monitoring PLANT *and* ANIMAL populations

Caryl L. Elzinga
Daniel W. Salzer
John W. Willoughby
James P. Gibbs

b
Blackwell Science

monitoring PLANT *and* ANIMAL populations

Caryl L. Elzinga
Owner and Principal Scientist
Alderspring Ecological Consulting
Tendoy, Idaho

Daniel W. Salzer
Coordinator of Research and Monitoring
The Nature Conservancy of Oregon
Portland, Oregon

John W. Willoughby
California State Botanist
U.S. Department of the Interior
Bureau of Land Management
Adjunct Faculty, California State University
Sacramento, California

James P. Gibbs
Assistant Professor
Department of Environmental and Forest Biology
College of Environmental Science and Forestry
State University of New York
Syracuse, New York

b
Blackwell Science

monitoring PLANT *and* ANIMAL populations

monitoring PLANT *and* ANIMAL populations

Caryl L. Elzinga
Owner and Principal Scientist
Alderspring Ecological Consulting
Tendoy, Idaho

Daniel W. Salzer
Coordinator of Research and Monitoring
The Nature Conservancy of Oregon
Portland, Oregon

John W. Willoughby
California State Botanist
U.S. Department of the Interior
Bureau of Land Management
Adjunct Faculty, California State University
Sacramento, California

James P. Gibbs
Assistant Professor
Department of Environmental and Forest Biology
College of Environmental Science and Forestry
State University of New York
Syracuse, New York

b

Blackwell Science

©2001 by Blackwell Science, Inc.

Editorial Offices:
Commerce Place, 350 Main Street, Malden, Massachusetts 02148, USA
Osney Mead, Oxford OX2 0EL, England
25 John Street, London WC1N 2BL, England
23 Ainslie Place, Edinburgh EH3 6AJ, Scotland
54 University Street, Carlton, Victoria 3053, Australia

Other Editorial Offices:
Blackwell Wissenschafts-Verlag GmbH, Kurfürstendamm 57, 10707 Berlin, Germany
Blackwell Science KK, MG Kodenmacho Building, 7-10 Kodenmacho Nihombashi, Chuo-ku, Tokyo 104, Japan

Distributors:

USA

Blackwell Science, Inc.
Commerce Place
350 Main Street
Malden, Massachusetts 02148
(Telephone orders: 800-215-1000 or
781-388-8250; fax orders: 781-388-8270)

Canada

Login Brothers Book Company
324 Saulteaux Crescent
Winnipeg, Manitoba R3J 3T2
(Telephone orders: 204-837-2987)

Australia

Blackwell Science Pty, Ltd.
54 University Street
Carlton, Victoria 3053
(Telephone orders: 03-9347-0300;
fax orders: 03-9349-3016)

Outside North America and Australia

Blackwell Science, Ltd.
c/o Marston Book Services, Ltd.
P.O. Box 269
Abingdon, Oxon OX14 4YN, England
(Telephone orders: 44-01235-465500;
fax orders: 44-01235-465555)

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without permission in writing from the publisher, except by a reviewer who may quote brief passages in a review.

Acquisitions: Nancy Whilton
Development: Jill Connor
Production: Irene Herlihy
Manufacturing: Lisa Flanagan
Marketing Manager: Carla Daves
Production Service: Andover Publishing Services
Cover and interior design by Eve Siegel
Typeset by Pine Tree Composition

00 01 02 03 5 4 3 2 1

The Blackwell Science logo is a trade mark of Blackwell Science Ltd., registered at the United Kingdom Trade Marks Registry

Library of Congress Cataloging-in-Publication Data

Monitoring plant and animal populations / by Caryl L. Elzinga . . . [et al.].
p. cm.
Includes bibliographical references (p.).
ISBN 0-632-04442-X
1. Environmental monitoring. 2. Animal populations. 3. Plant populations. I. Elzinga, Caryl L.

QH541.15.M64 M72 2001
577.8'8'0287--dc21

ISBN: 9780632044429

00-050771

Table of Contents

Preface		vii
Chapter 1	Introduction to Monitoring	1
Chapter 2	Monitoring Overview	11
Chapter 3	Selecting Among Priorities	21
Chapter 4	Qualitative Techniques for Monitoring	37
Chapter 5	General Field Techniques	49
Chapter 6	Data Collection and Data Management	65
Chapter 7	Basic Principles of Sampling	75
Chapter 8	Sampling Design	101
Chapter 9	Statistical Analysis	149
Chapter 10	Analysis of Trends	185
Chapter 11	Selecting Random Samples	195
Chapter 12	Field Techniques for Measuring Vegetation	205
Chapter 13	Specialized Sampling Methods and Field Techniques for Animals	231
Chapter 14	Objectives	247
Chapter 15	Communication and Monitoring Plans	271
Appendix I	Monitoring Communities	283
Appendix II	Sample Size Equations	299
Appendix III	Confidence Interval Equations	319
Appendix IV	Sample Size and Confidence Intervals for Complex Sampling Designs	329
Literature Cited		339
Index		353

PREFACE

We have designed this book as a practical handbook for field biologists responsible for developing monitoring studies. Our intended audience includes students planning to work in natural resource management and current natural resource managers, such as wildlife biologists, range ecologists, botanists, forest ecologists, and preserve managers. While written especially for those dealing with monitoring plant and animal populations, many of the methods are also applicable to community monitoring, and an appendix describes strategies and adaptation of methods for monitoring plant and animal communities.

Monitoring in natural resource management is the tool used to determine whether a management effort is having the desired effect of meeting management objectives. Monitoring is a powerful warning tool for identifying potential crises while cost-effective solutions remain available. Monitoring data can also demonstrate the success or failure of a management strategy. This is especially important in natural resource management, where many management actions are essentially implemented on an experimental basis because our understanding of ecosystems, even simple ones at small scales, remains very incomplete. Monitoring is the means to record the results of these often unreplicated experiments, advancing our understanding of system function and response.

Nearly all people involved in natural resource management are also involved in one of the stages of monitoring: designing monitoring studies, implementing the study in the field, analyzing the results, and applying the results. In the year 2000, the three largest land management agencies in the United States spent a combined estimated $125 million on monitoring of plant and animal populations and communities.[1]

In spite of this effort, monitoring often fails to provide the information needed to evaluate the success of management. Inconclusive or ambiguous monitoring results are expensive in terms of the resources wasted on monitoring projects, the loss of potentially valuable information, and the potential costs of incorrect actions. Typical sources of failure include ambiguous management objectives, poor study design, low statistical precision or power to detect change, lack of commitment to implement monitoring plans, and failure to communicate the results of monitoring. Because of these problems, monitoring results are often not incorporated into the decision-making or policy process, and monitoring fails to achieve its ultimate purpose.

We have tried to address the pitfalls that commonly derail monitoring efforts, and provide the tools for designing effective and useful studies. These include development of measurable objectives, application of proper field techniques, use of sampling design tools, identification of correct analysis approaches, and completion of monitoring by reporting and using results. We have avoided standardized techniques, believing that the best monitoring is specifically and locally designed for a particular management and natural resource situation.

This book would not exist without the effort of many people. First our thanks to students of the monitoring courses offered by the Bureau of Land Management and The Nature Conservancy. Our efforts to answer the questions raised by these field practitioners is largely what you hold in your hands. We also wish to thank other instructors and collaborators from whom many

---

[1] *This estimate includes monitoring of all biological resources by the U.S. Forest Service, U.S. Bureau of Land Management, and the U.S. Park Service. The Forest Service also includes inventory activities as part of its monitoring budget. The methods in this handbook, while particularly addressing issues of monitoring plant and animal populations and communities, are also applicable to inventory and monitoring of any natural resource.*

of the ideas in this book have come: Jim Alegria, Lisa Croft, Phil Dittberner, Sam Droege, Doria Gordon, Cheryl McCaffrey, Scott Melvin, John Randall, Roger Rosentreter, Nathan Rudd, Howard Snell, Rob Sutter, Bob Unnasch, and Rick Young. Janine Koselak, visual information specialist for the U.S. Bureau of Land Management's National Applied Resource Sciences Center, formatted nearly all the illustrations used in this book as part of an earlier publication *Measuring and Monitoring Plant Populations* (Elzinga et al. 1998) and provided a smooth transition of computerized materials between publications.

The pages of this book are graced by artwork from several excellent biological illustrators. We thank D. Andrew Saunders for the animal illustrations and Jennifer Shoemaker and Glenn Elzinga for several plant illustrations. Permission was granted by Brigham Young University Herbarium to use Kaye Thorne's lovely drawing of the dwarf bear-claw poppy, and by the Nevada Bureau of Land Management to use several illustrations that Jeanne R. Janish had prepared for the manual *Threatened and Endangered Plants of Nevada* (Mozingo and Williams 1980). Our thanks also to the Roosevelt Wildlife Station at the State University of New York College of Environmental Sciences and Forestry for assistance in preparing these illustrations for publication.

Finally, but definitely not least, we thank our spouses and families who graciously tolerated the extra work hours we devoted to this book and who supported our efforts throughout the process.

Caryl Elzinga
Daniel Salzer
John Willoughby
James Gibbs

The material in this book is supported by a Web site, kindly arranged by Sam Droege of the U.S. Geological Survey, Biological Resources Division, Patuxent Wildlife Research Center:

URL: http://www.mp1-pwrc.usgs.gov/monpop

On this site you will find the following, arranged by chapter and appendix:

- New information and references
- Links to other useful sites, including those cited in the text
- Problems and answers that can be used as part of a self-study approach or as classroom work
- Errata

CHAPTER 1
Introduction to Monitoring

Thelypodium repandum
Wavy-leaf Thelypody
Endemic to Challis volcanic soils in east-central Idaho
Artist: Glenn A. Elzinga

The root of the word *monitoring* means "to warn," and one essential purpose of monitoring is to raise a warning flag that the current course of action is not working. Monitoring is a powerful tool for identifying problems in the early stages, before they become dramatically obvious crises and while cost-effective solutions remain available. For example, an invasive species that threatens a rare plant or animal population is much easier to control at the initial stages of invasion, compared with eradicating it once it is well established at a site. Monitoring is also critical for measuring management success. Effective monitoring can demonstrate that the current management approach is working and can provide evidence supporting the continuation of current management.

For monitoring to function as a warning system or a measure of success, we must understand what monitoring is, as well as the close relationship between monitoring and improved natural resource management decision making.

DEFINITION OF MONITORING

In this handbook we define monitoring as the collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective. Good objectives are critical to successful monitoring. What is measured, how well it is measured, and how often it is measured are design features that are defined by how an objective is articulated. The objective describes the desired condition. Management is designed to meet the objective. Monitoring is designed to determine if the objective is met. Management is changed if monitoring reveals a failure to meet the objective. Objectives form the foundation of the entire monitoring project.

Monitoring is only initiated if opportunities for change in management exist. If no alternative management options are available, expending resources to measure a trend in a species population is futile. What can you do if a population is declining other than document its demise? Because monitoring resources are limited, they should be directed toward species for which management solutions are available. Fortunately, for most species there are management options available, although some may be politically difficult or very expensive to implement.

