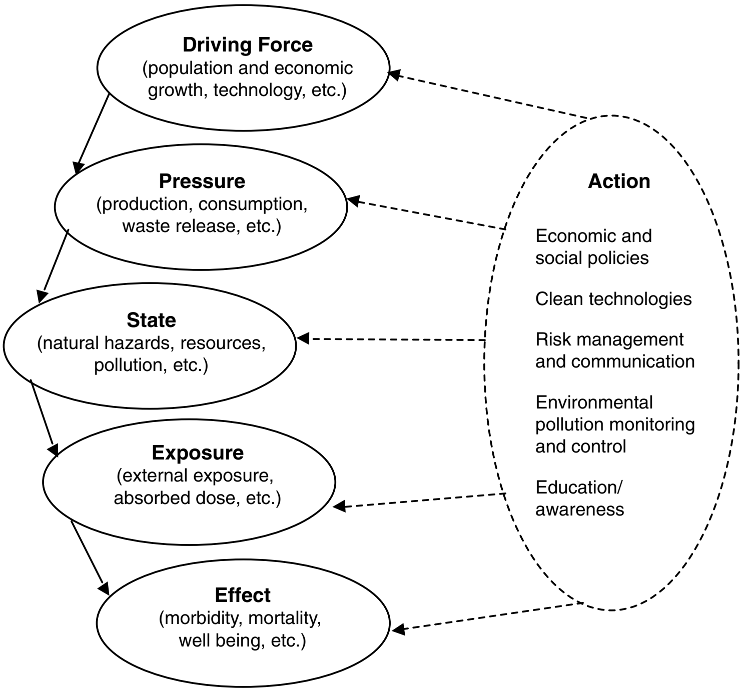


FIGURE 27.5 The Driving Forces–Pressure–State–Exposure–Effect–Action (DPSEEA) model of WHO.

- Impacts on human health, population, economy, ecosystems (e.g., water unsuitable for drinking)
- Response of the society (e.g., regulations, information, and taxes)

27.3.5 Driving Forces–Pressures–State–Exposure–Effects–Action (DPSEEA) Framework

A model developed at the World Health Organization (WHO) (Corvalan et al. 1996) took a broader approach to include macrodriving forces in the pressures on health and the environment. The DPSEAA model (Figure 27.5) is useful as it covers the full spectrum of potential forces and resulting actions and brings together professionals, practitioners, and managers from both environmental and public health fields to help orient them in the larger scheme of the problem.

27.4 EXAMPLES OF CURRENT MAJOR ENVIRONMENTAL MONITORING NETWORKS

This section will provide a brief overview of representative major environmental monitoring networks. The following chapters will provide detailed discussion of additional networks.

27.4.1 International Networks

There are several international networks that have been established to monitor systems and processes that are global (e.g., climate change and oceanic circulation) or cross-national borders (e.g., bird migrations).

27.4.2 Global Observing Systems (GOS)

In the 1980s several United Nations agencies began to develop a comprehensive, long-term, global monitoring program to observe phenomena related to climate change. They took advantage of ongoing international, regional, and national programs in this area and in 1990 proposed a long-term global monitoring system of coastal and near-shore phenomena related to global climate changes. Today the GOS is made up of three interrelated observing systems:

1. The **Global Climate Observing System (GCOS)** was established in 1992 to provide the comprehensive observations required for detecting and attributing climate change, for assessing the impacts of climate variability and change, and for supporting research toward improved understanding, modeling, and prediction of the climate system. It addresses the total climate system including physical, chemical, and biological properties, and atmospheric, oceanic, hydrologic, cryospheric, and terrestrial processes. It is sponsored by the World Meteorological Organization (WMO), the Intergovernmental Oceanographic Commission (IOC) of the United Nations Educational, Scientific, and Cultural Organization (UNESCO), the International Council for Science (ICSU) and the United Nations Environment Programme (UNEP). GCOS consists of three main networks:
   - The **GCOS Surface Network (GSN)** formalized in 1999 is a global network of approximately 1000 stations with approximately 50 member states of the WMO collecting surface meteorological observations including temperatures, pressures, winds, clouds, and precipitation.
   - The **GCOS Upper-Air Network (GUAN)** is a subset of 150 stations (of a total of about 900) in the upper air network of World Weather Watch Global Observing System. It was developed to address the requirement for a consistent baseline of homogeneous measurements for global climate. GUAN observations include measurements of pressure, wind velocity, temperature, and humidity from just above the ground to heights of up to 30 km (using radiosondes attached to free-rising balloons).
   - The **WMO Global Atmosphere Watch (GAW)** is considered the atmospheric chemistry component of the GCOS. The network of measurement stations consists of GAW Global and Regional measurement stations which are operated by the National Meteorological Services or other national scientific organizations of the host countries. More than 65 countries actively host GAW stations. In recent years satellite programs have produced important measurements of atmospheric

compounds and related parameters that complement the GAW network measurements. When highly accurate local measurements from GAW ground-based stations are coupled with the near global coverage of satellite measurements it results in a more complete picture of atmospheric composition and processes on global scales, and provides complementary checks of instrument calibrations.

2. The **Global Terrestrial Observing System (GTOS)** was established in 1996 to provide observations of long-term changes in land quality, availability of freshwater resources, pollution and toxicity, loss of biodiversity, and climate change. It is sponsored by the Food and Agriculture Organization of the United Nations (FAO), ICSU, UNEP, UNESCO, and WMO. GTOS employs satellite and supporting ground observations on global ecosystems. The system of observation networks includes thematic networks (e.g., net primary production, terrestrial carbon observation), habitat type (e.g., coastal, mountains) or region (e.g., southern Africa, Central Europe).
3. The **Global Ocean Observing System (GOOS)** is a sustained, coordinated international system for gathering information about the oceans and seas of the Earth. The IOC created GOOS in 1991 in response to the desire of many nations to improve management of seas and oceans and to improve climate forecasts for both of which it is necessary to establish observations dealing with physical, chemical, and biological aspects of the ocean in an integrated way. It is sponsored by the IOC, ICSU, UNEP, and WMO.

   Monitoring activities include 6500 volunteer merchant vessels observing meteorology and surface oceanography, 120 volunteer vessels observing subsurface temperature and salinity, 1400 drifters observing meteorology and surface oceanography, hundreds of moored ocean buoys, 10 to 20 volunteer merchant and research vessels making upper atmosphere vertical soundings, 3000 diving profilers which collect vertical profiles of upper ocean temperature and salinity, 400 tidal stations for sea level, and arctic and Antarctic ice monitoring. Together, these monitoring activities provide the information needed by governments, industry, science, and the public to deal with marine-related issues, including the present and future condition of the seas and oceans, and the effects of the ocean upon climate (GOOS 1998).

   At this time, these measurement activities are concerned primarily with physical observations. However, consideration is now being given to what chemical and biological information is required and how to integrate it with physical data. The challenge is to develop a high-quality, integrated approach to coastal monitoring and forecasting, taking into consideration the needs of resource managers. Examples of observing systems currently under consideration include:

   - Harmful Algal Bloom (HAB) program of the IOC
   - International Mussel Watch program
   - Marine Pollution and Monitoring Program (MARPOLMON)
   - Continuous Plankton Recorder (CPR) program

> Data from these systems will be supplemented by global coverage of the surface skin of the ocean from satellites, though none of these is operational (GOOS 1998).

It is envisioned that there will be regional centers that will process the global data streams and distribute products tailored to interested clients. This facilitates data collection that supports the global scale GOS, GCOS, and GTOS efforts and also transmits derived products and information to regional and local clients. For the monitoring networks to be maintained it is essential that there be global, regional, and local clients. An example is the Gulf of Maine Ocean Observing System (GoMOOS) which is designed to provide real-time and archived weather and oceanographic information to a variety of users in the Gulf of Maine region—shippers, fishermen, recreational boaters, managers, researchers, educators, and the general public.

In 1980, the U.S. National Science Foundation (NSF) established the **Long Term Ecological Research (LTER)** Network to support research on long-term ecological processes. The goals of the LTER Network are to understand ecological phenomena that occur over long temporal and broad spatial scales; create a legacy of well-designed and documented ecological experiments; conduct major syntheses and theoretical efforts; and provide information necessary for the identification and solution of environmental problems.

The conceptual frameworks of the LTER sites are broadly focused around five core areas: pattern and control of primary production, spatial and temporal distribution of populations selected to represent trophic structure, pattern and control of organic matter, pattern of inorganic input and nutrient movement through the system, and patterns, frequency, and effects of disturbance to the research site (NSF 2002). The LTER has developed 24 sites: 19 in the lower 48 states plus two sites in Alaska, two in Antarctica, and one in Puerto Rico.

In 1993, the LTER Network hosted a meeting on international networking in long-term ecological research. Representatives of scientific programs or networks that focus on ecological research over long temporal and large spatial scales formed the International Long-Term Ecological Research (ILTER) Network. As of January 2003, 25 countries have established formal national LTER Network programs and joined the ILTER Network (Australia, Brazil, Canada, China, Costa Rica, Colombia, Czech Republic, France, Hungary, Israel, Korea, Mexico, Mongolia, Namibia, Poland, Slovakia, South Africa, Switzerland, Ukraine, United Kingdom, United States, Uruguay, Venezuela, and Zambia). Several more are actively pursuing the establishment of national networks (Argentina, Austria, Ireland, Italy, Japan, Tanzania, and Vietnam) and many others have expressed interest in the model (Chile, Croatia, Ecuador, India, Indonesia, Kenya, Norway, Portugal, Romania, Slovenia, Spain, and Sweden) (Krishtalka et al. 2002). The ILTER is an example monitoring network of GTOS.

The North American Breeding Bird Survey (BBS), a large-scale survey of North American birds, is jointly coordinated by the USGS Patuxent Wildlife Research Center and the Canadian Wildlife Service, National Wildlife Research Center. It is a roadside survey, primarily covering the continental U.S. and southern Canada, although survey routes have recently been initiated in Alaska and northern Mexico.

The BBS was started in 1966, and over 4100 routes are surveyed in June by experienced birders.

The BBS was designed to provide a continent-wide perspective of population change. Routes are randomly located in order to sample habitats that are representative of the entire region. Other requirements such as consistent methodology and observer expertise, visiting the same stops each year, and conducting surveys under suitable weather conditions are necessary to produce comparable data over time. A large sample size (number of routes) is needed to average local variations and reduce the effects of sampling error (variation in counts attributable to both sampling technique and real variation in trends) (Sauer et al. 1997).

Participants skilled in avian identification collect bird population data along roadside survey routes. Each survey route is 24.5 mi long with stops at 0.5-mile intervals. At each stop, a 3-min point count is conducted. During the count, every bird seen within a 0.25-mi radius or heard is recorded. Surveys start one half hour before local sunrise and take about 5 h to complete. The survey routes are located across the continental U.S. and Canada; however, the density of BBS routes varies considerably across the continent, reflecting regional densities of skilled birders. The greatest densities are in the New England and Mid-Atlantic states, while densities are lower elsewhere (USGS 2001).

Once analyzed, BBS data provide an index of population abundance that can be used to estimate population trends and relative abundances at various geographic scales. Trend estimates for more than 420 bird species and all raw data are currently available via the BBS Web site (Sauer et al. 2002). The BBS is not a component of any GOS network.

27.4.3 National Networks

Canada established the Ecological Monitoring and Assessment Network (EMAN) in 1994 as a national network to provide an understanding and explanation of observed changes in ecosystems. When formally established, the mandate was to coordinate integrated ecosystem monitoring and research to provide an understanding and explanation of observed changes in ecosystems (Environment Canada 2002).

EMAN is a cooperative partnership of federal, provincial, and municipal governments, academic institutions, aboriginal communities and organizations, industry, environmental nongovernment organizations, volunteer community groups, elementary and secondary schools, and other groups/individuals involved in ecological monitoring. EMAN is a voluntary collaboration built upon existing programs, which applies standardized monitoring protocols and data management standards. This standardization allows for data integration and assessment.

Canada has developed a hierarchical classification framework that correlates across all political boundaries and provides several levels of generalization nationally (e.g., ecozones, ecoregions, ecosections). Ecological land classification is a process of delineating and classifying ecologically distinctive areas which have resulted from the mesh and interplay of the geologic, landform, soil, vegetative, climatic, wildlife, water, and human factors present. This holistic approach can be applied incrementally on a scale-related basis from site-specific ecosystems to very broad ecosystems (Wiken 1986).

Results of site monitoring, environmental assessment, or inventories can be compared with data from similar ecosystems in other parts of the country and can be reported at a variety of scales. The framework is being used by the Canadian Environmental Assessment Act (CEAA) and Canadian Environmental Protection Act (CEPA) to organize and report information requirements at the broad ecozone level. Protected areas strategies developed by both government and nongovernment organizations are being structured on an ecosystem, rather than political jurisdiction basis (Environment Canada 2003).

EMAN will provide better understanding of ecological change in response to stressors, and will contribute to the design of scientifically defensible pollution control and management programs and the evaluation of their effectiveness.

27.4.4 Monitoring and Research in the U.S.

The U.S. federal government spends more than $600 million per year collecting environmental data and, through regulatory requirements, imposes additional costs on the private sector for monitoring of emissions and effluents (National Science and Technology Council 1997b). A variety of organizations collect data, using different methods. Often the data are collected only for limited geographic areas or are not collected in a statistically representative manner. The lack of standardization of data or monitoring methods makes correlation across databases or across spatial scales problematic.

The need for an integrated monitoring system has been recognized by the CEQ in 1991 (CEQ 1992), the CENR in 1996 (National Science and Technology Council 1997b), and most recently by the Heinz Center (Heinz 2003). CENR proposed a three-tiered monitoring and research framework that would be:

- A collaborative effort (federal, state, tribal, local, private, and international)
- Built upon existing networks and programs
- Cost-efficient
- Continuous over the long run
- Adaptive so that it can evolve and innovate without losing the value of historical datasets
- Accessible to all public and private sectors

The proposed three monitoring tiers are those described earlier in the chapter, and as implemented in Canada. Such a network has not yet been implemented in the U.S. Existing environmental monitoring networks generally fall into two categories: regulatory or natural resource.

27.4.5 Example U.S. Regulatory Programs

27.4.5.1 National Air Quality

U.S. air quality programs are designed primarily to protect public health (Savage 2002). Eight federally funded national air quality monitoring networks are currently operating in the U.S.: Interagency Monitoring of Protected Visual Improvements

(IMPROVE), State and Local Air Monitoring Stations/National Air Quality Monitoring Stations (SLAMS/NAMS), Photochemical Assessment Monitoring Stations (PAMS), National Atmospheric Deposition Monitoring Stations/National Trends Network (NADP/NTN), National Atmospheric Deposition Monitoring Stations/Mercury Deposition Network (NADP/MDN), Clean Air Status and Trends Networks (CASTNet), Atmospheric Integrated Research Monitoring Network (AIRMoN), and Gaseous Polluted Monitoring Program (GPMP).

While support for these air monitoring networks has been sporadic over the past 25 years, the combined data from these networks provide documentation of trends in air quality and support evaluation of the effectiveness of control strategies (National Science and Technology Council, 1999). Data collected include visibility, nitrogen species, sulfur species, lead, particulate matter, carbon monoxide, ozone and/or precursors, and atmospheric deposition (wet and dry). These programs show that the U.S. air quality is generally improving (U.S. Environmental Protection Agency, 2003b).

More detailed information about CASTNet follows in Chapter 31.

27.4.5.2 National Water Quality

National water quality is significantly more difficult to measure than air quality. There is no single body for the entire planet but instead an intricate, branching web of brooks, ponds, bogs, groundwater, beaches, springs, swamps, streams, seeps, wetlands, rivers, marshes, estuaries, lakes, bays, etc. Water quality programs have to not only protect public health but also preserve ecosystems and ecosystem values and functions (Savage 2002).

The responsibility to monitor water quality rests with many different agencies. State pollution control agencies and Indian tribes have key monitoring responsibilities and conduct vigorous monitoring programs. Interstate commissions may also receive grants and maintain monitoring programs. Many local governments such as city and county environmental offices also conduct water quality monitoring. The EPA administers grants for water quality monitoring and provides technical guidance on how to monitor and report monitoring results.

States have applied a range of monitoring and assessment approaches (e.g., water chemistry, sediment chemistry, biological monitoring) to varying degrees, both spatially and temporally, and at varying levels of sampling effort. It is not uncommon for the reported quality of a water body (i.e., attainment or nonattainment) to differ on either side of a state boundary. Although some differences can be attributed to differences in water quality standards, variations in data collection, assessment methods, and relative representativeness of the available data contribute more to differences in assessment findings. These differences adversely affect the credibility of environmental management programs (U.S. EPA 2003a).

In July 2002, EPA issued guidance on the basic elements of a state water monitoring program, which serves as a tool to help EPA, and the states determine whether a monitoring program meets the prerequisites of the Clean Water Act. The ten basic elements of a state water quality monitoring program are:

- A comprehensive monitoring program strategy that serves its water quality management needs and addresses all state waters, including streams, rivers, lakes, the Great Lakes, reservoirs, estuaries, coastal areas, wetlands, and groundwater.
- Monitoring objectives critical to the design of a monitoring program that are efficient and effective in generating data which serve management decision needs.
- An integration of several monitoring designs (e.g., fixed station, intensive and screening-level monitoring, rotating basin, judgmental, and probability design) to meet the full range of decision needs.
- Core indicators (physical/habitat, chemical/toxicological, and biological/ecological endpoints as appropriate) selected to represent each applicable designated use, plus supplemental indicators selected according to site-specific or project-specific decision criteria.
- Peer-reviewed quality management plans and quality assurance program/project plans.
- An accessible electronic data system for water quality, fish tissue, toxicity, sediment chemistry, habitat, biological data, with timely data entry (following appropriate metadata and state/federal geo-locational standards) and public access. In the future, EPA will require all states to directly or indirectly make their monitoring data available through the new STORET system.
- A methodology which includes compiling, analyzing, and integrating all readily available and existing information for assessing attainment of water quality standards based on analysis of various types of data (chemical, physical, biological, land use) from various sources.
- Timely and complete water quality reports and lists called for under the Clean Water Act and the Beaches Act.
- Periodic reviews of each aspect of its monitoring program to determine how well the program serves its water quality decisions.
- Identification of current and future resource needs (funding, staff, training, laboratory resources, and upcoming improvements) to fully implement its monitoring program strategy (U.S. EPA 2003a).

This guidance should improve how monitoring is conducted, how information is shared, and how decisions based on monitoring are made. The intended end result will be a comprehensive, integrated national water quality monitoring network.

27.4.6 Example: U.S. Natural Resource Programs

There are a multitude of monitoring networks collecting data about natural resources in the U.S. Several national-scale networks are discussed below:

The *National Wetlands Inventory (NWI)* of the U.S. Fish and Wildlife Service produces information on the characteristics, extent, and status of the nation's wetlands habitats. The digital NWI maps are generated every 10 years from satellite images. NWI has mapped 90% of the lower 48 states, and 34% of Alaska. About 44% of the lower 48 states and 13% of Alaska are digitized. In 1982, the NWI produced the first comprehensive and statistically valid estimate of the status of the Nation's wetlands and wetland losses, and in 1990 produced the first update. Future national updates have been scheduled for 2000, 2010, and 2020. NWI maintains a database of metadata containing production information, history, and availability of all maps and digital wetlands data produced by NWI. This database is available over the Internet (U.S. F&WS 2003).

The *National Water Quality Assessment (NAWQA) Program* of the USGS collects and analyzes data and information in more than 50 major river basins and aquifers across the U.S. The goal is to develop long-term consistent and comparable information on streams, groundwater, and aquatic ecosystems to support sound management and policy decisions. NAWQA study units frequently cross state boundaries and usually encompass more than 10,000 km^2 (about 3,900 mi^2). Geographic areas were selected to represent a variety of important hydrologic and ecological resources; critical sources of contaminants, including agricultural, urban, and natural sources, and a high percentage of population served by municipal water supply and irrigated agriculture. NAWQA does not provide full coverage of the nation's fresh surface waters.

Each NAWQA study unit adheres to a nationally consistent study design and uniform methods of collecting data on pesticides, nutrients, volatile organic compounds, ecology, and trace elements. This consistency facilitates data aggregation and synthesis. Studies are long-term and cyclical. One third of all study units are intensively investigated at any given time for 3 to 4 years, followed by low-intensity monitoring. Trends are assessed about every 10 years and synthesized into national reports (Gilliom et al. 2001).

The *National Resources Inventory (NRI)* is a statistically based sample of land use and natural resource conditions and trends on U.S. nonfederal lands—about 75% of the total land area. It is the most comprehensive database of its kind ever attempted anywhere in the world.

The NRI is unique because

- It provides a nationally consistent database for all nonfederal lands.
- It features data gathered and monitored in 1982, 1987, 1992, and 1997 by thousands of technical and natural resource data collection experts.
- It has a direct correlation with soils data, which permits analysis of resources in relation to the capability of the land and in terms of soil resources and conditions.

Conducted by the U.S. Department of Agriculture's Natural Resources Conservation Service (NRCS) in cooperation with the Iowa State University Statistical Laboratory, this inventory captures data on land cover and use, soil erosion, prime farmland soils, wetlands, habitat diversity, selected conservation practices, and related resource attributes. Data are collected every 5 years from the same 800,000 sample sites in all 50 states, Puerto Rico, the U.S. Virgin Islands, and some Pacific Basin locations. The NRI is a statistically based survey that has been designed and implemented using scientific principles to assess conditions and trends of soil, water, and related resources.

The NRI is conducted to obtain scientific data that is valid, timely, and relevant on natural resources and environmental conditions. Through legislation — the Rural Development Act of 1972, the Soil and Water Resources Conservation Act of 1977, and other supporting acts — Congress mandates that the NRI be conducted at intervals of 5 years or less (USDA 2002).

In 1984, NOAA initiated the *National Status and Trends (NS&T) Program* to determine the current status of, and to detect changes in, the environmental quality of the nation's estuarine and coastal waters. The NS&T Program is comprised of three main projects: The Benthic Surveillance, The Mussel Watch, and Bioeffect Assessments (NOAA 1998).

The NS&T (1) conducts long-term monitoring of contaminants (analytes) and other environmental conditions at more than 350 sites along U.S. coasts, (2) studies biotic effects intensively at more than 25 coastal ecosystems, (3) partners with other agencies in a variety of environmental activities, and (4) advises and participates in local, regional, national, and international projects related to coastal monitoring and assessment. Both spatial and temporal trends can be obtained from the data.

The NS&T Program has a performance-based quality assurance program, and no analytical methods are specified. As a result, methods are allowed to change with time as newer, more efficient, and less costly analytical technologies become available. The minimum detection limits used in the NS&T Program have been thoroughly defined.

The *Gap Analysis Program (GAP)* is sponsored and coordinated by the USGS Biological Resources Division (BRD). GAP is designed to provide state- and national-level information on the status of ordinary species (those not threatened with extinction or naturally rare) and their habitats (Jennings and Scott 1997). GAP is mapping existing natural vegetation, predicted distribution of native vertebrate species, ownership of public land and private conservation lands. Vegetation is mapped from satellite imagery and other records using the National Vegetation Classification System (FGDC 1996).

GAP is currently made up of over 445 contributing organizations in 44 states. Contributors include business, universities, state and federal agencies, tribes, and nongovernmental organizations. National GAP programs are emerging in both Mexico and Canada. The U.S. GAP has been assisting counterparts in both of these countries and expects to continue to foster the development of standardized biogeographic data as well as standardized TM imagery for North America.

Multi-Resolution Land Characteristics (MRLC) Consortium. Due to the escalating costs of acquiring satellite images, in 1992 several federal agencies agreed to operate as a consortium in order to acquire satellite-based remotely sensed data for their environmental monitoring programs. Primary administration of the consortium is performed by the USGS Earth Resources Observation Systems Data Center (EROS) in Sioux Falls, South Dakota. Image processing and land-cover classification are performed at EROS, while accuracy assessments of the classification are primarily handled by member federal agencies (Vogelmann et al. 2001).

During the 1990s, the MRLC resulted in several successful mapping programs, including the Coastal Change Analysis Project (C-CAP) administered by NOAA, Gap Analysis Program (GAP) directed by the USGS-BRD, and the National Land Cover Data (NLCD) project directed by both the USGS and EPA. The data developed by these projects are available publicly either as download or by contacting the agencies involved. The consortium is currently pursuing a second nationwide acquisition of images (Wickham et al. 2002).

27.4.7 Monitoring and Research in Europe

There are a variety of environmental monitoring programs which cover multiple countries in Europe. One such program, the Cooperative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (EMEP) was implemented in 1978 under the United Nations Economic Commission for Europe (UNECE) and, today, covers more than 100 air quality monitoring stations in 35 countries.

Initially EMEP was focused on the effects of acidification, and has been expanded to include nitrogen, base cations, and ozone. Each country runs its own network and reports its results to a central data and assessment center.

The EMEP monitoring network provides high quality data on the state of the environment for model validation and national air quality assessments, national involvement, and for independent validation of abatement measures. Measurements are made using comparable and reliable methodologies that are fairly simple, robust, and cost-efficient. In addition, these have the advantage of being fairly easy to assess in terms of precision and reproducibility (Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe 2002).

Over time, new topics and priorities in air pollution policies emerged but standard monitoring protocols were not yet established. In April 2003 the EMEP Task Force on Measurements and Modeling developed a draft monitoring strategy (UNECE 2003) to respond to these shortcomings. One major objective of the new strategy is to more clearly identify the monitoring requirements needed to improve air quality in Europe by defining minimum monitoring requirements for each country and also to more directly describe how the EMEP monitoring efforts relate to other ongoing European monitoring efforts.

27.5 HUMAN HEALTH MONITORING

Public health agencies draw upon a population tracking approach that considers human health impacts in a broader context. This approach integrates public health sciences such as epidemiology, pharmacology, toxicology, and relies on hazard, exposure, and health outcome tracking data to guide decisions. Historically, public health agencies have focused their public health surveillance systems on the collection, analysis, and dissemination of data to prevent and control disease. Public health surveillance is particularly important for early detection of epidemics, as well as monitoring and evaluating disease prevention and control programs. The recent SARS outbreak and the AIDS epidemic underscore this essential role of public health surveillance.

Surveillance data is generally provided by health care professionals. A promising new approach, syndromic surveillance, is an investigational approach to early detection of outbreaks through the monitoring of real-time, electronic data that are screened for indicators of disease as early in the course of illness as possible. Although promising, this approach to public health surveillance has not undergone rigorous evaluation and validation for its usefulness and value (Fleming 2003).

WHO is working with its partners to create a comprehensive global network that detects and controls local outbreaks before they grow into worldwide pandemics. Currently, there are programs in 30 countries throughout the world that support disease detection activities and provide an essential link in global surveillance. Additionally, there is a concerted effort to develop and expand other fledgling regional disease surveillance networks that include less developed nations as members. Currently, there are networks in the Caribbean, South America, Africa, and Southeast Asia. As additional networks are developed, there will ultimately be a global health surveillance system that monitors priority diseases of global concern, including pandemic influenza, drug-resistant diseases, and diseases caused by biological agents. This global health surveillance system will be the human health equivalent of the GOS.

Improved disease surveillance does not entirely fill the environmental health gap identified by the Pew Environmental Health Commission (Pew 2000): the lack of basic information that could document the possible links between environmental hazards and chronic disease and the lack of critical information needed to reduce and prevent these health problems. The Environmental Public Health Tracking Network (described previously) will focus on chronic disease and other noninfectious health effects that may be related to environmental exposures (Qualters 2003). In the U.S., CDC and EPA have formed a collaborative partnership to increase collaboration and formal partnerships between traditional health and environmental entities at the national, state, and local levels.

While the Environmental Public Health Tracking Network will focus on monitoring health effects (disease and illness), additional information is needed to effectively link environmental condition and human health. Ideally, human health monitoring would collect data on each component of the DPSEEA.

Air and water quality monitoring programs can provide some of this information; however, until recently there were no national surveillance programs to track exposure to pollutants. The CDC's National Health and Nutritional Examinations Survey (NHANES) is beginning to help fill that gap.

Since 1999 NHANES has been an annual survey that employs a complex, stratified, multistage, probability-cluster design to select a representative sample of the civilian, noninstitutionalized population in the U.S. As part of the examination component, blood and urine samples are collected from participants and measured for environmental chemicals. Most of the environmental chemicals were measured in randomly selected subsamples within specific age groups to obtain an adequate quantity of sample for analysis and to accommodate the throughput of the mass-spectrometry analytical methods.

The first use of this sophisticated survey for environmental purposes occurred with NHANES II which operated from 1976 to 1980 (Murphy et al. 1983 and Kutz et al. 1992b). NHANES III (1988 to 1994) collected exposure information for selected pesticides and toxic substances. The results indicate a higher apparent pesticide burden for respondents in the South and Midwest (Allen et al. 2003). Because samples were mainly obtained from highly urban locations, exposures among the rural population living in agricultural areas were difficult to estimate.

Furthermore, sampling did not always occur in high-pesticide-use seasons in all areas surveyed, limiting the ability to make regional comparisons and possibly underestimating exposures in some areas.

CDC published the *National Report on Human Exposure to Environmental Chemicals (First Report)* in March 2001. The *Second National Report on Human Exposure to Environmental Chemicals* was published January 2003 and presents biomonitoring exposure data for 116 environmental chemicals for the noninstitutionalized, civilian U.S. population over the 2-year period of 1999 to 2000.

The use of biomonitoring by NHANES represents a significant step forward in capturing environmental exposure data. Biomonitoring, which involves taking samples from people to measure individual exposure, provides one approach to documenting the links between environmental hazards and disease. Residues of xenobiotic chemicals in human tissue and fluids have been used in the traditional risk assessment paradigm to assist in the estimation of exposure for both the general population and subpopulations (infants, elderly, etc.).

Environmental justice issues also are highlighted by human monitoring information. According to EPA (1995), the concept of environmental justice denotes that no population of people because of their race, ethnicity, income, national origin, or education should bear a disproportionate share of adverse environmental burdens or hazards.

Monitoring data, particularly from human studies, have been used to set regulatory priorities. Chemicals found to accumulate in adipose tissue have been regulated (Kutz et al. 1991). Monitoring data also can be utilized to determine the efficacy of environmental policies. In the years following the regulation of DDT, a national survey found a reduction of DDT residues in human adipose tissue (Kutz et al. 1977). Likewise, reduction of polychlorinated biphenyls residues in human adipose tissue was observed following the regulation of these chemicals under the Toxic Substances Control Act (U.S. EPA 1985).

The recently released Draft Report on the Environment (U.S. EPA 2003b) confirms that "more monitoring is required, along with more effective means to link ambient exposures to health effects."

27.6 ADDITIONAL FACTORS CRITICAL TO EFFECTIVE MONITORING NETWORKS

In order for managers to most effectively utilize a well-designed, well-implemented monitoring network, they must invest in several other factors, including information management and communication.

27.6.1 Information Management

Environmental managers face a wide array of challenges and need diverse and detailed information. No single organization can conduct wall-to-wall multidisciplinary studies or maintain a comprehensive data management system with all the necessary data. Sharing environmental monitoring data maximizes its usefulness by providing greater spatial, temporal, and disciplinary coverage than individual organizations can offer.

The Internet has made it easier to publish, find, and access data. Standards for data and metadata, essential for data exchange and integration, are emerging. These include standard names for chemical substances and biological organisms, and state names and abbreviations. The best standards are those that will be maintained by organizations willing to make long-term commitments (Hale et al. 2003).

Data integration is inextricably linked to program planning and objectives, all aspects of sampling and analysis, and the various methods and procedures employed in the analysis and interpretation of the resulting data. During the program-planning phase, requirements for databases and data management should be developed, including acceptable data and metadata characteristics, formats, data ownership, and accessibility. Continuous feedback should be obtained from users of the resulting databases with regard to their accessibility, utility, and any problems encountered, and appropriate changes should be made as required (NRC 1995).

27.6.2 Communications

The complexity of scientific data often limits its usefulness as a tool for affecting policy. In the courtroom, the meeting room, Congress or Parliament, or the voting booth, the messages carried by scientific evidence are effective only when they are conveyed clearly and simply. The need to display intricate scientific environmental data in ways that the public can grasp has become urgent enough that we need better ways to help the nonscientist understand the science (Bradley et al. 2000). Key points to remember are:

- Target your information.
- Emphasize comparisons and show trends.
- Use headlines.
- Focus on findings and implications.
- Provide geographic context.
- Include recommendations.
- Use graphics.

Chapter 28 on the South African River Health Program illustrates how these factors have contributed to a successful major monitoring network.

27.7 FUTURE POTENTIAL DEVELOPMENTS

Individual monitoring networks are evolving to provide information needed by decision-makers. "Recent trends in environmental health, ecology and health, and human ecology all suggest that the interface between sustainability, ecosystems, social systems, and health is fertile ground for optimizing environmental health interventions and maximizing public health gain" (Parkes et al. 2003).

Ecosystem health (the functioning and performance of ecosystems) and human health (the functioning and well-being of the dominant species in the current global ecosystem) should be the goal of international monitoring efforts in the future. Success depends on linking the many existing monitoring networks into a larger global network that can provide information on the complex linkages and interdependencies between biodiversity (the species and genetic and landscape diversity that make up the structure of ecosystems) and the health of ecosystems and humans.

27.8 SUMMARY

This chapter has covered the challenges managers face when developing and implementing a major monitoring network: the complexity of many interacting components; the need for a monitoring design that is statistically sound, effective, and efficient; and the need to link environmental monitoring and human health. Several conceptual models were presented: the CENR Framework, CDC's Framework for the Environmental Public Health Network, and the PSR, DPSIR, and DPSEEA indicator frameworks.

The chapter also provided a brief overview of international and national-scale environmental monitoring networks and human health monitoring. The need for sound information management and communications was emphasized, and possible future developments were briefly described.

Environmental data are available, but the quality of some is unknown, and large data gaps exist both in geographic area and scientific substance. The future appears bright and exciting for further integration of human and environmental monitoring data and other information into decision-making.

ACKNOWLEDGMENT

Although the research described in this chapter has been funded in part by the U.S. Environmental Protection Agency, it has not been subjected to Agency level review. Therefore, it does not necessarily reflect the views of the Agency. This chapter is EPA/ORD/NHEERL Contribution No. AED-03-084.

REFERENCES

Allen, R.H., M. Ward, G. Gondy, D.T. Mage, and M.C. Alavanja. 2003. Regional, Seasonal, and Ethnic Differences in the NHANES III Pesticide Epidemiology Study. Seminar at the National Center for Health Statistics May 28, 2003. (Contact: Charles Croner at cmc2@cdc.gov)

Bradley, M.P., B.S. Brown, S.S. Hale, F.W. Kutz, R.B. Landy, R. Shedlock, R. Mangold, A. Morris, W. Galloway, J.S. Rosen, R. Pepino, and B. Wiersma. 2000. Summary of the MAIA working conference. *Environ. Monit. Assess.* 63: 15–29.

Burke, T. 1997. Impediments to a Public Health Approach. Comments made at the Symposium on a Public Health Approach to Environmental Health Risk Management. Washington, D.C.

Centers for Disease Control. 2003a. Programs in Brief. The Public Health Information Network. URL: http://www.cdc.gov/programs/research8.htm. February 2003.

Centers for Disease Control. 2003b. CDC's Environmental Public Health Tracking Program — Background. URL: http://www.cdc.gov/nceh/tracking/background.htm. February 05, 2003.

Corvalan, C., D. Briggs, and T. Kjellstrom. 1996. Development of environmental health indicators. In *Linkage Methods for Environment and Health Analysis*. Briggs, D., Corvalan, C., and Nurminen, M., Eds., Office of Global and Integrated Environmental Health, World Health Organization, Geneva, Chap. 2.

Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (EMEP). 2002. What is EMEP? URL: http://www.emep.int/index_facts.html.

Council on Environmental Quality. 1992. Environmental Quality, 22nd Annual Report. Council on Environmental Quality, Washington, D.C., 382 pp.

Environment Canada. 2002. The ecological monitoring and assessment network: all about EMAN, in Information for Improving Europe's Environment. URL: http://eqb-dqe.cciw.ca/eman/program/about.html.

Environment Canada. 2003. A National Ecological Framework for Canada. http://www.ec.gc.ca/soer-ree/English/Framework/default.cfm.

European Environmental Agency. 1999. Conceptual Framework: How We Reason. URL: http://org.eea.int/documents/brochure/brochure reason.html

Federal Geographic Data Committee (FGDC). 1997. FGDC vegetation classification and information standards—June 3, 1996 draft. Federal Geographic Data Committee Secretariat, Reston, VA.

Fleming, D., Statement on CDC's Public Health Surveillance Activities before the United States House of Representatives Committee on Government Reform, Subcommittee on National Security, Emerging Threats, and International Relations. May 5, 2003.

Gilliom, R.J., P.A. Hamilton, and T.L. Miller. 2001. U.S. Geological Survey. The National Water Quality Assessment Program—Entering a New Decade of Investigations. USGS Fact Sheet 071-01.

Global Ocean Observing System. 1998. What is GOOS? A first update. URL: http://ioc.unesco.org/goos/docs/whatis01.htm.

Heinz, H. J. III Center for Science, Economics and the Environment. In *The State of the Nation's Ecosystems: Measuring the Lands, Waters, and Living Resources of the United States.* Cambridge University Press, New York.

Hale, S.S., A.H. Miglarese, M.P. Bradley, T.J. Belton, L.D. Cooper, M.T. Frame, C.A. Friel, L.M Harwell, R.E. King, W.E. Michener, D.T. Nicholson, and B.G. Peterjohn. 2003. Managing troubled data: coastal data partnerships smooth data integration. *Environ. Monit. Assess.* 81: 133–148.

Institute of Medicine. 1988. The Future of Public Health. National Academies Press, Washington, D.C.

Jennings, M. and J.M. Scott. 1997. Official description of the national Gap Analysis Program. Biological Resources Division, U.S. Geological Survey.

Krishtalka, L. et al. 2002. Long-Term Ecological Research Program: Twenty-Year Review. A report to the National Science Foundation, NSF, Arlington, VA, 39 pp.

Kutz, F.W., A.R. Yobs, and S.C. Strassman. 1977. Racial stratification of organochlorine insecticide residues in human adipose tissue. *J. Occup. Med.* 19: 619–622.

Kutz, F.W., P.A. Wood, and D.P. Bottimore. 1991. Levels of organochlorine pesticides and polychlorinated biphenyls in human adipose tissue. *Rev. Environ. Toxicol. Contam.* 120: 1–82.

Kutz, F.W., R.A. Linthurst, C. Riordan, M. Slimak, and R. Frederick. 1992a. Ecological research at EPA: new directions. *Environ. Sci. Technol.* 26(5): 860–866.

Kutz, F.W., B.T. Cook, O. Carter-Pokras, D. Brody, and R.S. Murphy. 1992b. Selected pesticide residues and metabolites in urine from the U.S. general population. *J. Toxicol. Environ. Health*, 37: 277–291.

Murphy, R.S., F.W. Kutz, and S.C. Strassman. 1983. Selected pesticide residues or metabolites in blood and urine specimens from a general population survey. *Environ. Health Persp.*, 48: 81–86.

National Oceanic and Atmospheric Administration (NOAA). 1998. NOAA's state of the coast report. Monitoring the coastal environment. URL: http://state_of_coast.noaa.gov/bulletins/html/mcwq_12/mcwq.html.

National Research Council. 1990. *Managing Troubled Waters: The Role of Marine Environmental Monitoring*. National Academies Press, Washington, D.C.

National Research Council. 1995. *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data.* National Academies Press, Washington, D.C.

National Science Foundation. 2003. International Long Term Ecological Research Network. URL: http://ilternet.edu.

National Science and Technology Council. 1997a. Integrating the Nation's Environmental Monitoring and Research Networks and Programs: A Proposed Framework. URL: http://www.epa.gov/monitor/pubs.html.

National Science and Technology Council, Committee on the Environment and Natural Resources. 1997b. National Environmental Monitoring and Research Workshop: Draft Proceedings. Sept. 25–27, 1996, S. Dillon Ripley Center, Smithsonian Institution, Washington, D.C. URL: www.epa.gov/cludygxb/html/pubs.htm

National Science and Technology Council, Committee on the Environment and Natural Resources. 1998. Fact Sheet. National Environmental Monitoring and Research Initiative. Integrating the Nation's Environmental Monitoring and Related Research Networks and Programs. URL: http://clinton3.nara.gov/WH/EOP/OSTP/Environment/html/fac_cenr.html.

National Science and Technology Council. Committee on Environment and Natural Resources, Air Quality Research Subcommittee. 1999. The Role of Monitoring Networks in the Management of the Nation's Air Quality. URL: http://clinton4.nara.gov/WH/EOP/OSTP/html/papers9899.html

Organisation for Economic Cooperation and Development. 1993. OECD Core Set of Indicators for Environmental Performance Reviews. A Synthesis Report by the Group on the State of the Environment. Paris.

Parkes, M., R. Panelli, and P. Weinstein. 2003. Converging paradigms for environmental health theory and practice. *Environ. Health Perspect.* 111: 5.

The Pew Environmental Health Commission. 2000. America's Environmental Health Gap: Why the Country Needs a Nationwide Health Tracking Network. Companion Report.

Qualters, J.R., M.A. McGeehin, A. Niskar, W.R. Daley, and S. Reynolds. 2003. Developing a National Environmental Public Health Tracking Network in the United States. Euroheis SAHSU Conference. Abstract for Oral Presentation. URL: http://www.med.ic.ac.uk/divisions/60/euroheis/abstracts/qualters.doc

Sauer, J.R., J.E. Hines, G. Gough, I. Thomas, and B.G. Peterjohn. 1997. The North American Breeding Bird Survey: Results and Analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J.R., J.E. Hines, and J. Fallon. 2002. The North American Breeding Bird Survey: Results and Analysis 1966–2001. USGS Patuxent Wildlife Center, Laurel, MD.

United Nations Economic Commission for Europe. 2003. Draft EMEP monitoring strategy 2000–2009. EB/AIR/GE.1/2003.3.Add.1.

U.S. Department of Agriculture. 2002. National Resources Inventory. URL: http://www.nrcs.usda.gov/technical/NRI.

U.S. Environmental Protection Agency. 2003a. Elements of a State Water Monitoring and Assessment Program. EPA 841-B-03-003.

U.S. Environmental Protection Agency. 2003b. Draft Report on the Environment 2003. EPA-260-R-02-006.

U.S. Fish & Wildlife Service. National Wetlands Inventory. 2003. URL: http://www.nwi.fws.gov.

U.S. Geological Survey. 2001. North American Breeding Bird Survey. PWRC Fact Sheet 2001–27.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel. 2001. Completion of the 1990s national land cover data set for the conterminous United States from landsat thematic mapper data and ancillary data sources. *Photogramm. Eng. Remote Sens.* 67(6): 650–662.

Wickham, J.D., K. Hegge, C. Homer, and J.T. Morisette. 2002. National Land Cover Data Base Update—2000. Earth Observer (EOS) Newsletter, 14(2): 13–15. (URL: http://eospso.gsfc.nasa.gov/eos_observ/3_4_02/mar_apr02.pdf)

Wiken, E.B. (compiler). 1986. Terrestrial ecozones of Canada. Ecological land classification series No. 19.

Williams, G. 2001. Inventory and Prototype Monitoring of Natural Resources in Selected National Park System Units 1999–2000. URL: http://www.nature.nps.gov/TR2001-1/TitlePage.htm.

28 From Monitoring Design to Operational Program: Facilitating the Transition under Resource-Limited Conditions

D.J. Roux

CONTENTS

28.1 INTRODUCTION

No matter how good the technical design of a monitoring program, the intended benefits or value can only be realized if the program is effectively implemented. For the purpose of this paper, implementation is defined as putting a theoretical concept or a new product, program, or service into practice. "Putting into practice" can be

1-56670-641-6/04/$0.00+$1.50

described as carrying out, executing, achieving, or accomplishing. Taylor[1] warns that "… nothing is more powerful than a great idea. And nothing is more deadly than its poor execution." Because of the universal elusiveness of putting a new idea into practice,[2] the implementation challenge has been the topic of many studies. These studies are generally in the context of organizational transformation and change management.[3–5] However, a study that documents the 10 most important lessons for implementing an integrated watershed approach to protecting wetlands, streams, rivers, and estuaries[6] shows that many of the principles that apply to effective implementation are generic.

The South African River Health Program (RHP) has, over a period of 9 years (1994 to 2003), grown from a mere idea to a national operation. This is especially significant when considering that adoption and implementation of the RHP are largely voluntary. To add to the achievement, program implementation is taking place in an environment characterized by limited financial resources, a multitude of competing social and economic priorities, and a severe scarcity of appropriately skilled people.

The RHP was designed in response to a specific information need, namely, to assess the ecological state of riverine ecosystems in relation to all the anthropogenic disturbances affecting them. It is a screening-level monitoring program operating on a low sampling frequency and a low resolution of sites scattered semi-randomly across catchments. The program's assessment philosophy is based on the concept of biological integrity[7] and use is made of biological indices (fish, invertebrates, riparian vegetation), as well as indices for assessing in-stream and riparian habitats. The RHP is geared to assess the general ecological state of rivers rather than site-specific impacts or conditions. A description of the design criteria and process is presented in Roux.[8]

While the design, development, and standardization (concepts, methods, processes) of the RHP is coordinated at a national level, implementation activities largely take place at the provincial level. Due to the relatively flexible and learn-by-doing approach that has been advocated for provincial adopters of the RHP, a diversity of implementation models has developed across the country. As a result, there are nine provincial "case studies" regarding the implementation of the RHP. These implementation models have resulted in varied levels of success — from two provinces being nearly self-sufficient in conducting routine surveys, health assessments, and reporting to two provinces still needing to take the basic step of establishing an implementation team.

Reflecting on the RHP successes and failures in South Africa provides an opportunity to better understand the transition of environmental monitoring programs from theoretical design to sustainable operation, particularly in resource scarce environments. A previous paper has used the RHP as case study to focus on this transition from a technological maturation perspective.[9] This communication focuses on three semi-social themes that appear to be key drivers of successful implementation of the RHP. The three themes are:

- A compelling vision and strategic conversations
- Shared ownership by means of virtual governance
- Creative packaging of scientific messages

28.2 A COMPELLING VISION AND STRATEGIC CONVERSATIONS

Collins and Porras[10] pointed out that companies that enjoy enduring success have a core purpose that remains fixed while their strategies and practices endlessly adapt to a changing world. An ability to effectively manage this balance between continuity and change is equally important for sustaining a national monitoring initiative. This ability is closely linked to the development of a vision and is based on the following pillars[10]:

- A *philosophical foundation* that defines why a particular venture exists and what it stands for (its reason for being)
- An *envisioned future* that radiates what we aspire to achieve
- The *strategic conversations* that form wide and consistent communication of both the philosophical foundation and the envisioned future to capture the imagination of people

28.2.1 Philosophical Foundation

A key step in enabling effective implementation is to understand what needs to be implemented. The "what" is commonly described by a core set of objectives or principles. However, the robustness and timelessness of the "what" is often a function of how well the "why" is understood. During the embryonic phase of the RHP, a substantial effort went into deliberating why this program is necessary, in addition to discussing what it will achieve. A deeper analysis of the latent needs of water resource managers, a scrutiny of motives, and an effort to understand the underlying concepts characterized this early phase. The result was a committed nucleus of thought leaders who shared a deep understanding of why developing this new monitoring program was important as well as what could be achieved with it.[9]

From this philosophical foundation emerged the overall objective or purpose of the program, namely, to measure, assess, and report on the ecological state of rivers in order to improve decision making regarding the sustainable use and protection of these systems. The methods and processes used to achieve this purpose may evolve, but the purpose remained unchanged ever since it was first published.[11] This was important for protecting the focus of the program. Having a program with a clear and rather simple focus is necessary to build critical capacity around the program, whereas the same number of people involved in a more diverse program will have to spread their attention too wide and thin to make real impact.

When the success of the program became visible, there were a number of temptations to dilute this focus. It was suggested that the program should also take on the monitoring of estuaries. However, these ecosystems fall outside the boundaries of the purpose. Similarly, since the program was in the process of establishing a network of monitoring sites, associated infrastructure, and human capacity, there was pressure to include variables that relate to human health, such as fecal coliforms, as part of the program. Adding a human health perspective could have skewed the true purpose, as high fecal coliform counts may be undesirable from a human health perspective but may be perfectly natural from an ecological perspective (for example, downstream from a pool inhabited by hippopotami). In both cases, the articulated purpose provided guidance for deciding against taking these additional components on as part of the program. In retrospect, staying true to the purpose promoted the development of a strong program identity as well as long-term loyalty among collaborators and stakeholders. In addition, a clear and simple purpose has a better chance of being implemented, whereas an all-inclusive purpose may easily become an end in itself with very little emerging beyond.

28.2.2 Envisioned Future

Where the purpose or objectives provide focus, the envisioned future or goal provides direction. Contrary to the purpose of the program, the envisioned future or goal has a restricted lifespan and should evolve to reflect improving understanding of both capabilities and constraints. This has certainly been the case with the RHP. The program essentially followed a phased design — through scoping, conceptualization, pilot application and testing, and operational rollout phases. Each of these phases started off with a goal, a target date for achieving the goal, and a vivid description of the future reality once this goal has been achieved.

An advantage of a phased approach is that the end of each phase can be used for reflection and strategic review. The goal for the next phase can be scaled up or down according to the insights gained during the previous phase. As an example, an ambitious 3- to 6-year goal was set during 1997 to implement and maintain the RHP on all major rivers of South Africa and to expand program implementation to other key rivers within southern Africa.[12] A reality check during a subsequent pilot application exercise[13] led to the setting of a more realistic goal for the next phase, namely, to achieve successful implementation on one river per province by 2003.

28.2.3 Strategic Conversations

Communication of the program purpose and vision provides the glue that holds an initiative together as it grows, decentralizes, and diversifies. In the RHP model of national development, standardization and quality control, and local ownership and implementation, it was important that the vision effectively dispersed to the multiple stakeholder levels. The RHP stakeholder system includes national government departments, R&D organizations, universities, as well as many provincial agencies responsible for nature conservation and environmental management (see Section 28.3). In order to compete for and direct the attention and resources of these groups, the RHP vision had to be effectively communicated.

The most critical success factor in dispersing the RHP vision was that a number of committed leaders took ownership of the message. Collectively, these leaders had influence in government, the academia, and conservation agencies, and their direct communication and endorsement were key to gathering wider support for the program. The influence of opinion leaders started with preexisting personal networks, extending outward to motivate other key groupings to get involved and allocate priority time and funding to the associated work.

In addition to having appropriate human carriers of the message, the content and tailoring of the message are of utmost importance. On the political front, and in a country where short-term social and economic needs override conservation aspirations, it was important to communicate very clearly the rationality of a monitoring program designed to diagnose the ecological state of rivers. For example, we can make plain the practicality of research by asking a series of simple "whys." If we analyze why we want to measure river health, we say it is to know whether rivers are healthy or not. Why? So that rivers can be effectively managed. Why? So that people can have sustained benefit from the services that these ecosystems provide. Why? Because these ecosystem services contribute to societal well-being and economic prosperity. Therefore, we do not monitor river health for the sake of aquatic biota but rather as an ecological means to a socio-economic end.

Annual symposia and specialist workshops contributed significantly to developing a common language, cohesion, and a sense of belonging among all those involved in RHP activities. In addition to people-to-people communication, a range of products was produced to make the river health message as pervasive as possible. These products carry subtle branding characteristics, such as the omnipresent picture of a dragonfly, to make them recognizable as part of the same program. Products include technical reports, implementation manuals, newsletters, popular articles for magazines, a coloring book for school children that was translated into four languages (Finny Fish tells about "my home, a healthy river"), generic posters explaining how the RHP works, an Internet Home Page (www.csir.co.za/rhp), and State-of-Rivers (SoR) posters and reports (see Section 28.4 on creative packaging of scientific messages).

In a demographically highly diverse country, multiple communication formats and distribution media are advisable. For example, the Internet provides a mechanism for collating formidable amounts of information. Yet, access to the Internet is not readily or reliably available in some institutions and parts of the country. For some of these users, including a university, the complete RHP home page was stored and distributed on CD.

A final learning point is that patient and persistent communication gets rewarded. One water resource manager listened somewhat skeptically to talks on the value and benefits of biomonitoring for close to 2 years. Then he suddenly became one of the most passionate proponents of the RHP. It just took a while for this individual to internalize the message and its implications into his personal reference framework. The need for strategically directed communications never comes to an end. There is always the risk that regression may set in among some adopters, while an ever-evolving vision must continuously be entrenched in the hearts, minds, and budgets of old and new "subscribers."

28.3 SHARED OWNERSHIP BY MEANS OF VIRTUAL GOVERNANCE

When human and financial resources are at a premium, *networking* (reaching out and getting in touch with others) and *collaboration* (to work in combination with others) become key success factors in bridging capability/capacity gaps and achieving demanding goals. Advantages associated with collaborative ventures include such initiatives leading to wider acceptance and quicker implementation of projects and programs. Exposure to collaborators can also provide an element of peer review of R&D functions and challenge in-house researchers and managers with new ideas. In addition, concepts, tools, and methods developed through collaboration will carry more weight in promoting a uniform standard, increasing goodwill across government and public sectors, and positively influencing future legislation.[14]

The true value of networking and collaboration probably lies in the formation of informal arrangements and relationships. In this regard, the RHP was particularly successful. This section looks at some of the interventions that resulted in the formation of a virtual network of developers and implementers across the country, with the key objectives being to:

- Unite researchers and implementers into one team
- Promote collaboration through regional implementation networks
- Progress from individual enthusiasm to organizational capability and accountability

28.3.1 Uniting Researchers and Implementers into One Team

The national Department of Water Affairs and Forestry, as the legal custodian of water resources in South Africa, has played the leading role in initiating and designing the RHP. This department is particularly strong in policy development and managing water resources in the conventional areas of quantity and chemical quality. They had the foresight to drive the development of a biological-response monitoring program but realized that they did not have the expertise and capacity to implement such a program across the country. The RHP could only become an operational reality given the collective resources of a number of national and regional (provincial) agencies and organizations.

A model of national development and coordination and provincial or local implementation (operational ownership) was adopted. The one side (national) was characterized by visionary thinking, concept and method development, and quality assurance; the other (provincial) by pragmatic considerations. This dual focus (scientific rigor and practical feasibility) was not merely a convenient arrangement but increasingly became a key factor for the sustainable implementation of the program. Based on the two foci or value propositions, four possible future scenarios can be delineated (Figure 28.1):

- *Scenario 1*: Both the scientific credibility and the value that the RHP presents to its stakeholders are lowered, and the program has no future. Increasing regression of efforts will eventually lead to the disappearance of the program, and both river managers and researchers will pursue more relevant options.
- *Scenario 2*: Resources are primarily directed towards technical design and ongoing improvement through research and development. The RHP is recognized for its scientific and technical excellence but stakeholders are not experiencing benefits from the program. Too little attention is given

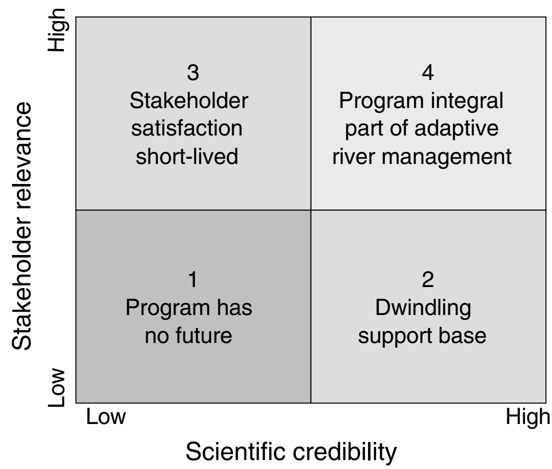


FIGURE 28.1 Two key factors influencing the continued relevance and impact of a monitoring program.

to understanding and satisfying the needs of the nontechnical stakeholder community. These end-users of river health information lose their enthusiasm for the RHP and redirect their support to other initiatives. The RHP largely remains of academic interest and will not become an operational program.

- *Scenario 3*: All attempts are made to understand and satisfy stakeholder needs but insufficient resources are allocated to scientific development, testing, verification, and ongoing improvement. Initial support by stakeholders is replaced by skepticism as the gaps in the program's science-base become evident. The end result is a program that will have ever-decreasing support and that will not be able to contribute to ecologically sound management of rivers.
- *Scenario 4*: The importance of adding real value for stakeholders as well as remaining technically relevant is recognized and pursued with sufficient resources. Scientists, river managers, and environmental policy makers interact frequently, which results in reconciliation of perspectives, development of a deep understanding of each other's needs and limitations, and adaptive improvement of the program over time to ensure continued scientific and managerial relevance. The program impacts positively on decision-making and on the health of rivers.

Early and ongoing interaction between researchers and perceived end users of a research product is the surest way to increase the likelihood that the product will be used. From the earliest stages of conceptualizing the RHP, a dedicated effort was made to seek the real end users (not necessarily the same as the client paying for the program design) and to uncover their real needs. These end users included both organizations responsible for environmental policy and agencies that would actually conduct river surveys. Researchers helped to shape these needs and ensured that a clear scope and design criteria were defined.

The inclusive style adopted for the development of the RHP resulted in virtually all groups, organizations, and authorities that would ultimately be involved with or responsible for the implementation of the RHP becoming involved at an early stage.[9] This approach led to the design of a user-oriented and pragmatic program where the operational manual was shaped through the collective expertise and expectations of a large and diverse group of stakeholders.

The value of an inclusive developmental approach lies in reducing the inherent lag between knowing what to do and actually applying this knowledge.[2] Scientists came to learn firsthand that the adoption of a new scientific tool is not driven by the scientific status of the tool or its underlying concepts but rather by convention, past practice and experience, social and economic considerations, and perceived value (determined by the user) and availability of required infrastructure. In addition, appropriate user skills and logistical support and a sufficiently knowledge-intensive environment must be in place before a new program can be deployed successfully.[15] In turn, river managers were able to experience the new protocols in action during pilot applications which promoted user "readiness." A lack of user readiness is regarded as a common constraint to the adoption of R&D outcomes.[16] The result of implementing readiness is that a natural progression is fostered, among all parties, from research to design to adoption and subsequent implementation. Part of this progression is the gradual creation of capacity in participating organizations.

28.3.2 Provincial Implementation Networks

The most acute challenge that had to be overcome was (and still is) to achieve critical capacity and endorsement of the program at operational levels. The provincial scale was selected for deploying the program, primarily due to the presence of agencies with relevant expertise and equipment that operate at these levels. Provincial conservation boards and provincial departments of environmental affairs were typically the organizations that could contribute the required expertise and equipment — for example, fish biologists and nets for sampling.

During a consultation meeting held in 1996, an "implementation champion" was elected for each province.[17] These champions were tasked with establishing and mobilizing a provincial implementation team (PIT) who would be responsible for provincial scale implementation and demonstration of the program. Although the responsibility of implementation was decentralized to provincial level, no financial resources accompanied this delegated responsibility. Success or failure of initiating a provincial initiative was largely a function of every champion's enthusiasm, ownership, ability to influence others and to mobilize funding, as well as the degree to which his/her organization would endorse river health activities. The latter is often somewhat dependent on the presence of the first mentioned qualities.

No single organization in any province could master all the expertise required for implementing the RHP. During a rather comprehensive pilot implementation and demonstration exercise,[13] a theoretical model for an interorganizational PIT was

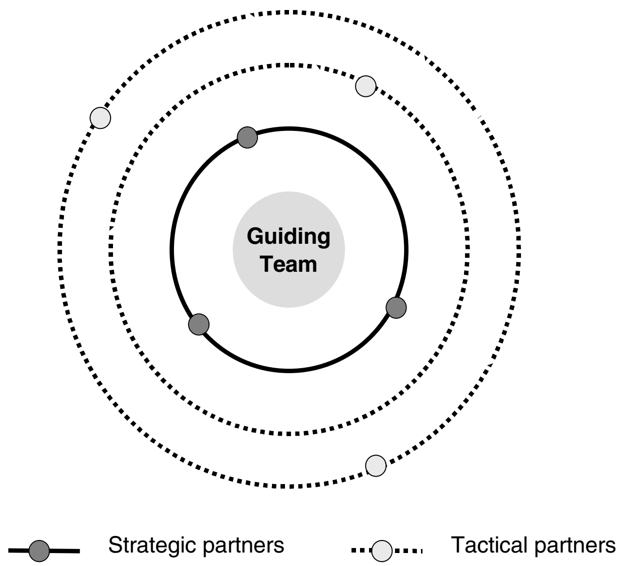


FIGURE 28.2 Positions for collaboration in a networked implementation team.

suggested (Figure 28.2). This model relates to the concept of communities-of-practice[18] which provides insight into how informal networking can be applied in support of a formal goal. The suggested networked PIT allowed for three distinct positions, based on the commitment, resources, and knowledge that a specific organization is willing to contribute as well as the relative permanence with which a network position is occupied:

- The guiding or *core group* essentially fulfills the leadership function and is constituted by the provincial lead agents of the study. These are the relatively permanent members of the network, who participate actively and largely determine the agenda and activities of the whole team. Ideally, a number of statutory bodies should be represented in the core group, and individual members should display strong leadership characteristics.
- *Strategic partners* constitute those individuals and organizations with whom a long-term relationship will be mutually advantageous. This may include the lead agents of a neighboring province with which a catchment is shared or a university that agrees to provide strategic support in method development, training, and student involvement with monitoring activities.
- *Tactical partners* would have a relatively short residence time in the network, based on the temporary requirement of a specific skill or expertise. These partners may be professional service providers that would be contracted to fill a temporary skills gap—for example, to coordinate a first river survey, assist with once-over selection of monitoring sites, develop a data management system, or compile a report.

Where the core group is intended to provide stability, the tactical partners provide flexibility as these members can be substituted as specific needs arise. Although the generic positions can be recognized in every provincial initiative in South Africa, every one of these initiatives are also distinctly different in composition, management style, and operational culture. However, those that have applied the networking principles of effective and inclusive communication and purposeful development of interorganizational relationships and trust have generally experienced a significant increase in workforce diversity and strength, access to a larger pool of capabilities, and a stronger standing both in their provinces and at national forums.

28.3.3 From Individual Enthusiasm to Organizational Accountability

Every province that participates reasonably successfully in the RHP finds itself somewhere on a maturation trajectory with three chronological stages, starting with individual enthusiasm, progressing to informal networking, and ending in organizational endorsement. Champions started off armed only with enthusiasm for the task ahead. Most of these champions have a background in the environmental sciences. The primary reason why they agree to championing their provincial initiative is because they care about rivers and they believe that the RHP would help them to generate the information that would contribute to sound river management.

The initial responsibility of the champion is to bring a group of people together to accomplish something collectively that they could not accomplish separately. The lobbying for team members is usually based on the need for certain basic skills as well as for having the representation of key organizations in the province. The alliance is still completely informal and individual members join based on their perception of the value that the initiative brings to them and to their organizations. The opportunity to expand professional networks, exchange knowledge, and make new friends is commonly cited as reasons for joining the PIT.

Either before or after joining the PIT, individual members would request official approval from their organizations for getting involved in RHP activities. Their case is strengthened if they can show examples of what the program produces and how this relates to their organizational mandates. Organizations would then consider whether and to what degree they would endorse the program. This decision may be in the form of allowing a staff member to spend a certain percentage of his or her time on program activities.

The predominantly bottom–up approach described above is supplemented with a top–down approach where, for example, the national Department of Water Affairs and Forestry would extend an official request to heads of key provincial agencies to commit resources to the RHP. Accountability for executing the RHP is likely to be much clearer and legitimized when and where Catchment Management Agencies (CMAs) come into being. These agencies are likely to have the delegated mandates — from national government — and statutory powers to coordinate monitoring and reporting on the ecological state of rivers. For South Africa, the establishment of CMAs is foreseen to take place systematically over the next two decades.

28.4 CREATIVE PACKAGING OF SCIENTIFIC MESSAGES

The overall goal of communicating natural resource information should be to change the behavior of the recipients of the information.[19,20] In the case of the RHP, the program must (1) communicate ecologically sound management of rivers in South Africa and (2) inform and educate the people of South Africa regarding the health of their rivers. Changed behaviors relate to the degree to which resource managers incorporate river health information in their decision-making processes. Similarly, a positive change in civil society's perception and appreciation of rivers would testify to effective communication. To achieve these goals, RHP practitioners had to rethink the formats used for packaging information as well as the strategies used for disseminating information.

Three communication strategies are highlighted in this section, namely:

- Reducing the complexity of scientific messages
- Developing a flagship communication product
- Uncovering and utilizing tacit knowledge

28.4.1 Reduce the Complexity

Scientists are often very well trained in packaging their work for, and disseminating it to, other scientists; for example, by means of peer reviewed papers. However, this does not help the cause of spreading the message widely through diverse audiences.[21] Ultimately, effective dissemination of resource information is about ensuring that information becomes available to those that might best use it, at the time they need it, in a format they can use and find comprehensible, and which reflects appropriate spatial and temporal scales.

In an era of information overload there is a major demand for products that are simple yet credible. This reminds us of an Albert Einstein quote: "Everything should be made as simple as possible, but no simpler." Due to the wide audience to whom the RHP needs to communicate, it was inevitable that the normal complexity associated with science had to be reduced. This is reflected in the evolution of the name of the program, which started as the National Aquatic Ecosystem Biomonitoring Program. In an attempt to be correct, scientists played with words such as *integrity*, *aquatic ecosystems*, and *biological monitoring*. It was quite a breakthrough when the name River Health Program received consensus approval. "River health" is readily interpreted by most people and, as such, is quite liberating terminology from a communication perspective.[22]

In early communication attempts, it became clear that decision makers are not all that interested in scientific explanations, references, graphs, and diagrams of aquatic invertebrates. A map showing the river of interest with color-coded dots that indicated the relative health of the river at various monitoring stations would commonly receive the most attention. This realization gave rise to the development of a river health classification scheme (Table 28.1) to allocate a specific category of health to each river reach. The health categories used by the RHP are simply called

TABLE 28.1
The River Health Classification Scheme Used for Reporting Information Generated from Findings of River Surveys

River Health Category	Ecological Perspective	Management Perspective
Natural	No or negligible modification of in-stream and riparian habitats and biota	Relatively untouched by human hands. No discharges or impoundments
Good	Ecosystem essentially in good state. Biota largely intact	Some human-related disturbances but mostly of low impact potential
Fair	A few sensitive species may be lost; lower abundances of biological populations are likely to occur or, sometimes, higher abundances of tolerant or opportunistic species	Multiple disturbances associated with need for socio-economic development, e.g., impoundment, habitat modification, and water quality degradation
Poor	Habitat diversity and availability have declined; mostly only tolerant species present; species present are often diseased; population dynamics have been disrupted (e.g., biota can no longer breed or alien species have invaded the ecosystem)	Often characterized by high human densities or extensive resource exploitation. Management intervention is needed to improve river health, e.g., to restore natural flow patterns, river habitats, or water quality

natural, *good*, *fair*, and *poor*. This classification system and the associated protocols used to assess data in their regional reference contexts allow the health of rivers to be directly comparable across the country.

The classification scheme provides a simplified “front end” to a much more intricate assessment process. This front end provides a tool for communicating technical concepts to nontechnical audiences. At the same time, stakeholders can use the river health classes in catchment visioning exercises to arrive at a desired state for their river. This desired state or goal could be decomposed into measurable management objectives which, in turn, relate to the same biological and habitat indices that were used to derive the present state.

28.4.2 Develop a Flagship Product

As part of national developments in the RHP, the design of an effective reporting format for river health information was seen as a priority. In the process, a number of alternative communications and dissemination media were experimented with. It was realized that, to effectively compete for attention amid the multitude of messages and an overall information bombardment that most people are exposed to, ecological messages must be communicated in highly attractive and professional formats.

Gradually, the State of Rivers (SoR) reporting concept emerged and matured to form the flagship products of the RHP.

Aligned with national State of Environment reporting, the RHP's SoR reporting initiative makes use of the Driving Force–Pressure–State–Impact–Response (DPSIR) Framework[23] to explain what is causing environmental degradation, how good or bad the conditions are, and what we can and are doing about it. Whereas the RHP focuses primarily on the present state and trends in river health, an effort is made to link the present state to specific driving forces and pressures on the river as well as to specific policies and management actions in place to manage the rivers.

The SoR reports are essentially brochure-style reports in full color, usually less than 50 pages in length and of quality print. Posters contain highly synthesized information presented in A0 size for display against walls. These products are primarily distributed in hard copy format. A simplified presentation of a SoR map is indicated in Figure 28.3.

A flagship product can only have the desired impact if it reaches its intended audience. In several instances where a batch of reports were dispatched to a specific contact person for further distribution, it was found that the reports remained in the first recipient's office—sometimes for many months. Personalization of report distribution is the ideal, where key recipients receive a hand-delivered report with a brief contextual explanation.

Even where people do receive a personal copy of an SoR report, there is no guarantee that they will make the time to read it and internalize the information, let alone initiate a required management intervention. An important element in the evolution of SoR reporting is a continuous process of assessing reader satisfaction. Based on actual feedback, the structure, style, and specific presentation features used for reporting information are updated and improved. As an example, feedback is used to refine the balance between text, graphics, information boxes, and white

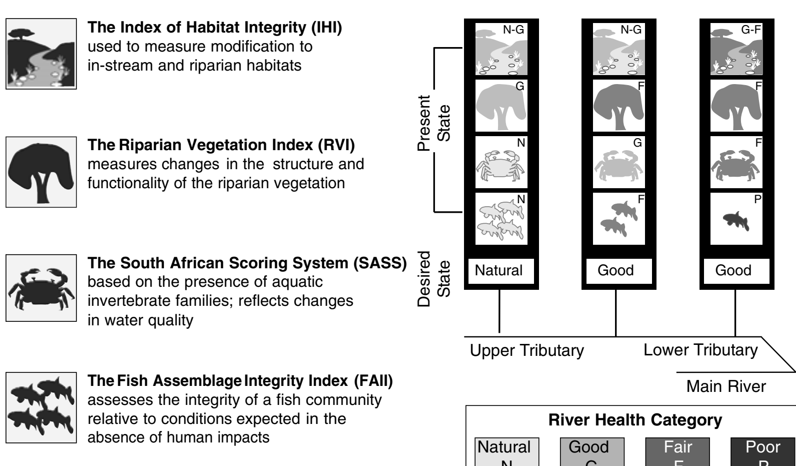


FIGURE 28.3 Format used in State-of-Rivers reports to summarize the present ecological state in relation to a future desired state for rivers.

space, as well as to address evolving information needs of stakeholders such as river managers. Satisfaction reviews for two subsequent reports[24,25] have indicated an increase of 27% in the proportion of readers that read more than 60% of the content. This improvement could largely be ascribed to adjustments related to content, presentation format, and style.[26]

28.4.3 Uncovering and Utilizing the Richness of Tacit Knowledge

While the design and structure of a communication product are important, the substance of the material that is available to make up the content is just as critical. In compiling SoR products, a conscious decision was made to utilize the knowledge that resides with relevant people to complement formally collected data and derived information. For some rivers, very little scientific or formal information was available prior to conducting the first RHP survey and producing the subsequent SoR report. In other instances, ecologists and river managers may have collected relevant information over extended periods of time. In all instances, the informal knowledge possessed by scientists, managers, farmers, or people from local communities has the potential to present a much more holistic and comprehensive picture of the river than data available in formal databases or publications. This section documents some of the process learning acquired in extracting the tacit and often latent knowledge of individuals and converting this knowledge into explicit form.

Explicit knowledge can be expressed in words and numbers and shared in the form of data, scientific formulae, specifications, and manuals. This form of knowledge can be readily transferred among individuals and within organizations. In contrast, tacit knowledge is highly personal and hard to formalize, making it difficult to communicate or share with others.[27] To uncover tacit knowledge is inherently difficult since even those with knowledge may not be conscious of what they know or what its significance is. Knowledge has an intrinsic tendency to stay where it was first internalized. Three issues were found to be particularly significant factors in influencing whether and how people would share their knowledge: trust, as the bandwidth for getting knowledge to flow from one person to the next; the environment (place and time) in which people are comfortable to share knowledge; and the degree of overlap in personal aspiration or professional goals between the people involved in communicating.

Nothing can compare with long-term personal relationships for cultivating the required levels of trust that get people to freely share their knowledge. In the context of river monitoring, these relationships often start as a result of sharing a common interest (e.g., an endemic species) or solving a common problem (e.g., controlling alien weeds). This issue reflects on the importance of the composition of the reporting team. Some form of overlap in the social networks or professional interests of the reporting team with those that are knowledgeable regarding the river in question represents a significant advantage. As an additional or alternative resort, the reporting team has to spend time with identified stakeholders in order to get them to share their tacit knowledge relevant to the report.

A second issue is that different individuals have different time and place requirements for sharing their tacit knowledge. In general, we have found that resource managers of relative seniority have a high premium on time. Once they understand and buy into the objectives of the RHP, they are willing to share their knowledge in a time-efficient manner. A short meeting in his or her office may prove sufficient, whereas some prefer to be away from their offices and associated demands on their attention. Some share more freely in small groups and one-to-one meetings. For field practitioners, "field meetings" have generally resulted in better returns. These meetings may take the form of a one-day visit to some monitoring sites followed by a day of work-shopping the results of the river surveys. The second day should preferably also take place in an informal environment, and a meeting facility on the banks of a river of concern may work well.

For the second type of knowledge exchange/conversion, the editorial team is responsible for interpreting the tacit knowledge that was shared among the stakeholder groups and for capturing this in explicit form—that is, the conversion from mind to report. This step requires multiple iterations of draft version between editors and knowledge contributors to ensure that context-specific knowledge has been captured correctly. The outcome is a report that provides much more context than could be derived from purely using collected and interpreted data. As an example, observations regarding the decline and subsequent recovery of a hippopotamus population as a result of a drought that happened almost a decade prior to the river survey in question, the occurrence of rare bat species and fish owls in the riparian forest, and the exceptional abundance of crocodiles in a particular river reach are all bits of information that were not found in a database but that surfaced during knowledge-sharing sessions.[25] It was felt that such tacit knowledge has the potential to significantly increase reader interest and the contextual orientation provided by SoR reports.

28.5 CONCLUDING REMARKS

The long-term vision of the RHP is that the information generated through its river surveys and SoR reports will eventually cover all the main rivers of South Africa to allow a qualified statement on the overall health of the nation's rivers. Repetitive surveys and reporting would provide a scientific indication of whether the ecological condition of rivers is deteriorating or, in fact, improving over time. Such information would be useful to "audit" the effectiveness of the policies, strategies, and actions of both the national custodian department and the decentralized agencies responsible for the sustainable management of river ecosystems.

The RHP is often lauded as an example of an environmental program that achieved the transition from being a good design to becoming an operational practice. Many factors played a role in stimulating the popularity, support, and growth that the program enjoyed. The three primary factors noted previously have played a significant role in the development, character, and dispersal of the program. However, the challenge is far from over and even these three factors need continued nurturing in order to leverage limited resources towards achieving future goals.