The management framework in which monitoring functions has been termed "adaptive management." In this framework, monitoring measures progress toward or success at meeting an objective and provides the evidence for management change or continuation (Holling 1978; Ringold et al. 1996). Because the term "adaptive management" has been adopted as a buzzword, its definition and meaning have become muddied by widespread use. In this handbook we define adaptive management as a process in which management activities are implemented in spite of uncertainty about their effects, the effects of management are measured and evaluated, and the results are applied to future decisions (Nyberg 1998).

Adaptive management is "learning by doing" (Lee 1999). It is a way of thinking about and implementing natural resource management that recognizes our understanding of ecosystems (even simple ones at small scales) is very incomplete and that any management we impose on the system is essentially an experiment (Gunderson 1999; Walters and Green 1997). There are three goals of adaptive management: 1) manage currently to the best of our knowledge, 2) learn from management, and 3) improve management in the future. In adaptive management, learning is as important as doing — monitoring is as important as management.

The adaptive management cycle is illustrated in Figures 1.1 and 1.2. In the first figure, monitoring successfully completes the adaptive management cycle. In the second figure, because the monitoring data are inconclusive, the cycle is incomplete and the management response is unknown. The monitoring effort is ineffective.

A successful adaptive management cycle involves the following steps:

1. A model of the system or species is developed. Models range from complex computer models designed to describe a complex system to simple doodling on a sheet of paper.

All are simply tools to help summarize, think about, and communicate with others concerning the system. We use only very simple models in this handbook, but encourage their use to help you think about your system and problem.

2. An objective is developed to describe the desired condition. We stress simplification and careful selection of one or a few objectives for each problem or species.
3. Management is designed and implemented to meet the objective.[1]
4. The resource is monitored. Monitoring techniques selected depend on the objectives. For example, if the objective is to increase cover of *Primula alcalina*, the technique selected would be one of those available to measure cover, not density or population size.
5. Monitoring data are analyzed to determine if objectives are reached. These results are summarized in a form accessible to decision-makers and stakeholders.
6. Management is adapted (changed) if objectives are not reached. We recommend identifying the proposed alternative management before monitoring is initiated, so that all parties understand how the monitoring data will be used to adapt management. If monitoring data provide new insights into the species or problem, the model is improved and new objectives are developed.

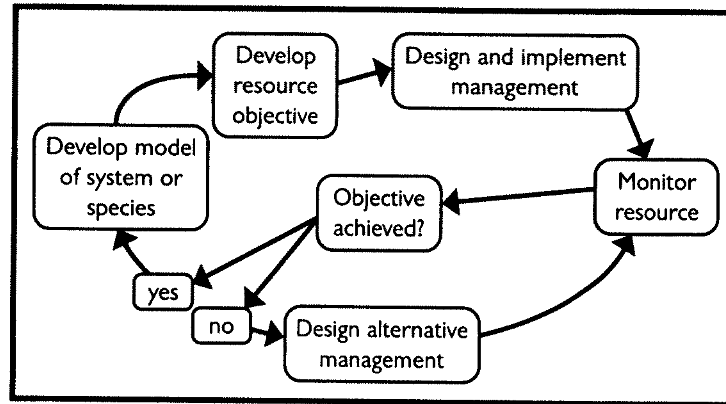


Figure 1.1. Diagram of a successful adaptive management cycle. Note that monitoring provides the critical link between objectives and adaptive (alternative) management.

Adaptive management has been promoted as a valuable tool for addressing large-scale, complex, natural resource management issues. Many authors further define adaptive management in terms of experimentation, in which the design of the management and monitoring incorporate a statistical design (research design) that allows changes to be attributed to (shown to be caused by) management (Lee 1993; Lee 1999; Nyberg 1998; Walters and Holling 1990). In this handbook, we focus on the simplest type of adaptive management: observational studies of single variables. Some authors classify this approach as the "monitor and modify" method, rather than true adaptive management (Johnson 1999). Rather than quibbling over terms, we contend that observational monitoring can be most effectively applied in an adaptive management framework. We also suggest that, realistically, given the technical skills available to most natural resource managers and the local scale of many management actions, observational studies will be the most common type of monitoring used.

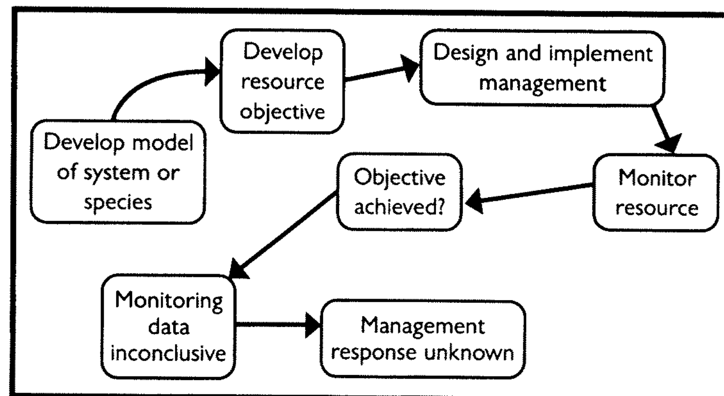


Figure 1.2. Diagram of monitoring that fails to close the adaptive management cycle. Because monitoring data are inconclusive, the management response is unknown and the cycle is unsuccessful.

[1] *Most descriptions of adaptive management recommend that management be designed not only to meet a resource objective, but also to learn more about the system, i.e., design management as an experiment. In this handbook, we focus on observational studies rather than experiments.*

What is the difference between an observational study and research? Both are information-gathering activities, and the field techniques used may be quite similar; the difference really is more one of degree than kind. Because of this, confusion exists about the difference between an observational study (especially one that applies sampling design and statistical analysis) and research.

Observational monitoring and research are ends of a continuum (Fig. 1.3). The confidence of attributing a change to a particular cause increases along the continuum, but the cost of acquiring the needed data also increases. Statistical significance is often erroneously equated with cause. In Figure 1.3, statistically significant differences were found in several scenarios, but only

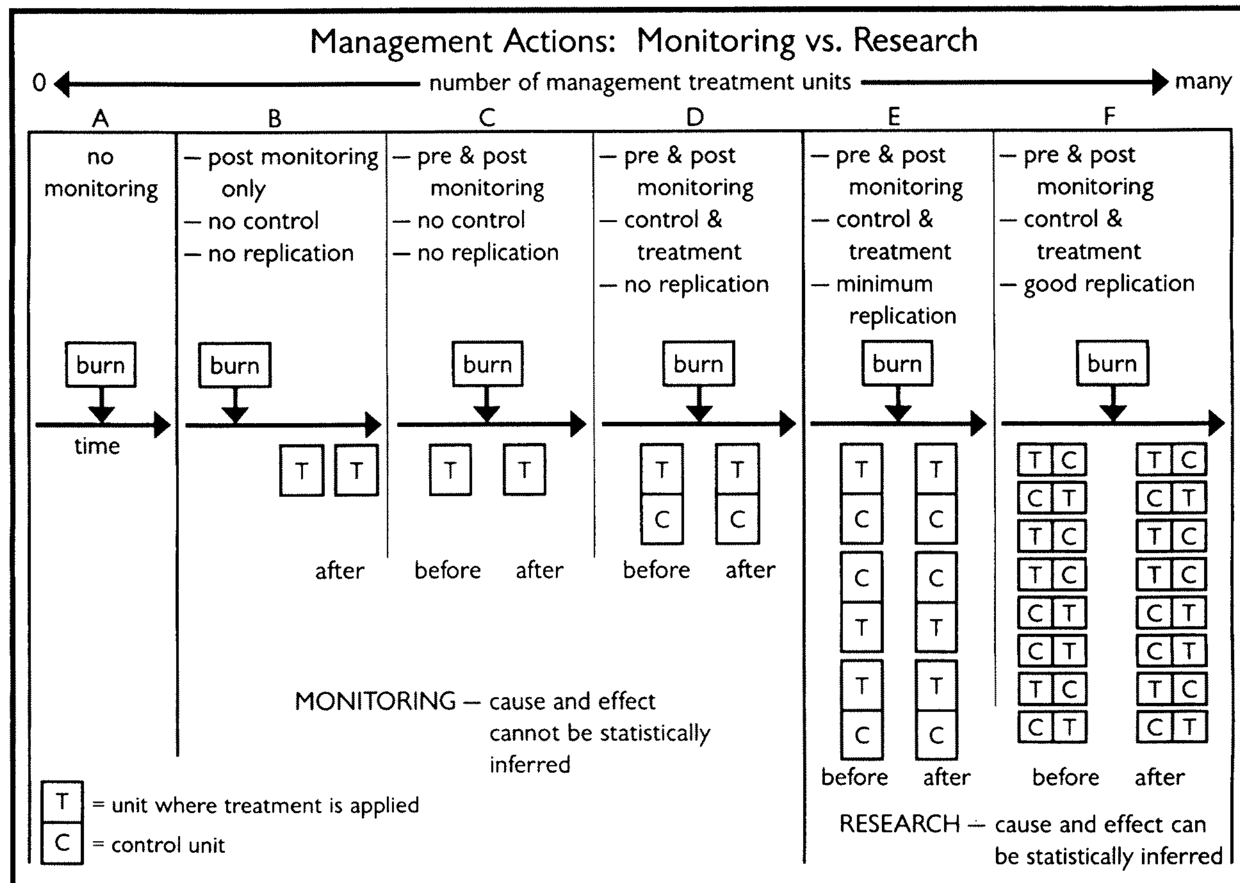


Figure 1.3. A comparison of monitoring and research approaches for detecting a treatment effect from a prescribed burn. For each of the scenarios shown in columns B-F above, statistical comparisons can be made between different time periods and a decision can be made as to whether or not a statistically significant difference occurred. However, the interpretation of that difference can be confounded by factors that are independent of the treatment itself. The diagram and the following examples illustrate a continuum of increasing confidence in determining likely causation as you move from left to right in the diagram. In column B, there is no pre-treatment measurement but you may see differences between years one and two after the burn. There is no way of knowing the conditions prior to treatment, and changes may be due to the burn, or they may be the result of some other factor such as lower precipitation. In column C, where data was gathered both before and after the burn, you still don't know if changes were due to the burn or some other factor that differed between the two time periods. In column D, there is a single treatment unit and a single control unit. Perhaps you see a change occur in the burned area but not the control area. The change could be caused by the burn or there may be some other factor that differentially affects the treatment area compared to the control. The burn unit, for example, could have a slightly lower water table than the control unit, a factor independent of the burn but not apparent. Other factors such as disease, insect infestation, and herbivory often occur in concentrations, affecting one area but not adjacent areas. Any of these factors could be the cause of observed differences. In the last two columns, the treatment and control are replicated in space; thus there is a possibility of attributing differences to the treatment. Since ecological systems are variable, the example in column E with three replicates may have inadequate statistical power to detect differences. The differences due to the treatment may be hidden by differences that occur due to other factors. The larger number of replicates in column F greatly increases the likelihood of detecting treatment differences due to the higher statistical power associated with 8 replicates as compared to 3 replicates.