During its life, the RHP has developed a strong identity that is well entrenched in the minds of a wide stakeholder group. The effective diffusion of the vision has been a key success factor in drawing the human and financial resources that the program has achieved to date. This in turn is a function of the program leadership at all levels over the past 9 years. Tom Peters said: "The only constant that correlates with success is top leadership." Continued nurturing of an appropriate leadership cadre is probably the most critical investment that can be made to ensure the future success of the program.

A significant concern is the lack of redundancy in program leadership. In an environment with a small pool of skilled human resources and an overwhelming list of developmental imperatives, it has been extremely difficult to do succession planning in the RHP leadership group. At both national and provincial levels, the program often relies on individuals, where the loss of a provincial champion or a national task leader (through promotion or needs/opportunities elsewhere) may render a particular initiative vulnerable to regression. The fact that any long-term program will, and should to some degree, experience turnover of key role players must be recognized and managed. Without significant overlap between old and new leaders, the memory and knowledge base associated with early developments can only erode.

From a program governance perspective, a model of fostering collaboration based on informal relationships and networks, promoting shared ownership, and allowing flexibility to cater for a diversity of resource realities and capabilities proved to be most successful. However, the relatively high degree of institutional flux and people mobility that prevails in the country remains a constraint to institutionalizing the program. Mandates, roles, responsibilities, and agendas of organizations and individuals change more rapidly than is desirable and informal arrangements leave the program vulnerable. It is perhaps time to introduce a more formal model of program governance where institutional responsibilities are made explicit and the advance towards covering all the main rivers of the country can be managed in a more systematic fashion.

On the technical side, SoR reporting epitomizes many of the underlying technical components that make up the RHP. Through testing and refinement, the SoR reporting initiative has developed into a state-of-the-art capability. Stable prototype tools make it possible to accelerate the rate at which rivers are incorporated in the monitoring and reporting cycle. However, the need for a monitoring program to continuously and dynamically evolve and improve should not be neglected. Internal learning, international benchmarking, and changing needs of key stakeholders need to be incorporated on an ongoing basis to ensure long-term relevance. The danger is that a program that is perceived to be successful and relatively mature in terms of its technical development may have difficulty in securing resources for further developmental work in the face of competing national and regional priorities. The reality is that, in order to capitalize on its successes to date, substantial funding and leadership are required to continue with the coverage of all the main rivers of South Africa over the next 5- to 7-year period.

ACKNOWLEDGMENT

Heather MacKay provided helpful editorial comments on a draft version of the manuscript, and Wilma Strydom prepared Figure 28.3.

REFERENCES

1. Taylor, W.C., Control in an age of chaos, *Harvard Bus. Rev.*, 72, 64, 1994.
2. Pfeffer, J. and Sutton, R.I., *The Knowing–Doing Gap: How Smart Companies Turn Knowledge into Action*, Harvard Business School Press, Boston, 2000.
3. Kotter, J.P., *Leading Change*, Harvard Business School Press, Boston, 1996.
4. Ulrich, D., A new mandate for human resources, *Harvard Bus. Rev.*, 76, 124, 1998.
5. Day, G.S., Creating a market-driven organization, *Sloan Manage. Rev.*, Fall, 11, 1999.
6. U.S. EPA, Top 10 Watershed Lessons Learned, EPA 840-F-97-001, 1997. Online at http://www.epa.gov/owow/lessons/
7. Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., and Schlosser, I.J., Assessing biological integrity in running waters—a method and its rationale, Illinois Natural History Survey, Special Publication 5, Champaign, IL, 1986.
8. Roux, D.J., Design of a national programme for monitoring and assessing the health of aquatic ecosystems, with specific reference to the South African River Health Programme, in *Biomonitoring of Polluted Water—Reviews on Actual Topics*, Gerhardt, A., Ed., Trans Tech Publications, Zurich, Switzerland, 1999, 13 pp.
9. Roux, D.J., Strategies used to guide the design and implementation of a national river monitoring programme in South Africa, *Environ. Monit. Assess.*, 69, 131, 2001.
10. Collins, J.C. and Porras, J.I., Building your company's vision, *Harvard Bus. Rev.*, 74, 65, 1996.
11. Hohls, D.R., National Biomonitoring Programme for Riverine Ecosystems: Framework Document for the Programme, Institute for Water Quality Studies, Department of Water Affairs and Forestry, Pretoria, South Africa, 1996.
12. Roux, D.J., National Aquatic Ecosystem Biomonitoring Programme: Overview of the Design Process and Guidelines for Implementation, Institute for Water Quality Studies, Pretoria, 1997, available online at: http://www.csir.co.za/rhp/reports/report6/report6.pdf
13. Roux, D.J., Ed., Development of Procedures for the Implementation of the National River Health Programme in the Province of Mpumalanga, WRC Report No. 850/1/01, Water Research Commission, Pretoria, South Africa, 2001.
14. Tidd, J., Bessant, J., and Pavitt, K., *Managing Innovation: Integrating Technological, Market and Organizational Change*, John Wiley & Sons, New York, 1997.
15. Steele, L.W., *Managing Technology: The Strategic View*, McGraw-Hill, New York, 1989.
16. Grayson, R.B., Ewing, S.A., Argent, R.M., Finlayson, B.L., and McMahon, T.A., On the adoption of research and development outcomes in integrated catchment management, *Aust. J. Environ. Manage.*, 7, 24, 2000.
17. DWAF, National Biomonitoring Programme for Riverine Ecosystems: Proceedings of consultation planning meeting, Institute for Water Quality Studies, Department of Water Affairs and Forestry, Pretoria, South Africa, 1996.
18. Wenger, E., McDermott, R. and Snyder, W.M., *Cultivating Communities of Practice*, Harvard Business School Press, Boston, 2002.
19. Denisov, N. and Christoffersen, L., Impact of Environmental Information on Decision-Making Processes and the Environment, Occasional Paper 01, UNEP/GRID-Arendal, Norway, 2000.
20. Cullen, P. et al., Knowledge-Seeking Strategies of Natural Resource Professionals, Technical Report 2/2002, Cooperative Research Centre for Freshwater Ecology, Canberra, Australia, 2001.
21. Saywell, D.L. and Cotton, A.P., Spreading the Word—Practical Guidelines for Research Dissemination Strategies (Interim Findings), Water, Engineering, and Development Centre, Loughborough University, Leicestershire, U.K., 1999.

22. Norris, R.H. and Hawkins, C.P., Monitoring river health, *Hydrobiologia*, 435, 5, 2000.
23. Smeets, E. and Weterings, R., Environmental Indicators: Typology and Interview, Technical Report No. 25, European Environment Agency, Copenhagen, Denmark, 1999. Online http://amov.ce.kth.se/courses/1b1640/docs/soetypology.pdf
24. RHP, State of Rivers Report: Crocodile, Sabie-Sand & Olifants River Systems, Report TT 147/01, Water Research Commission, Pretoria, South Africa, 2001.
25. RHP, State of Rivers Report: Letaba & Luvuvhu River Systems, Report TT 165/01, Water Research Commission, Pretoria, South Africa, 2001.
26. Strydom, W.F., Van Wyk, E., Maree, G. and Maluleke, T.P., The evolution of State-of-Rivers reporting in South Africa, Proceedings of the 7th International Conference on Public Communication of Science and Technology, Cape Town, December 5–7, 2002. Online: http://www.fest.org.za/pcst/programme/papers.html
27. Nonaka, I., Konno, N., and Toyama, R., Emergence of "Ba"—a conceptual framework for the continuous and self-transcending process of knowledge creation, in *Knowledge Emergence: Social, Technical, and Evolutionary Dimensions of Knowledge Creation,* Nonaka, I. and Nishiguchi, T., Eds., Oxford University Press, New York, 2001, Chap. 2.

29 The U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program

M. McDonald, R. Blair, D. Bolgrien, B. Brown, J. Dlugosz, S. Hale, S. Hedtke, D. Heggem, L. Jackson, K. Jones, B. Levinson, R. Linthurst, J. Messer, A. Olsen, J. Paul, S. Paulsen, J. Stoddard, K. Summers, and G. Veith

CONTENTS

29.1 INTRODUCTION

The broad goal of the Clean Water Act (CWA, Federal Water Pollution Control Act of 1972, as amended in 1977) is to maintain and restore the chemical, physical, and biological integrity of the nation's waters. However, following the passage of the CWA, the U.S. Environmental Protection Agency (U.S. EPA), states, and Indian tribes focused mainly on chemical stressors in the aquatic environment. The lack of more comprehensive aquatic information has, in part, reinforced calls for significant advances in the way federal agencies monitor environmental conditions (National Research Council (NRC) 1977; U.S. General Accounting Office (GAO) 1981, 1986,

1-56670-641-6/04/$0.00+$1.50

2000; USEPA, 1987). In 1988, the U.S. EPA Science Advisory Board (SAB) concluded that more research was needed on relating the effects of cumulative, regional, and long-term anthropogenic disturbances to ecosystem responses (USEPA, 1988). Increased research was also needed to develop ecological indicators and protocols for monitoring and to analyze and quantify uncertainty in assessments resulting from monitoring data. The goals of such research were improved detection of ecosystem status and trends, and improved change predictions for ecosystems, as the stresses on them change. The SAB urged EPA to develop a monitoring approach for determining status and trends in ecological conditions and to broaden the range of environmental characteristics and contaminants that were measured.

The SAB recommendations, the emerging vision of ecological risk assessment within EPA (Messer, 1989), and an increasing emphasis in government on "monitoring for results" (Reilly, 1989) led to the creation of the Environmental Monitoring and Assessment Program (EMAP) within EPA's Office of Research and Development (ORD). EMAP's challenge was to develop the tools necessary for measuring the condition of many types of ecological resources and the designs for detecting both spatial and temporal trends (Messer et al., 1991). EMAP used a tiered monitoring approach, which moved from remotely sensed national land cover to regional geographic surveys to index sites. Initially, the focus was on consistent, remotely sensed land cover for the U.S. and monitoring designs and indicators of ecological conditions for major classes of natural resources (e.g., surface waters, estuaries, forests, and wetlands).

By the mid-1990s it became apparent that EMAP alone could not implement and maintain a national monitoring program for the nation's resources without substantial additional funding. Since states and tribes have statutory responsibilities within the CWA to monitor and report on the condition of all surface waters in their jurisdiction, EMAP began to establish partnerships with them. The approaches developed by this partnership would allow the states to more effectively monitor the condition of their aquatic ecosystems in the future.

States and tribes historically dealt with "point source" pollutants and monitored chemical and physical attributes at targeted sites. This approach works well for what it is intended to do but it did not allow statistically valid assessments of water quality in unsampled waters (GAO, 2000). Only three scientifically defensible monitoring approaches will provide coverage of all waters of a state or tribe. The waters can be censused, predictive models can be used, or a probability survey can be made. In the census approach every single waterbody or stream segment within a state or tribal nation has to be visited and the condition measured. Conditions may also be inferred from physical, biological, and chemical process models of an aquatic system when extensive field and laboratory data are available. In a probability survey, a statistical approach is used to sample a subset of all waters, which allows an estimate of the condition of all waters to be made, along with an estimate of the confidence interval surrounding the estimate. In order to meet the CWA requirements for assessing all waters of states or tribes and determining their biological condition in a cost-effective manner, EMAP has focused on the development of unbiased statistical sampling frameworks. Within these frameworks, carefully selected indicators of biological, physical, and chemical conditions are combined with landscape characterization to support scientifically defensible statements about aquatic resources at local, state, regional, and national scales.

Remotely sensed land cover can provide an efficient stratification within a national monitoring design. EMAP initiated a partnership with other federal programs to deliver processed land cover imagery from across the conterminous U.S., resulting in a 30-m resolution land cover dataset (National Land Cover Data, http://www.epa.gov/mrlc/nlcd.html).

29.2 FRAMEWORK FOR A NATIONAL DESIGN

Using a statistical survey design to make unbiased statements on the overall condition of national resources from selected samples is feasible and economically attractive. To make representative statistical inferences requires a design for sampling across diverse ecosystems, indicators that are sensitive to stress and can be measured to determine the condition of the resource at a wide range of sampling locations, reference conditions to act as benchmarks against which to interpret the indicators, and a statistical analysis that incorporates the survey design. The foundational elements of a national monitoring design are the statistical survey design and the ecological indicators.

Statistical Survey Design: Probability-based sampling within a statistical survey design (Cochran, 1977) provides the only unbiased estimate of the condition of an aquatic resource over a large geographic area from a small number of samples. The principal characteristics of a probabilistic design are: (1) the population being sampled is unambiguously described (e.g., all lakes >1 ha in the coastal plain of a state), (2) every element in the population has the opportunity to be sampled with a known probability, and (3) sample selection is carried out by a random process. This approach allows statistical confidence levels to be placed on the estimates and provides the potential to detect changes and trends in condition with repeated sampling. Our design also specifies the information to be collected and at what locations. The validity of the inference rests on the design and subsequent analysis to produce regionally representative information.

EMAP does not rely on a single survey design but tailors the survey design to match the requirements specified by the monitoring objectives, institutional constraints, limitations of available sample frames, and characteristics of the aquatic resource. A common feature of all designs is spatial balance across the aquatic resource over the geographic region for the study. Spatial balance, in conjunction with random selection, improves the "representativeness" of the sample. Stevens (1997) and Stevens and Olsen (in press, a, b) describe the statistical theory and methodology for spatially balanced survey designs. Peterson et al. (1999) discuss the importance of representativeness and the use of probability surveys for lakes in the northeastern U.S. Olsen et al. (1999) have reviewed many national natural resource programs in the U.S., discussing their strengths and weaknesses, and the advantages of probability surveys for monitoring.

The EMAP approach uses classification to better define different strata of aquatic systems for which similar expectations exist. This allows for improved information about each of the strata and can improve statistical estimates of condition. At the coarsest level, EMAP divides aquatic resources into different water body or system types, such as lakes, streams, estuaries, and wetlands. Our rationale is that the biological, chemical, and physical characteristics for these systems are fundamentally different from one another, and we expect different indicators and designs will be necessary. Aquatic resources may be classified as discrete (point), linear extensive, or areal extensive. All small lakes and reservoirs within the U.S. are an example of a discrete resource. That is, it is reasonable to expect that an entire lake can be characterized as a unit. The Great Lakes and estuaries are examples of areal extensive resources, i.e., a single characterization is unlikely to apply to an entire Great Lake or to a large estuary. Wadeable streams and rivers are examples of linear extensive resources. For the survey design they are considered to have only length, and their characterization at a site results in a single value for each indicator. Wetlands are more complex systems from a survey design perspective. They may be an areal extensive resource such as the Everglades, or they may be a discrete resource such as the prairie potholes.

A discrete resource consists of distinct natural units such as small- to medium-sized lakes. Our population inferences for a discrete resource are based on the numbers of units that possess a measured property (e.g., 10% of the lakes are acidic). Extensive resources, on the other hand, extend over large regions in a more or less continuous and connected fashion (e.g., rivers, estuaries) and do not have distinct natural units. Our population inferences here are based on the length or area of the resource. The distinctions between discrete and extensive are not always clear, and in some cases a resource may be viewed as both at different times (e.g., discrete: small estuaries in a region; extensive: sample units within an estuary). Each aquatic resource has its own unique sampling characteristics. The approach we use depends on the nature of the resource and the available information for that resource.

Typically each survey design incorporates other ecosystem and hydrological characteristics to capture important differences in the aquatic resource in the region of interest. We use ecoregions for inland waters and biogeographic provinces for estuaries and coastal waters to capture regional differences in water bodies. Ecoregions are areas which have generally similar ecosystem characteristics: (geology, physiography, vegetation, climate, soils, land use, wildlife, and hydrology (see Omernik, 1995). Ecoregion maps for use in strata development have been compiled (Wiken, 1986; Omernik, 1987, 1995). Inland survey designs may also incorporate U.S. Geological Survey (USGS) hydrological units to define river basins within a state. For estuarine and coastal ecosystems we use biogeographic provinces analogous to those used by the U.S. Fish and Wildlife Service (Terrell, 1979) and the National Oceanographic and Atmospheric Administration (NOAA) (Beasley and Biggs, 1987) for strata development.

For lakes, streams, and estuaries, the distributions of their size is typically very skewed. That is, most lakes are less than 5 ha, most streams are headwater (Strahler 1st order) streams, and most estuaries are small. All of these characteristics may be used to stratify the sample or to select sample sites with unequal probability. Stratification and unequal probability of selection are common in survey designs implemented by EMAP.

Monitoring studies by states, tribes, and other organizations are conducted across years. One approach (rotating basins) is to partition the state into a number of "river basins," sample one or more basins per year, and complete all basins within a few years (typically 4 to 5 years). Other states take 5 years to complete the survey but do so with a state-wide sample each year, accumulating the desired number of sample sites at the end of the 5th year. The latter is achieved in the survey design by defining five panels of sites that are each a probability sample of the streams. More complex panel designs have been used for long-term monitoring programs (Stevens, 2002). Stratification may also be used to sample a subregion of the study area more intensely than others. For example, a special study of a single estuary may be included as part of a state-wide monitoring program for all estuaries.

An aquatic probability survey that incorporates stratification, unequal probability of selection, panels, and intensive study subregions can be complex to design. Herlihy et al. (2000) describe a stream survey design for the mid-Atlantic states that incorporates many common features for streams. EMAP has a process to create these designs and is currently creating statistical software for general use. Complex survey designs must be matched with the appropriate survey analysis procedure to make scientifically defensible estimates for all waters in a study. EMAP has software available to do so. The software is in the form of a library of functions in R statistical software (available at http://www.epa.gov/nheerl/arm). For each site in the sample, the minimum design information required for the statistical analyses is: (1) the stratum for the site, (2) the weight for the site, and (3) the geographic location of the site. More complex two-stage designs require additional information.

The weight for a site is the inverse of the probability of selecting the site. For a design with no stratification and equal probability of selection, the selection (or inclusion) probability depends on the type of resource: discrete, linear, or areal. For example, a sample of 50 lakes from a sample frame list of 1000 lakes results in an inclusion probability of 50/1000 and a weight of 20. That is, each sample lake represents 20 lakes. Similar calculations are done for linear and areal extensive resources with the weight being in units of length (e.g., km) for linear and units of area (e.g., km^2) for areal. The weights are used in the statistical analysis to produce estimates for the entire resource in the study.

Discrete resources such as small- to medium-sized lakes can be selected for sampling by associating each distinct unit with a point in space, e.g., the lake centroid. Each point is assigned a unit length and all these lengths are randomly placed on a single line with length equal to the number of units in the population to be sampled. For example, assume 1000 lakes are in the population, and a sample size of 50 is desired. Then a random number between 0 and 20 is selected, say, 12.3. The first sample lake selected is the one which occurs at 12.3 units down the line. Subsequent lakes in the sample are selected every 20 units thereafter (32.3, 52.3, etc.). This is a simple random sample. To achieve spatial balance in the process, we use hierarchical randomization which employs a hierarchical grid structure with randomization to map the units from geographic space to the line (Stevens and Olsen, in press, b).

A sample for a linear extensive resource such as streams uses essentially the same process as that used for discrete resources. However, the stream segment lengths in the population to be sampled are placed on a line using hierarchical randomization. The stream segments to be sampled are determined using the systematic selection process to select a point on the segment. Any point on the stream network may be selected; it is these points and not the stream segments themselves that are selected. A sample for an areal extensive resource, such as estuaries or large lakes, begins with the generation of a set of potential sample sites within the resource. The potential sample sites are generated by randomly placing a grid over the resource and randomly selecting a point within each grid cell. If a random point in a grid cell does not fall within the resource, no potential point is generated for that cell. The grid density is chosen to ensure that the number of potential sample points generated is greater than the desired sample size. We have found that more stability in expected sample size and spatial balance results when substantially more potential points are generated than the sample size. The potential points are then treated as if they are discrete units and the process follows that for a discrete resource.

The selection process is completed independently for each stratum when the design has more than one stratum. Unequal probability selection is accomplished by differentially "stretching" the units or segments placed on the line. For a discrete resource, if some units are to be selected with twice the probability of other units, then instead of giving them a unit length, they are given two-unit lengths. For linear resources, a segment length is multiplied by two to double the probability of selecting a point on the segment. If subregions of an areal resource are to be sampled with double the probability, then the potential sample points in that subregion are multiplied by two.

Our EMAP designs provide the unbiased statistical framework for where we take our sample, and define how we extend our inference from our samples to the entire population of interest. It is within this context that we have developed indicators which establish the condition of an aquatic ecosystem.

Ecological Indicators: An indicator is one (or more) measure(s) or model(s) that contributes to the description of the condition of the system in question (Jackson et al., 1999). Historically, state and federal agencies have used chemical indicators as surrogates for biological condition when monitoring the condition of aquatic systems but these chemical measures do not necessarily provide a complete picture of the ecological condition. While EMAP does include chemical indicators, we have focused our research on ecological indicators.

Streams: Many of the current EMAP stream community or assemblage indicators come from analysis of multimetric indices formed by combining biological indicators within a taxon (e.g., those indices related to fishes, benthic macroinvertebrates, and algal/plant communities). Fish assemblages contain species representing a variety of trophic levels (omnivores, herbivores, invertivores, planktivores, and piscivores), and require intact ecosystems to thrive. Fish can also reflect multiyear and broad-scale environmental conditions. Thus, they often reflect overall ecosystem condition. Fish are of immense interest to the public, and many species are declining. Stream fishes have been well studied and have been shown to be suitable indicators (Matthews and Heins, 1987; McAllister et al., 1986; Miller and Rolbison, 1980; Minckley, 1973; Moyle, 1976; Plafkin et al., 1989; Platts et al., 1983; Rankin and Yoder, 1991). EMAP uses an index of biotic integrity (IBI, a multimetric biological indicator, similar to Karr et al., 1986) to evaluate the overall fish assemblage as a measure of condition. The IBI approach remains consistent across regions, even though species will change.

Stream macroinvertebrates (a heterogeneous assemblage of animals inhabiting the stream bottom) are important as components of stream ecosystems and for their role in the food web. They integrate changes in the lower trophic levels and because they are not as mobile as fish, they often can provide more information on specific stressors. Stream benthic macroinvertebrates have a long history of use as a biomonitoring tool. Changes in the assemblage structure (abundance and composition) and function can indicate water resource status and trends (Cummins and Klug, 1979; Plafkin et al., 1989). Different stressors on stream ecosystems have been studied and have been found to result in different assemblages or functions (Armitage, 1978; Hart and Fuller, 1974; Hilsenhoff, 1977; Metcalfe, 1989; Resh and Unzicker, 1975). Again, EMAP currently uses an IBI approach for stream macrobenthos in order to evaluate condition.

Plants are the base of the food chain and can act as the primary link between the chemical constituents in aquatic ecosystems and the higher trophic levels. Changes in their abundance or species composition can greatly affect food availability to macroinvertebrates and herbivorous fishes. The algal component of stream periphyton (a mixture of organisms attached to substrates including algae, bacteria, microinvertebrates, and associated organic materials) has been used extensively in the analysis of water quality for several decades (Lange-Bertalot, 1979; Patrick, 1968; Stevenson and Lowe, 1986; Watanabe et al., 1988). It has proven to be a useful indicator for environmental assessments because algal species have rapid reproduction rates and short life cycles and, thus, respond quickly to perturbation (Stevenson et al., 1991). Attached stream algae are therefore valuable as direct indicators of stream condition, as well as for their role in the food chain.

Estuaries: To establish estuarine biological conditions, we currently use benthic invertebrates (e.g., polychaetes, molluscs, crustaceans) and fishes. Benthic invertebrates are reliable indicators of estuarine conditions because they are sensitive to water quality impairments and disturbance (Rakocinski et al., 1997). We have developed a benthic index of estuarine conditions (similar to the stream invertebrate IBI) that incorporates changes in diversity, structure, and abundance of selected estuarine benthic species. The EMAP estuarine benthic index has been used successfully in estimating the ecological condition of Atlantic coast estuaries from Cape Cod, MA, to central Florida, and of estuaries in the Gulf of Mexico (Engle et al., 1994; Engle and Summers, 1999; Paul et al., 2001; Van Dolah et al., 1999). Unlike streams, fishes in estuaries are typically difficult to sample quantitatively. Because these fishes are important commercially and recreationally, we have chosen to use a generalized fish health assessment in EMAP. This assessment consists of examining fishes for gross pathological disorders and pathologies of the spleen (specifically, splenic macrophage aggregates). High rates of gross pathological and/or spleen abnormalities can be associated with environmental contamination.

Landscape: Land-use activities within watersheds are major causes of nonpoint source pollution problems affecting aquatic ecosystems (Comeleo et al., 1996; Jones et al., 1997; Paul et al., 2002; Hale et al., submitted). By using satellite-based remote-sensing techniques, a complete census of the landscape and determination of the different land-cover types in an area can be made. Changes in landscape composition and pattern can influence water quality, the biological condition of streams, and hydrology. As relationships are established between landscape pattern indicators and aquatic ecosystem conditions, a comprehensive low-cost alternative to field monitoring can be developed.

EMAP has used the comprehensive and internally consistent spatial data coverage (NLCD) to develop indicators of status (Jones et al., 1997) and changes in landscape composition and pattern (Jones et al., 2001). Spatial correlations may be made between a watershed's terrestrial components and its aquatic ecosystems. For example, streams draining areas with agriculture on slopes greater than 3% may be at increased risk from sedimentation and potentially from other nonpoint source pollution (Jones et al., 1997). Areas with high amounts of urban development or agriculture (e.g., >60% of the land surface) may be sufficiently degraded so that variation in landscape pattern has little influence on the area's aquatic ecosystem condition. Conversely, areas with very low human occupancy may possess such significant resilience and variation in landscape pattern that these areas will also have little effect on aquatic ecological processes and conditions. It is the area with moderate levels of human occupancy where the pattern of use may have the greatest influence on aquatic ecological conditions. As remote sensing is used to identify and screen areas where pattern matters to aquatic ecological condition, a large reduction in the monitoring costs for aquatic systems may be possible (Jones et al., 2001).

EMAP has expended considerable research effort to understand how natural and anthropogenic variability in our indicators affects status and trend detection under different design considerations (Larsen et al., 1995). For our indicators we have been able to: (1) summarize variance components for several indicators of aquatic condition, (2) evaluate the influence of sources of variability on trend estimation, (3) evaluate the sensitivity of different sample survey design options for trend detection, (4) demonstrate the ability of sample survey designs to detect trends in indicators of condition, given the summarized variance components (McDonald et al., 2002).

Even with efficient designs and indicators, only relative measures of change can be determined, unless there is some benchmark against which current conditions can be measured. These benchmarks or reference conditions, defined in EMAP as the condition of a resource under minimal contemporary human disturbance, allow a scientifically defensible basis upon which to detect impairment. Since minimal disturbance is often not available, we measure least disturbed conditions as a best approximation, which is often associated with models and historical information to estimate minimal disturbance.

Currently, we use a combination of best professional judgmental and probability-based selection for the development of reference conditions. Judgement is used to select sites from excellent to poor along a predefined condition gradient; these are then sampled for the indicator of interest. Indicator values from the reference sites are plotted on a frequency distribution, along with those derived from probability sampling in the same area of interest. Best professional judgement is again used to determine the percentile of the probability sample associated with the lowest level of excellent conditions as defined by the reference sites, and that becomes the nominal threshold for excellent conditions for the indicator within the area of interest. In all cases, we provide the explicit criteria by which we judged reference sites to be classified along the condition gradient. The criteria are sufficiently rigorous that the sites can be classified consistently.

EMAP also has academic partners that focus their research on many of these indicator and design issues through EPA's STAR Grants Program (www.epa.gov/ncer). Examples of recent STAR-funded research supporting EMAP include new and better indicators of coastal and riverine condition through the EaGLes (Atlantic coast, Pacific coast, Gulf coast, Great Lakes, and Mississippi River), developing new and improved statistical design and analysis approaches for aquatic ecosystem status and trend determination through statistical centers, and new and better methods for regional scaling and for classification schemes for aquatic ecosystems and landscapes.

29.3 GEOGRAPHIC APPLICATIONS

The first regional scale geographic study in EMAP was conducted in the mid-Atlantic area (EMAP's Mid-Atlantic Integrated Assessment or MAIA) and served as the proof of concept for the EMAP approach. We focused on characterizing both the ecological condition of the region and assessing the important environmental stressors associated with impairment. This resulted in the first State of the Region baseline assessments for individual types of ecological systems [estuarine (Paul et al., 1998; Kiddon et al., 2002), stream (USEPA, 2000), and landscape (Jones et al., 1997)].

The EMAP probabilistic monitoring design for estuaries (Stevens, 1994) considered estuaries an extensive resource and used three strata based on the subpopulations of interest — large estuaries (>260 km^2), small estuaries (2.6 to 260 km^2), and large tidal rivers (>260 km^2) (Paul et al., 1999) — to establish condition. EMAP conducted surveys of estuarine resources in the northeast Atlantic, including MAIA, between 1990 and 1993 (Strobel et al., 1995, Paul et al. 1999). These data were combined with data from multiple agencies and programs to assess the condition of MAIA estuaries (Figure 29.1, USEPA, 1998). From 1997 to 1998 EMAP returned to the MAIA estuaries to: (1) improve characterization of the resource at multiple scales, (2) supplement the early core indicators with additional indicators (e.g., eutrophication), and (3) demonstrate the utility and value added of common designs and indicators being used by multiple agencies (Kiddon et al., 2002).

We also used a probabilistic sampling survey to characterize the condition of the wadeable streams in MAIA. To ensure characterization of the stream reaches in the region it was necessary to classify streams on the basis of their size, using stream order as a size surrogate. We also linked our EMAP sample survey data with data from the more deterministic, temporally intensive effort of the USGS National Water Quality Assessment Program (NAWQA, Hirsch et al., 1988). The information generated in MAIA has helped identify areas of concern and the most prevalent stressors (USEPA, in press). In addition, the study provided baseline information for comparing ecological conditions of wadeable streams (Figure 29.2, USEPA, 2000).

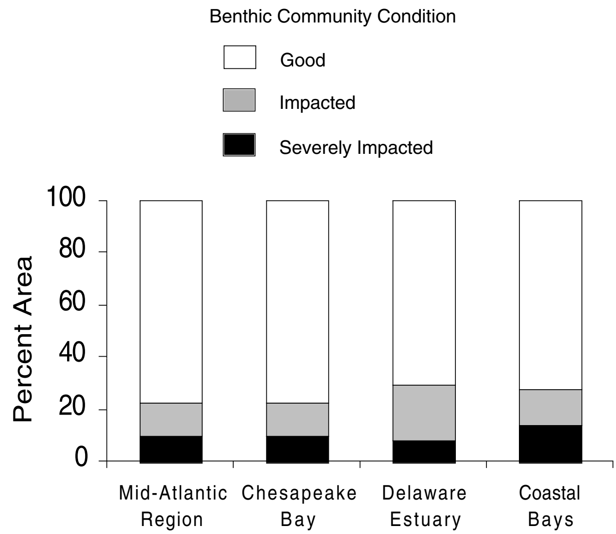


FIGURE 29.1 Estuarine benthic community condition in MAIA (USEPA. 1998. Condition of the Mid-Atlantic Estuaries, EPA/600-R-98-147, U.S. Environmental Protection Agency, Washington, D.C.).

In MAIA satellite-derived imagery was used to develop a land cover map for the mid-Atlantic region (Jones et al., 1997). This, combined with spatial data bases on biophysical features (e.g., soils, elevation, human population patterns), was used to produce an ecological assessment of the impact of changing landscape conditions. Using fine-scale spatial resolution (e.g., 30 to 90 m) to census MAIA, we were able to analyze and interpret environmental conditions of the 125 watersheds in the mid-Atlantic region based on 33 landscape indicators (Jones et al., 1997).

Current Geographic Research: In MAIA, the feasibility of the EMAP approach was extensively tested. The ecosystems contained in MAIA are relatively homogeneous and representative primarily of the mid-Atlantic and the southeastern U.S. For the EMAP approach to be applicable to other regions of the country, we needed to reduce the uncertainty associated with our indicators and designs in more ecologically diverse areas of the country, in order to demonstrate that the EMAP approach can be used nationally. Therefore, EMAP is currently focusing a major portion of its efforts on design and monitoring research on Western aquatic resources and the potential stressors affecting them (e.g., habitat modification, sedimentation, nutrients, temperature, grazing, timber harvest). This effort, Western EMAP, is the largest comprehensive study conducted by EPA on the ecological condition of the West. Western EMAP is a cooperative venture involving 12 western states, tribes, universities, and the western EPA regional offices. Initiated in 1999, Western EMAP is poised to establish the condition of aquatic ecosystems throughout the West, and establish the feasibility of our approach nationally.

Biological, physical habitat, and water chemistry indicators are being used in our assessment of stream conditions in order to characterize the biological communities,

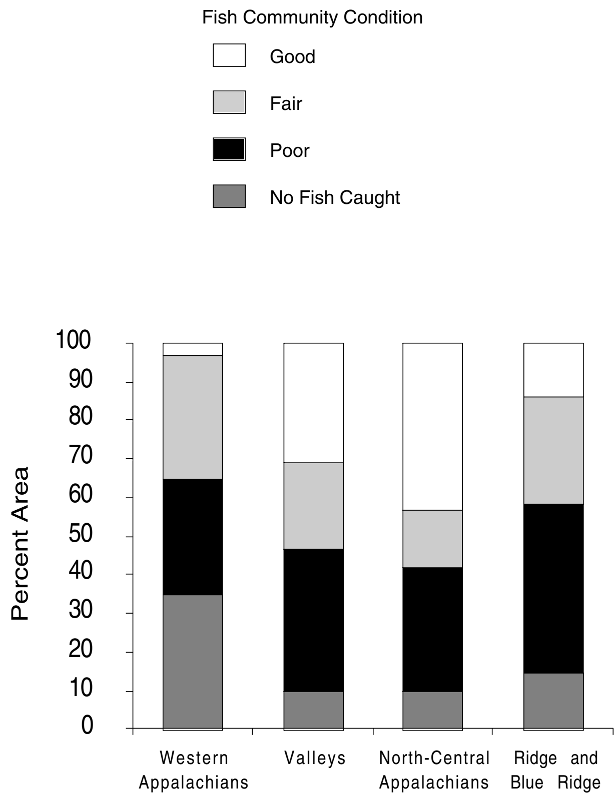


FIGURE 29.2 Geographic differences in the fish index of biotic integrity for the mid-Atlantic Highlands area of MAIA (USEPA. 2000. Mid-Atlantic Highlands Streams Assessment. EPA/903/R-00/015. U.S. Environmental Protection Agency, Washington, D.C.).

TABLE 29.1
Core EMAP Surface Water Indicators

Core	Additional Indicators	
Conventional water quality parameters	Fish tissue chemistry/toxics	Amphibians
Fish assemblage	Sediment metabolism	Bacteria
Macroinvertebrate assemblage	Sediment chemistry	Biomarkers
Periphyton assemblage	Sediment toxicity	Riparian birds
Physical habitat structure	Water column toxicity	
Riparian vegetation		

habitat attributes, and levels of stress. All of our state partners in the Western EMAP have agreed to sample a core group of stream indicators (Table 29.1). An additional set of research indicators (riparian condition, continuous temperature, microbial organisms, sediment toxicity, additional tissue contaminants, etc.) are also being experimentally tested on a smaller scale in focused areas.

Through the use of remote-sensing techniques and the availability of the NLCD data, land cover for the entire western U.S. is available. The western U.S. has presented a challenge to developing and interpreting landscape indicators. There are hardware limitations to processing the data for an area this large. There are also some unique stresses on western landscapes, including grazing and timber harvest, that do not result in changes in land-cover types but rather in substantially altered states of land cover conditions. We are assessing spatial variability in landscape pattern and the degree to which landscape patterns influence the conditions of Western estuaries and inland surface waters. By linking watershed-level aquatic resource condition with landscape patterns, we are attempting to evaluate how landscape data can contribute to the assessment of the condition of aquatic resources at many scales across the western U.S.

Our initial hierarchical design for western coastal estuarine resources included locations in the small estuaries of Washington, Oregon, and California, as well as sites in Puget Sound, the Columbia River Estuary, and San Francisco Bay. Our intent was to modify the existing state programs as little as possible, but still meet our probabilistic requirements, and provide estimates of estuarine condition and guidance for determining estuarine reference conditions.

Our Western EMAP estuarine work has now been merged with our National Coastal Assessment (NCA). In NCA we demonstrate a comprehensive national assessment approach to establishing the condition of the nation's estuaries and nearshore coastal environments. There are 24 marine coastal states (including Alaska and Hawaii) and Puerto Rico that are sampling their estuaries using a consistent EMAP design and a consistent set of biological and stressor indicators (Table 29.2) to estimate the condition of their estuarine resources. Other partners in the EMAP NCA are the EPA Regions, EPA's Office of Water, USGS, and NOAA.

On a smaller scale, EMAP partners with each of the ten EPA Regions and sponsors one to two Regional-EMAP (R-EMAP) projects which focus on more local problems (USEPA, 1993; http://www.epa.gov/emap/remap/index.html). These smaller-scale projects complement our larger geographic assessments by allowing us to develop and test new indicators and designs across wide biogeographic and political boundaries. Our R-EMAP studies also provide additional opportunities for capacity building and technology transfer to potential partners.