*the term "significant" means that a statistical test was carried out and the difference was significant according to the test.

for the last two scenarios (Columns E and F) can you attribute significant differences to a cause. Observational monitoring data are usually of limited value in determining causes of change, and you must be careful to not misrepresent monitoring data as information on cause and effect. For example, if one were to simply note a decline in a species population after logging, this would support the hypothesis that logging negatively impacts the species, but it does not prove that logging is the cause of the decline. The decline has to be consistently found at several logging sites and not found in uncut areas to confidently determine logging activities as the cause of the decline. Only with several replications of treatment and control can you confidently attribute changes to a treatment or cause. Your design must incorporate control (minimize the differences between the treatment and nontreatment areas except for the treatment itself) and replication (measure the difference between treatment and nontreatment consistently over several-to-many independent units).

Natural resource managers must decide during the development of a project whether proving causal relationships is important. If demonstrating causality is required, the cost of obtaining that information must be evaluated. In many cases of resource management, a research approach may not be feasible. Some typical problems with incorporating a research design are the complexity of the system, the nonlinear response of organisms to causal mechanisms, and the lack of available replicates because only one "treatment" area is available (Thomas et al. 1981).

Box 1.1. *COMMON FAILURES IN MONITORING PROGRAMS*

TECHNICAL PROBLEMS

1. *Poor design leads to inconclusive results.*
2. *The use of multiple observers or unreliable data collectors complicates interpretation of results.*
3. *Data are lost, either physically (because of poor documentation or storage) or institutionally (because no one can decipher the data sheets).*
4. *Data are not analyzed because the biologist lacks the skills or the will to do so.*
5. *Natural system fluctuation obscures change caused by management.*

INSTITUTIONAL PROBLEMS

1. *Lack of institutional support results in premature termination of monitoring because of the loss of the project advocate (usually the biologist), budget decreases, changes in priorities, or increased politicization of the situation.*
2. *Lack of institutional support limits resources needed to implement monitoring as planned. Often data are collected, but inadequate resources are allocated to analyze, interpret, and communicate results.*
3. *Managers refuse to use monitoring data to make decisions because internal (other staff specialists) or external (stakeholders) antagonists question the data.*
4. *Failure to place monitoring within a management framework encourages the perception that the data are interesting but not directly applicable to a management decision.*

This handbook is not designed as a guide for developing research projects. Good design is essential to the success of a research project and often requires specialized skills. We suggest you consult with a statistician who is well versed in sampling design, especially if the treatments are expensive (such as eradication of a predator or prescribed burns). Underwood (1997), Hairston (1989), and Manly (1992) all provide an excellent introduction to good, effective research designs in ecology.

WHY DOES MONITORING OFTEN FAIL?

Biologists and botanists are often frustrated because their monitoring efforts fail to result in management changes or improvement in the resources they are concerned about. Box 1.1 lists common reasons why monitoring projects fail. Generally, these can be classified as technical failures (poor data collection methods, loss of data) and social and institutional failures (ending the monitoring program prematurely because of new crises or lack of support). All of these are linked to a fundamental problem: the failure to place monitoring within the context of a management framework. Successful monitoring is much more than counting plants and animals in plots over time.

MONITORING SPECIES, HABITATS, AND THREATS

We can monitor a species directly through counts or measures of performance, such as cover or reproductive output, or we can monitor some indicator of species success, such as habitat indicators, another species whose success is related to the success of the species of interest, threats to the species, and, for animals, indirect indices of abundance. Often monitoring a species directly is quite difficult and indicators may be much more easily measured (MacDonald and Smart 1993). For example, annual plant populations are difficult to monitor directly because so much of the population exists in a cryptic form as seeds stored in the soil. Similarly, many animals are secretive and difficult to observe directly. An indicator may be the only reasonable way to monitor such species. For some species such as very short-lived species with fluctuating populations or very long-lived populations that exhibit change slowly, monitoring an indicator may be more sensitive to detecting undesirable change than monitoring the species directly.

When management efforts designed to conserve a species are focused primarily on improving habitat conditions for the species, monitoring habitat indicators may be a more direct measure of management success than monitoring the population. For example, we may know that successful reproduction of woodcock (*Scolopax minor*), a small earthworm-eating game bird, requires (among other things) meadow openings in forested areas as courtship grounds. We may also know that fire suppression has resulted in a reduction in size and quality of these meadows on a preserve that we manage. We plan to introduce prescribed burning into this preserve to improve these meadows for use by the woodcock. A reasonable monitoring strategy is measuring the change in structure of the meadows (size, number, quality), with perhaps some corollary qualitative observations on use of the meadows by the bird, rather than trying to monitor the response of the woodcock population itself to our management action. Because population size may be affected by a number of factors, monitoring the woodcock population is not only more difficult and expensive than monitoring the meadows, but the data probably will not provide much feedback on whether our prescribed burning activity was beneficial.[2]

---

[2]*Monitoring that incorporates a research design may allow you to determine if prescribed burning actually did benefit the population, but this approach may not be affordable. In many management situations, a research design may not even be possible. It is unlikely in this example that you could isolate the effects of burning on the population within the habitat area because of the movement of the birds and the lack of a large enough area to implement independent treatment and control sites.*

Monitoring the change in extent or intensity of a threat may also be a more direct measure of the effectiveness of management than measuring the species itself (Salafsky and Margoluis 1999). For example, you may know that a riparian plant species disseminates nondormant seed in the fall (does not store seed in the soil) and that herbivory by cattle of the seed stalks can reach up to 80% within a particular population. Changing the grazing and monitoring the success of the grazing management by monitoring stubble height of riparian vegetation may be less expensive but just as effective as monitoring herbivory on seedstalks or measuring overall population change. Salafsky and Margoluis (1999) argue that monitoring changes in threats can often be done effectively through a qualitative rating assessment completed by those involved in the management project without special data collection activities.

Two final benefits of monitoring of habitat and threats are that data are often more immediately forthcoming and may be gathered using techniques already familiar to many resource managers. In the woodcock example above, population response may lag several years behind the management action. Assessing changes in threats, especially when using a qualitative approach like that of Salafsky and Margoluis (1999), can often be done on an annual basis (or even more often), whereas measuring the response of habitats or populations to a change in threat may take several years. Monitoring populations using habitat or threat indicators often relies on methods that are relatively familiar such as vegetation measurements, while monitoring the populations themselves may require more elaborate and unfamiliar techniques such as mark/recapture methods for animals or demographic monitoring for plants. People with the specialized skills needed for these techniques may not be available.

In exchange for ease, low cost, and immediacy, one must accept limitations and risks. A serious criticism of monitoring habitats or threats as indicators of species condition is that your selected indicator may not really be indicative of changes in the species population. Habitat monitoring is most effective when research has demonstrated a relationship between a habitat parameter and the condition of a species, but for most plant and animal populations, these data are lacking, and the relationship between a habitat parameter and a species must be inferred from hypothesized and known ecological relationships. You may have chosen an indicator that is not well correlated with species success or one that is correlated in an unexpected way. Some indicators, for example, are fairly robustly correlated with population changes at low densities, but become less sensitive as the population reaches higher densities.

Because monitoring an indicator may give a false sense of security, one must examine these trade-offs explicitly and design the monitoring project based on available monitoring resources and with an awareness of the risks of being wrong. For these reasons, when threats, habitat attributes, indicator species, or abiotic variables are used as surrogates for tracking individual populations, it is advisable to periodically assess the population itself to ensure the validity of the surrogate relationship.

These risks should not deter use of indicators, however. The risk of being wrong exists even when monitoring populations directly. As discussed earlier, without a research design, we do not know if changes in a population are the result of management or if they result from changes in weather patterns, insect infestations, rates of herbivory or predation, prey base or nutrient levels, incidence of disease, or other factors. All monitoring data should be interpreted with caution, recognizing sources of uncertainty. Using indicators, however, introduces the additional source of uncertainty in the assumed relationships between the indicator and the species. This risk must be assessed against the benefits of using indicators for monitoring.

MONITORING COMMUNITIES

Habitat monitoring is closely allied with monitoring of communities, and those who are interested in monitoring communities will benefit from many of the concepts described in this handbook. Appendix 1 provides additional information specific to community monitoring.

RELATED ACTIVITIES

The term "monitoring" has been applied to a variety of data-gathering activities. We have defined monitoring in this handbook as driven by objectives and implemented within a management context. This differs from many activities described below that are often implemented under the general term "monitoring." While we know that many of these activities will benefit from applying the technical concepts described in later chapters, throughout this handbook we will maintain our narrower definition of monitoring as objective-based.

Inventories are point-in-time measurements typically used to determine location or condition. Inventories may be designed for the following purposes:

- Locate populations of a species.
- Determine the total number of individuals of a species.
- Assess ages, sizes, and conditions of individuals within a population.
- Locate all populations of a species within a specific area (often a project area).
- Locate all species (or species of a certain type such as listed species) occurring within a specific area (project area, habitat type).
- Assess and describe the habitat of a species (e.g., associated species, soils, aspect, elevation).
- Assess existing and potential threats to a population.

Data collected during an inventory may be similar to those collected during monitoring. For example, the number of individuals in each population may be counted during an inventory. Similarly, a monitoring project may require counts of a single or several populations every year for several years.

Information collected during an inventory is most useful in developing models of the species of interest and in generating reasonable objectives. Inventory data may provide a baseline, or the first measurement, for a monitoring study, but do not assume that information collected during inventory will always be directly useful for monitoring the specific objective you develop. The following is a typical example:

During inventory for the rare mustard *Physaria didymocarpa* var. *lyrata*, qualitative estimates of population size and various habitat parameters were noted. Exhaustive inventory, however, only identified four populations. All are over 500 individuals, but all are restricted to small areas of extremely steep scree slopes. Management conflicts are severe at all four sites. Because of the demonstrated rarity of the species and conflicting uses, public and agency concern over management is intense. Qualitative estimates of population size are considered inadequate for monitoring this species, and quantitative objectives and monitoring are recommended.

Natural history studies investigate basic ecological questions. For animals, these questions may concern food habits; breeding, resting, and foraging sites; timing of reproduction; mortality and reproductive rates; home range; dispersal and migration behavior; predators; and disease. For plants, these questions concern pollination ecology, life history, seed viability, seed-bank longevity, herbivory, and seed predation. These questions often must be answered before effective monitoring can be designed, but such studies are not monitoring.

Implementation monitoring assesses whether the activities are carried out as designed. For example: Was the fence built in the right location according to specifications so it will effectively protect the plant population from deer? Was the off-highway vehicle (OHV) closure maintained? Were the cows moved on the right date to allow the rare plant to successfully produce seed? While such monitoring does not measure a population, it does provide critical feedback on whether the planned management is being implemented. Implementation monitoring can also identify which variables are most likely to be causing a change in the resource and thus will help eliminate from consideration some hypothetical causes of change. This type of monitoring, although critical to successful management, is not discussed further in this handbook.

Measuring change over time is a main characteristic of monitoring, but simply measuring change does not meet the definition of monitoring in this handbook. Studies that measure change can be implemented in the absence of an identified need for decision making. In contrast, monitoring is characterized primarily by objectives and by being part of an adaptive management cycle in which monitoring data are used to evaluate management and make decisions (Perry et al. 1987).

Studies measuring change in the absence of a management context have been collectively termed "surveillance" (Perry et al. 1987), but three types are recognized and described here: *trend studies, baseline studies, and long-term ecological studies.* The distinction among the types is blurred, and resource managers have frequently used the terms interchangeably.

Trend studies are designed to learn how the resource is changing over time (some authors call this "baseline monitoring" — see next section). An example of a study objective for measuring trend is as follows:

Study objective: Determine if the density of *Primula alcalina* is increasing, decreasing, or remaining stable at the Texas Creek Population over the next 5 years.

While this is important information, the trend study could be placed into a management framework by developing this study objective into a management objective:

> *Management objective*: Allow a decrease of no more than 20% from the current density of *Primula alcalina* at the Texas Creek population over the next 5 years.
>
> *Management response* (if cause is unknown): If *Primula* density declines by more than 20% over the 5-year period, more intensive monitoring or research will be initiated to determine the cause of the decline; or
>
> *Management response* (if cause is suspected): If *Primula* density declines by more than 20% over the next 5 years, grazing at the site will be limited to late fall to allow seed set and dissemination.

A subtle but fundamental difference exists between monitoring for trend and monitoring for management, even though the actual measurements and analysis may be the same. The second approach places the measurements within the adaptive management cycle and identifies the changes in management that will occur if the monitoring has a certain result. At the time the study begins, we do not know whether the population is stable, declining, or increasing. By conducting the study within the framework of an objective and a management response, the course of action at the end of 5 years is known before monitoring begins. If monitoring shows the population is increasing or stable, current management may continue. If populations are declining, an alternative management approach is outlined. If the study is done simply to detect change, the course of action at the end of the 5 years will be unclear. What will likely occur is continuation of existing management and the trend study to determine if the decline of the rare plant continues.

Biologists and ecologists are often hesitant to develop objectives and management responses because of a lack of information on the desired condition of the population and the relationship of management to that condition. At a minimum, however, an objective to maintain the current condition can be established and a commitment made to respond with more extensive monitoring, study, or research if a decline is measured.

Baseline monitoring is another type of activity implemented as monitoring. This is the assessment of existing conditions to provide a standard, or "baseline," against which future change is measured. Commonly, many variables are measured in hopes of capturing within the baseline dataset the ones that turn out to be important later. Baseline monitoring is sometimes termed "inventory monitoring" (MacDonald et al. 1991) because it often involves the collection of data to describe the current condition of a resource. Measurement again at a later date may be intended, but a commitment or plan for periodic measurement is usually lacking. Periodic measurement is integral to a monitoring study. The problem with baseline studies, and using inventory data as baseline data, is that the design of the study may be inadequate to detect changes. This inadequacy usually results from including too many variables and using too small of a sample size.

If the study is implemented with scheduled periodic measurements, a baseline study may be termed a *long-term ecological study*. The most common goal of these studies is to learn about the natural range of temporal variability of the resource by documenting the rates and types of changes that occur in response to natural processes such as succession and disturbance. The term "long-term ecological monitoring" usually is used to describe the measurement of community variables to determine change over the long term, 50 to 200 years or more. In most studies, many variables are measured on a few, large, permanent plots (usually greater than 0.1 hectare). Commonly measured variables include cover or density of all plant species, demographic parameters of important species, soil surface conditions, fuel loads, and animal signs (Greene 1984, 1993; Dennis 1993; Jensen et al. 1994; Schreuder and Geissler 1999; Scott et al. 1999; Stohlgren 1999). To add confusion to our classification, the term "baseline monitoring" is also sometimes used for this activity.

Two key differences exist between baseline and long-term ecological studies and the monitoring described in this handbook:

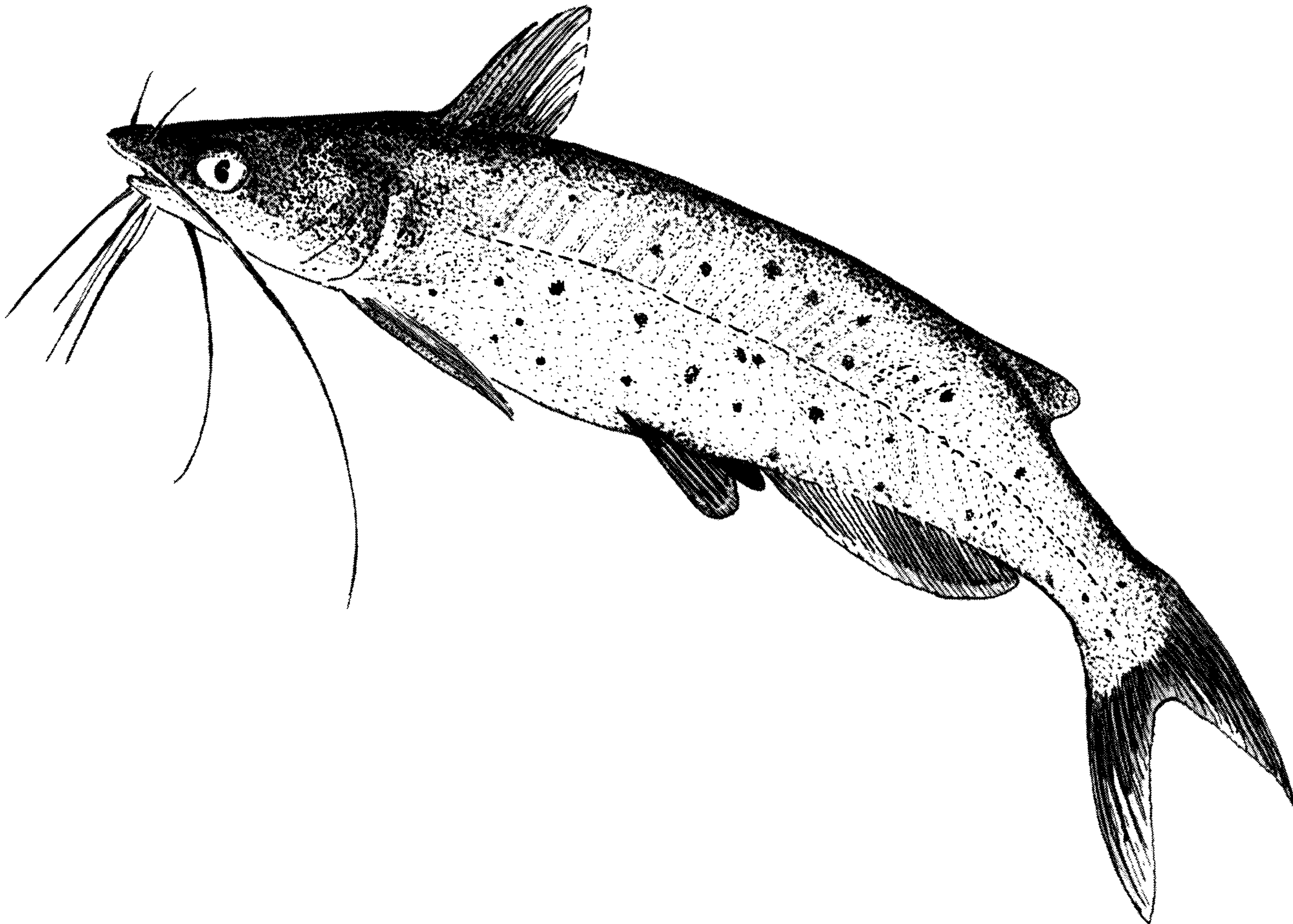
1. Baseline and long-term ecological studies do not specifically evaluate current management nor result in a management decision, although they may provide important information for management direction in the future by describing system functions and fluctuations (Perry et al. 1987). In monitoring, the application of the data to management is identified before the measurements are taken because monitoring is part of the adaptive management cycle.
2. These studies often attempt to maximize the number of characteristics and species measured because those most sensitive for measuring change are not known. In contrast, in this handbook, we advocate the explicit selection of one or a few measurable variables to be monitored.

One type of monitoring explicitly involves the measurement of a "baseline" and is sometimes termed "baseline monitoring." In this monitoring design a series of measurements are taken prior to the initiation of a management activity and are used for comparison (a "baseline") with the series of measurements taken afterward (Green 1979; MacDonald et al. 1991). This type of situation is common in water-quality monitoring. For example, measurements of water column sediment in a river may be taken for 5 years prior to the construction of a power plant and then for 5 years afterward to determine the background, or baseline, level of sediment and to determine whether the pollution controls of the plant are adequate to prevent elevated sediment levels. When measurements are made at both treatment and control areas, this type of monitoring design is termed the before-after, control-impact (BACI) design (Bernstein and Zalinski 1983; Faith et al. 1995; Long et al. 1996; Schwarz 1998). It is unusual in resource management to have several years' notice before initiating an activity during which a baseline can be measured, but if the opportunity arose, such a monitoring design can be very effective.

MANAGEMENT IMPLICATIONS

Monitoring must be placed in the context of a management framework to effectively improve or validate management. This framework, called adaptive management, involves 1) developing a model of the system; 2) describing the desired condition of the resources with an objective; 3) monitoring the response of the resource; 4) analyzing monitoring data; and 5) adapting management based on the monitoring information. Objectives are the foundation of the monitoring process. These may be based on the species itself (e.g., population size, reproductive success) or on a change in a habitat attribute or a threat. This objective-based character of monitoring differentiates it from other data-gathering activities such as inventory and survey studies, natural history investigations, and long-term ecological studies.

CHAPTER 2
Monitoring Overview



Ictalurus punctatus
Channel catfish
Artist: D. Andrew Saunders

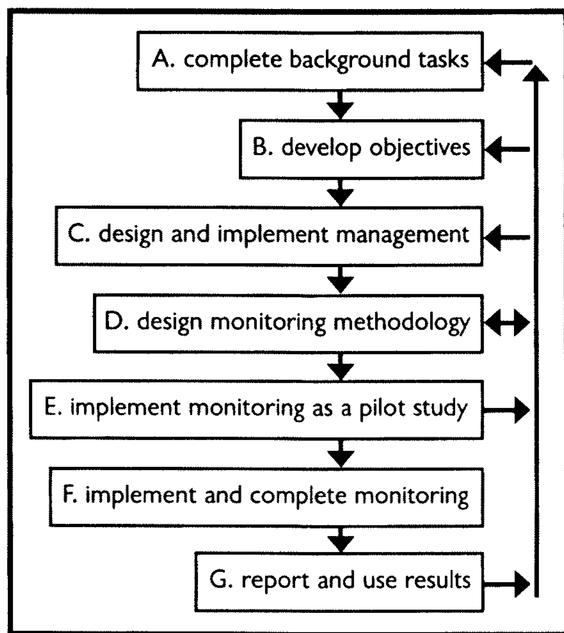


Figure 2.1. These seven major steps are broken into sub-steps and illustrated in Figures 2.2–2.5.

This chapter provides an overview of the development of objectives and monitoring methods and briefly addresses the development of management strategies. The steps described below and illustrated in the flow diagrams in Figures 2.1 through 2.5 provide an overview of the development of an adaptive management cycle (Fig. 1.1).