TABLE 29.2
Core EMAP Coastal Indicators

Water Column	Sediments	Biota
Dissolved oxygen	Grain size	Benthic community structure and abundance
Salinity	Total organic carbon	Fish community structure and abundance
Temperature	Sediment chemistry	Fish pathologies
Depth	Sediment toxicity	Fish tissue residues
pH		Submerged vegetation
Nutrients		
Chlorophyll-*a*		

29.4 EMAP INFORMATION MANAGEMENT

As with any large environmental effort, effective information management systems are critical (NRC, 1995), along with rigorous quality assurance and quality control procedures. EMAP information management is an integral part of the landscapes, surface waters, and coastal groups and is designed to support programmatic and policy objectives (Hale et al., 1998, 1999). Stringent quality assurance procedures for sample collection and processing in streams (Lazorchak et al., 1998; Peck et al., 2000) and estuaries (Heitmuller, 2001) assures that data of the highest quality are incorporated into our databases. These data are collected with a consistent design by consistent methods, are of known quality, and are well described by metadata. EMAP data are used primarily by study participants to fulfill study objectives of assessing environmental conditions (Hale and Buffum, 2000), but also have other scientific and management uses.

Our data management approach is guided by both considerations for conducting EMAP statistical analyses and for sharing the data. Use of common standards such as the Integrated Taxonomic Information System for species names and the Federal Geographic Data Committee metadata standard facilitates data exchange with other programs. EMAP data are publicly accessible on the EMAP Web site (www.epa.gov/emap). Long-term archival and data sharing is being ensured through periodic uploads to EPA's STORET system (www.epa.gov/storet) and other archives, such as the Ecological Society of America's *Ecological Archives* (e.g., Riitters et al., 2000; Hale et al., 2002).

29.5 MEASURES OF CURRENT SUCCESS

Implementation and routine use of EMAP's statistical design and ecological indicators by states and other monitoring agencies is an unmistakable sign of acceptance. To date, more than 30 states have implemented or are testing EMAP protocols. In addition, the U.S. Forest Service has adopted EMAP's rotating panel design for its national Forest Health Monitoring program. The EMAP probability design approach has been incorporated into USEPA Office of Water guidance on Consolidated Assessment and Listing Methodology for implementation under the CWA. In part, state acceptance has likely been based on sampling cost-effectiveness. For example, to determine the condition of lakes in the northeastern U.S. with respect to phosphorus levels, the northeastern states sampled 4219 lakes (out of approximately 11,076). However, their findings could not be applied to the remaining lakes that were not sampled. Using a probability survey, only 344 lakes had to be sampled to make an estimate of the condition of all 11,076 lakes and to provide the statistical uncertainty for the estimate (Whittier et al., 2002). By the states' own admissions, their estimates were biased toward problem lakes, applied only to the lakes reported, and were not directly comparable from state to state. The probability survey estimates were objective, representative, accurate, and cost-effective (only 8% as many lakes needed to be sampled).

As the EMAP approach becomes more accepted, we expect an increase in the use of program results in environmental decisions and policies. EMAP data on streams in the mid-Atlantic region have provided the justification for EPA to require a full Environmental Impact Statement for proposed mountain-top removal for mineral extraction in West Virginia. And, in Maryland, the Department of Natural Resources used an EMAP-type approach for monitoring data collection reported in The State of Maryland's Freshwater Streams (EPA/903/R-99/023), which was being used to make planning decisions in support of the Governor's Smart Growth Initiative.

The ability to detect trends over time will complete validation for EMAP's statistical design. It will require long-term implementation of EMAP methods, as continued sampling over multiple years will be necessary to analyze for trends. However, we have begun looking at change detection between sampling periods as a prelude to having sufficient data to detect trends. We have been able to detect significant change in the invertebrate benthic community in the Chesapeake Bay between 1991 and 1993 and from 1997 to 1998 (Kiddon et al., 2002).

29.6 FUTURE DIRECTIONS FOR EMAP MONITORING RESEARCH

Integrated Monitoring: Historically, EMAP has focused its research on establishing effective ways for states to meet the CWA goal of assessing the overall quality of their waters (Section 305b) and whether that quality is changing over time. However, the CWA also requires states to identify problem areas and to take the management actions necessary to resolve those problems. These include nonpoint source control (Section 319) issues, total maximum daily load (TMDL) allocation (Section 303d) and NPDES permitting (Section 402). Effective monitoring information is certainly needed to effectively address these CWA goals.

EMAP is developing an integrated monitoring process that can address these issues (Brown et al., submitted). This framework is predicated on the states' using a probability-based survey for an accurate, broad-scaled assessment of the quality of all waters. From this, a staged approach can be taken which then refines the assessment through screening models to determine which areas have a high probability of impairment. The development of these probability of impairment models is a key future research area of EMAP.

Probability survey data can be analyzed to provide estimates of areal extent of condition (e.g., 20% of the population has a fish IBI <3). However, this information can be combined with remotely sensed landscape data to form empirical relationships (e.g., percent of agricultural land on steep slopes vs. stream sediment concentration). These relationships can then be used to predict the probability of stream impairment by sediments in unmonitored areas. As additional data are collected, these simple empirical models can be further refined and validated. Ultimately, such models could be sufficiently accurate to allow 303(d) listing and delisting decisions to be made prior to the TMDL development.

We are currently conducting multiple watershed studies to determine quantitative relationships between landscape metrics (e.g., riparian wetlands, urbanization, agricultural land cover) and aquatic condition indicators (e.g., stream total nitrogen concentration, sediment contamination, benthic invertebrate community condition) (Comeleo et al., 1996; Paul et al., 2002, Hale submitted). From this, we will begin to establish empirical relationships for other prevalent stressors affecting aquatic resource condition.

This new, combined monitoring approach builds upon the data collected at each stage. By incorporating additional layers of information at each point we can cost-effectively refine where impairments of aquatic resources are occurring and reduce the uncertainty in management's decision-making. The process is iterative, both within the process and through time. However, the entire process is predicated on state and tribal implementation of a probability survey design for their ambient monitoring. This new combined approach can give states and tribes enormous flexibility when developing their monitoring programs to include not only condition but ways to deal with nonpoint source listing of impaired waters for TMDL development.

Large Rivers: Large river systems are regulated for multiple, and often conflicting, uses without reliable information on their long-term condition. These large rivers are the inland receiving waters for the majority of our landmass and can serve to organize TMDL efforts for thousands of smaller watersheds nationwide. Large rivers are the link between small upland streams and our estuaries, but they are currently too large and complex for conventional environmental monitoring. Consequently, the large rivers have routinely been ignored and now represent a scientific gap in our understanding of the flowing waters of the U.S.

As part of Western EMAP we have begun developing the designs and indicators necessary to determine the baseline condition of the Upper Missouri River. We are currently expanding our geographic focus to the Lower Missouri, Mississippi, and Ohio Rivers. Our intent is to reduce the remaining scientific uncertainties by developing new and better biological indicators for use on these large rivers, and establishing the necessary probabilistic framework for unbiased and representative sampling. This work will be a research partnership between EPA, states, tribes, river authorities, and other federal agencies.

As the condition of the large rivers of the Mississippi River Basin are established, the groundwork will be laid for exploring integrated watershed processes. We will be able to examine the condition of flowing waters from the headwater streams (and their associated stressors) in the upper Midwest to the nearshore areas of the Gulf of Mexico (an area draining approximately 41% of the conterminous U.S.). From these baselines, a statistically valid means of documenting the improvements in condition will be available and will allow an unbiased assessment of policy and program performance.

29.7 SUMMARY

A scientifically rigorous determination of the condition of an aquatic resource is fundamental to all subsequent research, modeling, protection, and restoration issues. Environmental risk characterization is predicated on a knowledge of condition and the rate at which that condition is changing (USEPA, 1996). Through a probability-based sampling design with ecological indicators, the EMAP approach does provide a statistically valid basis for determining aquatic ecological condition; and when implemented over time, change detection, and likely trends in condition. The use of defensible estimates of environmental condition provides states and EPA a better means of identifying water quality problems, setting priorities, and carrying out key management and regulatory activities. Additionally, the benefits derived from a state's or EPA's protection and restoration strategies can be quantified and documented. Currently, this is the only practical approach to large-scale, performance-based aquatic environmental reporting that is required by Congress.

REFERENCES

Armitage, P.D. 1978. Downstream changes in the composition, numbers, and biomass of bottom fauna in the Tees below Cow Green Reservoir and an unregulated tributary, Maize Beck, in the first five years after impoundment. *Hydrobiologia* 58:145–156.

Beasley, B. and R. Biggs. 1987. Near coastal water segmentation. Report by College of Marine Studies. University of Delaware.

Brown, B.S., N. Detenbeck, and R. Eskin. Submitted. Integrating 305(b) and 303(d) monitoring and assessment of state waters. *Environ. Monit. Assess.*

Cochran, W.G. 1977. *Sampling Techniques.* 3rd ed. John Wiley & Sons. New York. 428 pp.

Comeleo, R.L., J.F. Paul, P.V. August, J. Copeland, C. Baker, S.S. Hale, and R.L. Latimer. 1996. Relationships between watershed stressors and sediment contamination in Chesapeake Bay estuaries. *Landscape Ecol.* 11(5):307–19.

Cummins, K.W. and M.J. Klug. 1979. Feeding ecology of stream invertebrates. *Annu. Rev. Ecol. Syst.* 10:147–172.

Engle, V.D. and J.K. Summers. 1999. Refinement, validation, and application of a benthic condition index for the Gulf of Mexico estuaries. *Estuaries* 22:624–635.

Engle, V.D., J.K. Summers, and G.R. Gaston. 1994. A benthic index of environmental condition of the Gulf of Mexico estuaries. *Estuaries* 17:372–384.

GAO (U.S. General Accounting Office). 1981. Better Monitoring Techniques Are Needed to Assess the Quality of Rivers and Streams. Volume 1. U.S. General Accounting Office, Washington, D.C.

GAO. 1986. The Nation's Water: Key Unanswered Questions About the Quality of Rivers and Streams. U.S. General Accounting Office, Washington, D.C.

GAO. 2000. Water Quality: Identification and Remediation of Polluted Waters Impeded by Data Gaps. U.S. General Accounting Office, Washington, D.C.

Hale, S.S. and H.W. Buffum. 2000. Designing environmental monitoring databases for statistical analyses. *Environ. Monit. Assess.* 64:55–68.

Hale, S.S., M.H. Hughes, J.F. Paul, R.S. Mcaskill, S.A. Rego, D.R. Bender, N.J. Dodge, T.L. Richter, and J.L. Copeland. 1998. Managing scientific data: The EMAP approach. *Environ. Monit. Assess.* 51:429–440.

Hale, S.S., M.M. Hughes, C.J. Strobel, H.W. Buffum, J.L. Copeland, and J.F. Paul. 2002. Coastal ecological data from the Virginian Biogeographic Province, 1990–1993. *Ecology* 83:2942 and *Ecol. Arch.* E083–057.

Hale, S.S., J.F. Paul, and J.F. Heltshe. Submitted. Using landscape characteristics of watersheds to find impaired bottom communities in estuaries. *Estuaries*.

Hale, S., J. Rosen, D. Scott, J. Paul, and M. Hughes. 1999. EMAP Information Management Plan: 1998–2001. EPA/600/R-99/001a. U.S. Environmental Protection Agency, Washington, D.C.

Hart, C.W., Jr. and S.L.H. Fuller. 1974. *Pollution Ecology of Freshwater Invertebrates*. Academic Press, New York.

Heitmuller, T. 2001. National Coastal Assessment. Quality Assurance Project Plan, 2001–2004. EPA/620/R-01/002. U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL.

Herlihy, A.T., D. Larsen, S. Paulsen, N. Urquhart, and B. Rosenbaum. 2000. Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP Mid-Atlantic Pilot Study. *Environ. Monit. Assess.* 63:95–113.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin 100. Wisconsin Department of Natural Resources, Madison, WI.

Hirsch, R.M., Alley, W.M., and Wilber, W.G. 1988. Concepts for a National Water-Quality Assessment Program. U.S. Geological Survey Circular 1021. U.S. Geological Survey. Denver, CO.

Jackson, L., J. Kurtz, and W. Fisher, Eds. 1999. Evaluation Guidelines for Ecological Indicators. EPA/620/R-99/005. U.S. Environmental Protection Agency. Washington, D.C.

Jones, K.B., K.H. Riitters, J.D. Wickham, R.D. Tankersley, Jr., R.V. O'Neill, D.J. Chaloud, E.R. Smith, and A.C. Neale. 1997. An Ecological Assessment of the United States Mid-Atlantic Region: A Landscape Atlas. EPA/600/R-97/130. U.S. Environmental Protection Agency. Washington, D.C.

Jones, K.B., A.C. Neale, M.S. Nash, R.D. Van Remotel, J.D. Wickham, K.H. Riitters, and R.V. O'Neill. 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States mid-Atlantic Region. *Landscape Ecol*. 16:301–312.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and L.J. Schlosser. 1986. Assessment of Biological Integrity in Running Water: A Method and its Rationale. Special Publication 5. Illinois Natural History Survey. Champaign, IL.

Kiddon, J.A., J.F. Paul, C.S. Strobel, B.S. Brown, H.W. Buffum, J.L. Copeland, P.M. Mantel, and M.M. Hughes. 2002. Mid-Atlantic Integrated Assessment (MAIA) — Estuaries 1997–98 Summary Report. EPA/620/R-02/003, U.S. Environmental Protection Agency, Narragansett, RI.

Lange-Bertalot, H. 1979. Pollution tolerance as a criterion for water quality estimation. *Nova Hedwigia* 64:285–304.

Larsen, D.P., N.S. Urquhart, and D.L. Kugler. 1995. Regional scale trend monitoring of indicators of trophic condition of lakes. *Water Resour. Bull.* 31:117–140.

Lazorchak, J.M, D. J. Klemm, and D.V. Peck. 1998. Environmental Monitoring and Assessment Program — Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94-004F. U.S. Environmental Protection Agency, Washington, D.C.

Matthews, W.J., and D.C. Heins. 1987. *Community and Evolutionary Ecology of North American Stream Fishes*. University of Oklahoma Press, Norman.

McAllister, D.C., S.P. Platania, F.W. Schueler, M.E. Baldwin, and D.S. Lee. 1986. Ichthyofaunal Patterns on a Geographic Grid. In: *The Zoogeography of North American Freshwater Fishes*. Hocutt, C.H. and E.O. Wiley, Eds. Wiley & Sons, New York. p. 1751.

McDonald, M.E., S. Paulsen, R. Blair, J. Dlugosz, S. Hale, S. Hedtke, D. Heggem, L. Jackson, K.B. Jones, B. Levinson, A. Olsen, J. Stoddard, K. Summers, and G. Veith. 2002. Environmental Monitoring and Assessment Program Research Strategy. EPA/620/R-02/002. U.S. Environmental Protection Agency, Washington, D.C.

Messer, J.J. 1989. Keeping closer watch on ecological risks. *EPA J.* 15:34–36.

Messer, J.J., R.A. Linthurst, and W.S. Overton. 1991. An EPA program for monitoring ecological status and trends. *Environ. Monit. Assess.* 17:67–78.

Metcalfe, J.L. 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. *Environ. Pollut.* 60:101–139.

Miller, R.J., and H.W. Rolbison. 1980. *The Fishes of Oklahoma*. Oklahoma State University Press, Stillwater.

Minckley, W.L. 1973. *Fishes of Arizona*. Sims Printing, Phoenix.

Moyle, P.B. 1976. *Inland Fishes of California*. University of California Press, Berkeley.

National Research Council. 1977. *Environmental Monitoring*. Volume IV. National Academy of Sciences. Washington, D.C. 150 p.

National Research Council. 1995. *Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data.* National Academies Press, Washington, D.C.

Olsen, A.R., J. Sedransk, D. Edwards, C.A. Gotway, W. Liggett, S. Rathbun, K.H. Reckhow, and L.J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. *Environ. Monit. Assess.* 54:1–45.

Omernik, J.M. 1987. Ecoregions of the conterminous United States (map supplement). *Ann. Assoc. Am. Geogr.* 77:118–125.

Omernik, J.M. 1995. Ecoregions — A Framework for environmental management. In: *Biological Assessment and Criteria — Tools for Water Resource Planning and Decision Making*. Davis, W.S. and T.P. Simon, Eds. Lewis Publishers, Boca Raton, FL. p. 49–62.

Patrick, R. 1968. The structure of diatom communities in similar ecological conditions. *Am. Nat.* 102:173–183.

Paul, J.F., R.L. Comeleo and J. Copeland. 2002. Landscape metrics and estuarine sediment contamination in the mid-Atlantic and southern New England regions. *J. Environ. Qual.* 31:836–845.

Paul, J.F., J.H. Gentile, K.J. Scott, S.C. Schimmel, D.E. Campbell, and R.W. Latimer. 1999. EMAP–Virginian Province Four-Year Assessment (1990–93). EPA/620/R-99/004. U.S. Environmental Protection Agency, Washington, D.C.

Paul J.F., K.J. Scott, D.E. Campbell, J.H.Gentile, C.S. Strobel, R.M. Valente, S.B. Weisberg, A.F. Holland, and J.A. Ranasinghe. 2001. Developing and applying a benthic index of estuarine condition for the Virginian Biogeographic Province. *Ecol. Indicat.* 1:83–99.

Paul J.F., C.J. Strobel, B.D. Melzian, J.A. Kiddon, J.S. Latimer, D.E. Campbell, and D.J. Cobb. 1998. State of the estuaries in the mid-Atlantic region of the United States. *Environ. Monit. Assess.* 51:269–84.

Peck, D.V., J.M. Lazorchak, and D.J. Klemm. 2000. Environmental Monitoring and Assessment Program — Surface Waters: Western Pilot Study Field Operations Manual for Wadeable Streams. U.S. Environmental Protection Agency, Corvallis, OR.

Peterson, S.A., N.S. Urquhart, and E.B. Welch. 1999. Sample representativeness: A must for reliable regional lake condition estimates. *Environ. Sci. Technol.* 33:1539–1565.

Plafkin, J.L., M.T. Barbour, K.D. Proter, S.K. Gross, and R.M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. EPA 444/4-89/001. U.S. Environmental Protection Agency, Washington, D.C. 172 pp.

Platts, W.S., W.F. Megahan, and G.W. Minshall. 1983. Methods for Evaluating Stream, Riparian, and Biotic Conditions. General Technical Report INT-138. U.S. Forest Service, Ogden, UT.

Rakocinski, C.F., S.S. Brown, G.R. Gaston, R.W. Heard, W.W. Walker, and J.K. Summers. 1997. Macrobenthic responses to natural and contaminant-related gradients in northern Gulf of Mexico estuaries. *Ecol. Appl.* 7:1278–1298.

Rankin, E.T., and C.O. Yoder. 1991. The Nature of Sampling Variability in the Index of Biotic Integrity (IBI) in Ohio Streams. Ohio Environmental Protection Agency, Columbus, OH.

Reilly, W.K. 1989. Measuring for environmental results. *EPA J.* 25:2–4.

Resh, V.H. and J.D. Unzicker. 1975. Water quality monitoring and aquatic organisms: The importance of species identification. *J. Water Pollut. Control. Fed.* 47:9–19.

Riitters, K.H., J.D. Wickham, J.E. Vogelmann, and K.B. Jones. 2000. National land-cover pattern data. *Ecology* 81:604 and *Ecol. Arch.* E081-004.

Stevens, D.L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics* 8:167–195.

Stevens, D.L., Jr. 2002. Sampling design and statistical analysis methods for the integrated biological and physical modeling of Oregon streams. OPSW-ODFW-2002-07. Oregon Department of Fish and Wildlife, Portland, OR.

Stevens, D.L., Jr. and A.R. Olsen 1999. Spatially restricted surveys over time for aquatic resources. J. *Agricul. Biol. Environ. Stat.* 4(4):415–428.

Stevens, D.L., Jr. and A.R. Olsen (in press, a). Variance estimation for spatially balanced samples of environmental resources. *Environmetrics*.

Stevens, D.L., Jr. and A.R. Olsen (in press, b). Spatially-balanced sampling of natural resources in the presence of frame imperfections. *J. Am. Stat. Assoc.*

Stevenson, R.J. and R.L. Lowe. 1986. Sampling and interpretation of algal patterns for water quality assessments. In: *Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems*. Isom, B.G. Ed. ASTM STP 894:118–149. American Society for Testing and Materials, Philadelphia, PA.

Stevenson, R.J., C.G. Peterson, D.B. Kirschtel, C.C. King, and N.C. Tuchman. 1991. Density-dependent growth, ecological strategies, and effects of nutrients and shading on benthic diatom succession in streams. *J. Phycol*. 27:59–69.

Strobel, C.J., D.J. Keith, H.W. Buffum, and E.A. Petrocelli. 1995. Statistical Summary: EMAP-Estuaries Virginian Province — 1990–1993. EPA/620/R-94/026. U.S. Environmental Protection Agency, Washington, D.C.

Terrell, T.T. 1979. Physical Regionalization of Coastal Ecosystems of the United States and its Territories. Office of Biological Services, U.S. Fish and Wildlife Service, FWS/OBS-79/80.

USEPA (U.S. Environmental Protection Agency). 1987. Surface Water Monitoring: A Framework for Change. Office of Water and Office of Policy, Planning, and Evaluation. U.S. Environmental Protection Agency, Washington, D.C.

USEPA. 1988. Future Risk: Research Strategies for the 1990s. Science Advisory Board. SAB-EC-88-040. U.S. Environmental Protection Agency, Washington, D.C.

USEPA. 1993. R-EMAP — Regional Environmental Monitoring and Assessment Program. EPA 625/R-93/012. U.S. Environmental Protection Agency. Washington, D.C.

USEPA. 1996. Strategic Plan for the Office of Research and Development. EPA/600/R-96/059. U.S. Environmental Protection Agency. Washington, D.C.

USEPA. 1998. Biological Criteria: Technical Guidance for Streams and Small Rivers. EPA-882-B-98-003. Office of Water. U.S. Environmental Protection Agency, Washington, D.C.

USEPA. 1998. Condition of the Mid-Atlantic Estuaries. EPA 600-R-98-147. U.S. Environmental Protection Agency, Washington, D.C.

USEPA. 2000. Mid-Atlantic Highlands Streams Assessment. EPA/903/R-00/015. U.S. Environmental Protection Agency, Washington, D.C.

USEPA. in press. Mid-Atlantic Integrated Assessment (MAIA): State of the Flowing Waters Report. U.S. Environmental Protection Agency, Corvallis, OR.

Van Dolah, R.F., J.L. Hyland, A.F. Holland, J.S. Rosen, and T.R. Snoots. 1999. A benthic index of biological integrity for assessing habitat quality in estuaries of the southeastern USA. *Marine Environ. Res*. 48:269–283.

Watanabe, T., K. Asai, and A. Houki. 1988. Numerical Water Quality Monitoring Of Organic Pollution Using Diatom Assemblages. In: *Proceedings of 9th Diatom Symposium*, September 1986. Frank Round Press, Bristol, U.K. pp. 123–141.

Wiken, E. 1986. Terrestrial Ecozones of Canada: Ottawa. Environment Canada, Ecological Land Classification Series No. 19. 26 pp.

Whittier, T.R., S.G. Paulsen, D.P. Larsen, S.A. Peterson, A.T. Herlihy, and P.R. Kaufmann. 2002. Indicators of ecological stress and their extent in the population of northeastern lakes: A regional-scale assessment. *BioScience* 52:235–247.

30 The U.S. Forest Health Monitoring Program

K. Riitters and B. Tkacz

CONTENTS

30.1 INTRODUCTION

Historically, forest monitoring systems were built to meet the information needs for timber harvest scheduling, insect and disease control, and other forest management concerns.[1] In the past 25 years, the demand for new information has led to new monitoring systems.[2] Forests are increasingly viewed as holistic systems that can only be monitored through an integrated approach to sustainable forest management that considers the ecological and social aspects of forests.[3,4] Some of the new information requirements have been addressed through the FHM program, a cooperative and integrated approach to collecting data and reporting on many aspects of forest health. Here we provide an overview of the FHM program, beginning with a brief history and summary of the conceptual approaches to forest health monitoring. We then describe current operations and development efforts, and give several examples of how the program is addressing forest health issues in the U.S.

30.2 HISTORY AND MANAGEMENT OF FHM

Forest Health Monitoring grew from two related seeds that were sown in the 1980s in response to concern for the effects of air pollution on forest vegetation. As part of the National Acid Precipitation Assessment Program (NAPAP), the Forest Service established the National Vegetation Survey (NVS) to conduct field surveys of acid rain and ozone impacts on forests.[5] Several years later, the Environmental Protection

1-56670-641-6/04/$0.00+$1.50

Agency established the Environmental Monitoring and Assessment Program (EMAP)[6] that included the EMAP-Forests component.[7] Within a few years, the NVS and EMAP-Forests were combined with additional federal and state partners to form the cooperative FHM program.[8]

Early efforts focused on reviewing existing forest inventory programs,[9] candidate indicators,[10] sample designs,[11,12] and auxiliary data.[13] There were many field tests of proposed procedures.[14–21] The tests facilitated the development of field manuals,[22] quality assurance plans,[23,24] and information management systems.[25]

The first implementation of FHM in 1990 was by six northeastern states: Maine, New Hampshire, Vermont, Massachusetts, Connecticut, and Rhode Island.[26,27] States in the southern region joined in the following year[28] and reports of tree and crown conditions were produced.[29–32] Today, with full implementation of FHM, the Forest Service (including the State and Private Forestry (S&PF) program, National Forest System (NFS), and Research and Development divisions) and states are the primary cooperators.

The FHM program initially established plots and conducted surveys in parallel with existing Forest Service programs such as Forest Inventory and Analysis (FIA) and S&PF. In the mid-1990s, there was an effort to integrate FHM with those other programs,[3,33] and this was largely achieved by the year 2000. Since 1999, FIA is responsible for field plot establishment and most ground-based measurements. Forest health measurements are made on a plot network known as Phase 3 of an expanded FIA program.[34] Phase 3 consists of a subset of the FIA timber inventory plot network (Phase 2) where plots are visited to collect an extended suite of ecological data.

As part of S&PF, the Forest Health Protection (FHP) program has long coordinated an extensive survey effort aimed at identifying forest health problems.[35] These surveys provide maps of problem areas, and they are supplemented by directed ground surveys in some cases.[36] The FHP program also conducts follow-up investigations to evaluate changes in forest health that are observed on the plot network or in surveys.[37] Most of this survey work was integrated with FHM in 1998. The integration of FHM with other programs has resulted not only in efficiency for full implementation but also in the standardization of protocols across states and regions, which, in turn, has allowed the delivery of consistent databases for forest health assessments.

While early FHM objectives addressed air pollution impacts on forests, subsequent development has addressed new concerns including the goal of sustainable forest management as embodied in the Montréal Process Criteria and Indicators.[38,39] The Montréal Process is an agreed-upon national basis for strategic forest planning,[40] national resource assessments,[41] and forest health monitoring.[42,43] The criteria and indicators address social and economic goals as well as ecological goals. Together with FIA and FHP, the FHM program delivers data and assessments pertinent to three criteria — conservation of biodiversity, maintenance of forest ecosystem health, and conservation and maintenance of soil and water resources. Biodiversity indicators in the Montréal Process address forest extent, protected status, and fragmentation. Forest ecosystem health indicators address air pollution impacts, forest disturbance regimes, and biological functioning. Soil indicators include erosion, compaction, and other physical and chemical properties. Adoption of the Montréal

Process with its set of common indicators has made it easier to assess and report on FHM data for a diverse set of stakeholders.

The FHM program has three levels of internal management. A national Steering Committee is comprised of two state members appointed by the National Association of State Foresters and three federal members from the NFS, S&PF, and Research and Development divisions of the Forest Service. The Steering Committee sets broad strategic goals and directions to be implemented by the FHM National Program Manager. The National Program Manager is responsible for the overall management of the program budget and implementation of FHM.

The second level of management is provided by the FHM Management Team which includes 15 rotating state and federal members with operational responsibilities in implementing various aspects of the program in different regions. The members provide a variety of expertise including data collection, research, and forest management. The FHM Management Team works closely with the National Program Manager to implement all aspects of the FHM program nationwide. The third level of management consists of *ad hoc* groups organized to address specialized needs—the design of a rapid-response field survey, for example, or the development of a new measurement protocol. The *ad hoc* groups typically include disciplinary specialists and are closely coordinated with data collection specialists from FIA, FHP, and state agencies.

30.3 CONCEPTUAL APPROACHES TO FHM

Forests are continually exposed to a changing array of natural and anthropogenic stresses, producing both normal and abnormal changes in forest health over time. The response to a given stress varies among biophysical regions and according to local circumstances with a region. Stresses also interact with each other and change over time, and forest responses to stresses can occur at multiple scales and may be delayed rather than immediate. These and other factors make it very difficult to establish baselines of forest health and to detect important departures from normal forest ecosystem functioning. The conceptual approach to forest health monitoring must also take into account the fact that many ecological processes are only poorly understood.

The primary objective of monitoring is to identify ecological resources whose condition is deteriorating in subtle ways over large regions in response to cumulative stresses. This objective calls for consistent, large-scale, and long-term monitoring of key indicators of health status, change, and trend. A second objective is to define the extent of resources whose condition is deteriorating rapidly or is at risk of rapid deterioration, from specific stresses, and to develop mitigation and management strategies for those events. This objective calls for more focused surveys and monitoring.

To address both objectives, the FHM program adopted a tiered strategy based on the detection of unusual conditions on a regional scale, followed by progressively more detailed studies to explore the causes and consequences of the observed changes. In the *detection tier* of monitoring, the forests are systematically sampled in space and time, and a small set of integrative health indicators is used to classify the status of forest resources and to gauge the stresses placed on those resources. Repetition of the indicator measurements provides a basis for periodic reporting of the health status and trends of forests and establishes a baseline for future comparisons. Like a human health survey, the forest health survey provides statistically credible information about status and trends. It can suggest plausible mechanisms for observed changes, but by itself cannot resolve many important questions such as the causes of change or the ecological and social significance of change.

The routine, long-term, and large-scale monitoring of selected indicators is supplemented by an *evaluation tier* that provides for intensive surveys and research when warranted by observations. The details naturally depend entirely on circumstances and therefore the evaluation component is not fully defined in advance. Included in the evaluation tier are focused surveys to address the second objective of FHM.

Sometimes a particular question about forest health must be answered in a very short period of time. An early example was the concern for air pollution impacts on forests, which eventually led to the inclusion of several field measurements of tree crown condition, lichen abundance, and ozone injury in the long-term monitoring design. A recent example that will be described in more detail later is the concern regarding the spread of sudden oak death, first observed in 1995 in the San Francisco Bay region. The potential impacts of some phenomena are so large that it makes sense to immediately conduct an evaluation of them, and not wait for signs and symptoms to be manifested through the detection tier of the monitoring system.

An indicator-based approach to detection monitoring employs a multidimensional suite of indicators to monitor several aspects of forest health. Many measurements are needed to comprehensively characterize forest ecosystem structure, function, and process, but only a few can be realistically employed in a long-term national program. Ideally, a small set of indicators addresses many dimensions of forest condition such as sustainability, productivity, aesthetics, contamination, utilization, diversity, and extent. If only a few aspects of forest condition are monitored, important changes could be overlooked. Another way to miss changes is to focus attention only on diagnosing known cause–effect relationships because that requires highly specific measures of conditions that are more appropriate for evaluating known problems than for detecting unknown health problems. The emphasis on detecting without necessarily explaining regional changes in health leads to the emphasis on integrative indicators of forest health. Detection monitoring accepts a high rate of false positives (i.e., a high Type I error rate) as the price of not overlooking change (i.e., a low Type II error rate) and resolves the false positive errors through closer evaluations of the observations.

Whether or not a change will be detected depends partly on the scale of the indicator measurements and sample design relative to the scale of the change phenomena of interest.[44] The time and space scales of surveys should be linked[45] so that detection monitoring utilizes annual or longer measurement cycles and the measurements are sparsely distributed over very large areas. Knowledge of finer scale temporal or spatial variability typically contributes little information about long-term and large-scale changes.[46] For example, the long-term and regional impacts of climate change, air pollution, and urbanization are best monitored on a long-term and regional basis because model-based extrapolation from a few intensively studied research sites cannot reliably detect regional changes, and small-scale intensive surveys can say nothing about areas not included in the surveys. At the same time, research sites and intensive surveys are key elements of the evaluation tier of monitoring because they provide detailed information that cannot be provided by the detection tier of the system.

In summary, the FHM strategy has a component to detect long-term regional changes, a component to assess the practical importance of observed changes and to develop options for mitigation and management, and a component to implement intensive surveys and research to rapidly deliver information about particular changes and concerns. Detection monitoring is largely statistical and relies on integrative indicators of condition that are expected to yield a high rate of false positives. Evaluation monitoring is designed to clarify the information, to increase the "signal-to-noise" ratio, and to focus attention on important health problems. The research tier is reserved for conditions that are known to affect large regions in a practical, important way when detailed information is needed about the causes and consequences of poor health and when options for prevention and mitigation are required.

30.4 OPERATION OF FHM

This section will describe the data collection and processing for the detection tier of forest health monitoring. These procedures are more or less fixed and are conducted in a consistent fashion nationwide. In other tiers of monitoring, data that are collected for evaluation monitoring and research purposes are typically decided according to the specific project requirements. Detection monitoring includes field plot measurements, aerial surveys, and assessments of the data. Ancillary data obtained from supplemental sources are used to interpret the FHM measurements and to estimate some indicators. For example, tree crown condition information provided by FIA is interpreted in light of regional weather patterns as reported by the National Oceanic and Atmospheric Administration. Data from the U.S. Geological Survey are used to measure forest fragmentation, and the U.S. Environmental Protection Agency provides data to estimate air pollution exposure.

The field plots that are measured by the FIA program are located according to a systematic national grid. A systematic grid is appropriate for sampling extensively distributed resources like forests and makes it easier to aggregate the resulting data according to states, ecological regions, or some other geographic partitioning required for particular reporting purposes. The basic design[34,47] identifies one Phase 2 (FIA timber inventory) sample plot location for every 6,000 acres with a total of about 125,000 possible plot locations in the lower 48 states. The Phase 3 forest health measurements are made on a 1/16 subset of the plots (~8,000 plots nationwide). Depending on a schedule for each state, each plot is measured once every 5 to 10 years through a rotating panel design (Figure 30.1).

Table 30.1 describes the measurements that are made in and around Phase 3 field plots by state and federal field crews.[34] The measurements are supported by national training and quality assurance programs to ensure the quality and consistency of the data. The field plot design (Figure 30.2) includes a cluster of four 0.042-acre

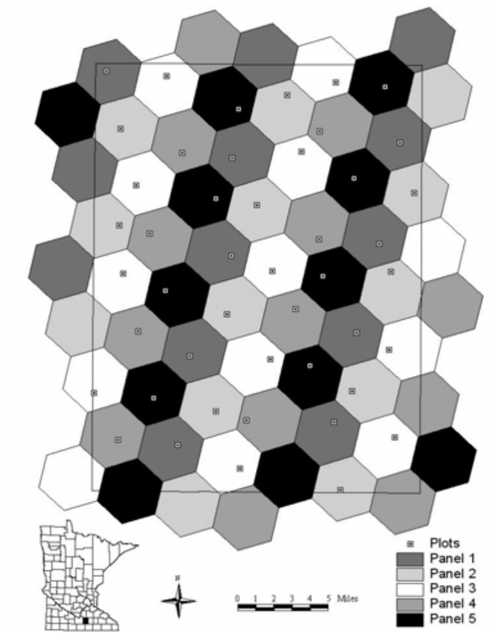


FIGURE 30.1 The FIA sample design is based on a tiling of hexagons. A timber inventory plot is located within each hexagon, and a forest health plot is located within one of every sixteen hexagons. Plot measurements are scheduled according to a rotating panel design as indicated by the shading of hexagons. (From U.S. Forest Service, Sampling and Plot Design Fact Sheet, Forest Inventory and Analysis, Washington, D.C., 2003.)

circular subplots with subplot centers located 120 ft apart. All large trees are measured within each of the subplots. Each subplot contains smaller fixed-area plots and line transects that are used for sampling smaller vegetation and woody material on the forest floor. Soils, lichens, and tree crowns are measured in the area between the subplots. Additional measurements of ozone injury are made at sample locations near the plot; the specific site and species required for such measurements may not occur in the plot.

Aerial and ground-based surveys are conducted by federal and state forest health specialists. Each state and Forest Service administrative region is responsible for conducting an annual survey of forested lands within their jurisdiction. In most states, the data are collected by flying over the prescribed area in a systematic fashion, drawing polygons on a map to show the locations of affected areas, and making notes of the observed signs and symptoms. Maps are digitized into geographic information systems and the data are forwarded to a national processing center for compilation and reporting (Figure 30.3). An increasing amount of sampling uses an automated sketch mapping system that allows observers to digitize polygons directly into a computer linked with the aircraft global positioning system. Like the plot measurements, the aerial survey data are supported by national training and assurance programs to ensure the quality and consistency of the data.

TABLE 30.1
Description of Forest Health Measurements on FIA Field Plots

Indicator	Measurements Included
Crown condition	Amount, condition, and distribution of foliage, branches, and growing tips of trees (crown ratio, crown density, foliar transparency, dieback, and crown width)
Tree damage	Type, location, and severity of injury caused by diseases, insects, storms, and human activities
Tree mortality	Type, location, and severity of injury caused by diseases, insects, storms, and human activities
Vegetation diversity and structure	Type, abundance, and vertical position of vascular plant species (includes an inventory of small trees, herbs, grasses, vines, ferns, and fern allies)
Down woody material	Species, size, and stage of decay of fallen trees, dead branches, and large fragments of wood on the forest floor
Ozone injury	Symptoms, species, and severity of foliar injury on ozone bioindicator species
Lichen communities	Lichen species abundance, diversity, and community composition
Soil condition	Physical and chemical soil properties of the litter, O-horizon, and mineral soil, including erosion and soil compaction

Field plot measurements are reported with timber inventory statistics by the FIA program according to the schedules established in each FIA region. Similarly, the survey data are analyzed and reported by the FHP program for states and Forest Service administrative regions. Additional analysis and reporting conducted by the FHM program is focused on two topics. First, interpretive reports at state, regional, and national levels augment the routine reporting of status and trends statistics by FIA and FHP. Second, the FHM program compiles statistics from FIA, FHP, and other sources to produce annual forest health summaries at the national level. FHM is the only entity whose entire function is to integrate forest health information from many data collection agencies to produce reports of forest health.