The major steps from completing background tasks to reporting and using results are shown in Figure 2.1. Each of these seven steps is broken down into its components and is described and illustrated. Steps are shown roughly in the order in which they occur in developing an adaptive management project, but recognize that feedback loops and reviews are many, as shown by the multidirectional arrows in the flow diagrams. At nearly any point in the process of developing a project, earlier decisions may have to be revisited and changes made.

COMPLETE BACKGROUND TASKS (FIG. 2.2)

Compile and Review Existing Information

Compile relevant information on the species and/or populations. For those monitoring projects where the target species and/or population are predetermined, you will only need the information specific to the species. For management programs that are just beginning, you will likely want to assemble the information needed to set priorities among all the target species occurring in your conservation/management area. If you manage many species, you may wish to start with a short list of species that are high priority, perhaps because of legal reasons such as nationally listed rare species and species of concern (see Chapter 3).

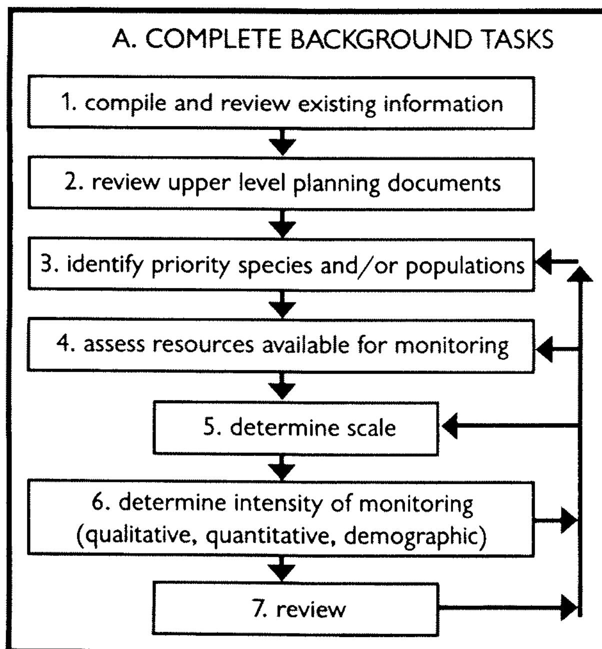


Figure 2.2. Flow diagram of the monitoring process, continued. Steps associated with completing background tasks are illustrated in detail.

Review Upper-Level Planning Documents

Consistent, local, land management depends on following upper-level planning documents, especially in management of State and Federal lands. These documents describe to the public the agency's planned activities. Because managers are accountable for implementing these plans, specific management activities for rare or target species should demonstrate progress toward meeting the goals and objectives described in them. Even if you believe your agency's land use plan provides little specific direction for management of a particular species (many of the older ones do not), you will increase support for your specific project if you can show a clear relationship between it and the general directives outlined in the planning documents (see Chapter 3). Stewards of private conservation areas may also wish to consider upper-level government planning documents to increase consistency of their conservation actions with those on adjacent or nearby Federal or State lands.

Identify Priority Species and/or Populations

Prioritize the species for monitoring, and document the process. This documentation will be immediately useful for review by the other parties that are involved in setting priorities, and it will also be useful to you and your successor if managers and other parties question the priority ranking at a later date. For priority species, select priority populations. These priorities may periodically require reassessment because of changes in threats, management, conflicts, and the interest of outside parties (see Chapter 3).

Assess the Resources Available for Monitoring

Resources for monitoring depend on management support, priorities, and the people and equipment available. Has management placed a priority on this monitoring project, or is support and funding limited? You may need to promote the importance of the project before you begin working on it. Are qualified personnel available to do the work? Do you have the necessary field equipment such as vehicles and measuring tapes? Is any high-tech equipment available (e.g., geographic information systems, global positioning systems, survey or forestry equipment)? Are people willing to give reviews and help sharpen your thinking? Do you have access to people with specialized skills? The types and amounts of resources will limit the extent and complexity of a monitoring project (see Chapter 3).

Determine Scale

Identify the scale of interest for monitoring (e.g., the range of the species, the populations within a certain watershed, populations in certain types of management units, a single population, a portion of a single population such as a key area or macroplot). Decide the scale of interest early in the monitoring process because it will influence later decisions and design. If, for example, the scale of interest is the species across its entire range, you will need to coordinate with various administrative units to develop a network of monitoring studies (see Chapter 3).

Determine Intensity of Monitoring

Will qualitative monitoring be adequate? Do you need quantitative data? Does the rarity of the species, the degree of threats, the uncertainty of management effects, or the political sensitivity of potential decisions warrant the use of an intensive monitoring approach? You may need to reevaluate the selected intensity of monitoring as you work through the remaining monitoring decisions (see Chapter 3).

Review

At this point, management should be briefed, and opinions and review should be solicited. For small projects, you could complete these steps on your own and then solicit internal and possibly external review. For larger programs or highly controversial species and populations, you may need to assemble a team (see Chapter 15).

DEVELOP OBJECTIVES (FIG. 2.3)

Develop an Ecological Model

In this handbook we promote the use of narrative or diagrammatic summaries (models) of the ecological and management interrelationships of the species of interest (Chapter 14 gives examples). Completing a model will help develop objectives, focus your monitoring, and improve interpretation and application of the data.

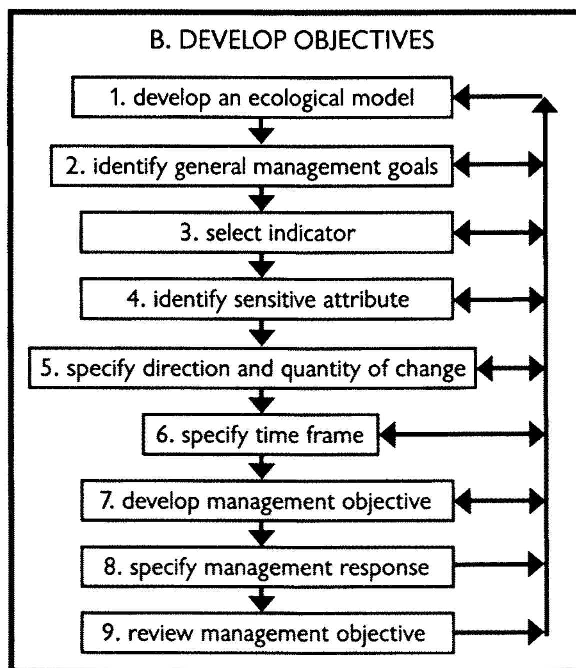


Figure 2.3. Flow diagram of the monitoring process, continued. Tasks associated with developing objectives are illustrated in detail.

Identify General Management Goals

Using your ecological model, try to refine conservation goals. Should the population size of the species be increased? Maintained? Recruitment increased? Mortality decreased? Describing these general management goals is the first step toward developing specific objectives.

Select Indicator

You may choose to monitor some aspect of the species itself or some indicator of species success. Effective indicators include other species, threats, habitat characteristics, or an indirect index of population abundance or success (such as animal track density). Other species that respond to management in a way similar to the target species but are easier to measure may be a cost-effective indicator. Monitoring threats can form an effective basis for management changes. Habitat indicators are especially useful for species that are difficult to measure or monitor directly (e.g., secretive or highly mobile animals, annual plants, long-lived species). For animals that are difficult to count, indirect indices of abundance may be the only way to assess population dynamics.

Identify Sensitive Attribute

Attributes may be measured values, such as population size, density, height, or age. Attributes also include qualitative and semiquantitative measures, such as presence or absence of the species, estimates of cover by cover class, and visual estimates of population size. Attributes for habitat characteristics or threat may also be measured variables or qualitative or semiquantitative. The attribute most sensitive and useful for monitoring depends on the management situation, the life history and morphology of the species, and the resources available to measure the attribute. Some species are so poorly known that you may have difficulty identifying a sensitive parameter. Make the best choice you can, or postpone monitoring until you know more about the natural history of the species.

Specify Direction and Quantity of Change

Will you monitor for a percentage change or an absolute change, a target value or a threshold value (see Chapter 14)? What amount of increase do you want to see, or what decrease will you tolerate? Can you specify a target population size? The quantity has to be measurable (confidently measuring a 1% change in average density is extremely difficult) and biologically meaningful (a 10% change in density of an annual plant species or some insect species is probably not important). Again, you may be limited by lack of information. You may also be limited by the amount of change you can detect in a sampling situation (see Chapters 7, 8, 9, and 10).

Specify Time Frame

How soon will management be implemented? How quickly do you expect the species to respond? How long do you want this monitoring program to continue if some threshold is not reached? The time frame should be biologically meaningful for the change you are anticipating. A 50% increase in the density of a long-lived woody plant, for example, is unlikely to occur over the next 3 years (although a decline of that magnitude may be possible and alarming).

Develop Management Objective

The priority species or population, the selected scale (location), the sensitive attribute, the quantity and direction of change, and the time frame of change are the critical components of the objective. Combine them into a simple, measurable, understandable objective (see Chapter 14).

Specify Management Response

Given the potential alternative results of monitoring, what management changes would be implemented in response to each alternative (see Chapter 14)? These management responses should be clarified before monitoring begins so all parties know the implications of monitoring results.

Review Management Objective

Preferably, a team of specialists and management would complete many or all of these steps, but sometimes the biologist will work alone through these steps. Before proceeding to the design of monitoring, solicit internal and external review, especially from parties that may be affected by management changes made in response to monitoring data (see Chapter 15). Do others have information about the biology or ecology of the species that you should incorporate into the model? Do all agree on the management objective? Do all agree with the proposed management response?

DESIGN AND IMPLEMENT MANAGEMENT

Depending on the situation, current management may be continued or new management proposed. Often current management is continued and monitored because little is known about the ecology and management requirements of a particular rare species. In some cases, however, previous monitoring data or natural history observations may suggest a need for management change. The ecological model may provide insight on needed changes as well. If new management is required, it must be completely described so it can be implemented effectively.

The design of conservation management strategies involves consideration of the ecology of the species, funding, management options, conflicting uses and activities, and communication and coordination with public and user groups. This complex and difficult step is unique to each situation and is a subject beyond the scope of this handbook.

DESIGN THE MONITORING METHODOLOGY (FIG. 2.4)

Qualitative Monitoring

Design General Methodology

Methods for qualitative monitoring include estimating quantity (e.g., ranked abundance, cover class) and quality (e.g., population stage class distribution, habitat condition) and using a permanent recording method such as a photopoint or a video sequence (see Chapter 4).

Design Methods to Reduce Variability Among Observers

The biggest drawback of using qualitative techniques is that estimates among observers can vary significantly. Between-observer variability can be reduced by several strategies described in Chapter 4.

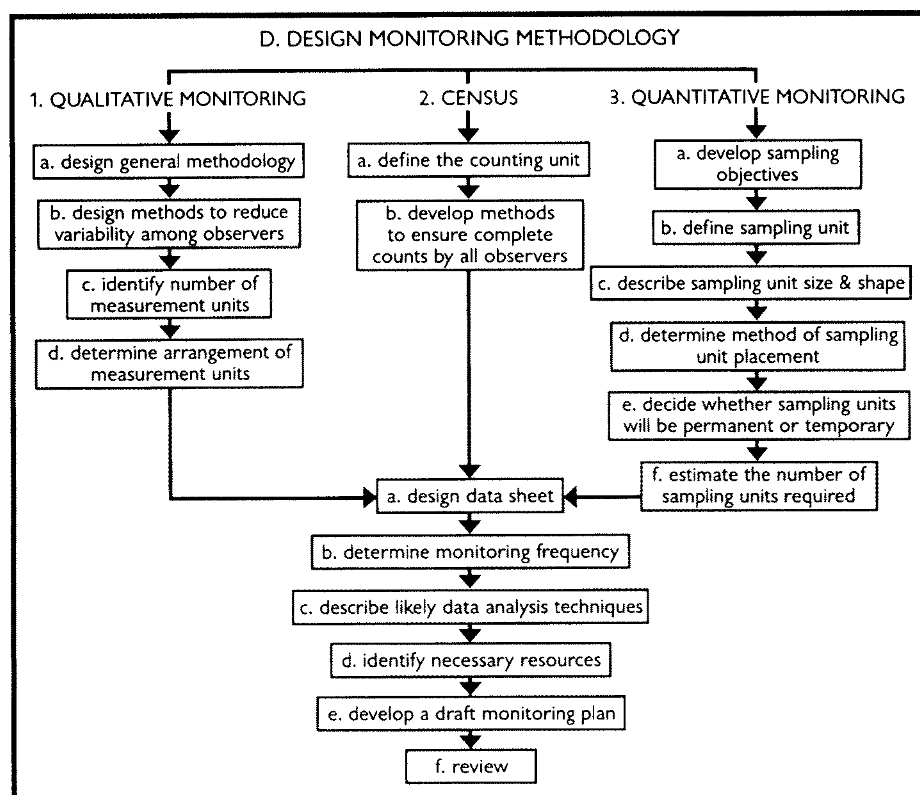


Figure 2.4. Flow diagram of monitoring process. The decisions required for each of the three types of monitoring—qualitative, census, and quantitative (sampling)—are summarized.

Identify Number of Measurement Units

Some qualitative monitoring situations may require several to many measurement units such as macroplots or photoplots. These are not sampling units, since they will not be combined and analyzed as a sample. Many design decisions, however, are similar to those required for sampling units and include selecting size, shape, and permanence.

Determine Arrangement of the Measurement Units

How will these measuring units be distributed in the population or across the landscape? Will you selectively place them based on some criteria such as threat or ease of access? Will you distribute these units evenly across the population to enhance dispersion and avoid bias?

Census

Define the Counting Unit

For plants you must decide if you will count individuals (genets), stems, clumps, or some other unit. For all species you must decide if you will count all individuals or only certain classes (such as juveniles)? If all individuals are not completely detectable and easily counted (usually the situation with animals), then a census will not be possible and some sampling protocol that adjusts for incomplete detectability will be required (see Chapter 13). These questions must be clearly addressed in the design to ensure that different observers conduct counts using the same criteria.

Develop Methods to Ensure Complete Counts

Will you have standardized methods (transects, plots, or grids)? Counts that are intended to be a complete census are often incomplete. What strategies will you use to ensure that small or cryptic individuals are not overlooked? For many animal species, counts are incomplete because of cryptic or elusive individuals.

Quantitative Studies with Sampling

Develop Sampling Objectives

If you are using sampling to estimate population sizes or mean values (such as density, cover, or frequency), you must also identify an acceptable level of precision of the estimate. If you are sampling and determining the statistical significance of changes over time, you must identify the size of the minimum detectable change (previously specified in your management objective), the acceptable false-change error rate, and the missed-change error rate (or statistical power level). What is the risk to the species if your monitoring fails to detect a real change (missed-change error), and how confident must you be of detecting a change over time (statistical power)? What is the risk to alternative uses/activities if your monitoring detects a change that is not real (false-change error)? (See Chapters 7, 8, and 9.)

Define the Sampling Unit

Will sampling units be plots, line transects, a collection of plots or points placed along a line or in a cluster, individuals, or parts of individuals (in plants such as number of seedpods)? Will all individuals be equally detectable? If not, adapt sampling techniques to adjust for incomplete detectability (see Chapter 13 for specialized sampling methods for animals). The sampling unit must be explicitly identified to ensure that the selected units are random and independent (see Chapter 8).

Describe Unit Size and Shape

The most efficient size and shape of the sampling unit depend on the spatial distribution of the species you are sampling. Most plants and animals are spatially arranged in clumps (i.e., individuals are not randomly dispersed across the landscape). Unless careful consideration is made of sampling-unit size and shape, many units may fail to intersect clumps of the target species. Many sampling units will be required in such a design to meet the specified precision and power of the sampling objective. Efficient sampling design using sampling units of appropriate size and shape can dramatically reduce the number of sampling units that must be measured, thus reducing the time and resources required for the field work and data entry. The size and shape of the sampling unit may be the most important decision affecting the success of projects where sampling is used (see Chapter 8).

Determine Sampling-Unit Placement

Sampling units must be positioned without bias. There are several methods described in Chapter 8.

Decide Whether Sampling Units Will Be Permanent or Temporary

Permanent sampling units are suitable for some situations, while temporary ones are more suitable for others (see Chapter 8). If the sampling units are permanent, monumentation or another method of relocation becomes critical and will require additional field time for plot establishment during the first year of the monitoring project (see Chapter 5).

Estimate the Number of Sampling Units Required

Data from a pilot study are the most reliable means to estimate the number of sampling units required to meet the targets of precision and power established in the sampling objective (see below). Chapter 8 and Appendix II describe estimation of sample size based on pilot data, as well as some alternative methods.

Design Issues Common to All Three Types

Design Data Sheet

While some studies may use electronic tools to record data, in most studies the researcher will record measurements on a data sheet. A well-designed data sheet can simplify rapid and accurate data recording and later computer data entry (see Chapter 6).

Determine Monitoring Frequency

How often should the parameter be measured? Will you be monitoring annually? Every 3 years? The frequency varies with the life-form of the species and the expected rate of change (e.g., long-lived plants or animals may require infrequent measurement), the rarity and trend of the species (the risk of loss for very rare or very threatened species is higher), and the resources available for monitoring.

Describe the Likely Data-Analysis Techniques

For all projects, describe how the data will be evaluated and analyzed. If you are using quantitative sampling, identify the statistical tests appropriate for the data you are planning to collect so the assumptions of the tests can be considered in the design stage (see Chapter 9). Do not assume that you can collect data, give it to an "expert," and expect meaningful results. Useful data analysis starts with good field design and data collection. This is also a good point to check whether the data will actually address the objective, given the analyses you plan to use.

Identify Necessary Resources

Now that you have specifically designed the monitoring project, estimate the projected annual and total costs and compare needed versus available resources. Reevaluate equipment and personnel required to successfully implement your project, and ensure that they are available. Document the individual or team responsible for implementation of the monitoring, the source and amount of the funding for monitoring (annually and over the life of the project), and the necessary equipment and personnel.

Develop a Draft Monitoring Plan

If all of these steps have been documented and reviewed, many components of your monitoring plan have been completed. The draft monitoring plan provides four important benefits: 1) it focuses the thinking of the author by forcing articulation, 2) it provides a vehicle for communication and review, 3) it documents approval and acceptance when finalized, and 4) it provides a history of the project and guards against the untimely end of the monitoring project if the primary advocate leaves (see Chapter 15). For those monitoring projects requiring minimal review from people outside the organization, the monitoring plan may be postponed until after data from the pilot stage have been analyzed.

Review Plan

Use the monitoring plan to solicit review of your proposed project (see Chapter 15). Do all reviewers agree with the methodology? Does the proposed methodology really monitor the objective? It may be necessary to revise the methodology, the objective, or both. For example, your objective may involve increasing cover of the target plant species; however, as you design the monitoring, you may realize that measuring cover of this particular species will be difficult. Treat the development of objectives and design as an interactive process. The objective drives the design of the monitoring, but the practical constraints of the morphology or biology of the species, the characteristics of the site, or the availability of monitoring resources may require reevaluation of the objective.

IMPLEMENT MONITORING AS A PILOT STUDY (FIG. 2.5)

Collect Field Data and Evaluate Field Methods

The first trial of a monitoring method in the field often exposes problems with the methodology (e.g., plots cannot be positioned because of dense vegetation; the proposed counting unit cannot be applied consistently; lacy vegetation proves a problem for measuring shrubs along a line intercept). This is why the pilot period is important for testing the feasibility of the proposed moni-

toring approach and for identifying improvements. You may find at this stage that the project cannot be implemented as planned and requires substantial revision, or even abandonment, in spite of all the work done to this point.

Analyze Pilot-Study Data

Analyze data from the pilot study. Do assumptions of the ecological model still appear correct? Are sampling objectives of precision and power met? If not, you may need to alter your monitoring design (add more sampling units or improve the efficiency) or the sampling objective (accept lower precision and/or power) or perhaps abandon the entire project. Does the level of change or difference you have specified seem realistic? Do changes caused by weather seem larger than you anticipated, thus swamping the quantity specified in your objective, or do the plants appear so slow growing that the proposed change is unrealistic? You may need to reassess the quantity or time-frame component of your objective.

Reassess Time/Resource Requirements

The pilot project should provide a better estimate of the resources required for monitoring. Your estimate of costs should include the amount of time it has taken to develop the monitoring to this point, as well as how much time it will take to continue the monitoring annually and to complete final data analysis and reporting.

Review

Solicit review of the results of the pilot period. Do all parties still agree to continue the monitoring and abide by the results? Are there resources available to implement monitoring throughout its life span? Make necessary changes to the monitoring design and the monitoring plan, and solicit final review.

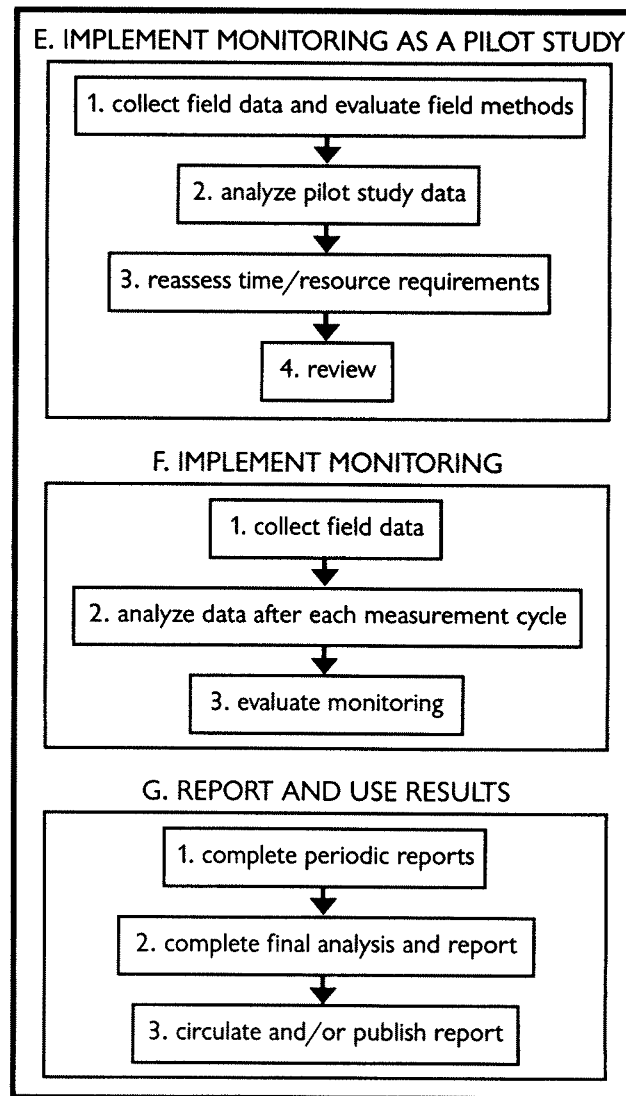


Figure 2.5. Flow diagram of the monitoring process, continued. Tasks associated with implementing monitoring as a pilot study, continuing monitoring, reporting and using results are illustrated.