30.5 DEVELOPMENT EFFORTS IN THE FHM PROGRAM

The development of the FHM program is currently focused on five major themes:

- Completing the implementation of plot measurements in all states through the FIA program
- Completing the integration of information and reporting systems through the FHP survey program
- Evaluating a possible expansion of FHM to urban forests, riparian forests, and other locations that are not included in FIA or FHP sample designs

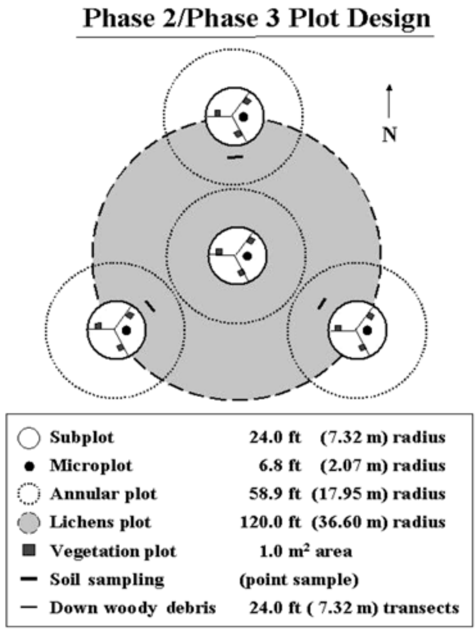


FIGURE 30.2 Field plot layout for FIA forest health measurements. (From U.S. Forest Service, Sampling and Plot Design Fact Sheet, Forest Inventory and Analysis, Washington, D.C., 2003.)

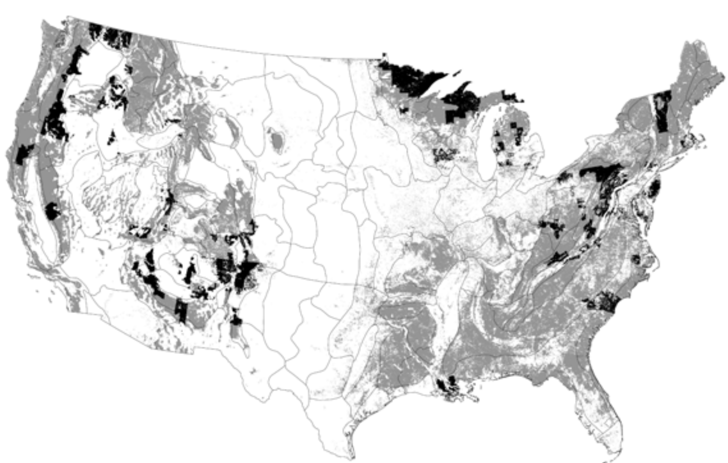


FIGURE 30.3 Example of the national integration of aerial survey data. The shades of gray indicate relative exposure of forests to defoliating agents (insects, diseases, etc.) for the years 1996–2000. Ecoregion boundaries are shown for comparison. (From Coulston, J.W. et al., 2002 Forest Health Monitoring National Technical Report, General Technical Report, U.S. Forest Service, Southern Research Station, Asheville, NC, in press.)

- Evaluating additional forest health indicators for possible deployment and improving the efficiency of current indicators
- Developing ways to use the field plot and aerial survey data along with supplemental data to produce state, regional, and national assessments of forest health

As of 2003, the plot network is operational in 47 of the 50 states, and full implementation is expected by 2005. The FIA program is nearing completion of a transition to a common national system employing the same plot design, measurement protocols, and reporting standards. The survey component of FHP (including all states and territories) was linked with FHM in 1998, and integration of the resulting maps is now part of the normal reporting within FHP. Initiated at the same time, the evaluation tier of FHM now involves annual selection of follow-up studies in all regions of the nation.

In response to growing information needs, FHM is evaluating an extension or intensification of the plot network and surveys in two special-interest populations. Recent concerns for the condition of the forest–urban interface surfaced mainly because of the catastrophic fire seasons from 1999 to 2002 when many homes were lost. At the same time, municipalities are placing increasing importance on reserves of forestland within their boundaries and are attempting to manage them in sustainable ways. Since 2000, FHM has sponsored prototype tests of urban forest monitoring in five states and is approaching a decision on operational deployment. There has also been a preliminary investigation of intensification of monitoring in riparian forests, the second special interest population. This effort is motivated by the requirements of the NFS for integrated watershed-based monitoring of forest and water resources.

Research continues to develop new indicators of forest health that directly address the Montréal Process framework for sustainable forest management. The initial focus of FHM was on air pollution, insects, and diseases, all of which are part of the framework, and the integration with timber inventory also made it possible to address indicators of forest extent and protected status. Research and prototype tests are underway for other criteria and indicators. Included are assessments of forest fragmentation as part of the biodiversity criterion and of invasive species and fire risk as part of the forest disturbance indicator. One research theme is to develop remotely sensed indicators of forest spatial patterns and fragmentation, and models for interpreting them. Other research is developing field procedures to survey and assess invasive species, and analytical procedures to estimate fuel loads and carbon sequestration from plot data.

30.6 FHM REPORTS

The data flowing through the FHM program are now sufficient to produce meaningful statistical and interpretive reports for states, regions, and the nation. FHM cooperators including FIA and FHP are generally responsible for state and regional reporting. In addition, the FHM program produces national level reports supporting the periodic Forest Service Strategic Planning and Resource Assessments, the Montréal Process, and a variety of other interagency and nongovernmental assessments.[42,43] We can give only a few examples of the many reports that have been developed using FHM data.

The first two examples illustrate the operation of the detection and evaluation tiers of FHM. The 2001 FHM National Technical Report[42] used measurements from the plot network to identify apparently increased levels of dieback and mortality in some forests in Indiana in comparison to national averages. In a follow-up evaluation study, Stephen Krecik and Philip Marshall of the Indiana Department of Natural Resources analyzed the Indiana FHM data geographically, by species group and by forest type.[48] The changes in dieback were found to be inconsistent among forest types, and the increased mortality in some forest types was a false positive attributed to statistical estimation procedures and a small sample size. The index was high because of the mortality of a small number of relatively large trees, probably caused by known problems such as Dutch elm disease (fungus *Ophiostoma ulmi*).

The 2001 national FHM report also identified high levels of crown dieback on softwood tree species in northwest Wisconsin. Sally Dahir and Jane Cummings Carlson of the Wisconsin Department of Natural Resources combined historical information from aerial surveys and FIA timber inventories to evaluate the association among the current softwood dieback, the distribution of tree species, and six previous years of jack pine defoliation from the jack pine budworm (*Choristoneura pinus*).[49] The study concluded that the dieback observed on FHM plots was consistent with historical defoliation patterns and identified specific site and stand conditions to predict where future budworm outbreaks are most likely to occur. These two examples show that the first tier of monitoring is capable of detecting a strong signal of apparently unusual forest conditions, but also that baselines are not well established from only a few years of monitoring and that follow-up evaluations are critical to resolve whether the apparently unusual conditions are of concern or not. The studies also demonstrate the value of combining FHM data with supplemental data to interpret the information on a regional basis.

FHM data are also combined with supplemental data to address national reporting requirements. For example, the 2003 National Report on Sustainable Forests[50] required an assessment of the area of different forest types that are exposed to different levels of sulfate and nitrate deposition. The results are critical in assessing if air pollution is adversely affecting forests over large regions. FHM does not monitor air pollution exposure because that is done by the U.S. EPA. The FHM program prepared wet deposition maps from the EPA data and then combined them with a national forest type map. The map-based approach enabled the tabulation of the area of each forest type that was exposed to different levels of wet deposition, as required for national reporting purposes.[51]

The next example demonstrates the ecological interpretation of FHM data with respect to a regional forest health issue. Over much of the Rocky Mountain region, there is a concern for the perpetuation of the aspen (*Populus tremuloides*) forest type.[52] A team of researchers in the Forest Service Interior West region used FHM and FIA plot data to document a regional pattern of aspen stand maturation and subsequent loss by natural succession. In-depth investigations suggested that the absence of stand disturbance and fire over several prior decades had resulted in an unbalanced age–class distribution of aspen stands, and that the total area of the aspen type was likely to be dramatically reduced in the future unless mitigation efforts such as prescribed burning were initiated.[53] Over the past several years, an unusually large area of the Rocky Mountain forest has been burned by uncontrolled wildfires, and attention has now turned towards using the plot network for monitoring the possible recovery of the aspen forest.

We will close with an example of risk-based assessments, showing how FHM expertise and resources are helping to resolve critical issues in a timely and effective fashion. Sudden oak death (SOD) is one of the most important forest health issues today.[54] It is the common name of a disease caused by *Phytophtora ramorum*, a fungal pathogen closely related to *P. lateralis*, the cause of Port Orford cedar root rot.[55] The origin of SOD is not known with certainty, but it also occurs in Europe.[55] SOD has been identified as the cause of death for unusually large numbers of tanoak (*Lithocarpus densiflorus*) and oak trees (*Quercus* spp. including coastal live oak and black oak) since 1995 in near-coastal, nonurban locations from southern Oregon to Big Sur, California.[56] SOD has been characterized by rapid decline and development of bleeding or oozing cankers on the lower trunk of diseased trees, with mortality caused by phloem death and girdling of the tree.[54] In laboratory studies, saplings may be killed within weeks and the field survival of infected mature trees may be only a few years. SOD is a major health issue because the known or likely hosts include many commercially, aesthetically, and ecologically important tree and shrub species that are dominant or widespread throughout most of the nation.[54] The fungus is potentially transportable through ornamental nursery stock, creating the possibility of relatively rapid dissemination and infection of susceptible trees near horticultural centers. There is no known cure, and quarantine is the only effective control.

While quarantines have already been established, there is an immediate need to conduct a national risk assessment of SOD. FHM is conducting a national field survey to determine the extent of SOD and is preparing national risk maps to pinpoint the most likely locations where SOD might appear. FHM is working in cooperation with a large number of state, federal, and international agencies in this effort. In 2002, FHM developed a national risk map based on the historical occurrences of *P. ramorum*. National climate maps were combined with maps of the distribution of known and probable hosts and maps of nurseries to identify high- and low-risk regions for the field survey (Figure 30.4). Extensive aerial survey protocols were established for the low-risk regions, and intensive field survey protocols were established for high-risk regions. These surveys were designed to complement other surveys of horticultural nurseries and urban environments. As of the date of this writing, only preliminary results are available from the surveys, and so far the news is good. While new occurrences of SOD have been identified within the high-risk region already affected in Oregon and California, the presence of SOD has not been confirmed outside of those areas. This risk-based approach to detection of new infestations of invasive species will be applied to other exotic pests, including the emerald ash borer (*Agrilus planipennis*) which has recently emerged as a new threat to the ash forests (*Fraxinus* spp.) of North America.

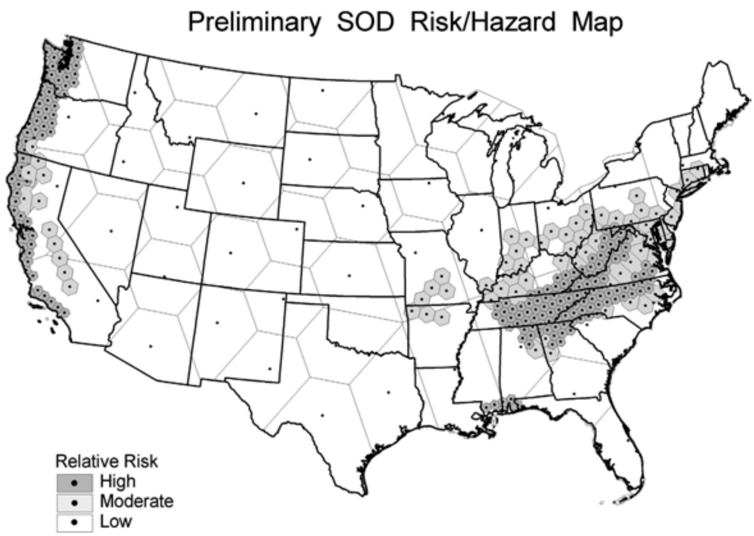


FIGURE 30.4 As part of a nationwide survey for the presence of sudden oak death, relative risk is estimated from climate, host species geographic range, and other risk factors. The darkest hexagons have the highest risk and a more intensive sample during the field survey. (From Smith, W., personal communication, 2003.)

30.7 CONCLUSION

As it matures, the Forest Health Monitoring program will become an increasingly important part of the overall environmental assessment effort in the U.S. After only a decade, the FHM program is producing meaningful reports for a variety of clients and purposes, and efficiencies have been gained by transferring technology to and gaining cooperation from other entities such as FIA and FHP. Now that the majority of data collection efforts are conducted by other entities, FHM is refocusing its resources on key forest health issues, national forest health assessments, and research that will identify additional protocols for possible future deployment in the field.

More information on the components and partners of FHM are available through the National Association of State Foresters, individual state forestry and natural resource agencies, and the U.S. Forest Service. The following sites on the World Wide Web (last accessed in May 2003) provide additional information and links to all the projects and programs described in this chapter:

- Forest Health Monitoring — http://www.na.fs.fed.us/spfo/fhm/
- National Association of State Foresters — http://www.stateforesters.org/
- Forest Health Protection — http://www.fs.fed.us/foresthealth/
- Forest Inventory and Analysis — http://www.fia.fs.fed.us/
- Forest Research and Development — http://www.fs.fed.us/research/
- Sustainable Resource Management — http://www.fs.fed.us/sustained/

REFERENCES

1. Frayer, W.E. and Furnival, G.A., Forest survey sampling designs: a history, *J. For.*, 97, 4, 1999.
2. Van Deusen, P.C., Prisley, S.P., and Lucier, A.A., Adopting an annual inventory system: user perspectives, *J. For.*, 97, 11, 1999.
3. Gillespie, A.J.R., Rationale for a national annual forest inventory program, *J. For.*, 97, 16, 1999.
4. Reams, G.A., Roesch, F.A., and Cost, N.D., Annual forest inventory: cornerstone of sustainability in the South, *J. For.*, 97, 21, 1999.
5. National Acid Precipitation Assessment Program (NAPAP), The U.S. National Acid Precipitation Assessment Program 1990 Integrated Assessment Report, NAPAP Office of the Director, Washington, D.C., 1991.
6. Messer, J.J., Linthurst, R.A., and Overton, W.S., An EPA program for monitoring ecological status and trends, *Environ. Monit. Assess.*, 17, 67, 1991.
7. Palmer, C. et al., Monitoring and Research Strategy for Forests: Environmental Monitoring and Assessment Program (EMAP), EPA/600/4-91/012, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., 1992.
8. Burkman, W.G. and Hertel, G.D., Forest health monitoring, *J. For.*, 90, 26, 1992.
9. Hazard, J.W. and Law, B.E., Forest Survey Methods Used in the USDA Forest Service, EPA/600/3-89/065, U.S. Environmental Protection Agency, Environmental Research Laboratory, Corvallis, OR, 1989.
10. Riitters, K.H. et al., A selection of forest condition indicators for monitoring, *Environ. Monit. Assess.*, 20, 21, 1992.
11. Schreuder, H.T. and Czaplewski, R.L., Long-term strategy for the statistical design of a forest health monitoring system, *Environ. Monit. Assess.*, 27, 81, 1993.
12. Scott, C.T., Cassell, D.L., and Hazard, J.W., Sampling design of the U.S. national forest health monitoring program, in *Proceedings of Ilvessalo Symposium on National Forest Inventories*, Nyyssonen, A., Poso, S., and Rautala, J., Eds., Finnish Forest Research Institute, Helsinki, Finland, 1993, 150.
13. Cooter, E.J. et al., Role of Climate in Forest Monitoring and Assessment: a New England Example, EPA/600/3-91/074, U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC, 1991.
14. Riitters, K.H. et al., Eds., Forest Health Monitoring Plot Design and Logistics Study, EPA/600/S3-91/051, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., 1991.
15. Alexander, S.A., Carlson, J.A., and Barnard J.E., The visual damage survey: a study to evaluate the eastern forest condition, in *Ecological Indicators*, Vol. 1, McKenzie, D.H., Hyatt, D.E., and McDonald, V.J., Eds., Elsevier Applied Science, New York, 1992, 361.
16. Baker, F.A. et al., Evaluation of the root disease indicator used in the forest health monitoring program, *Phytopathology*, 82, 1152, 1992.
17. Papp, M.L. et al., FY91 Forest Health Monitoring Western Pilot Operations Report, EPA/600/X-92/009, U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV, 1992.
18. Alexander, S.A. et al., Forest Health Monitoring: 1991 Georgia Indicator Evaluation and Field Study, EPA/620/R-94/007, U.S. Environmental Protection Agency, Environmental Research Laboratory, Corvallis, OR, 1993.

19. Lewis, T.E. and Conkling, B.L., Eds., Forest Health Monitoring: Southeast Loblolly/Shortleaf Pine Demonstration Interim Report, EPA/620/R-94/006, U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC, 1994.
20. Williams, M.S., Bechtold, W.A., and LaBau, V.J., Five instruments for measuring tree height: an evaluation, *Soc. J. Appl. For.*, 18, 76, 1994.
21. Stapanian, M.A., Cline, S.P., and Cassell, D.L., Evaluation of a measurement method for forest vegetation in a large-scale ecological survey, *Environ. Monit. Assess.*, 45, 237, 1997.
22. Tallent-Halsell, N.G., Ed., Forest Health Monitoring 1994 Field Methods Guide, EPA/620/R-94/027, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., 1994.
23. Palmer, C.J., The 1992 Quality Assurance Annual Report and Workplan for the Interagency Forest Health Monitoring Program, TIP # 92-295, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1992.
24. Pollard, J.E., Smith, W., and Palmer, C.E., Forest Health Monitoring 1999 Plot Component Quality Assurance Implementation Plan, U.S. Forest Service, National Forest Health Monitoring Program, Research Triangle Park, NC, 1999.
25. Liff, C.I., Riitters, K.H., and Hermann, K.A., Forest health monitoring case study, in *Environmental Information Management and Analysis: Ecosystem to Global Scales*, Michener, W.K., Brunt, J.W., and Stafford, S.G., Eds., Taylor and Francis, London, 1994, 101.
26. Brooks, R.T., Miller-Weeks, M., and Burkman, W.G., Summary Report Forest Health Monitoring New England 1990 (Northeastern Area Association of State Foresters), NE-INF-94-91, U.S. Forest Service, Radnor, PA, 1991.
27. Hyland, J., Forest health monitoring in Alabama, *Treasured Forests*, Fall, 1991.
28. Bechtold, W.A., Hoffard, W.H., and Anderson, R.L., Summary Report: Forest Health Monitoring in the South, 1991, General Technical Report SE-81, U.S. Forest Service, Southeastern Forest Experiment Station, Asheville, NC, 1992.
29. Brooks, R.T. et al., The New England Forest: Baseline for New England Forest Health Monitoring, Resource Bulletin NE-124, U.S. Forest Service, Northeastern Forest Experiment Station, Radnor, PA, 1992.
30. Georgia Forestry Commission, Georgia Forestry Commission Forest Health Report 1991–1993, Georgia Forestry Commission, Dry Branch, GA, 1993.
31. Gillespie, A.J.R. et al., Summary Report, Forest Health Monitoring, New England/mid-Atlantic, NE/NA-INF-115-R93, U.S. Forest Service, Radnor, PA, 1993.
32. Blunt, W.H. and Schomaker, M., Colorado Forest Health Report 1993, U.S. Forest Service, Lakewood, CO, 1994.
33. AFPA (American Forest and Paper Association), Forest Inventory and Analysis Program: The Report of the Second Blue Ribbon Panel, American Forest and Paper Association, Washington, D.C., 1998.
34. Stolte, K. et al., Forest Health Indicators: Forest Inventory and Analysis Program, FS-746, U.S. Forest Service, Washington, D.C., 2002, URL: http://fia.fs.fed.us/library/ForestHealthIndicators.pdf (accessed May 2003).
35. U.S. Forest Service, America's Forests: 1999 Health Update, U.S. Forest Service, Forest Health Protection, Washington, D.C., 1999, URL: http://www.fs.fed.us/foresthealth/fh_update/update99/index.html (accessed May 2003).
36. Mitchell, R. and Buffam, P., Patterns of long-term balsam woolly adelgid infestations and effects in Oregon and Washington, *West. J. Appl. For.*, 16, 121, 2001.

37. Kanaskie, A. et al., Ground Verification of Aerial Survey for Port Orford Cedar Root Disease in Southwest Oregon, U.S. Forest Service, Pacific Northwest Research Station, Portland, OR, 2002.
38. Anonymous, Sustaining the worldís forests: the Santiago declaration, *J. For.*, 93, 18, 1995.
39. Anonymous, First Approximation Report on the Montreal Process, The Montreal Process Liaison Office, Natural Resources Canada, Canadian Forest Service, Ottawa, Canada, 1997.
40. U.S. Forest Service, USDA Forest Service Strategic Plan (2000 revision), FS-682, U.S. Forest Service, Washington, D.C., 2000.
41. U.S. Forest Service, 2000 RPA Assessment of Forest and Range Lands, FS-687, U.S. Forest Service, Washington, D.C., 2001.
42. Conkling, B., Coulston, J., and Ambrose, M., Eds., 2001 Forest Health Monitoring National Technical Report, General Technical Report, U.S. Forest Service, Southern Research Station, Asheville, NC, in press.
43. Coulston, J.W. et al., 2002 Forest Health Monitoring National Technical Report, General Technical Report, U.S. Forest Service, Southern Research Station, Asheville, NC, in press.
44. Allen, T.F.H., OíNeill, R.V., and Hoekstra, T.W., Interlevel relations in ecological research and management: some working principles from hierarchy theory, *J. Appl. Syst. Anal.*, 14, 63, 1987.
45. Meentemeyer, V., Geographical perspectives of space, time, and scale, *Land. Ecol.*, 3, 163, 1989.
46. O'Neill, R.V. et al., *A Hierarchical Concept of Ecosystems*, Princeton University Press, Princeton, NJ, 1986.
47. U.S. Forest Service, Sampling and Plot Design Fact Sheet, U.S. Forest Service, Forest Inventory and Analysis, Washington, D.C., 2003.
48. Krecik, S.G., Marshall, P.T., and Smith, W.D., Indiana Hardwood Dieback and Mortality: Evaluation of the FHM National Technical Report 1996–1999, Poster presented at the FHM annual meeting, Monterey, CA, 2003, URL: http://www.na.fs.fed.us/spfo/fhm/posters/(accessed May 2003).
49. Dahir, S. and Cummings, C.J., Softwood dieback and jack pine budworm defoliation in Wisconsin, Poster presented at the FHM annual meeting, Monterey, CA, 2003, URL: http://www.na.fs.fed.us/spfo/fhm/posters/(accessed May 2003).
50. U.S. Forest Service, National Report on Sustainable Forests—2003, FS-766, U.S. Forest Service, Washington, D.C., in press.
51. Coulston, J.W., Riitters, K.H., and Smith, G.C., A preliminary assessment of Montréal process indicators of air pollution for the United States, *Environ. Monit. Assess.*, in press.
52. Kay, C.E., Is aspen doomed?, *J. For.*, 95, 4, 1997.
53. Rogers, P., Using forest health monitoring to assess aspen forest cover change in the southern Rockies ecoregion, *For. Ecol. Manage.*, 155, 223, 2002.
54. Rizzo, D.M. and Garbelotto, M., Sudden oak death: endangering California and Oregon forest ecosystems, *Front. Ecol. Environ.*, 1, 197, 2003.
55. Werres, S. et al., *Pytophthora ramorum* sp nov: a new pathogen on *Rhododendron* and *Viburnum*, *Mycol. Res.*, 105, 1155, 2001.
56. Rizzo, D.M. et al., *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp and *Lithocarpus densiflorus* in California, *Plant. Dis.*, 86, 205, 2002.

31 Clean Air Status and Trends Network (CASTNet)—Air-Quality Assessment and Accountability

M. Kolian and R. Haeuber

CONTENTS

1-56670-641-6/04/$0.00+$1.50

31.1 INTRODUCTION

The 1990 CAAA established the Acid Deposition Control Program (Title IV) and mandated a significant reduction in the emissions of sulfur and nitrogen oxides (NO_x) from electric utilities burning fossil fuels. Title IV is intended to minimize air-quality related public health risks as well as to protect sensitive ecosystems from the adverse affects of acid deposition. Title IX (Clean Air Research) of the 1990 CAAA requires that the environmental effectiveness of the Acid Deposition Control Program be assessed through comprehensive research and air pollutant monitoring. Congress recognized the need to track real-world environmental results through continued acid rain research and monitoring as emission reductions were implemented. In response, the U.S. Environmental Protection Agency (EPA), in coordination with the National Oceanic and Atmospheric Administration (NOAA), established the Clean Air Status and Trends Network (CASTNet) with the goal of assessing the impact and effectiveness of Title IV through a large-scale air-quality network. Monitoring programs such as CASTNet represent a firm commitment to long-term monitoring which is critical in documenting air quality and deposition trends for determining regulatory accountability.

The developmental framework of CASTNet can be traced to the National Dry Deposition Network (NDDN) which began in 1986 with the goal of providing long-term estimates of dry deposition for the U.S. Since dry deposition was recognized as a principal component of total acid deposition (i.e., the sum of wet and dry deposition), NDDN was subsumed entirely into CASTNet. NDDN operated approximately 50 sites that became the core CASTNet sites when NDDN was incorporated into CASTNet in 1991. Although CASTNet technically came into existence in mid-1991, the CASTNet data record extends to 1987 when field measurements first began under NDDN.

Instrumental in the development of CASTNet was the knowledge that there was an inherent need for a better understanding of the science regarding the dry deposition component of total (wet + dry) acid deposition. Dry deposition is the transfer of particles and gases to the landscape through a number of atmospheric processes in the absence of precipitation. Although wet deposition rates of acidic species across the U.S. have been well documented over the last 20 years, comparable information had been unavailable for dry deposition rates. This lack of information on dry deposition increases the uncertainty in estimates of interregional, national, and international transport and confounds efforts to determine the overall impact of atmospheric deposition

(USEPA 1998a). The creation of CASTNet, however, would allow dry deposition rates to be used in conjunction with the wet deposition monitoring measurements of the National Atmospheric Deposition Program/National Trends Network (NADP/NTN) to accurately determine total acid deposition. This can be accomplished by locating dry deposition monitoring sites at or nearby sites measuring wet deposition.

The development of CASTNet was also influenced by the idea of supplementing the monitoring science and common interests of the Interagency Monitoring of Protected Visual Environments (IMPROVE) program. One of the goals of the IMPROVE network which began establishing rural long-term monitoring sites in 1987 was to measure the composition of visibility-reducing aerosols to help identify the source type and strength of fine particles and gaseous precursors to secondary particles (CENR 1999). Using similar instrumentation and monitoring protocols, CASTNet provided the opportunity to enhance the visibility measurements of the IMPROVE network by implementing and measuring ambient concentrations of gaseous-phase aerosols. Like the IMPROVE network, CASTNet monitoring stations are located in rural areas and collect data to establish site-specific measurements in the absence of local influences from area or local sources.

Since EPA initiated CASTNet, it has evolved into a robust, national, long-term monitoring program that measures changes in air quality and atmospheric deposition over broad geographic regions of the U.S. CASTNet currently represents the nation's primary source of atmospheric data on the dry deposition component of total acid deposition (wet + dry), continuous rural ground-level ozone (O_3), and meteorological variables. The network is continuing to expand and consists of 87 monitoring stations (as of January 2003). Table 31.1 provides a brief summary of the network.

TABLE 31.1
Network Summary

Air-quality network	CASTNet
Initiated	1986[a]
Number of sites (as of January 2003)	87
Lead federal agency	EPA
Sampling schedule	7-d (168 h)
Ambient measurements	Gaseous: SO_2 (sulfur dioxide), HNO_3 (nitric acid), ozone (O_3) Particulate: SO_4^{2-} (sulfate), NO_3^- (nitrate), NH_4^+ (ammonium), Ca^{2+} (calcium), Na^+ (sodium), Mg^{2+} (magnesium), and K^+ (potassium)
Meteorological measurements	Temperature at 2- and 9-m, solar radiation, relative humidity, precipitation, scalar wind speed, vector wind speed, wind direction, standard deviation of the wind direction within the hour, rate of flow through the filter pack, wetness
Information on land use and vegetation	Site survey and observations by site operator: vegetation type, leaf area index (LAI), and percent green leaf out

[a]The National Dry Deposition Network (NDDN) was established in 1986 and field measurements began in 1987. With the passage of the Clean Air Act Amendments of 1990, NDDN was entirely subsumed by CASTNet in mid-1991.

The fundamental objectives for CASTNet are to:

- Monitor the status and trends in regional air quality and atmospheric deposition
- Provide information on the dry deposition component of total acid deposition, ground-level ozone (O_3), and other forms of atmospheric gaseous and aerosol pollution
- Assess and report on geographic patterns and long-term, temporal trends in ambient air pollution and acid deposition

31.2 DESIGN RATIONALE

As a long-term monitoring program, CASTNet allows for characterizing trends in deposition levels and identifying relationships among emissions, atmospheric loadings, human health, and ecological effects. Atmospheric changes occur very slowly and trends are often obscured by the wide variability of measurements and climate; therefore, numerous years of continuous and consistent data are required to overcome this variability. To determine whether emission reductions are having their intended effect on atmospheric concentrations, it is critical for network sites to remain operational and provide a continuous or uninterrupted data record. This establishes a site-specific base of information helpful in the process of determining the status and trends in atmospheric chemistry over time. Furthermore, the network relies on the importance of stability and standardization of sites and protocols in order to achieve quality monitoring data. All CASTNet sites are located and installed according to strict siting criteria, with a selection process designed to avoid undue influences from point sources, area sources, and local activities (e.g., agriculture).

CASTNet is founded on the ability to measure seasonal and annual average concentrations and depositions over many years. The principal sampling design of CASTNet involves the measurement of rural representative concentrations of atmospheric sulfur and nitrogen species using single-filter pack weekly samples at each site in order to estimate dry deposition fluxes, detect and quantify trends, and define the spatial distribution of pollutants across the network. CASTNet monitoring stations also include continuous analyzers for the measurement of hourly average ozone (O_3) concentrations. In addition, the goal of estimating dry deposition requires the measurement of meteorological parameters together with supporting information on vegetation and land use.

Since dry deposition is difficult to measure directly, inferential models using routine meteorological measurements and vegetation observations have been developed to estimate the deposition velocity for monitored chemical species. The meteorological, vegetation, and land-use data are used as input to the multilayer model (MLM), a micrometeorological and mathematical model that simulates atmospheric dry deposition processes (Meyers et al., 1998 and Finkelstein et al., 2000). The MLM is used to calculate deposition velocities (V_d), which are combined with the concentration measurements to estimate dry deposition rates of gaseous and aerosol constituents. A new more accurate model called the multilayer biogeochemical model (MLBC) has been developed by the EPA's National Exposure Research Laboratory (NERL) and will serve as the next generation model for estimating deposition velocities (Wu et al., 2003). This model builds on the MLM but also accounts for plant photosynthesis and respiration in estimating deposition velocities. Additional information on CASTNet measurement data, current standard operating procedures, and the CASTNet Quality Assurance Project Plan (QAPP)(Harding ESE, 2002a) are available on the CASTNet website located at: http://www.epa.gov/castnet.

The monitoring sites of CASTNet are regionally distributed across the U.S., and many factors are associated with selecting their locations. CASTNet was designed to be a rural monitoring network collecting data to establish site-specific measurements of atmospheric concentrations and deposition. Rural areas typically possess air masses that are more chemically stable and less variable with regard to ambient pollutant concentrations. CASTNet sites are typically not impacted by local sources and are sufficiently isolated which is particularly important in providing a measure of continental ozone (O_3) background. Urban sites are avoided because of "spatial heterogeneity" and high variability of concentrations on a temporal and spatial basis due to many localized sources of pollution. Sites are selected to be representative of regional conditions of that locale in order to provide accurate deposition flux rates for an area within 1 km of the site. Furthermore, the current state of science does not allow accurate interpolations of deposition between sites.

Figure 31.1 presents a map containing CASTNet site locations for 2002. To determine the effects of emission reductions the network established a high concentration

FIGURE 31.1 CASTNet site map as of January 2003.

of sites in the eastern U.S. This is where the majority of Phase I and Phase II emission sources affected by the Acid Rain Program are located. As a result, eastern sites typically have the longest CASTNet monitoring data record. There are fewer CASTNet sites in the western and central U.S.; however, additional sites are planned to expand site coverage in both of these regions. A list of active CASTNet sites with associated site IDs and sponsoring agency is provided in Table 31.2. The CASTNet site selection criteria can be divided into two categories: (1) regional siting criteria and (2) local or site-specific criteria.

Regional siting criteria involve project-wide objectives such as regional representativeness, presence of critical habitat or natural resources, long-term availability, accessibility, and broad geographic distribution to determine meaningful nationwide status and trends information. Regional representativeness refers to the overall similarity of the site to a characteristic area (typically 100 km by 100 km) surrounding the site. Land use near the site should match, as much as possible, the dominant regional land use to make appropriate use of meteorological data measured at each site. This implies that land use, vegetation, and topography of the site must be of the region and that local sources of SO_2 and/or NO_x do not unduly influence the concentration measured at the site. For example, large point sources such as power plants must be a minimum 20 to 40 km from CASTNet measurement apparatus. Site selection may also be determined by where specific research issues can be addressed such as where natural resources or critical habitat are at risk (e.g., national parks, wildlife refuges). Another important siting consideration is the requirement of monitoring sites to be continuously available and accessible for extended periods (10 to 15 years). A long, continuous data record provides a more meaningful assessment of dry deposition trends and changes in air quality over time (USEPA, 1998a).

Site-specific criteria play an important role in the selection of CASTNet sites. These criteria concern local features in the immediate vicinity of a prospective site that may perturb air quality and meteorological observations. Local sources of air contaminants and local features have to be taken into consideration as they may influence wind speed and direction, turbulence, and deposition patterns. In addition to these criteria, it is often advantageous to consider monitoring sites already established under the auspices of another network when identifying candidate monitoring locales. CASTNet site selection must take into account the proximity to wet deposition measurements (<60 km) to estimate total deposition. Similarly, the collaboration with other monitoring networks such as IMPROVE is considered for feasibility. Installing CASTNet monitoring stations at sites of other monitoring networks creates the opportunity for network methods and research comparability. Furthermore, it maximizes the overall knowledge gained from a particular site and ultimately enhances the value of the network.

31.3 PARTNERSHIPS

The CASTNet program encourages implementation of monitoring activities in cooperation with other agencies and organizations. CASTNet operates in partnership with other rural, long-term monitoring networks, such as the NADP/NTN. Together these networks allow for regional assessment of total (wet + dry) acid deposition throughout the U.S. Not only does this strengthen the overall relationship of the environmental

TABLE 31.2
List of Active CASTNet Monitoring Stations

Site ID	Monitoring Station	Agency	State
SND152	Sand Mountain	EPA	AL
CAD150	Caddo Valley	EPA	AR
GTH161	Gothic	EPA	CO
ROM206	Rocky Mountain NP Collocated	EPA	CO
ABT147	Abington	EPA	CT
IRL141	Indian River Lagoon	EPA/SJRWMD	FL
SUM156	Sumatra	EPA	FL
GAS153	Georgia Station	EPA	GA
STK138	Stockton	EPA	IL
ALH157	Alhambra	EPA	IL
BVL130	Bondville	EPA	IL
VIN140	Vincennes	EPA	IN
SAL133	Salamonie Reservoir	EPA	IN
KNZ184	Konza Prairie	EPA	KS
CKT136	Crockett	EPA	KY
MCK131	Mackville	EPA	KY
MCK231	Mackville Collocated	EPA	KY
BWR139	Blackwater NWR	EPA	MD
CDZ171	Cadiz	EPA	KY
BEL116	Beltsville	EPA	MD
ASH135	Ashland	EPA	ME
HOW132	Howland	EPA	ME
UVL124	Unionville	EPA	MI
HOX148	Hoxeyville	EPA	MI
ANA115	Ann Arbor	EPA	MI
CVL151	Coffeeville	EPA	MS
CND125	Candor	EPA	NC
BFT142	Beaufort	EPA	NC
COW137	Coweeta	EPA	NC
PNF126	Cranberry	EPA	NC
WST109	Woodstock	EPA	NH
WSP144	Washington Crossing	EPA	NJ
CTH110	Connecticut Hill	EPA	NY
CAT175	Claryville	EPA	NY
HWF187	Huntington Wildlife Forest	EPA/SUNY	NY
CHE185	Cherokee Nation	EPA/CN	OK
LYK123	Lykens	EPA	OH
OXF122	Oxford	EPA	OH
QAK172	Quaker City	EPA	OH
DCP114	Deer Creek	EPA	OH
EGB181	Egbert	EPA	ON
EGB281	Egbert Collocated[a]	EPA	ON

(*Continued*)

TABLE 31.2
List of Active CASTNet Monitoring Stations (*Continued*)

Site ID	Monitoring Station	Agency	State
KEF112	Kane Exp. Forest	EPA	PA
LRL117	Laurel Hill	EPA	PA
ARE128	Arendtsville	EPA	PA
MKG113	M.K. Goddard	EPA	PA
PSU106	Penn State	EPA	PA
ESP127	Edgar Evins	EPA	TN
SPD111	Speedwell	EPA	TN
PED108	Prince Edward	EPA	VA
VPI120	Horton Station	EPA	VA
LYE145	Lye Brook	EPA	VT
PRK134	Perkinstown	EPA	WI
PAR107	Parsons	EPA	WV
CDR119	Cedar Creek	EPA	WV
PND165	Pinedale	EPA	WY
CNT169	Centennial	EPA	WY
DEN417	Denali NP	NPS	AK
POF425	Poker Flats, Yukon Flats NM	NPS	AK
GRC474	Grand Canyon NP	NPS	AZ
CHA467	Chiricahua NM	NPS	AZ
JOT403	Joshua Tree NM	NPS	CA
SEK402	Sequoia NP — Lookout Pt	NPS	CA
PIN414	Pinnacles NM	NPS	CA
YOS404	Yosemite NP — Turtleback Dome	NPS	CA
LAV410	Lassen Volcanic NP	NPS	CA
DEV412	Death Valley NM	NPS	CA
MEV405	Mesa Verde NP	NPS	CO
ROM406	Rocky Mtn NP	NPS	CO
EVE419	Everglades NP	NPS	FL
HVT424	Hawaii Volcanoes NP	NPS	HI
ACA416	Acadia NP	NPS	ME
VOY413	Voyageurs NP	NPS	MN
GLR468	Glacier NP	NPS	MT
THR422	Theodore Roosevelt NP	NPS	ND
GRB411	Great Basin NP	NPS	NV
GRS420	Great Smoky NP — Look Rock	NPS	TN
BBE401	Big Bend NP	NPS	TX
CAN407	Canyonlands NP	NPS	UT
SHN418	Shenandoah NP — Big Meadows	NPS	VA
VII423	Virgin Islands NP — Lind Pt	NPS	VI
OLY421	Olympic NP	NPS	WA
NCS415	North Cascades NP	NPS	WA
MOR409	Mount Rainier NP	NPS	WA

TABLE 31.2
List of Active CASTNet Monitoring Stations (*Continued*)

Site ID	Monitoring Station	Agency	State
YEL408	Yellowstone NP	NPS	WY
MAC426	Mammoth Cave NP	NPS	KY
PET427	Petrified Forest NP	NPS	AZ

[a]Site operates collocated dual day/night filter-pack sampling equipment.