IMPLEMENT MONITORING (SEE FIG. 2.5)

Collect Field Data

Complete data collection at specified intervals. Ensure that data sheets are completely filled out, duplicated, and stored in a safe place.

Analyze Data after Each Measurement Cycle

Complete data analysis soon after data collection. Data should not be stored over several years before analysis for a final report. Timely analysis identifies problems early, reduces the work associated with the final report, and ensures that questions requiring additional field visits can be addressed. In addition, questions that occur when field data sheets are entered into the computer can often be answered because the field work is still fresh in your memory.

Evaluate Monitoring

Evaluate field methods, costs, sample size, and relevancy of the monitoring project after each data collection. Recognize that at any time in the process a problem may arise that causes you to change or abandon your monitoring effort. All the steps preceding this one reduce that risk, but do not eliminate it.

REPORT AND USE RESULTS (SEE FIG. 2.5)

Complete Periodic Reports

Completing a summary report each time data are collected will yield the following benefits: 1) display the importance and usefulness of the monitoring to management, thus increasing continued support; 2) provide a summary for successors in the event of your departure; and 3) provide a document that can be circulated to other interested parties.

Complete Final Analysis and Report

At the end of the specified time frame (or earlier if objectives are achieved), prepare a final monitoring report and distribute to all interested parties (see Chapter 15). This final report presents and summarizes the data, analyses, and results and provides recommendations. If the monitoring project has been designed and documented as described above and data have been analyzed periodically, completion of this final report should not be onerous nor contain major surprises.

Circulate and/or Publish Report

Sharing the results of your monitoring increases the credibility of the organization, assists others in the design of their monitoring projects, enhances partnerships, contributes to the scientific enterprise, and reduces redundancy. Sharing the results in a technical forum such as a symposium or a journal article is also a good opportunity for professional development for you.

CHAPTER 3
Selecting Among Priorities

Cryptantha tumulosa
Mohave Cryptantha
Southeast Nevada and adjacent California
Artist: Jeanne R. Janish

Imagine you are a botanist, and you have just been hired by a 3.5 million-acre National Forest. The Forest lands are known to harbor at least 65 populations of eight rare plant species, some of them listed under the Endangered Species Act and others recognized as Species of Concern. What will you do? How will you decide which species to work on first? Or imagine you are a manager in charge of a 30,000-acre preserve. On your preserve are populations of four rare plant species, a threatened species of beetle, a stream containing a trout species that has been recognized as a Species of Concern in your state, one of the largest great blue heron rookeries in the state, and a cottonwood gallery forest that is occasionally used by transient bald eagles. What will you do? In this situation you not only must set priorities among species and populations, but you will also need to decide among life-forms.

Resources and funding for management and monitoring are limited. You will not likely be able to develop objectives, implement management, and monitor achievement of objectives for all the species and populations for which you are responsible. Priorities must be set, and the scale and intensity of monitoring these priorities must be determined. Scale describes the spatial extent. The scale of monitoring can range from a macroplot subjectively placed within a population to all populations of a species across its range. Intensity describes the complexity and cost of the monitoring. Intensity can vary from a single photopoint that is revisited every 5 years to a labor-intensive demographic technique that requires annual assessment of every individual in a population.

Clearly, as you increase the scale and intensity, you will know more about the species and its trend and status, but the monitoring will be more expensive. With limited funds, you can monitor one or a few species at a large scale and high intensity, or more species at a more limited scale and lower intensity. The setting of priorities is the first step in determining the importance and number of species and/or populations that require attention, the monitoring resources that should be allocated to each, and the complexity of the objective for each species or population that can be monitored.

In the absence of explicit priorities, species that are in need of monitoring because of their rarity and sensitivity may be ignored, while more common species may be addressed because they are controversial (such as wolves), associated with an urgent or high-profile issue (such as animals or plants found in riparian areas or old growth), or are public favorites (such as orchids, butterflies, and eagles). The narrow margin of existence of some species and the rapid rate of decline in others leave little room for misallocation of management and monitoring resources. Although you would expect that the rarest species are managed and monitored most intensively, a review of monitoring in the United States found that according to priority classifications used by The Nature Conservancy and the U.S. Fish and Wildlife Service, nearly half of the plant species monitored were of low-priority ranking (Palmer 1986, 1987). Surprisingly, nearly a third of the studies reviewed used a demographic approach, the most intensive method of monitoring and a choice that likely meant ignoring other species. No similar study has been conducted for animals, but it is likely that results would be similar. Explicit setting of priorities would alleviate the problem.

UPPER-LEVEL PRIORITIES

The first step in the setting of priorities is to identify any upper-level guidance that may exist. Priority species or populations may have already been identified in an accepted land use plan or activity plan. These documents provide overall management direction for large areas of land. A recent, interagency, regional plan that encompasses portions of several states occurring within an ecosystem boundary may also provide guidance (e.g., the Columbia Basin Project in the western United States). Other programs attempt to plan across even larger areas. For example, Partners in Flight is a coalition of individuals from many public and private interests concerned with the

conservation of birds in the Americas. This group identifies sets of priority species for monitoring in particular regions, states, and provinces. By doing so, Partners in Flight attempts to coordinate and integrate bird conservation activities around the continent. Many planners of bird monitoring programs take into consideration these higher-level directives from the Partners in Flight group. Following these existing guidelines ensures continuity of management across time (from successor to successor) and space (across administrative boundaries). In the absence of compelling reasons such as new information or the appearance of new threats, the priorities identified at these higher levels of coordination should be accepted and followed.

Many agency land use plans, especially older ones, provide little direction for management of rare species. If the plan lacks goals directly related to rare species, look for supporting goals such as "maintain a full complement of flora and fauna" or "maintain viable populations of native species." Occasionally, directions for conservation of rare species may be found in lists of standard operating procedures.

Another source of identified priorities may be state or regional lists of rare species. In the United States, for example, some Bureau of Land Management (BLM) State Offices issue a list of priority species, as do some Forest Service Regional Offices. The Fish and Wildlife Service assigns a listing priority to species based on threats and taxonomic status (Box 3.1). The Nature Conservancy and its associated Natural Heritage Programs[1] rank all rare species with a state and global rarity ranking based on the number of occurrences (Box 3.1). In some states, priorities are recommended at an annual meeting of representatives from federal and state agencies, universities, and private firms.

Box 3.1. *FOUR WIDELY USED SYSTEMS FOR RANKING SPECIES. SUCH APPROACHES MAY BE USEFUL FOR SETTING MONITORING PRIORITIES.*

LISTING PRIORITY, U.S. FISH AND WILDLIFE SERVICE

PRIORITY	THREAT MAGNITUDE	THREAT IMMEDIACY	TAXONOMIC STATUS
1.	High	Imminent	Monotypic genus
2.	High	Imminent	Species
3.	High	Imminent	Subspecies / Variety
4.	High	Not imminent	Monotypic genus
5.	High	Not imminent	Species
6.	High	Not imminent	Subspecies / Variety

Priorities 7–12 have the same approach, but for species with low magnitude of threat.

THE NATURE CONSERVANCY / NATURAL HERITAGE PROGRAM RATING SYSTEM

1. *Critically imperiled (5 or fewer occurrences or very few [<1000] individuals or few acres).*

2. *Imperiled (6–20 occurrences or few [<3000] remaining individuals, or few acres).*

3. *Very rare and local; found locally in a restricted range; or vulnerable to extinction or extirpation by outside factors (21–100 occurrences) or <1000 individuals.*

(Continued)

[1] *In many locations, Natural Heritage programs have been transferred to a government agency and go by a different name.*

Box 3.1. *FOUR WIDELY USED SYSTEMS FOR RANKING SPECIES. SUCH APPROACHES MAY BE USEFUL FOR SETTING MONITORING PRIORITIES. (Continued)*

4. Apparently secure, although may be rare in parts of its range.

5. Demonstrably secure, although may be rare in parts of its range.

6. Status uncertain, with the need for more information; possibly in peril.

These rankings are used either at the state scale (within the state only) or at a global scale (the entire range of the species) and are often presented, for example, as "S1" or "G1," for critically imperiled at the state or global level.

STATE LISTING CRITERIA[a]

- State Priority 1. *A taxon in danger of becoming extinct or extirpated from the state in the foreseeable future if identifiable factors contributing to its decline continue to operate; these are taxa whose populations are present only at critically low levels or whose habitats have been degraded or depleted to a significant degree.*

- State Priority 2. *A taxon likely to be classified as Priority 1 within the foreseeable future if factors contributing to its population decline or habitat degradation or loss continue.*

- Sensitive. *A taxon with small populations or localized distributions within the state that presently do not meet the criteria for classification as Priority 1 or 2, but whose populations and habitats may be jeopardized without active management or removal of threats.*

- Monitor. *Taxa that are common within a limited range, as well as those taxa that are uncommon, but have no identifiable threats (for example, certain alpine taxa).*

- Review. *Taxa that may be of conservation concern, but for which we have insufficient data on which to base a recommendation regarding their appropriate classification.*

- Possibly extirpated. *Taxa that are known in the state only from historic (pre-1920) records or are considered extirpated from the state.*

WORLD CONSERVATION UNION (IUCN) CATEGORIES

- Extinct *(EX), when there is no reasonable doubt that the last individual in a taxon has died.*

- Extinct in the Wild *(EW), when a taxon is known only to survive in captivity well outside the past range.*

- Critically Endangered *(CR), when a taxon is facing an extremely high risk of extinction in the wild in the immediate future.*

- Endangered *(EN), when a taxon is not Critically Endangered, but is facing a very high risk of extinction in the wild in the near future.*

---

[a]*Used for prioritizing state-listed species in several states in the United States.*

- Vulnerable *(VU), when a taxon is not Critically Endangered or Endangered, but is facing a high risk of extinction in the wild in the medium-term future.*
- Lower Risk *(LR), when a taxon has been evaluated and does not satisfy the criteria for any of the categories Critically Endangered, Endangered, or Vulnerable. Taxa included in the Lower Risk category are further separated into the following categories:*

  Conservation Dependent *(CD), taxa that are the focus of a continuing conservation program, the cessation of which would result in the taxon qualifying for one of the threatened categories above within a period of 5 years.*

  Near Threatened *(NT), taxa that do not qualify for Conservation Dependent, but that are close to qualifying for Vulnerable.*

  Least Concern *(LC), taxa that do not qualify for Conservation Dependent or Near Threatened status.*

SETTING LOCAL PRIORITIES

When upper-level direction is minimal or nonexistent, you will need to set priorities locally. Doing so requires information about the species, the people who are stakeholders in the management of the area, and a process to compare species and populations. Because establishing priorities in land management is a subjective process, different people will list the same species in different priority order. For a manager, the highest priority species may be the one that conflicts with the dominant commodity activity. For a land user, the highest priorities are species and populations that may affect their use of the area. For the botanist or wildlife biologist, the highest priority species may be the rarest or most-threatened ones. Legal direction, existing plans, and pet projects may all conflict with priorities that would result from a strict following of biological criteria.