Agency abbreviations: CN = Cherokee Nation; EPA = U.S. Environmental Protection Agency; FS = U.S. Forest Service; NPS = National Park Service; SUNY = State University of New York; SJRWMD = St. Johns River Water Management District.

monitoring and research community, but network collaboration also helps to fill gaps in monitoring coverage on a national level.

In the CAAA of 1977, Congress specifically addressed the need for increased protection and enhancement of air-quality related values (including visibility) from the "adverse effects of air pollution in national parks, national wilderness areas, and other areas of special value." Consequently, the National Park Service (NPS) has established air-quality monitoring stations as part of its responsibility to air-quality management. NPS has recognized that the protection of air-quality in national parks requires extensive knowledge about the origin, transport, and fate of air pollution, as well as its impacts on resources (DOI, 2002). In 1994, the NPS and the EPA entered into a partnership agreement to operate CASTNet sites. While the EPA is the lead federal agency and administers the CASTNet program, the NPS sponsors and operates approximately 1/3 of the network's sites, many of which are located in western national parks or wilderness areas designated as Class I areas. The NPS and the EPA are responsible for operating their sites, under a common set of quality assurance (QA) standards and similar monitoring and data validation protocols.

The EPA and other agencies depend on the data and information generated from these long-term monitoring networks to assess the effectiveness of several national air pollution control efforts, including the Acid Rain Program (Title IV of the 1990 CAAA), the NO_X State Implementation Plan (SIP) Call, and implementation of the National Ambient Air Quality Standards (NAAQS). In addition, along with other air-quality and deposition monitoring networks, CASTNet will be a critical component in the national accountability framework that will be necessary to assess progress of future air pollution control efforts, such as the various multipollutant emission reduction proposals currently before the 108th Congress.

Utilizing several existing CASTNet documents the remainder of this chapter describes the network in greater detail and provides results of CASTNet measurement data as examples of network output. CASTNet results include concentration trends over a 13-year period (1990 through 2002); 3-year average concentration trends for SO_2 and particulate SO_4^{2-}; annual and quarterly mean concentration data for atmospheric gaseous and particulate sulfur and nitrogen species for 2002; estimates of dry, wet, and total deposition of sulfur and nitrogen species, including trends over the 13-year period; and ozone concentrations based on the fourth highest daily maximum 8-h concentrations for 2002 as well as modeled ozone deposition flux.

31.4 NETWORK DESCRIPTION

The locations of CASTNet sites, as of January 2003, are shown in Figure 31.1. The alphanumeric site codes include three letters and three numbers. The letters provide an approximate description of the site name or location, e.g., IRL (Indian River Lagoon). Numerically, CASTNet sites are designated as 100-series sites for EPA-sponsored sites and 400-series for NPS-sponsored sites. In 2002, there were 87 CASTNet monitoring stations. The network includes 84 site locations with 30 NPS-sponsored sites, 54 EPA-sponsored sites, and three collocated sites.

Each CASTNet dry deposition station measures weekly average concentrations of gaseous sulfur dioxide and nitric acid and particulate sulfate, nitrate, ammonium, and four cations — sodium, potassium, magnesium, and calcium, using a three-stage filter pack (see Figure 31.2). Trained site operators visit their respective CASTNet site each Tuesday, coinciding with a 7-d sampling schedule, and among many other duties, retrieve the filter pack for chemical analysis. In addition to the filter-pack measurements, all CASTNet sites (with the exception of two) possess continuous analyzers to measure hourly average ozone (O_3) concentrations. Also present at each CASTNet site are continuous instruments that measure several meteorological parameters to provide input into the MLM. The measured ambient pollutant concentrations are then used in conjunction with meteorological measurements and surface characteristics to estimate dry deposition.

CASTNet has historically operated and maintained a collocated sampling program for the purpose of estimating the overall precision of dry deposition and supporting data. The program involves duplicate sets of dry deposition sampling instruments installed adjacent to existing instruments. The collocated sampling system

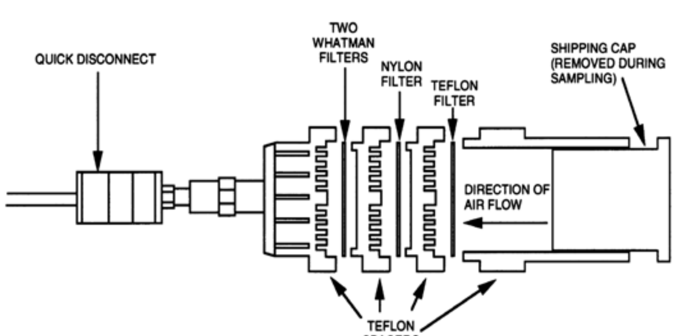


FIGURE 31.2 Diagram of a three-stage filter pack.

has been in operation for various periods at 11 collocated field sites, and all sampling and operations are performed using standard CASTNet procedures (Sickles and Shadwick, 2002). In the beginning of 2001, collocated monitoring systems were operated at Mackville, KY (MCK131/231), and Ashland, ME (ASH135/235). The collocated sampling has since terminated at ASH135, and a new collocated site was established at Rocky Mountain National Park, CO (ROM206/ROM406). The two national park sites are operated independently as ROM206 is operated on behalf of the EPA and ROM406 on behalf of the NPS—the first time in the history of the network that collocated sites with different operators have been used. The collocated site at ROM206/ROM406 provides a good measure for network comparability between EPA- and NPS-sponsored sites.

CASTNet also includes a collocated monitoring site located in Egbert, Ontario, Canada (EGB181/EGB281). Initiated in 1995, this site gathers results from day and night samples, collected weekly, along with a standard weekly CASTNet filter pack. This system provides the means for ongoing intercomparison between measurements made by CASTNet and the Canadian Air and Precipitation Monitoring Network (CAPMoN). CAPMoN is a rural network initiated in 1983 and currently consists of 19 monitoring sites in Canada and 1 in the U.S. The network measures wet deposition and (inferential) dry deposition as well as ambient concentrations of acid-forming gases and particles. Particle and trace gas concentrations are determined using 24 h integrated filter measurements. More information on CAPMoN can be obtained at: http://www.msc-smc.ec.gc.ca/capmon/index_e.cfm.

In addition to the collocation of CASTNet sites, NADP/NTN operates wet deposition sampling systems collocated at 15 EPA-sponsored and 28 NPS-sponsored CASTNet sites. NADP/NTN is responsible for and administers the analysis and reporting of precipitation chemistry samples utilizing the protocols developed by NADP's network operations subcommittee. Each sampling system is equipped with a precipitation chemistry collector and rain gauge. The collector is automated to ensure that the sample is exposed only during precipitation events. Site operators collect samples weekly and send them to the Central Analytical Laboratory (CAL) at the Illinois State Water Survey for analysis, data entry, and validation. NADP/NTN operates wet deposition sampling systems at other locations near virtually every CASTNet site. Detailed information on NADP can be found at: http://nadp.sws.uiuc.edu/. CASTNet had operated a network of 21 wet deposition sites which were transferred to NADP protocol in 1999 to promote nationwide consistency in wet-deposition monitoring. Data is available from the CASTNet Website at: http://www.epa.gov/castnet.

CASTNet had included measurements of fine mass ($PM_{2.5}$) and its chemical constituents at some sites with the objective to measure air quality and related parameters thought to affect visibility. The visibility network consisted of eight sites in the eastern U.S. The sites operated from October 1993 to May 2001. These measurements were discontinued in 2001, and the CASTNet sampling systems were replaced by systems operated by the IMPROVE network to ensure program consistency with sampling and QA procedures. Historical information on data collected by CASTNet operated visibility systems is available from the CASTNet Website. For more information regarding the IMPROVE network, visit: http://vista.cira.colostate. edu/improve/.

31.5 METHODS

This section provides a brief overview of methods employed for CASTNet as outlined in the CASTNet Quality Assurance Project Plan (QAPP)(Harding ESE, 2002a). Step-by-step protocols and additional details on these activities can be found in the QAPP and CASTNet annual reports both of which can be accessed on EPA's website: http://www.epa.gov/castnet.

31.5.1 Field Operations

Atmospheric sampling for all species except ozone is integrated over weekly collection periods using a three-stage filter pack as shown in Figure 31.2. In this approach, particles and selected gases are collected by passing air at a controlled flow rate through a sequence of Teflon®, Nylon®, and dual Whatman® filters, impregnated with potassium carbonate (K_2CO_3) used for the collection of SO_2. The Teflon filter removes all particulates greater than 0.01 Fm in diameter: sulfate, nitrate, and ammonium, and certain cations, and the Nylon filter removes nitric acid. The impregnated Whatman® filters are cellulose filters and are used for the removal of sulfur dioxide. In practice, a fraction (usually <20%) of ambient SO_2 is captured on the Nylon filter. The Nylon filter SO_2 and Whatman filters SO_2 are summed to provide weekly average concentrations. The Nylon filter HNO_3 is converted to NO_3 and added to the NO_3 collected on the Teflon filter to provide weekly total NO_3 concentrations.

Filter packs are prepared and shipped to the field weekly and exchanged at each site every Tuesday. Filter-pack sampling and O_3 measurements are performed at 10 m using a tilt-down aluminum tower. Filter-pack flow is maintained at 1.5 l/pm at eastern sites and 3 l/pm at western sites for standard conditions of 25ºC and 760 mmHg with a mass flow controller (MFC).

Both EPA-sponsored sites and NPS-sponsored sites operate continuous O_3 analyzers. Meteorological variables and O_3 concentrations are recorded continuously and reported as hourly averages. CASTNet quality assurance (QA) procedures for the EPA O_3 analyzers are different from the EPA requirements for State and Local Monitoring Stations (SLAMS) monitoring as described in 40 CFR Part 68, Appendix A (USEPA, 1998b). On the other hand, the QA procedures for the O_3 analyzers at NPS sites meet the SLAMS requirements. Consequently, not all O_3 data can be used to gauge compliance with National Ambient Air Quality Standards (NAAQS) for ozone.

Unlike urban O_3 measurements, the CASTNet O_3 measurements are considered regionally representative and, therefore, able to define geographic patterns of rural ozone across most of the U.S. These data are appropriate for use in establishing general status and trend patterns in regional O_3 levels and for making general statements regarding the extent to which rural areas exceed the concentration levels mandated by the NAAQS. Ambient O_3 concentrations are measured via ultraviolet (UV) absorbance. Zero, precision (90 ppb), and span (400 ppb) checks of the O_3 analyzer are performed every Sunday using an internal O_3 generator.

Each CASTNet site employs meteorological equipment to measure temperature, delta temperature, relative humidity, solar radiation, scalar and vector wind speed, sigma theta, surface wetness, and precipitation. CASTNet meteorological measurements are described in greater detail in the CASTNet QAPP (Harding ESE, 2002a). The data obtained are archived as hourly averages. All field equipment is subject to semiannual inspections and multipoint calibrations using standards traceable to the National Institute of Standards and Technology (NIST). Results of field calibrations are used to assess sensor accuracy and flag, adjust, or invalidate field data. In addition, field audits are performed annually by Air Resource Specialists, Inc. (ARS).

31.5.2 Laboratory Operations

All filter-pack samples (both EPA-sponsored sites and NPS-sponsored sites) are loaded, shipped, received, extracted, and analyzed by the EPA's contractor, MACTEC Engineering and Consulting (formerly, Harding ESE), at their Gainesville, FL, laboratory. Following receipt from the field, exposed filters and blanks are extracted and then analyzed for SO_4^{2-} and NO_3 by ion chromatography (IC); for NH_4^+ by the automated indophenol method; and for four cations (Ca^{2+}, K^+, Mg^{2+}, and Na^+) using inductively coupled argon plasma–atomic emission spectroscopy (ICAP-AE). All analyses are completed within 72 h of filter extraction. Results of all valid analyses are stored in the laboratory data management system (Chemical Laboratory Analysis and Scheduling System [CLASS™]). Atmospheric concentrations are calculated based on volume of air sampled, following validation of hourly flow data.

31.6 METHODS OF DATA ANALYSIS

31.6.1 Modeling Dry Deposition

The original network design was based on the assumption that dry deposition or flux could be estimated as the linear product of ambient concentration (C) and deposition velocity (V_d):

$$Flux = \overline{C} \times \overline{V_d}$$

where the overbars indicate an average over a suitable time period (e.g., 1 h).

The influence of meteorological conditions, vegetation, and chemistry is simulated by V_d. Dry deposition processes are modeled as resistances to deposition:

$$R = R_{\mathrm{a}} + R_b + R_c = 1/V_d$$

where R_a signifies aerodynamic resistance or the resistance to turbulent vertical transport; R_b is the boundary layer resistance to vertical transport in a very shallow layer adjacent to the surface; and R_c is the canopy resistance or the resistance to pollutant uptake by the vegetative canopy. R_c simulates several physical and chemical processes.

The meteorological, vegetation, and land-use data are used as input to the MLM, a micrometeorological, mathematical model that simulates atmospheric dry deposition processes. MLM calculations are not one-dimensional but applied through a 20-layer canopy in which model parameters are modified by the redistribution of heat, momentum, and pollutants. The MLM software code is updated according to established version control procedures (the most recent software is designated as Version 2.3). The meteorological variables used to determine R_a and R_b are obtained from the 10-m meteorological tower at each of the sites, which is normally located in a clearing over grass or another low vegetative surface. Data on vegetative species, leaf area index (LAI), and percent green leafout are obtained from site surveys and observations by the site operator. LAI measurements were taken during 1991, 1992, and 1997 at times of summer maximum leafout. LAI values used in the MLM were extrapolated from these measurements using percent leafout observations. The resistance terms (R_a, R_b, and R_c) are calculated for each chemical species and major vegetation/surface type for every hour with valid meteorological data. The V_d for a site is then calculated as the area-weighted V_d over vegetation types within 1 km of the site.

31.6.2 Deposition Flux Calculations and Aggregations

Hourly deposition fluxes are calculated as the product of the hourly deposition velocity (V_d) obtained from the MLM and the corresponding hourly concentration. Hourly concentrations are obtained from the weekly filter-pack results and measured hourly O_3 concentrations; all hourly concentrations during a filter-pack sampling period are assumed to be equal to the filter-pack sample concentration and constant for the duration of the sample.

Weekly deposition fluxes are the sum of the valid hourly fluxes for a standard deposition week, divided by the ratio of valid hourly fluxes to the total number of hours in the standard week to account for missing or invalid values. A standard deposition week is defined as the 168-h period from 0900 Tuesday to 0900 the following Tuesday. For some weeks, the filter-pack sampling period does not correspond exactly with the standard deposition week, resulting in some deposition weeks being derived from hourly concentrations from more than one filter-pack sample. A weekly deposition flux is considered valid if it is comprised of valid hourly values for at least 70% of the 168-h week (i.e., 118 h).

Similarly, quarterly fluxes are calculated from weekly values, and are considered valid if they are comprised of valid weekly values for approximately 70% of the weeks of the 13-week period. Also, the midpoint of the sampling week had to occur in the quarter to be included as part of the respective quarterly average. Annual values are calculated from quarterly values and are considered valid if they are comprised of at least three valid quarters. Quarterly and annual mean concentrations are aggregated based on the same requirements as the flux aggregations. However, the concentrations are averaged while the fluxes are summed.

31.7 QUALITY ASSURANCE

At the beginning of NDDN and continuing with CASTNet, the EPA established rigorous quality guidelines for program operations and data. CASTNet has a fully documented quality assurance (QA) program which is compliant with American National Standards Institute's "Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs" (ANSI/ASQC E4-1994). The program includes collocated sites for determining network precision, interlaboratory comparisons, field blanks, system audits and a standardized laboratory quality control program. The program also incorporates corrective action and reporting procedures in an internal audit system to provide project-wide assessments of the field, laboratory, and data reporting operations.

Data quality indicators (DQI) have been formulated to gauge the achievement of CASTNet overall data quality objectives (DQO). The DQI, which are quantitative and qualitative descriptors used to interpret the degree of acceptability and utility of the collected data such as completeness, accuracy, precision, and comparability are applied to the hourly, weekly, and annual data to ensure that the data are of known and documented quality. For example, the CASTNet DQI regarding completeness (i.e., the percentage of valid data points relative to total possible data points) requires a minimum completeness of 90% for every measurement for each quarter. In addition, the data aggregation requires approximately 70% data completeness for hourly fluxes and weekly concentrations/fluxes in order to calculate weekly and quarterly values, respectively. Among the DQO developed for the CASTNet program has been the determination of seasonal and annual trends in ambient concentrations of sulfur and nitrogen species and rural ground-level ozone for a minimum change of 10% over a period of 10 years with a 95% level of confidence. The CASTNet quality program is reviewed and revised on an annual basis or more often if necessary. It is discussed in detail in the CASTNet Quality Assurance Project Plan (QAPP) (Harding ESE, 2002a).

31.8 CASTNET DATABASE

The CASTNet database contains archives of continuous meteorological, ozone, and flow data, concentrations measured on exposed filters, and MLM output of hourly, weekly, quarterly, and annual dry deposition fluxes. The CASTNet database is available to the public via the EPA's CASTNet data Web page: www.epa.gov/castnet/data.html.

The Web site provides archives of the concentration and deposition data in delimited ASCII files compressed using the PKZIP compression utility. Fully validated data are generally available approximately 10 months following collection. Documentation describing the content, format, and codes of the data file is included in the compressed ZIP file. In addition, a table is provided on the Web page which describes each of the available files and lists them by table name. Other documentation for the network, including information about all CASTNet sites, can be found at the CASTNet website: www.epa.gov/castnet/.

31.9 LIMITATIONS

CASTNet dry deposition is determined using an inferential method of measured ambient concentrations and modeled deposition velocity which creates a source of uncertainty in these estimations. Routine monitoring of deposition of pollutants by dry processes is difficult to measure directly and, consequently, has proven not practical for regional networks. CASTNet is designed to be a rural monitoring network, collecting data to establish site-specific measurements of total deposition. Continued research on how best to estimate dry atmospheric deposition and accurately extrapolate measurements from a single monitoring location to entire ecosystems is necessary. Inferential model flux calculations for CASTNet are generally biased low due to the weekly integrated sampling protocol. A higher temporal resolution for ambient concentration measurements and dry deposition flux calculations may provide an improved understanding in aerosol process science and air-quality events (i.e., diurnal temperature fluctuations).

31.10 CONCENTRATION TRENDS

As previously mentioned, one of the major goals of CASTNet is to monitor trends in air quality and deposition over time as emission reductions take place. The emission reductions of Title IV only began in 1995, coinciding with Phase I of the Acid Rain Program, and air concentrations and deposition responses to those reductions varies. The sites in Figure 31.1, depicted by a solid symbol, indicate the locations of the 34 eastern sites used to perform trend analyses of pollutant concentrations measured during the period 1990 through 2002. The reference sites in Figure 31.1 were selected using criteria similar to those used by the EPA in its National Air Pollutant and Emissions Trends Report (2000). Sites possessing complete data for at least 10 of the 13 years were selected. Missing quarterly data were interpolated from adjacent quarterly data (e.g., first quarter 1996 data were interpolated from 1995 and 1997 first quarter data). Missing quarterly means for 1990 or 2002 were assumed equal to adjacent quarterly values. A valid quarterly mean was based on at least 9 valid weeks (69%). Annual means were based on the four quarterly means for the year.

Figure 31.3 to Figure 31.6 (box plots) depict an annual trend analysis based on this reference site determination for SO_2, SO_4^{2-}, HNO_3, and NH_4^+, respectively. The

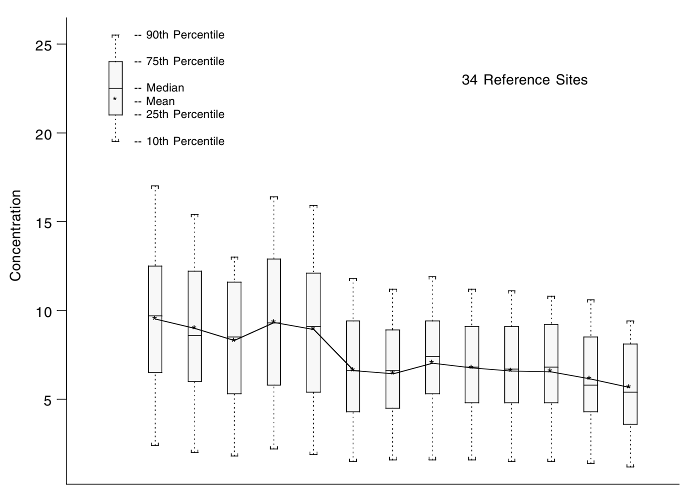


FIGURE 31.3 Trend in annual SO_2 concentrations ($\mu g/m^3$) — eastern U.S.

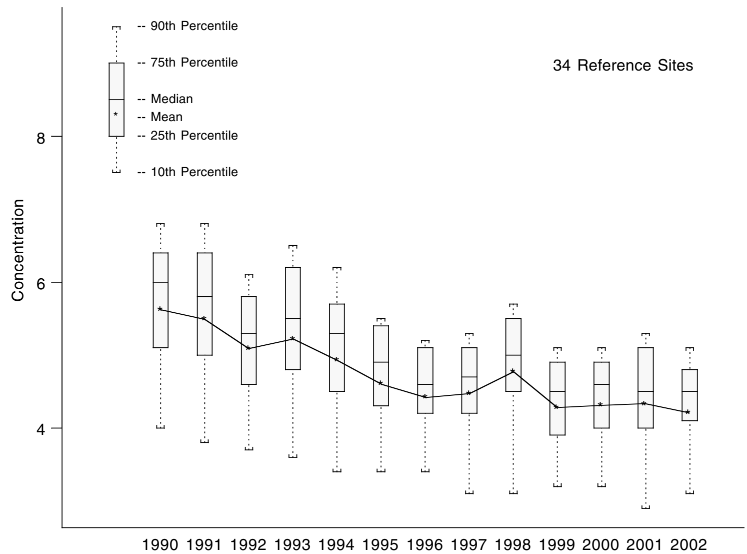


FIGURE 31.4 Trend in annual SO_4^{2-} concentrations (μg/m^3)—eastern U.S.

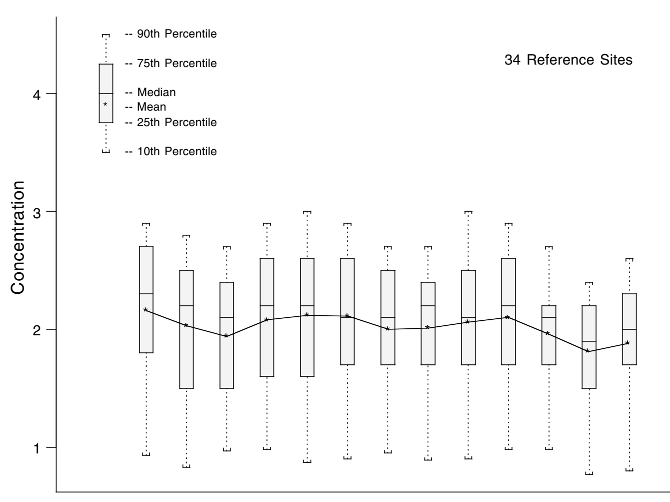


FIGURE 31.5 Trend in annual HNO_3 concentrations (μg/m^3)—eastern U.S.

figures illustrate one example of identifying trends in air pollutants utilizing CASTNet concentration measurements from the 34 reference sites. The intersite variability among the 34 sites is shown graphically by the mean, median, and percentile values of annual mean concentrations for each year.

Additional analysis using 3-year concentration averages for the years 1989 to 1991 and 1999 to 2001 provides the trend for sulfur dioxide and particulate sulfate

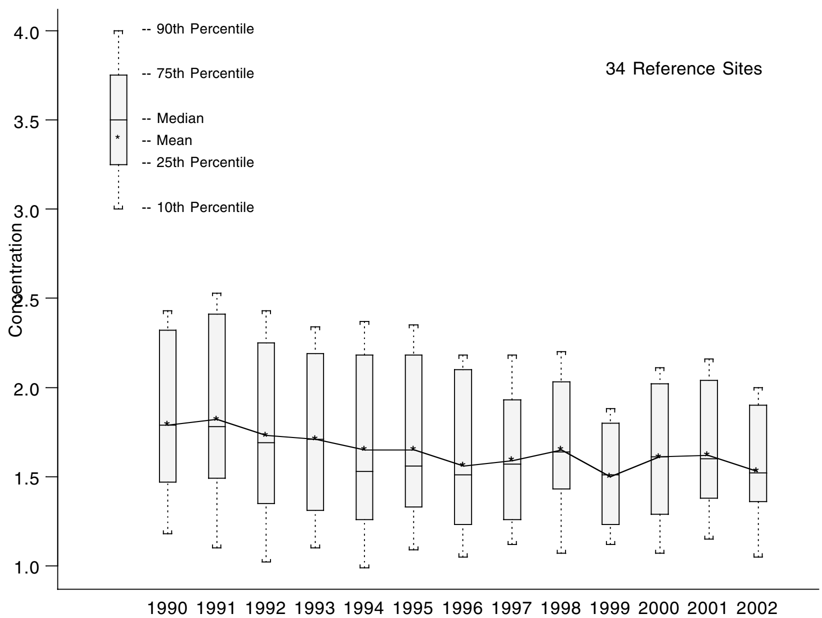


FIGURE 31.6 Trend in annual NH_4^+ concentrations (μg/m^3) — eastern U.S.

as measured by eastern CASTNet sites (see Figure 31.7a, Figure 31.7b, Figure 31.8a, and Figure 31.8b).

31.10.1 Sulfur Dioxide

Figure 31.3 provides a box plot that illustrates the trend in annual mean SO_2 concentrations (μg/m^3) over the 13-year period, 1990 to 2002. The graph shows a sharp reduction in annual mean SO_2 concentrations in 1995 in response to the Phase I emission reductions, a small increase in the median value in 1997, and a subsequent decline since then. The difference between 3-year means from the beginning to the end of the 13-year period is 28%.

31.10.2 Particulate Sulfate

The trend in annual mean SO_4^{2-} concentrations (μg/m^3) is shown in the box plot in Figure 31.4. Sulfate concentrations have declined less rapidly than SO_2 concentrations; however, the plot shows a meaningful decline in measured sulfate over the last 13 years. The difference between the 3-year means at the beginning and end of the 13-year period is 20%.

31.10.3 Nitric Acid

The box plot shown in Figure 31.5 illustrates the trend in nitric acid concentrations (μg/m^3). The figure does not signify a trend over the 13 years, although, the data does suggest a slight decline in annual HNO_3 concentrations over the last 3 years.

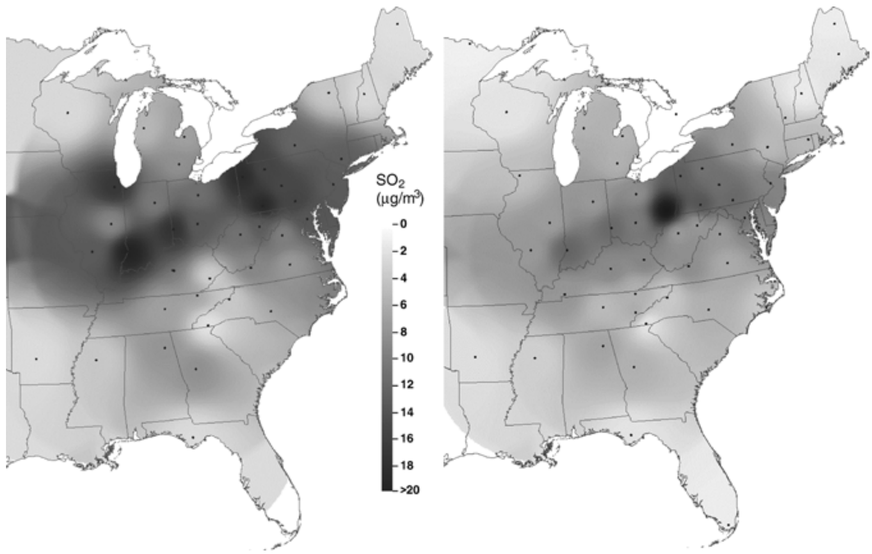


FIGURE 31.7 Sulfur dioxide concentration average of 1989 to 1991/1999 to 2001.

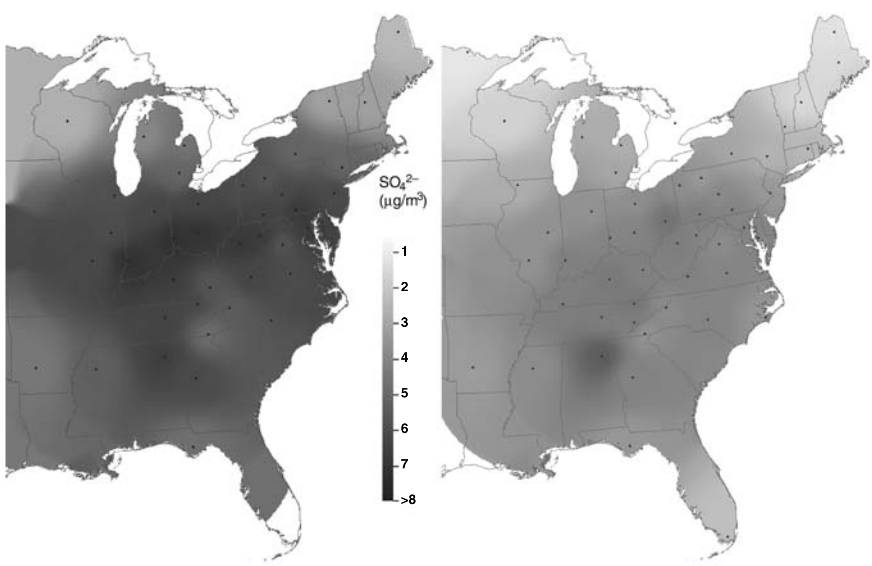


FIGURE 31.8 Sulfate concentration average of 1989 to 1991/1999 to 2001.

31.10.4 Particulate Ammonium

Figure 31.6 presents a box plot of annual mean NH_4^+ concentrations ($\mu g/m^3$). The data show a slight reduction in annual mean NH_4^+ concentrations over the 13-year period.

31.11 2002 CONCENTRATIONS OF SULFUR AND NITROGEN

CASTNet filter packs obtain weekly measurements of atmospheric concentrations of six pollutants — SO_2, SO_4^{2-}, HNO_3, NO_3^-, total nitrate (HNO_3 plus particulate NO^{3-}), NH_4^+, and four cations — Ca^{2+}, Mg^{2+}, Na^+, and K^+. This section presents findings for the six pollutants, including maps of 2002 mean concentrations (Figure 31.9 to Figure 31.14). The maps were prepared using concentration shading to illustrate the magnitude of concentrations for each of the pollutants measured at CASTNet sites across the continental U.S. The shading was prepared using an

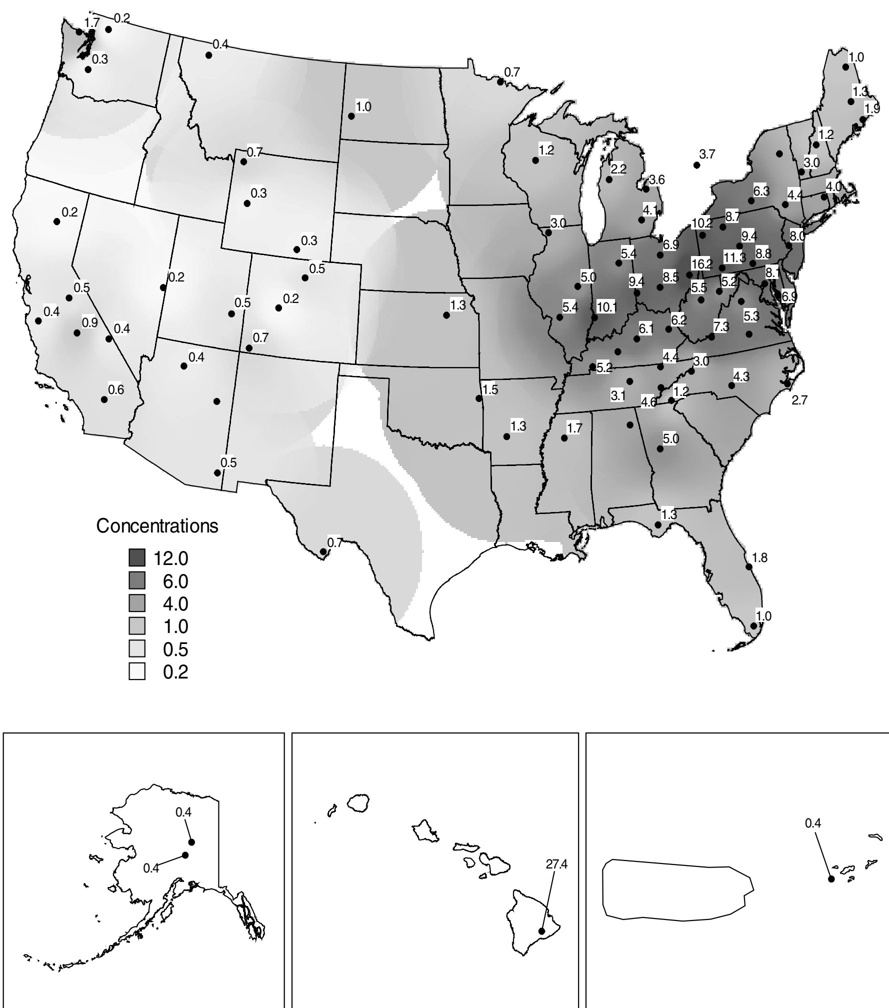


FIGURE 31.9 Annual mean SO_2 concentrations ($\mu g/m^3$) for 2002.

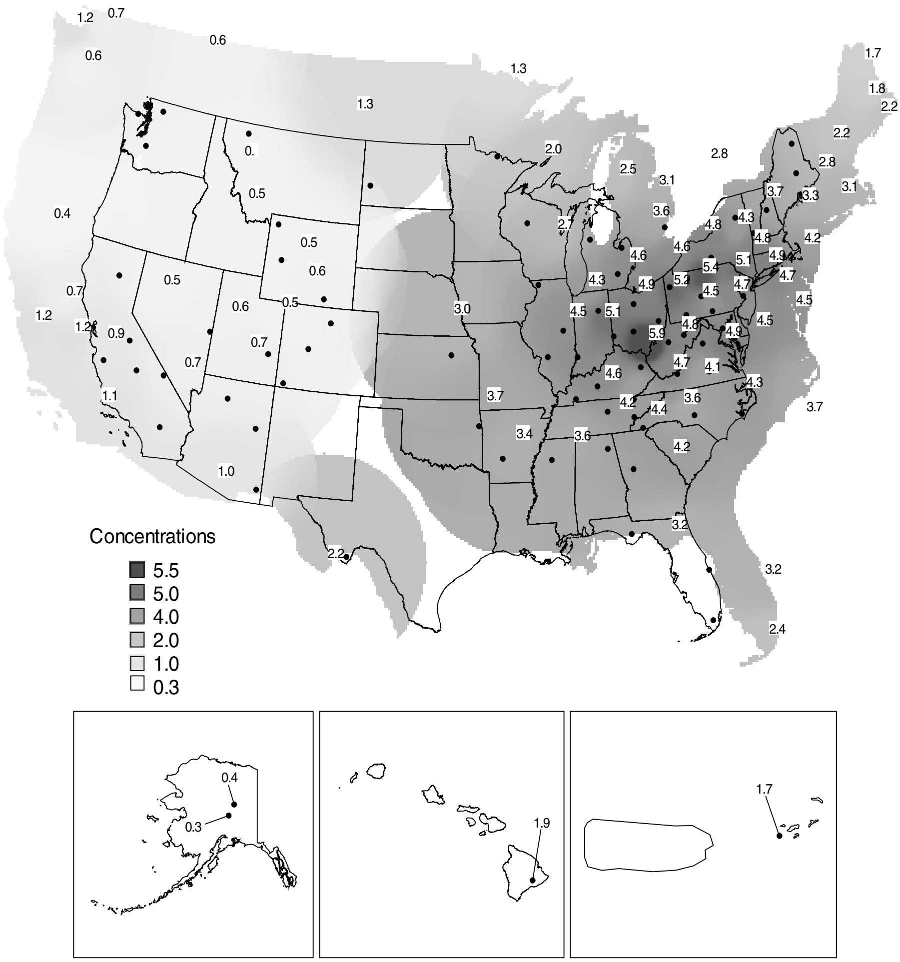


FIGURE 31.10 Annual mean SO_4^{2-} concentrations (μg/m^3) for 2002.

algorithm based on inverse distance cubed weighting with a radius of influence of 500 km. Shading was not prepared for Alaska, Hawaii, and the Virgin Islands.

31.11.1 Sulfur

The concentration maps illustrate the differences in primary (e.g., SO_2) and secondary (e.g., SO_4^{2-}) pollutants and their geographic relationship to major SO_2 sources (i.e., Phase I affected units), particularly in the Ohio River Valley region of the eastern U.S. There is a more direct relationship between source strength and downwind ambient concentration for primary air pollutants than for secondary pollutants (Holland et al., 1999). A map of 2002 mean SO_2 concentrations is provided in

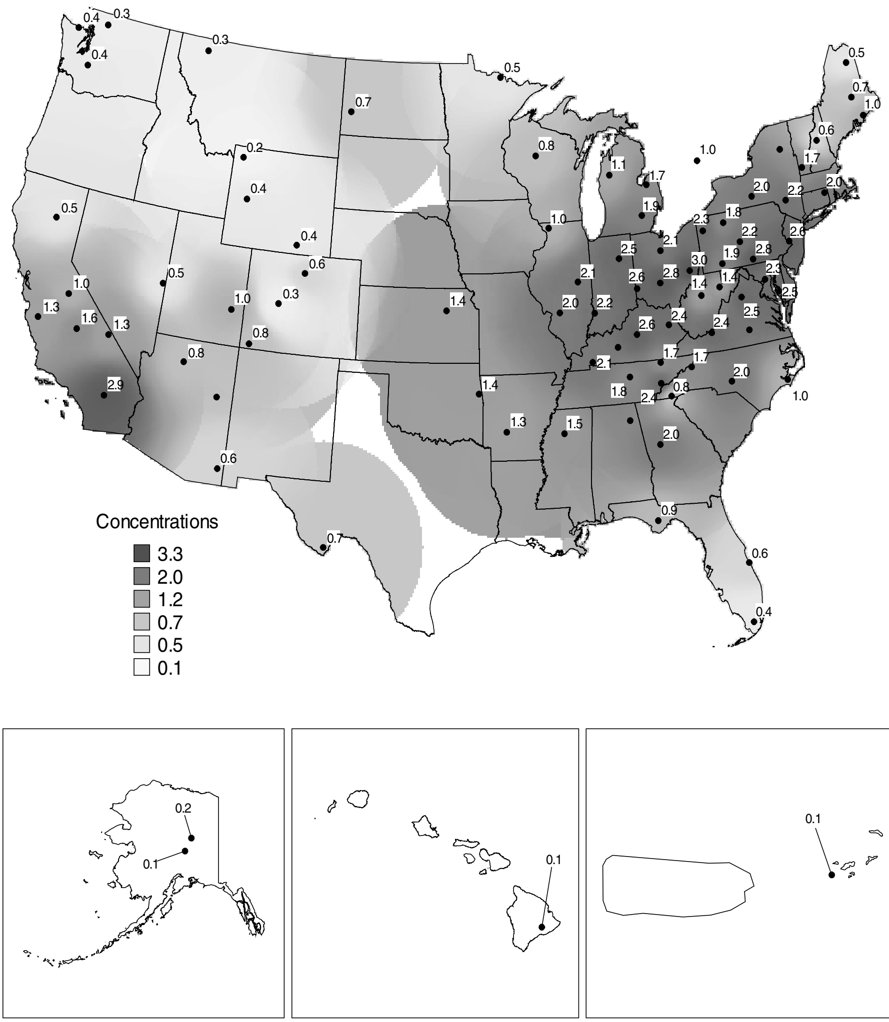


FIGURE 31.11 Annual mean HNO_3 concentrations (μg/m^3) for 2002.

Figure 31.9 and annual mean particulate sulfate (SO_4^{2-}) concentrations (g/m^3) observed during 2002 are presented in Figure 31.10. The primary pollutant (sulfur dioxide) is higher in the source region whereas the secondary pollutant (particulate SO_4^{2-}) which is formed from SO_2 through photochemical reactions in the atmosphere is more uniformly distributed in and around the source region.

31.11.1.1 Sulfur Dioxide

The SO_2 map shows a region of relatively high concentrations extending from southwestern Kentucky and Illinois to New Jersey. This region corresponds to the major SO_2 source region of the Ohio River Valley. The single highest annual mean

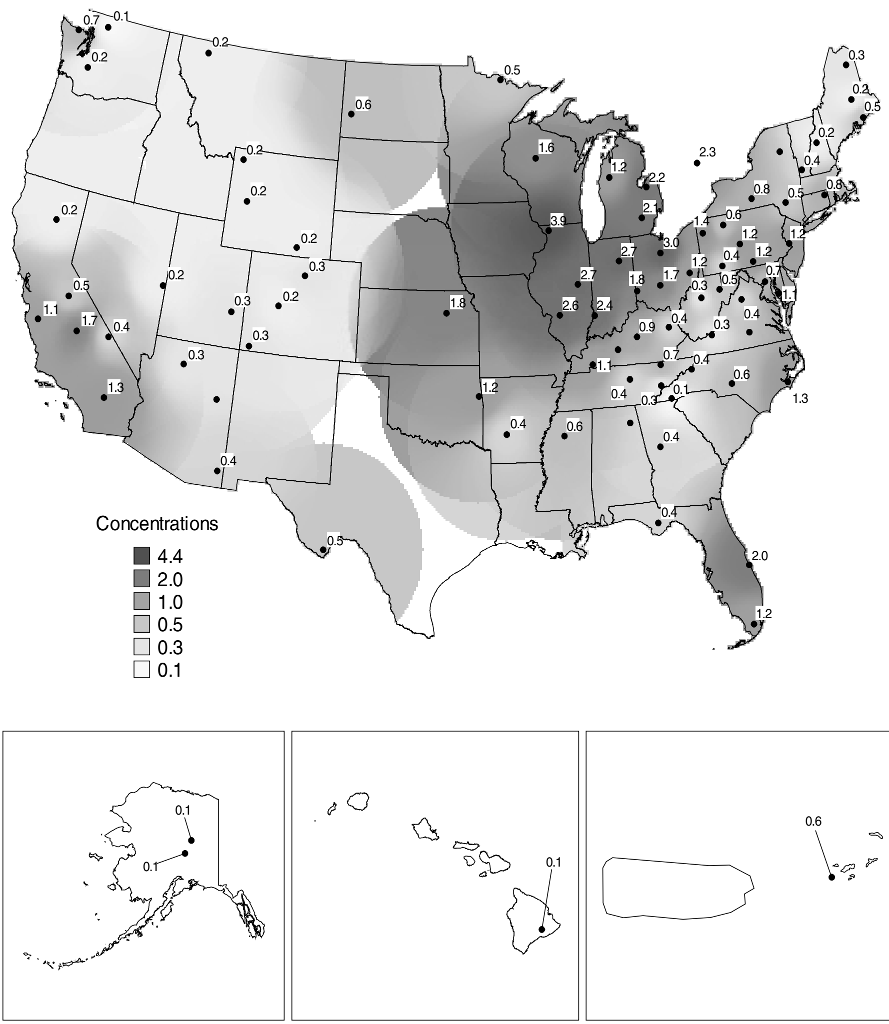


FIGURE 31.12 Annual mean NO_3 concentrations (µg/m^3) for 2002.

concentration (16.2 µg/m^3) at CASTNet sites in the continental U.S. was observed at Quaker City (QAK172) in eastern Ohio. According to CASTNet measurement data, QAK172 has registered the highest annual SO_2 concentration from 2000 to 2002). Concentrations observed at western CASTNet sites were significantly lower than those measured in the East with only a few sites measuring levels above 1 µg/m^3.

31.11.1.2 Particulate Sulfate

The SO_4^{2-} map indicates a large region with measured annual mean concentrations above 4.5 g/m^3 extends from Kentucky to eastern Maryland. In 2002, the single

FIGURE 31.13 Annual mean NH_4^+ concentrations (μg/m^3) for 2002.

highest annual mean concentration of 5.9 μg/m^3 was measured in central Kentucky at Mackville (MCK131). Annual mean SO_4^{2-} concentrations for 2002 showed an overall decline from 2001 at all reference sites over the period 1990 to 2002 as illustrated in Figure 31.4.

31.11.2 Nitrogen

31.11.2.1 Nitric Acid

A map of annual mean HNO_3 concentrations for 2002 is presented in Figure 31.11. A large region in the eastern U.S. with concentrations at or above 2 μg/m^3 extends from western Kentucky to Connecticut. The highest annual mean concentration

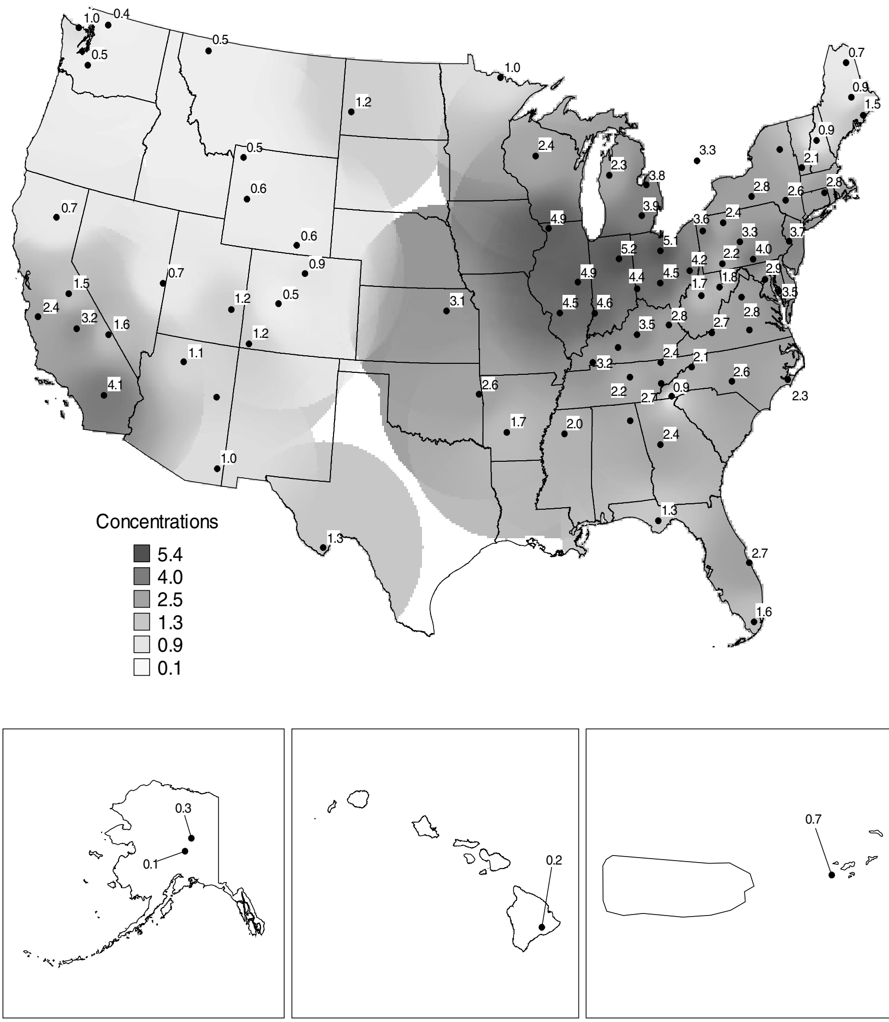


FIGURE 31.14 Annual mean total NO_3^- concentrations (μg/m^3) for 2002.

(3 μg/m^3) was measured in eastern Ohio at Quaker City (QAK172). Joshua Tree National Monument, CA (JOT403) measured the highest concentration (2.5 μg/m^3) of gaseous nitric acid in southern California and the West.

31.11.2.2 Particulate Nitrate

Figure 31.12 presents 2002 annual mean concentrations of particulate nitrate. The map depicts a region of relatively high levels (i.e., 2 μg/m^3 or greater) from Wisconsin and Illinois northeastward to Ontario, Canada. The single highest value (3.9 μg/m^3) was observed in northern Illinois at Stockton (STK138). The intersite variability in the concentrations that were observed among the midwestern sites is significant.

While most western sites measured annual mean concentrations of 0.5 μg/m^3 or lower, three sites in southern California measured values above 1 μg/m^3.

31.11.2.3 Total Nitrate

Annual mean concentrations of total nitrate (HNO_3 plus particulate NO_3^-) for 2002 are shown in Figure 31.13. The map displays a region in the midwestern states (Wisconsin, Illinois, Indiana, and Ohio with concentrations above 4 μg/m^3. The highest annual mean concentration was observed at northern Indiana at Salamonie Reservoir (SAL133) with a value of 5.2 μg/m^3. Western CASTNet sites in southern California also measured relatively high total nitrate concentrations.

31.11.2.4 Particulate Ammonium

Figure 31.14 presents a map of 2002 annual mean particulate NH_4^+ concentrations. Sites in Illinois, Indiana, northern Kentucky, Ohio, observed concentrations greater than or equal to 2 μg/m^3. With the exception of sites in Minnesota, northern New England, and Florida, most eastern sites measured levels above 1 μg/m^3.

31.12 DEPOSITION OF SULFUR AND NITROGEN

Sulfur and nitrogen transport and deposition into ecosystems are primarily accomplished through dry and wet atmospheric processes. One measure of policy and regulatory effectiveness is whether sustained reductions in the amount of atmospheric deposition over broad geographic regions are occurring. This section presents monitoring results for atmospheric deposition of sulfur and nitrogen species utilizing network maps to illustrate total deposition for 2001. Total (wet + dry) sulfur deposition (as S) and total (wet + dry) nitrogen deposition (as N) for 2002 are presented in order to highlight the relative contributions of wet and dry deposition to total sulfur and nitrogen deposition. In addition, the trends in dry and wet deposition for sulfur and nitrogen as estimated by CASTNet are presented using data obtained from the 34 reference sites for the years 1990 to 2001.

Wet deposition measurements were obtained from NADP/NTN to estimate wet deposition at CASTNet sites. Wet deposition at each site was obtained by multiplying concentrations and measured precipitation amounts from those sites. Concentrations were estimated through an inverse-distance weighting function using a combination of NADP/NTN sites and CASTNET wet deposition sites. Dry deposition processes were simulated using the MLM described by Meyers et al. (1998) and Finkelstein et al. (2000). In this section, the MLM was run using CASTNet meteorological measurements and information on land use, vegetation, and surface conditions to calculate deposition velocities for SO_2, SO_4^{2-}, HNO_3, NO_3^-, and NH_4^+. Deposition velocities were calculated for each hour with valid meteorological data for each CASTNet site for the entire period 1990 to 2001.

The MLM has been evaluated for a limited number of vegetation types, terrain settings, and time periods (summarized by Baumgardner et al. 2002). The model underestimates SO_2 dry deposition, especially for forested settings. The bias of the model for SO_2 deposition simulations is small for crops and grass. Overall, the MLM exhibits a small positive bias for HNO_3 and a small negative bias for particle nitrate. The model has not been evaluated rigorously for particles. The figures associated with trends in total wet and dry deposition of sulfur and nitrogen were created using the 34 eastern U.S. reference sites.

31.13 RELATIVE CONTRIBUTIONS TO TOTAL ATMOSPHERIC DEPOSITION

Deposition estimates represented in the pie graphs (Figure 31.15) were derived from 2001 CASTNet data using the mean of the annual deposition values for all network sites. The compiled information indicate that wet sulfate deposition was the major

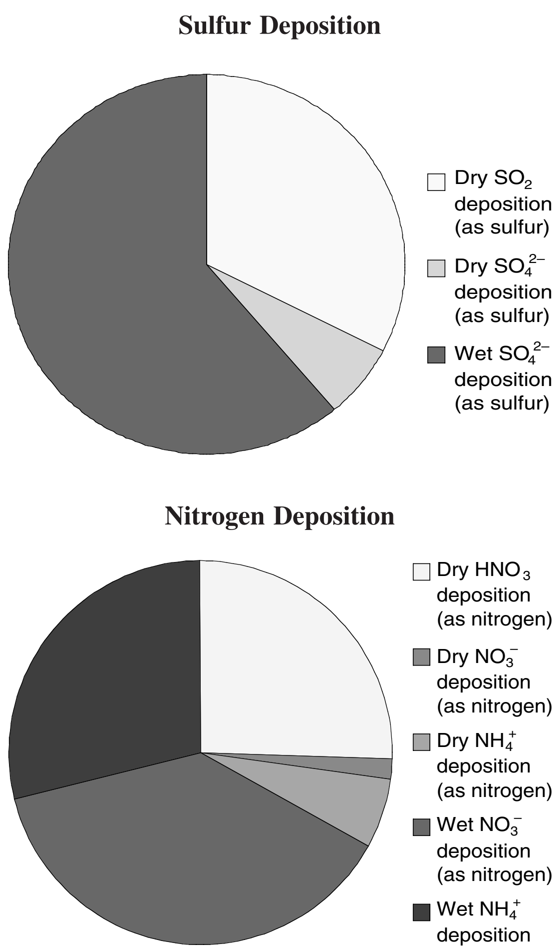


FIGURE 31.15 Wet vs. dry deposition.

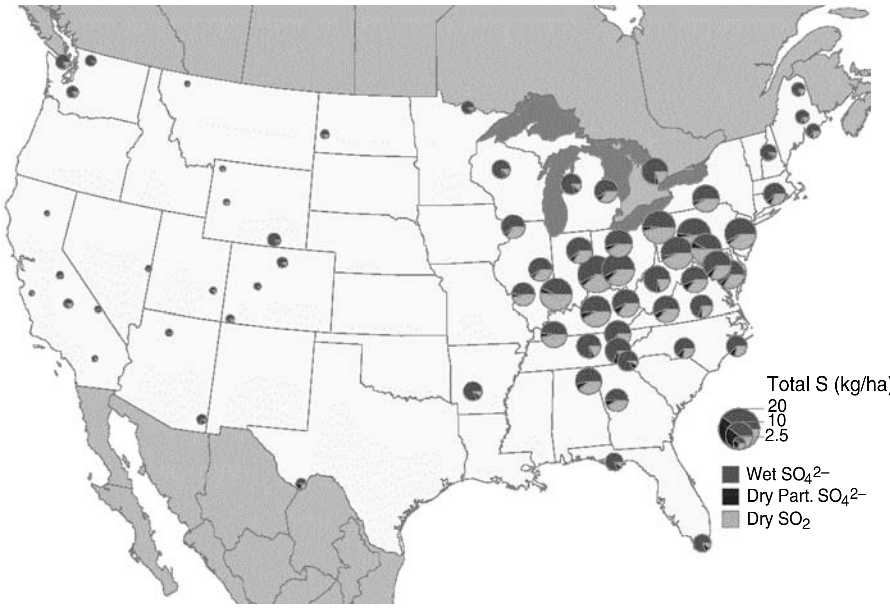


FIGURE 31.16 Total (wet + dry) sulfur deposition (as S) (kg/ha/yr) for 2001.

contributor to total sulfur deposition, followed by dry SO_2 and a much smaller contribution from dry SO_4^{2-}. Dry deposition contributed about 40% of the total sulfur deposition. Wet NO_3 deposition was the major contributor to total nitrogen deposition followed by wet NH_4^+, dry HNO_3, dry NH_4^+, and dry NO_3^-. Dry deposition contributed approximately one third of total nitrogen deposition.

31.13.1 Sulfur Deposition

Figure 31.16 illustrates the total sulfur deposition (as S) (kg/ha/yr) for 2001 as obtained from a combination of NADP/NTN and CASTNet measurements. The circles in the figure illustrate the magnitude of total sulfur deposition and also the relative species contributions to wet and dry deposition at each CASTNet site. Comparison of the deposition rates of sulfur showed that wet sulfur deposition (SO_4^{2-}) was the major contributor to total sulfur deposition, followed by dry SO_2 and then dry SO_4^{2-} deposition. Not pictured are sites in Alaska, Hawaii, and Virgin Islands. Figure 31.17 summarizes total wet and total dry sulfur deposition over the 12-year period (1990 to 2001) using the 34 eastern reference sites. Both wet and dry sulfur deposition rates decreased significantly. Dry deposition of sulfur accounted for approximately 40% of the total (wet + dry) deposition.

31.13.2 Nitrogen Deposition

Figure 31.18 presents a map of total dry nitrogen (HNO_3 + NO_3^- + NH_4^+, as N) fluxes for 2001. This map was constructed by summing the individual MLM simulations for the three species measured at CASTNet sites: HNO_3, particulate NO_3^-,

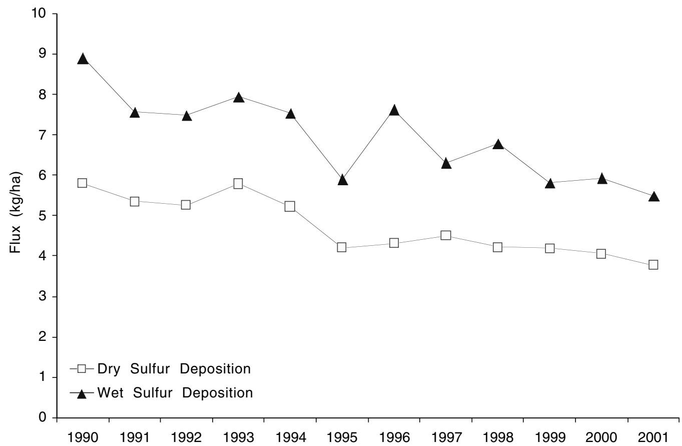


FIGURE 31.17 Trend in total dry and total wet sulfur deposition — eastern reference sites.

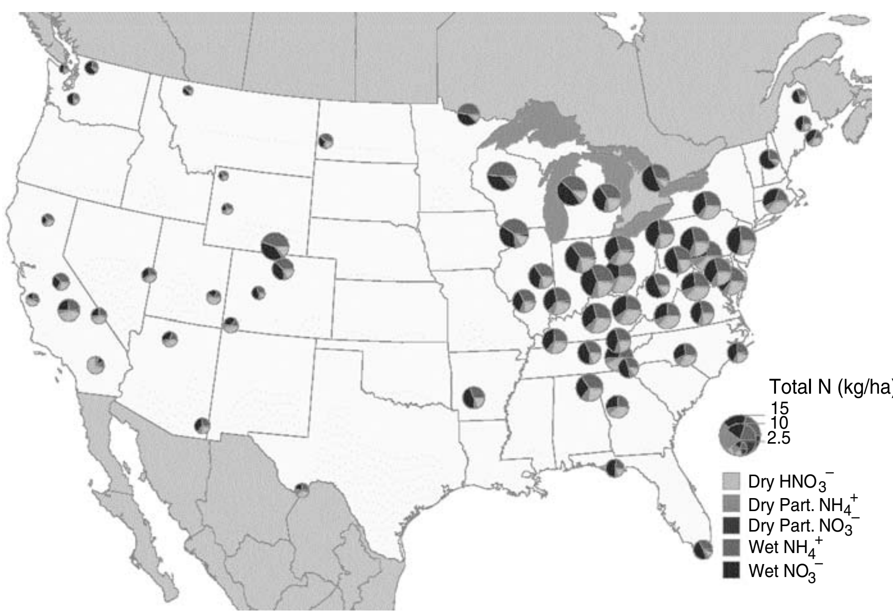


FIGURE 31.18 Total (wet + dry) nitrogen deposition (as N) (kg/ha/yr) for 2001.

and particulate NH_4^+, and wet deposition measurements obtained from NADP/NTN for NO_3^- and NH_4^+. The circles in the figure illustrate the magnitude of total nitrogen deposition and also the relative species contributions to wet and dry deposition at each CASTNet site. Not pictured are sites in Alaska, Hawaii, and Virgin Islands. Figure 31.19 provides the trend in total dry and wet nitrogen deposition over the 12-year period (1990 to 2001) using the 34 eastern reference sites.

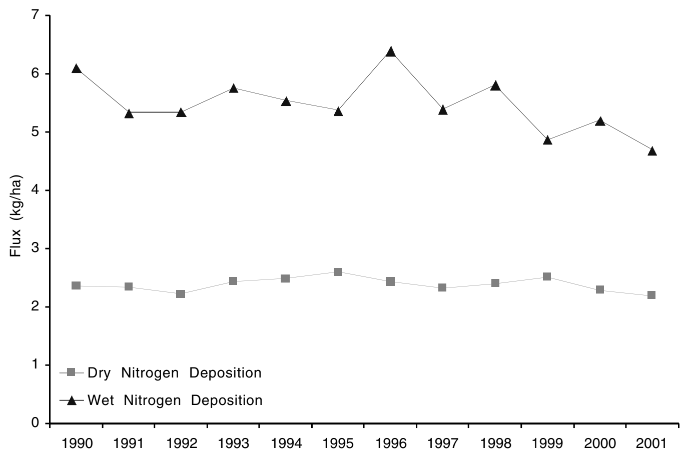


FIGURE 31.19 Trend in annual total (wet + dry) nitrogen deposition (as N) (kg/ha/yr) — eastern reference sites.

31.14 OZONE CONCENTRATIONS AND DEPOSITION

Regionally representative continuous ozone concentrations are measured at virtually every CASTNet site. The quality assurance (QA) procedures for some of the CASTNet ozone analyzers are different from the EPA-approved State and Local Monitoring Stations (SLAMS) network, consequently not all of the O_3 data can be used to gauge compliance with NAAQS. However, all of these data are appropriate for use in establishing general status and trend patterns in regional O_3 levels across the U.S. CASTNet O_3 data can also be used to approximate the extent to which rural areas potentially exceed the concentration levels mandated by the NAAQS. Seasonal fluctuation in ozone concentrations can also be analyzed using quarterly aggregated CASTNet measurements. However, for the purposes of this section, network examples include a map with concentration shading of the fourth highest daily maximum 8-h concentrations based on 2002 CASTNet measurements. A map indicating modeled deposition of ozone across the network for 2002 is also included.

31.14.1 Eight-Hour Concentrations

The EPA has developed a new rule that outlines an 8-h standard based on scientific evidence demonstrating that ozone causes adverse health effects at lower ozone concentrations, over longer periods of time, than the current 1-h ozone standard. The 8-h standard is met when the fourth highest daily maximum ozone concentration is less than or equal to 0.08 ppm. Figure 31.20 presents the fourth highest daily maximum 8-h average O_3 concentrations in ppb measured by CASTNet ozone analyzers during 2002. Under the new 8-h standard the 3-year average of the annual fourth highest daily maximum must not exceed 85 ppb.

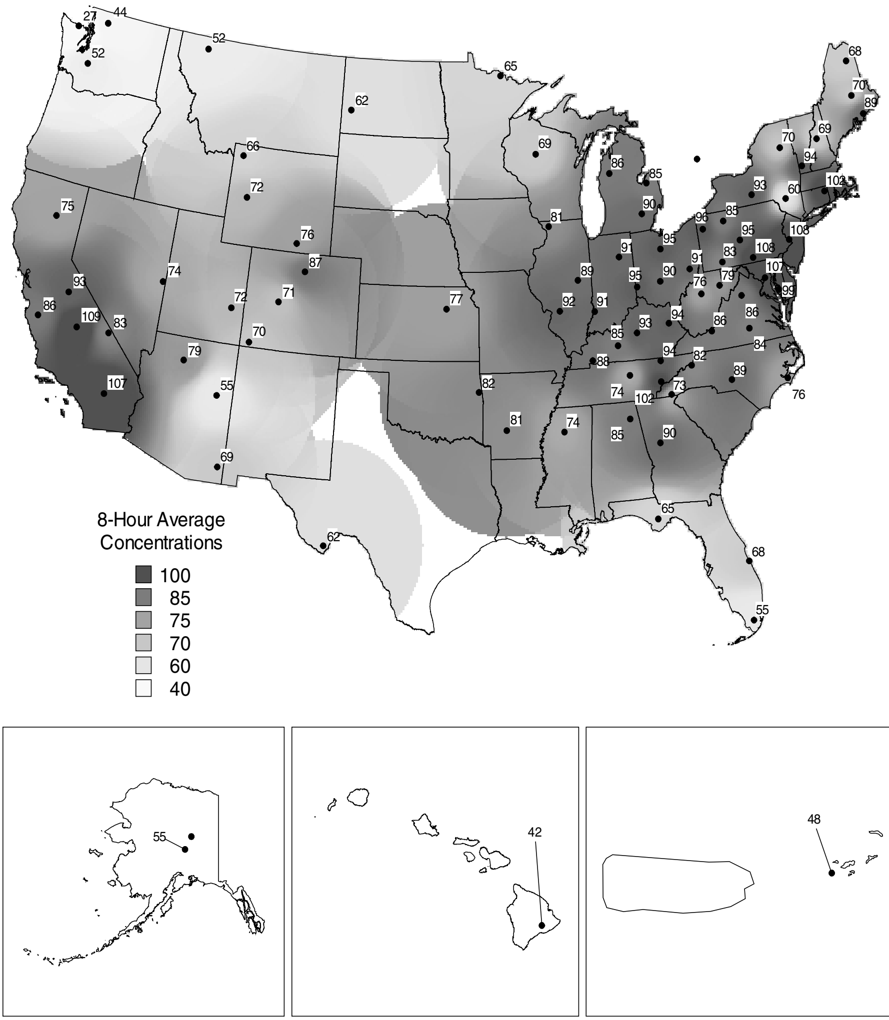


FIGURE 31.20 Fourth highest daily maximum 8-h O_3 concentrations (ppb) for 2002.

Figure 31.21 presents a map of modeled 2002 ozone-deposition fluxes. These estimates represent dry deposition of O_3 to the environment. The highest fluxes were estimated for Virginia, West Virginia, Kentucky, and North Carolina.

31.15 CONCLUSION

In order to evaluate the effectiveness of environmental policies and programs, a firm commitment to long-term monitoring is critical. Changes in the atmosphere happen very slowly and trends are often obscured by the wide variability of measurements and climate. Therefore numerous years of continuous and consistent data are

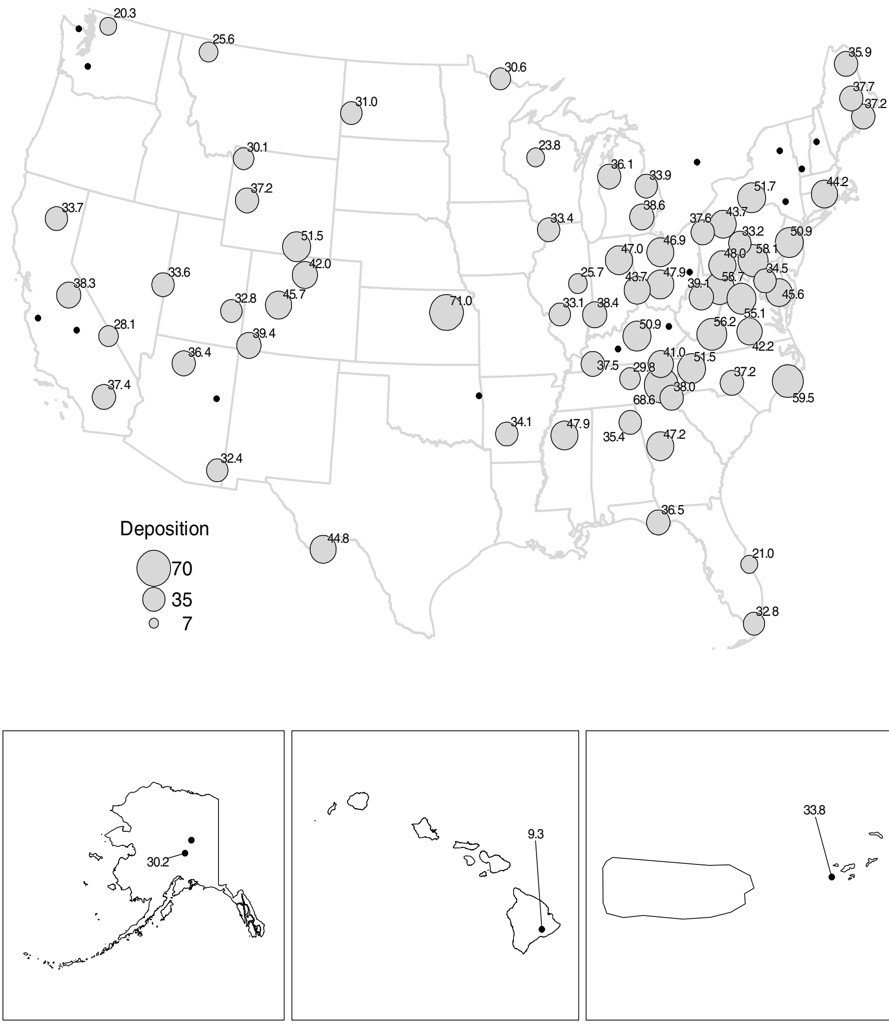


FIGURE 31.21 Modeled O_3 dry deposition fluxes (kg/ha/yr) for 2002.

required to overcome this variability, making long-term monitoring networks particularly important for characterizing deposition levels and identifying relationships among emissions, atmospheric loadings, and effects on human health and the environment. Environmental monitoring networks must use consistent procedures and quality-assured practices for observing long-term and significant changes in atmospheric composition. To accurately determine if control programs are indeed working and for assessing the ongoing costs and benefits of a given control program, an extensive data record is necessary. Monitoring programs such as CASTNet help in evaluating the status and trends in air quality, the health of ecosystems, and other important assessment endpoints. Comprehensive assessment of environmental policies and programs requires a full suite of monitoring capabilities, including tracking stack emissions, analyzing atmospheric concentrations of pollutants, estimating wet and dry deposition to land and water surfaces, and evaluating the ultimate environmental impacts through surface water chemistry and biological monitoring.

CASTNet monitoring site locations are predominantly rural by design to assess the relationship between regional pollution and changes in regional patterns in deposition. CASTNet also includes measurements of rural ozone and the chemical constituents of $PM_{2.5}$. Rural monitoring sites of the NADP and CASTnet provide data where sensitive ecosystems are located and provide insight into natural background levels of pollutants where urban influences are minimal. These data provide needed information to scientists and policy analysts to study and evaluate numerous environmental effects, particularly those caused by regional sources of emissions for which long-range transport plays an important role. Measurements from networks such as CASTNet are also important for understanding nonecological impacts of air pollution, such as visibility impairment and damage to materials, particularly those of cultural and historical importance.

REFERENCES

Baumgardner, R.E., Jr., Lavery, T.F., Rogers, C.M., and Isil, S.S. 2002. Estimates of the atmospheric deposition of sulfur and nitrogen species: clean air status and trends network, 1990–2000. *Environ. Sci. Technol.* 36(12): 2614–2629.

Finkelstein, P.L., Ellestad, T.G., Clark, J.F., Meyers, T.P., Schwede, D.B., Hebert, E.O., and Neal, J.A. 2000. Ozone and sulfur dioxide dry deposition to forests: observations and model evaluation. *J. Geophys. Res*. 105 (D12): 15,365–15,377.

Harding ESE, Inc. (Harding ESE). 2002a. Clean Air Status and Trends Network (CASTNet) Quality Assurance Project Plan. Prepared for the U.S. Environmental Protection Agency (EPA), Research Triangle Park, N.C., Contract No. 68-D-98–112. Gainsville, FL.

Harding ESE, Inc. (Harding ESE). 2002b. Clean Air Status and Trends Network (CASTNet) 2001 Annual Report. Prepared for the U.S. Environmental Protection Agency (EPA), Research Triangle Park, NC, Contract No. 68-D-98-112. Gainesville, FL.

Holland, D.M., Principe, P.P., and Sickles, J.E., II. 1999. Trends in atmospheric sulfur and nitrogen species in the eastern United States for 1989–1995. *Atmos. Environ.* 33: 37–49.

MACTEC Engineering and Consulting. Clean Air Status and Trends Network (CASTNet) 2002 Annual Report. Prepared for the U.S. Environmental Protection Agency (EPA), Office of Air and Radiation (OAR). Washington, D.C. Contract No. 68-D-03-052.

Meyers, T. P., Finkelstein, P., Clarke, J., Ellestad, T.G., and Sims, P.F. 1998. A multilayer model for interferring dry deposition using standard meteorological measurements. *J. Geophys. Res.* 103D17: 22,645–22,661.

National Science and Technology Council (NSTC), Committee on Environment and Natural Resources (CENR). 1999. The Role of Monitoring Networks in the Management of the Nation's Air Quality. Washington, D.C. March 1999.

National Science and Technology Council (NSTC), Committee on Environment and Natural Resources (CENR). 1998. National Acid Precipitation Assessment Program (NAPAP) Biennial Report to Congress: An Integrated Assessment. Silver Spring, MD, May 1998.

Sickles, J.E. and Shadwick S.D. 2002. Precision of atmospheric dry deposition data from the clean air status and trends network. *Atmos. Environ*. 36: 5671–5686.

U.S. Department of the Interior (DOI), National Park Service (NPS) — Air Resources Division. 2002. *Air Quality in the National Parks*, 2nd ed. D-2266 September.

U.S. Environmental Protection Agency (USEPA). 1998a. Clean Air Status and Trends Network (CASTNet) Deposition Summary Report (1987–1995). Office of Research and Development. EPA/600/R-98/027. July 1998.

U.S. Environmental Protection Agency (USEPA). 1998b. Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS). 40 CFR 50, Appendix L.

Wu, Y., Brashers, B., Finkelstein, P.L., and Pleim, J. E. 2003. A multilayer biochemical dry depostion model, Part 2: model evaluation. *J. Geophys. Res.* 108 (D1).

32 EPA's Regional Vulnerability Assessment Program: Using Monitoring Data and Model Results to Target Actions

E.R. Smith, R.V. O'Neill, J.D. Wickham, and K.B. Jones

CONTENTS

32.1 BACKGROUND

Until recently, ecological studies and management practices were conducted and implemented at local scales. During the past two decades, however, it has become clear that evaluations of environmental problems and management practices cannot be considered only at local scales. Increasingly, acidic deposition, global climate

1-56670-641-6/04/$0.00+$1.50

change, atmospheric contaminant transport, transformation and fate, forest fragmentation, biodiversity loss, and land use changes have been recognized as problems that have to be addressed at broader scales.[1] Local-scale assessments continue to provide valuable information, but expanded knowledge about broader-scale problems and their contribution to local-scale problems, as well as the cumulative effects of local-scale issues are needed. Unfortunately, many traditional approaches and tools are not applicable at broader scales. Approaches for collecting, analyzing, and interpreting information have to be modified or developed if efficacious management practices are to be implemented to ameliorate local, regional, and global scale problems. Drawing inferences requires more than just aggregating existing local site data.[2]

Multiple stressors affect multiple resources at these broader scales and identifying and partitioning the individual and cumulative effects of stressors across all resources represent a major research challenge.[3,4] In addition to addressing this research challenge, a regional approach is also needed to effectively target risk management activities and gain insight into the most cost-effective or socially acceptable ways to address the complex issues associated with multiple stressor–multiple resource interactions.[5] These research needs were highlighted by the U.S. Environmental Protection Agency (U.S. EPA) Science Advisory Board (SAB)[1] and were incorporated in the U.S. EPA Office of Research and Development (ORD) Ecological Research Strategy[6] as a high priority research area. Regional scale insight is critical if finer scale problems are to be put into perspective and management practices are to be effective.

The U.S. EPA's Regional Vulnerability Assessment (ReVA) program is designed to develop approaches that address the latter phases of an integrated ecological assessment, following development of specific assessment questions (problem formulation) and building on available monitoring data with a focus on integrating and synthesizing information on the spatial patterns of multiple exposures to allow comparison and prioritization of risks. ReVA is not designed to do complete regional assessments and assumes that assessment endpoints have already been identified and also that monitoring data representing these endpoints are available. As EPA ORD has learned through our 10-year involvement with the Mid-Atlantic Integrated Assessment (MAIA—a federal, state, and local partnership led by EPA Region 3), a comprehensive integrated regional assessment involves many steps and incorporates data and research that focus on understanding ecosystem processes at a variety of scales. As MAIA has evolved, five distinct, iterative steps toward improving environmental decision making have emerged: (1) monitoring to establish status and trends, (2) association analyses to suggest probable cause where degradation is observed, (3) prioritization of the role of individual stressors as they affect cumulative impacts and risk of future environmental degradation, (4) analysis of the trade-offs associated with future policy decisions, and (5) development of strategies to restore areas and reduce risk. The Environmental Monitoring and Assessment Program (EMAP) is developing approaches to steps one and two; ReVA is developing approaches to address steps three and four, and approaches to address step five will be developed by a new research program that is under development within EPA's National Risk Management Research Laboratory (NRMRL) (Figure 32.1). ReVA will demonstrate application of these approaches and provide guidance for this phase

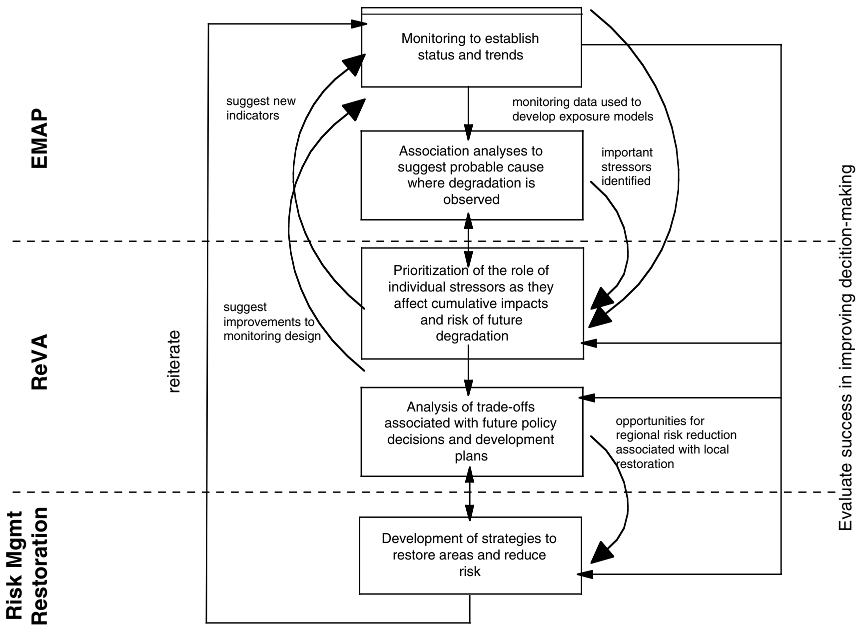


FIGURE 32.1 EPA Office of Research and Development's integrated assessment process and research programs developing approaches.

of the assessment process, but the full assessment of regional vulnerabilities is primarily the responsibility of regional decision makers. ReVA is designed to improve the methods and tools available to these decision makers.

32.2 INTEGRATED SCIENCE FOR ECOSYSTEM CHALLENGES

In 1999, the National Science and Technology Council (NSTC) and Committee for the Environment and Natural Resources, Subcommittee on Ecological Systems (CENR/SES) proposed a strategy to address major ecosystem challenges such as stresses that affect the integrity of ecosystem structure and function, and the ability of natural and managed systems to provide goods, services, and other benefits to society.[4] These stresses, which often act in concert to produce cumulative effects which are poorly understood, include: (1) changes in land and resource use, (2) introductions of invasive species, (3) inputs of pollutants and excessive nutrients, (4) extreme natural events, and (5) changes in atmospheric and climate conditions. The proposed strategy called for interagency efforts in the area of Integrated Science for Ecosystem Challenges (ISEC) to "develop, coordinate and maintain a national research infrastructure to provide the scientific information needed for effective stewardship of the nation's natural resources."

Three strategic priorities were established:

- Synthesize existing information and develop new knowledge about the structure, function, and resiliency of habitats and ecosystems
- Improve understanding of the effects of multiple stresses on habitat and the delivery of ecosystem goods and services
- Provide advanced models and information technologies to improve assessments and forecasts of habitat and ecosystem conditions under alternative policy and management options

ReVA funding was made available through ISEC in 2000, and these strategic priorities form the basis for ReVA's research priorities.

32.3 REGIONAL VULNERABILITY

A *region* is defined as a large, multistate geographic area such as the Mid-Atlantic, Northeast, Southeast, or Pacific Northwest regions within the U.S. An EPA Region is a useful representation of a geographic region because it reflects the size of the geographic area being considered in the ReVA Program and because strategic planning and management decisions are made at this scale.

Vulnerability has multiple elements in its definition but is most simply represented by the probability that future condition will change in a negative direction. We see the ecosystem as a relatively stable configuration of a number of species with the ability to resist and/or recover from the normal array of disturbances such as fire, flood, and drought that it has experienced over its evolutionary history. We assume stability, resiliency, adaptability, and resistance when we extract resources from the system and depend on it to purify wastes or to impose recreational impacts. However, these assumptions are no longer valid when the stresses we impose are outside the range that the organisms have evolved to resist, and also when we move that ecological system outside the normal range of variability. Thus, the vulnerability of an ecological system increases as the number, intensity, and frequency of stressors increase.

Vulnerability includes the interaction of multiple changes on a system. There is a long and well-documented history of synergistic effects in which one stressor lowers the resistance of the organisms to other stresses.[7] An example is provided by the high elevation spruce–fir forest in the southern Appalachians. It has been suggested that acid precipitation has lowered the resistance of the trees, permitting an epidemic outbreak of a nonindigenous pest, the balsam woolly adelgid, and the Fraser fir forests are being destroyed.[8] Even though each stressor could be resisted independently, superimposed stressors can cause unexpected unstable reactions. In most cases, we do not have sufficient research information to identify all of the potential synergistic effects, but we can at least point out that the vulnerability of an ecosystem increases as the number of different changes increases. The more different stressors there are, even at sublethal levels, the greater the risk that synergistic effects will occur and cause serious damage. Improved risk assessment of synergistic effects will be a ReVA research focus in out-years (beyond 2006).

Ecosystems are particularly vulnerable to human land use/cover changes—both urban and agricultural. In addition to altering the native ecosystem, land use conversion causes an array of side effects such as habitat fragmentation, nonpoint source pollution, and changes in water movement and water quality. Land use conversion both encourages and is influenced by road building, with its own array of environmental impacts.[9] Land use change can also drive changes in air quality with additional vehicular traffic and particulate matter associated with disturbance. Therefore, vulnerability must also consider the socioeconomic drivers that determine the probability that natural ecosystems will be converted to human uses.

In addition to the stressors on regional systems, vulnerability must also consider the sensitivity and adaptability of natural systems being stressed. Some systems may be particularly sensitive to a stressor or to cumulative effects from several stressors. Some of the sensitive resources can be identified from research results, e.g., some plant species are particularly sensitive to ozone, some fish species to chemical stressors, and small stream systems to removing riparian vegetation. Other sensitive resources can be deduced; plant species at the northern or southern edge of their distribution are already under climatic stress and may be sensitive to any additional stress.

Vulnerability also must consider the sensitivity or probability of extinction for rare species and rare habitat. Such species and habitats are vulnerable in the sense that the damage is irreparable. There are also unique habitats like wetlands, mountaintops, and caves that support a unique flora and fauna, which must also be considered vulnerable and irreplaceable.

Finally, there are ecological resources that are considered vulnerable which provide society with valued goods, services, and other benefits. This may involve resource extraction (e.g., forests and fisheries), food production (e.g., agricultural systems), recreation (e.g., forests and lakes), waste treatment, and nutrient recycling. Vulnerable ecological resources in this category are critical because damage to them can have an immediate impact on society.

Regional vulnerability means many things. It is rarity, synergy, sensitivity, and spatial context. No single question or approach will suffice. Regional vulnerability analysis will draw on many sources of data, will explore many different assessment methods, and will enable decision makers to ask many questions.

32.4 GOAL AND OBJECTIVES

The goal of ReVA is to develop and demonstrate approaches to comprehensive, regional-scale assessment that effectively informs decision makers of the severity, extent, distribution, and uncertainty of current and projected environmental risks.[10]

ReVA's objectives represent the sequential steps needed to achieve this goal:

1. Provide regional-scale, spatially explicit information on the extent and distribution of both stressors and sensitive ecological resources
2. Develop and evaluate techniques to integrate information on exposure and effects so that ecological risk due to multiple environmental changes can be assessed and compared and management actions prioritized

3. Project consequences of potential environmental changes and risk management strategies under alternative future scenarios
4. Effectively communicate economic and quality of life trade-offs associated with alternative environmental policies
5. Develop techniques to prioritize areas for ecological risk reduction
6. Identify information gaps and recommend actions to improve monitoring and focus research[10]

32.5 REVA'S RESEARCH HYPOTHESES

ReVA's working hypotheses or assumptions are:

- Spatial connections (upstream–downstream, transportation network, shortest air distance) are important in determining the total ramifications of local human activities. Therefore, cumulative ecological condition over large regions is related to large-scale patterns as well as small-scale decisions.
- Spatial variability reduces the efficiency of bottom–up approaches (and increases the efficiency of top–down approaches) when assessing ecological condition over large regions.
- Sustainability can only be achieved by maintaining regional variability. Some areas must be reserved to maintain regional biodiversity. Some areas are vulnerable to human disturbance.[10]

32.6 REVA'S APPROACH

ReVA is designed to improve environmental decision making by putting regional-scale issues in perspective. Vulnerability, as discussed earlier, depends on the cooccurrence of stressors, along with their sequence and magnitude and variation in the sensitivity of resources. ReVA's approach to assessing vulnerability will proceed by first mapping stressors, sensitive resources, and surrogate indicators of exposure (e.g., landscape indicators such as fragmentation, agriculture on steep slopes, etc.), integrating this information to map vulnerabilities, and finally, assessing the risk of impacts to valued resources (Figure 32.2). Through the development of alternative future scenarios, ReVA will ideally illustrate how policy decisions and the associated changes in vulnerability affect our quality of life. This illustration will help decision makers evaluate the trade-offs associated with alternative policy choices.

32.7 MAPPING EXPOSURES

32.7.1 Step One

The first step in ReVA will be the development of spatial datasets for the Mid-Atlantic assessment area (Figure 32.3). This will involve all relevant data that are available for the entire region and that can be represented spatially (i.e., a value for each point on the regional map). Spatially explicit data are required to compare risk

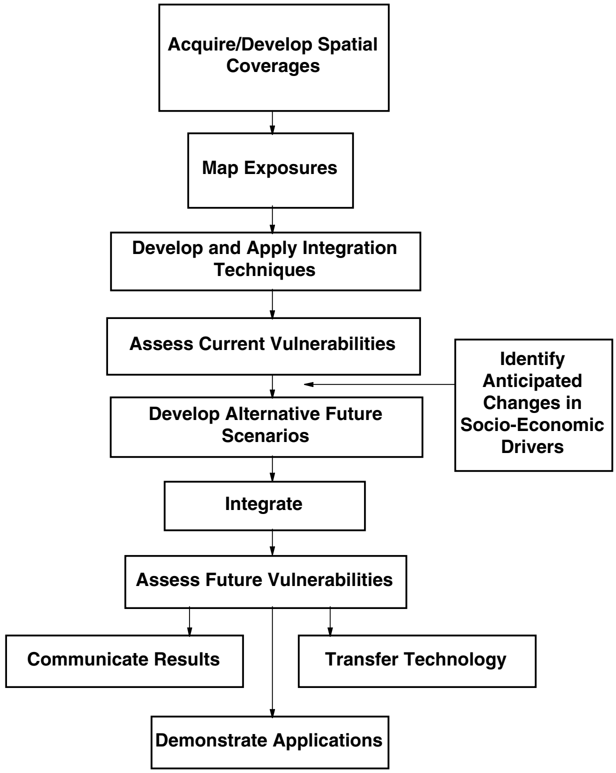


FIGURE 32.2 EPA's Regional Vulnerability Assessment Program process.

across the entire region.[11] The data will include infrastructure (e.g., roads), stressors (e.g., atmospheric deposition[12] and chemical inputs), landscape characterization,[13] sensitive resources (e.g., wetlands), and ecological endpoints (e.g., avian biodiversity[14]). Spatial data will be retained at the finest resolution available. This will keep options open for reporting assessments at a variety of spatial units.[15] Developing the spatial databases involves addressing three important issues: (1) which data to include, (2) reaggregation of data into consistent reporting units, and (3) how to estimate and retain the uncertainties in the data.

Recent research reported in the Landscape Atlas,[16] the ReVA Stressor Atlas,[17] and the Index of Watershed Indicators (IWI) (http://www.epa.gov/iwi/) indicates that comprehensive regional databases exist for a surprising number of environmental stressors and ecological resources. The emphasis in ReVA will be on utilizing the wealth of available data. Collection of new data will not be emphasized, particularly for the early phases.

Where critical gaps exist, ReVA will conduct indicator research. For example, none of the current landscape indicators addresses distance between patches. Since interpatch distance is known to be critical to the impact of habitat fragmentation on wildlife, the new indicator fills a critical gap.[18–20] This research will be conducted

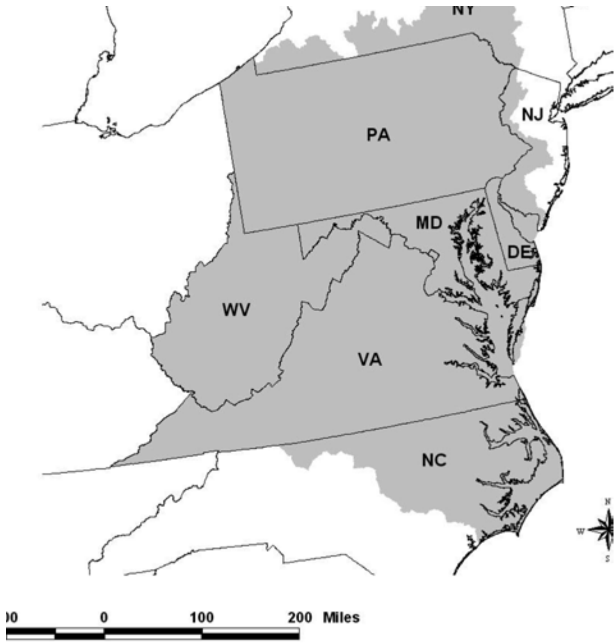


FIGURE 32.3 The mid-Atlantic region of the U.S.

as a joint project of ReVA and the Landscape Sciences Program (NERL, Environmental Sciences Division, ESD).

Some indicators will be direct measures of stress, such as projected land development or pesticide concentrations in streams. Other indicators will be the output of models, such as sulfur dioxide concentrations or nutrient loadings calculated from watershed properties.[11] Still other indicators will be extrapolations from finer-scaled mechanistic studies.[21] Surrogate indicators of exposure (e.g., fragmentation indices, agriculture on steep slopes) and other metrics (e.g., human use index) from the landscape sciences will supplement the existing information on stressors and allow a more comprehensive view of regional vulnerability.

All measurements have associated uncertainty or measurement error. There are at least two sources of error. First, the value assigned to the spatial unit has an associated uncertainty. Second, even if the indicator value were known with certainty, there is uncertainty associated with the impact actually occurring within that spatial unit.[22]

In addition, individual indicators are not necessarily independent. Agriculture on steep slopes, for example, is correlated with nutrient concentrations in freshwaters or with sediment loading to estuarine ecosystems. Spatial and autocorrelation will also be considered. For example, roads and particularly intersections of interstate highways are correlated with economic development activity, habitat fragmentation, and increased runoff from impervious surfaces. There does not appear to be any way to reduce the dataset to a subset of orthogonal indicators since, for example, the spatial network of roads is also needed as an indicator of routes for exotic species invasion and as barriers to dispersal between fragmented habitat patches.

Because risk assessment considers uncertainty, all sources of uncertainty must be retained. Measurement errors and correlations with other indicators are retained as a variance–covariance matrix. Coverages from statistical models have an associated goodness-of-fit estimate. Standard error analysis techniques will be used to estimate the uncertainty associated with coverages generated by process models. Research on combining uncertainty estimates for integrated rankings will be a high priority.

32.7.2 Step Two

The second step will be the development of additional coverages derived from the primary spatial data. ReVA will employ three approaches.

The first approach follows the ecological risk assessment paradigm.[6] Spatial distributions of stressors, resources, and exposure indicators are overlapped to represent exposure (Figure 32.4). The second approach builds on the work of NERL's Landscape Sciences Program (ESD) in developing and demonstrating the use of landscape indicators to assess ecological condition.[16] The approach uses statistical models to estimate impacts on valued resources such as water quality and wildlife habitat. For example, this approach allows us to move from available spatial data on nutrient deposition, roads crossing streams, intact riparian zones, and agriculture on steep slopes to an estimate of the risk of increased nutrient loading for every watershed across the region.

The third approach uses spatial process models to extrapolate from point monitoring data to the spatial distribution of stressors. Examples include the estimation of atmospheric deposition (nitrogen, ozone, sulfate, and some toxics) using monitoring data from systems such as the Clean Air Status and Trends Network (CASTNet, see http://www.epa.gov/ardpublc/acidrain/castnet/) combined with deposition models (e.g., Models3/CMAQ [Community Multiscale Air Quality], see http://www.epa.gov/asmdner/models3) and known occurrences of aquatic nonindigenous species combined

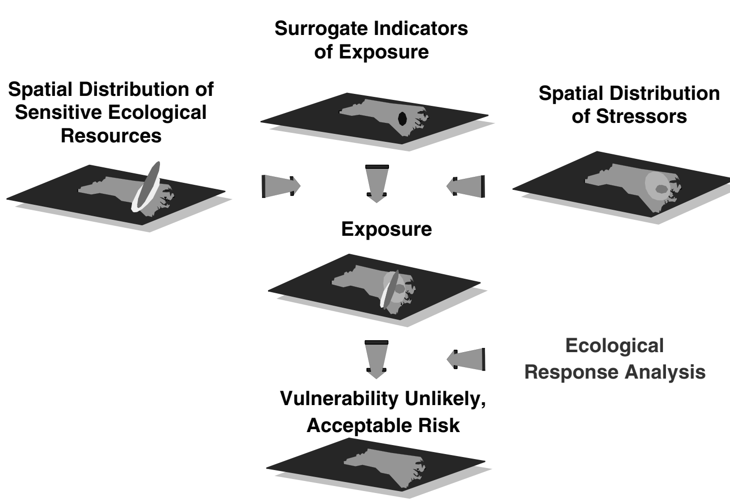


FIGURE 32.4 Illustration of spatial overlay of resources, stressors, and exposure indicators.

with geographic path analysis techniques to identify risk to native species. Survey and monitoring network data also will be used to formulate and test structured equation models using path analysis techniques.

Research will involve the development of models to link stressors to impacts on specific ecological resources. For example, metapopulation theory relates properties such as stability to the spatial configuration of a habitat. The theory can be used to relate changes in the spatial pattern of habitat to wildlife and endangered species. Development of such specific models would allow ReVA to translate fragmentation indicators into indicators of potential impact on specific species. All models used or developed by ReVA will be evaluated through sensitivity and error analyses. Where multiple models representing the same process or relationship are available, results will be compared as to where they agree or disagree and an assessment of these differences will be made and documented.

32.8 INTEGRATION AND EVALUATION

Information must be integrated across response variables and stressor distributions to assess overall environmental condition.[23] This synthesis is a complex process with many challenges, and developing the appropriate methodologies is a high-priority research task in ReVA. The integration methods developed and employed by ReVA will be evaluated as to sensitivities to different data issues (e.g., continuous or noncontinuous data and skewed distributions), to small changes in variables, and to the ease of use and understanding by the user.

Perhaps the most easily understood integration method is the weighted-sum approach (e.g., see Wade et al.[24]). Condition is the sum of the component indicators, weighted by their importance. Relative importance, however, involves subjective judgment. The National Park Service may have a very different view of watershed vulnerability from the state planner responsible for economic development. The weighted sums algorithm has the flexibility to allow the decision maker to vary the "importance" weights associated with individual indicators and develop individualized overviews of vulnerability.

Another evaluation approach might consider the risk of crossing thresholds established by federal law. This approach would use a suite of indicators to estimate the risk that a particular assessment unit was in violation of the Endangered Species Act, the Clean Water legislation, etc. Areas in potential violation of statutes or regulations could then be identified for further study.

A third assessment approach might be to represent each indicator value as a point in a multivariate state space.[25] The measure of vulnerability is then calculated as the Euclidean distance to the closest region of "unacceptable impact." This approach would also be valuable in identifying the magnitude of change in specific indicators which would produce the greatest reduction in risk. A measure of current environmental condition could also be calculated as the distance between the present state and an optimal point defined as the ideal or desired environmental condition.

The state space approach has the advantage of easily accounting for errors in both site and reference measurements. The approach uses a measure of distance derived from Fuzzy Set Theory.[26,27] This theory was developed by Zadeh[28] to deal with problems resulting from "fuzzy" or uncertain data. Each measured value is described by an interval, for example, the mean plus and minus a standard deviation. The site measurements and the reference site are represented by a fuzzy circle or polygon in the multidimensional state space. In essence, the distance measure integrates all the distances, from the nearest to the furthest and between the site and the desired state.

A similar approach can be taken using multivariate statistics. The statistical calculations are based on the variance–covariance matrix and therefore directly account for measurement errors and codependencies. This method estimates the probability that a site has departed significantly from the "natural" reference site and/or that the site is no longer significantly different from the "degraded" reference site. ReVA will also explore methods such as multiple objective decision theory,[29] artificial intelligence,[30] and spatial information integrating technology.[31]

Risks also can be estimated for economic objectives. This approach would directly consider the viewpoint of a decision maker responsible for economic development in the region. In general, the attractiveness of a specific geographic location to development and investment is determined by socioeconomic factors such as available labor, power, transportation, etc. But decisions are also based on the environmental condition of the location, involving factors such as clean air and water, available recreation, scenic values, etc. The ReVA indicators could be integrated into an overall estimate of the extent to which the economic attractiveness of an area might be degraded by environmental impacts. Such an overview would be invaluable to decision makers in setting planning and restoration priorities and developing land-use plans for zoning and other available planning tools.

32.9 DEVELOPING ALTERNATIVE FUTURE SCENARIOS

The development of alternative future scenarios is an important component of ReVA. Future scenarios will include increases in human population (e.g., see Campbell[32]) with its associated increases in pressures on the environment (e.g., air deposition, hydrologic modification), continued urban development,[33] and climate change.[34] Future scenarios for the Phase One assessment will focus primarily on extrapolating changes in land use but will also estimate potential changes in spatial distribution of pollution, nonindigenous species (NIS), and resource extraction to examine potential impacts to water quality, human health, etc. Future scenarios for Phase Two will incorporate projected land use changes as well as anticipated changes in other stresses (e.g., air emissions, spread of NIS, projected resource extraction) and changes in regional weather patterns and will be developed by predicting likely changes in socioeconomic drivers over the next 20 to 25 years.

As future scenarios are developed, the challenge is to translate the projected scenario into spatial changes in stressors and resources. In most cases, the changes can be extrapolated using the same models used in the assessment of current conditions. For example, population growth and urbanization will result in increased SO_2 that can be distributed spatially by the same air diffusion models. The most difficult task involves translating socioeconomic drivers into land use change. One approach will combine models from economic geography and statistical analysis to determine the probability of a pixel changing to agricultural or residential land use. The underlying geographic theory states that the probability of change will be determined by the access to roads (i.e., cost to transport products) as well as by the distance and size of accessible urban markets.[35]

Another approach is based on an analysis of landscape sensitivity to fragmentation. Some landscapes are more sensitive (to the loss of resource area) than others owing to differences in the spatial arrangement of their component resources. Models predict the likelihood of fragmentation as a result of different degrees of land-cover change in different spatial patterns of driving forces (e.g., roadbuilding vs. urban sprawl). Continental maps of current fragmentation sensitivity (in progress) measure the relative risks already taken by historical patterns of land use. Sensitivity can be reevaluated for any future scenario and map of land-cover change, including local remediation, to estimate the likely regional impacts on fragmentation-sensitive resources. The models have been developed by using the Multi-Resolution Land Characteristics (MRLC) land-cover maps, but the general procedure can be tuned for other data sources and scales.

Once the scenario is translated into spatial changes in stressors and response variables, the regional assessment can proceed with the same approach used in assessing current vulnerability. The result will be an overview of the impact of the proposed changes on environmental condition of the region. This will make possible a comparison of pre- and postscenario overviews including the identification of areas that are at greatest risk of degradation under alternative scenarios. These areas, therefore, might be the most vulnerable and targeted for risk management activities.

32.10 DEMONSTRATION OF FINER-SCALE APPLICATIONS

Recognizing that most environmental decision making occurs at scales that are less than regional, ReVA will partner with stakeholder groups to develop and demonstrate applications of local-scale decision making. These demonstrations will occur concurrently with the regional scale assessment as they can provide useful feedback that will improve the overall assessment product. Stakeholder groups may include nongovernmental groups (e.g., the Canaan Valley Institute), county planning organizations (e.g., Baltimore County), or private institutions (e.g., the Water Environment Research Foundation) that are interested in developing decision-support tools that enhance local planning or restoration efforts. An example is the development of a restoration tool that uses decision-tree analysis to identify and prioritize areas for riparian reforestation based on attributes associated with erodability and reducing sediment loads, habitat enhancement and regional connectivity for migratory species, aesthetics, and wetland protection. ReVA will provide data and technical assistance in the use of regional-scale information for these demonstrations. It is expected that the tools developed will use a combination of regional-scale information along with finer-scale models that provide additional detail. An evaluation of how these data and models work together may provide insights into the issue of scale and of identifying the limits of how far fine-scale information can be extrapolated and how far regional scale information can be drilled down to answer local management questions.

ACKNOWLEDGMENTS

The U.S. Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under Interagency Agreement DW89938037 with Oak Ridge National Laboratory and contract number 68-C98-187 with TN and Associates. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products does not constitute an endorsement or recommendation for use.

REFERENCES

1. USEPA. 1995. SAB Report: Futures Methods and Issues. USEPA Science Advisory Board, Environmental Futures Committee. EPA-SAB-EC-95-007A. Washington, D.C., 86 pp.
2. O'Neill, R.V., D.L. DeAngelis, J.B. Waide, and T.F.H. Allen. 1986. *A Hierarchical Concept of Ecosystems*. Princeton University Press, Princeton, NJ.
3. USEPA. 1988. Future Risk: Research Strategies for the 1990s. USEPA Science Advisory Board. Washington, D.C.
4. National Science and Technology Council. 1999a. Integrated Science for Ecosystem Challenges: A Proposed Strategy. Committee on Environmental and Natural Resources. Unpublished report, 31 pp.
5. Graham, R.L., C.T. Hunsaker, R.V. O'Neill, and B.L. Jackson. 1991. Ecological risk assessment at the regional scale. *Ecol. Appl.* 1: 196–206.
6. USEPA. 1998. Guidelines for Ecological Risk Assessment. Office of Research and Development, Washington, D.C., EPA/630/R-95/002F.
7. Foran, J. and S. Ferenc (Ed.). 1999. *Multiple Stressors in Ecological Risk and Impact Assessment*. Society of Environmental Toxicology and Analytical Chemistry. SETAC. Pensacola, FL.
8. Hain, F.P. and F.H. Arthur. 1985. The role of atmospheric deposition in the latitudinal variation of Fraser fir mortality caused by the Balsam Woolly Adelgid (*Adelges picea* (Ratz.) (Hemipt., Adelgidae)): a hypothesis. *Z. Angew. Entomol.* 99: 145–152.
9. Forman, R.T.T. and L.E. Alexander. 1998. Roads and their major ecological impacts. *Annu. Rev. Ecol. Syst.* 29: 207–231.
10. Smith, E.R., R.V. O'Neill, J.D. Wickham, K.B. Jones, L. Jackson, J.V. Kilaru, and R. Reuter. 2002. The U.S. EPA's Regional Vulnerability Assessment Program: a Research Strategy for 2001–2006, http://www.epa.gov/reva/reva-strategy.pdf, 43 pp.
11. Hunsaker, C.T.D., A. Levine, S.P. Timmins, B.L. Jackson, and R.V. O'Neill. 1992. Landscape characterization for assessing regional water quality. In *Ecological Indicators*, D.H. McKenzie, D.E. Hyatt, and V.J. McDonald, Eds. Elsevier Applied Science, New York.
12. Chang, J.S., R.A. Brost, I.S. Isaksen, S. Mandonich, P. Middleton, W.R. Stockwell, and C.J. Walcek. 1987. A three-dimensional Eulerian acid deposition model. *J. Geophys. Res.* 92: 681–700.
13. Jones, K.B., J. Walker, K.H. Riitters, J.D. Wickham, and C. Nicoll. 1996. Indicators of landscape integrity. In *Indicators of Catchment Health*, J. Walker and D.J. Reuter, Eds. CSIRO Publishing, Melbourne, Australia, pp. 155–168.
14. Flather, C.H., S.J. Brady, and D.B. Inkley. 1992. Regional habitat appraisals of wildlife communities: a landscape-level evaluation of a resource planning model using avian distribution data. *Landscape Ecol.* 7: 137–147.

15. O'Neill, R.V., C.T. Hunsaker, K.B. Jones, K.H. Riitters, J.D. Wickham, P.M. Schwartz, I.A. Goodman, B.L. Jackson, and W.S. Baillargeon. 1997. Monitoring environmental quality at the landscape scale. *BioScience* 47: 513–519.
16. Jones, K.B., K.H. Riitters, J.D. Wickham, R.D. Tankersley, Jr., R.V. O'Neill, D.J. Chaloud, E.R. Smith, and A.C. Neale. 1997. An Ecological Assessment of the United States Mid-Atlantic Region. USEPA, Office of Research and Development, Washington, D.C., EPA/600/R-97/130.
17. Lunetta, R.S., R. Araujo, S. Bird, L.A. Burns, R.O. Bullock, D.E. Carpenter, R. Carousel, K. Endres, B.H. Hill, K.B. Jones, D. Luecken, and R. Zepp. 1999. Mid-Atlantic Stressor Profile Atlas. URL: www.epa.gov/eimsreva.
18. Keitt, T.H., D.L. Urban, and B. Milne. 1997. Detecting critical scales in fragmented landscapes. *Conserv. Ecol.* 1: 4. URL: www.consecol.org.
19. Barabasi, A. and A. Reka. 1999. Emergence of scaling in random networks. *Science* 286: 509–512.
20. Sutherland, G.D., A.S. Harestad, K. Price, and K.P. Lertzman. 2000. Scaling of natal dispersal distances in terrestrial bird and mammals. *Conserv. Ecol.* 4: 16. URL: www.consecol.org.
21. Rastetter, E.B., A.W. King, B.J. Cosby, G.M. Hornberger, R.V. O'Neill, and J.E. Hobbie. 1991. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecol. Appl.* 2: 55–70.
22. Wickham, J.D., R.V. O'Neill, K.H. Riitters, T.G. Wade, and K.B. Jones. 1997. Sensitivity of selected landscape pattern metrics to land-cover misclassification and differences in land-cover composition. *Photogramm. Eng. Remote Sens.* 63: 397–402.
23. Wickham, J.D., K.B. Jones, K.H. Riitters, R.V. O'Neill, R.D. Tankersley, E.R. Smith, A.C. Neale, and D.J. Chaloud. 1999. An integrated environmental assessment of the Mid-Atlantic Region. *Environ. Manage.* 24: 553–560.
24. Wade, T.G., J.D. Wickham, and W.G. Kepner. 1995. Using GIS and a graphical user interface to model land degradation. *Geo. Info. Syst.* 5: 38–42.
25. Johnson, A.R. 1988. Diagnostic variables as predictors of ecological risk. *Environ. Manage.* 12: 515–523.
26. Zimmerman, H.J. 1987. *Fuzzy Sets, Decision Making, and Expert Systems*. Kluwer Academic Publishers, Boston, MA.
27. Tran, L. and L. Duckstein. 2002. Comparison of fuzzy numbers using a fuzzy distance measure. *Fuzzy Sets and Systems*, 130(3): 331–341.
28. Zadeh, L.A. 1965. Fuzzy sets. *Inform. Control* 8: 338–353.
29. Hipel, K.W. (Ed.). 1992. Multiple Objective Decision Making in Water Resources. American Water Resources Association, Monograph 18, Herndon, VA.
30. Rauscher, H.M. and R. Hacker. 1989. Overview of artificial intelligence applications in natural resource management. *J. Knowledge Eng.* 2: 30–42.
31. Osleeb, J.P. and S. Kahn. 1999. Integration of geographic information. In *Tools to Aid Environmental Decision Making*, V.H. Dale and M.R. English, Eds. Springer-Verlag, New York, pp. 161–189.
32. Campbell, P.R. 1997. Population Projections: States, 1995–2025. U.S. Bureau of the Census, Population Division, Washington, D.C., pp. 25–113.
33. Knox, P. (Ed.). 1993. *The Restless Urban Landscape*. Prentice-Hall, Englewood Cliffs, NJ.
34. Watson, R.T., M.C. Zinyowera, R.H. Moss, and D.J. Dokken. 1996. *Climate Change 1995*. Cambridge University Press, Cambridge, U.K.
35. Wickham, J.D., R.V. O'Neill, and K.B. Jones. 2000. A geography of ecosystem vulnerability. *Landscape Ecol.* 15(6): 495–504.

Index

B

D

E

F

G

H

I

J

M

N

P

Q

R

S

T

V

W

X

Y

Z

Environmental Science and Technology

Environmental Monitoring

The current rate and scale of environmental change around the world make the detection and understanding of these changes increasingly urgent. Subsequently, government legislation is focusing on measurable results of environmental programs, requiring researchers to employ effective and efficient methods for acquiring high-quality data.

Environmental Monitoring is the first book to bring together the conceptual basis behind all monitoring activities with specific approaches to the monitoring of air, water, and land. Coverage includes integrated monitoring at the landscape level, as well as case studies of existing monitoring programs. The book focuses on pollution issues and impacts resulting from human activities and includes results from monitoring efforts carried out in various countries on several continents. The multinational makeup of the contributors and the cases presented gives the book worldwide relevance.

Academics in all fields relating to environmental monitoring, as well as governmental agency employees and supporting contractor agencies, will find this book a comprehensive and useful reference.

Features

- Addresses the recent legislative focus on results, which can only be measured with high-quality data
- Includes results from environmental monitoring efforts around the world
- Combines the concepts of monitoring activities with specifics on conducting monitoring programs in different ecosystems and environmental media

L1641

ISBN 1-56670-641-6

90000

9 781566 706414

CRC PRESS

www.crcpress.com