Because the setting of priorities is subjective, we recommend using a team of stakeholders or, at a minimum, setting priorities with input from managers and other specialists. For more controversial areas, solicit input from others outside of the agency as well such as conservation and commodity groups. Setting priorities is a situation-specific activity.

You must design a process of setting priorities that encourages participation. It must summarize information about the species or populations and provide access for meaningful dialog and comments. We propose a simple matrix approach that prioritizes species based on additive ranking of a combination of criteria (Fig. 3.1). The criteria selected depend on the type of species and the management situation, and should incorporate both biological and management (social) criteria. The lists of criteria that follow are not meant to be exhaustive; there may be other criteria important to your specific situation that you should include.

Criteria for Species Comparisons

- *Rank.* Some approaches have utilized the conservation status or rank assigned a species such as from one of the systems described in Box 3.1. Note that in many systems this rank is already a composite of criteria. For example, the ranking used by the U.S. Fish and Wildlife Service combines taxonomic distinctness with the magnitude and imminence of the threat.
- *Rarity.* Rarity relates to population size, number of populations, and distribution of populations across the landscape. In comparing species, perhaps the most useful

		BIOLOGICAL CRITERIA							MANAGEMENT CRITERIA						
		rarity	taxonomic status	sensitivity	known decline	extent of threats	immediacy of threats	importance of local population(s)	existing conflict	monitoring difficulty	availability of management actions	recovery potential	public interest	potential for crisis	Total
SPECIES	**WEIGHTING**	**4**	**2**	**5**	**5**	**5**	**5**	**5**	**2**	**1**	**5**	**5**	**1**	**1**	
species A (a rare variety)	rating for species	3	1	3	3	2	3	3	1	3	1	2	1	1	27
	rating x weight	12	2	15	15	10	15	15	2	3	5	10	1	1	106
species B	rating for species	2	2	1	3	3	3	1	3	1	3	3	3	3	31
	rating x weight	8	4	5	15	15	15	5	6	1	15	15	3	3	110
species C	rating for species	1	3	3	2	2	1	1	1	3	3	3	1	3	27
	rating x weight	4	6	15	10	10	5	15	2	3	15	15	1	3	104
species D	rating for species	1	1	1	2	1	1	2	1	3	1	1	1	1	17
	rating x weight	4	2	5	10	5	5	10	2	3	5	5	1	1	58
species E	rating for species	3	1	3	3	3	3	3	1	1	3	3	1	1	29
	rating x weight	12	2	15	15	15	15	15	2	1	15	15	1	1	124

Figure 3.1. Completed matrix for setting priorities among five species. Criteria are given weighting values between 1 and 5, with the highest weighting values given to the criteria considered most important in this situation. Species are rated from 1 to 3, with species rated 1 of lowest importance. In this example, Species A is a very rare, rating a "3" for rarity, but, because it is only a variety, it is given a "1" for its taxonomic status. For monitoring difficulty, a low number means it is a difficult species to monitor; in this example, these species are considered of lower importance for monitoring. Rankings are summed across all criteria, both with and without weighting. In this example, Species D is lowest-ranked species (17), while Species B is the highest (31) based on unweighted rank and Species E is the highest based on weighted rank (124).

aspect of rarity for monitoring is the number of populations. A species restricted to a single, large population is at more risk than one with fewer total numbers distributed in several populations. Similarly, populations clustered in a small area may all be affected by the same threat, while populations that are widely distributed are less likely to be affected by a single impact.

- *Taxonomic distinctness.* A species that is the only representative in its family would rank above one that is the only representative in its genus, which would rank above a species that occurs in a genus with many species. The concept behind this ranking is that the taxonomic distinctness of a single-species family correlates with high genetic uniqueness. A subspecies or variety would for the same reasons be considered of less value. A drawback to this approach is that most current taxonomic divisions are still largely based on morphological differences and may not directly relate to evolutionary distinctness.
- *Sensitivity to threats.* Species vary in their sensitivity to threats, depending on their biology and ecology. Plant species with a long-lived seed bank are buffered from population declines because a single, good, germination year can function as a rescue. Plant species with populations that vary widely from year to year but lack a seed bank are more prone to local extinction. A similar situation occurs in animals. Some groups such as turtles, large mammals, and some long-lived birds have long generation times that can buffer their populations against short-term disruptions to reproduction, whereas others such as many endangered invertebrates (e.g., beetles and butterflies) or small mammals can quickly vanish because population persistence is dependent on high and continued levels of reproduction each year. Species that are limited to midsuccessional stages are vulnerable to both disturbance and succession caused by lack of disturbance. Species that have exacting habitat requirements are more sensitive than those that are more cosmopolitan in habitat.
- *Known declines.* Species with known declines based on monitoring or observation are more important for monitoring and management than species that are considered stable.
- *Extent of threats.* Threats can be evaluated in terms of scale and intensity. Scale describes the percentage of the populations affected and the distribution of threats across the landscape. Intensity describes the degree to which populations are affected by threats (e.g., extirpation of the population, mortality of a few individuals).
- *Immediacy of threats.* The rate at which threats may occur and populations decline is another important consideration. Species or populations with ongoing or immediate threats would rank higher than those with potential threats.
- *Importance of local population(s).* You likely only manage a portion of the range of a species. How important are the populations under your jurisdiction to the long-term success of the species as a whole? If you have one of the largest known populations of the species, or the only known population of the species, your management actions are very important. It may be more important to manage and monitor an excellent occurrence of a fairly common species (such as the largest great blue heron rookery in the state that happens to occur in your preserve) than a poor occurrence of a rarer species.
- *Conflict.* The degree of management conflict between potential conservation actions and existing or alternative uses (usually commercial) may be an important consideration in prioritizing populations. The degree of conflict may also dictate the intensity of monitoring (high-conflict situations may require quantitative monitoring or even research into cause and effect).

- *Monitoring difficulty.* Monitoring populations of some types of plants such as annuals and geophytes can be nearly impossible because of temporal and spatial variability. Many animals are also notoriously difficult to monitor. These include the smaller beaked whales of the great oceans, nocturnal birds such as spotted owls, secretive mammals like mustelids (e.g., fishers), species that make an unpredictable and ephemeral appearance as some amphibians and insects do, and species that occur only in inaccessible sites (e.g., seabirds breeding in burrows or fishes of deep-water areas). Some animals combine being highly secretive with being quite dangerous to handle, thereby limiting the opportunity to use trapping to track their numbers (e.g., venomous snakes or carnivorous mammals). Some species such as those found on cliffs are difficult to access. Monitoring species growing on fragile sites such as erosive slopes or semi-aquatic habitats may cause unacceptable investigator impacts.
- *Availability of management actions.* If no management options are available, monitoring resources should be directed toward other species with management alternatives.
- *Recovery potential.* Some species will only recover with a large expenditure of resources, while others have high recovery potential. You may choose to focus on the species with the highest potential, especially if several species could be managed for recovery with the same resources required for one.
- *Public interest.* The fact that birds and mammals (e.g., bald eagles and grizzly bears) corner vastly more recovery resources than amphibians or invertebrates is largely a result of this factor. Similarly, a common orchid will have more public support and interest than a rare moss or lichen.
- *Potential for crisis.* Crisis can be defined in biological terms (potential for extinction) and in management terms (potential for politically heated conflict).

Criteria for Population Comparison

Applying criteria for setting priorities among populations of the same species differs from that of species comparisons. For species, the goal is to protect and maintain as many species as possible (to combat extinction), so species most at risk are those most important to manage. For populations, your strategy may be to identify the ones that are in the best condition (largest, fewest threats, lowest sensitivity, etc.) and focus your management energy on those. Conversely, your strategy may be to concentrate on the populations that are in the most trouble and to assume that the ones in better condition will take care of themselves. Either strategy is legitimate, and both have been applied in conservation, but you need to think about your conservation philosophy explicitly.

- *Population size.* Investing in larger populations may be a better conservation strategy than salvaging small populations. Larger populations are better able to withstand annual variability, and they provide a larger buffer for decline. Conversely, it may be more important in some situations to monitor small populations because they are more prone to extinction and to assume that the larger ones are at less risk. These smaller populations, especially if disjunct or isolated, may be critical for conserving the genetic variability for the species.
- *Population viability.* A population with individuals distributed among all age or stage classes is more demographically "healthy" than one with obviously skewed age or stage distributions (e.g., all old or dying individuals). Monitoring may be concentrated on those populations with the best potential for long-term survival or on those that are obviously in trouble.

- *Population location.* Selecting populations on the fringe of the distribution of the species usually increases the range of genetic variability conserved. These populations may also occupy fringe habitats that are marginal and stressful, and they may express response to rangewide stresses such as climate changes before more central populations.
- *Habitat quality.* Depending on your conservation philosophy, higher priority for monitoring may be applied to populations found on degraded or disturbed habitat (because they are more at risk) or on stable or pristine habitat (because protection is a better conservation investment than restoration).
- *Unique habitat.* Populations located on unique habitat likely contain unique genetic combinations and are important for conserving the range of genetic diversity of the species.
- *Previous information/monitoring/research.* Populations with previous monitoring or natural history studies may be a higher priority if data suggest a decline or problem, or a lower priority if data suggest the population is stable or increasing.
- *Special management area.* Specially designated areas such as Research Natural Areas (RNAs) and purchased preserves represent a significant investment of resources. If a species population was an important justification for establishment or purchase, maintaining the population is a management priority. Monitoring of these populations would be a higher priority than populations in nondesignated areas. Conversely, it may be assumed that the protection afforded by designation reduces threats, as well as the need for monitoring.
- *Other.* Most of the criteria applicable to prioritizing among species are also applicable to prioritizing among populations (e.g., sensitivity to threats, extent of threats, monitoring difficulty, availability of management actions, recovery potential, public interest, and potential for crisis). Remember that, depending on your strategy of either conserving the best or saving the most in peril, you will apply these criteria differently.

LOCATING INFORMATION ABOUT SPECIES AND POPULATIONS

Reviewing existing species and population information serves five important functions in the development of a monitoring project:

1. The compilation and comparison of existing information are important for setting priorities. To allocate the monitoring resources, you must know about the relative rarity of and the threats to all the species you manage.
2. The response of a species to a management approach may have already been monitored elsewhere. An initial review may identify the need for immediate changes in management and thereby avoid monitoring a decline before the management action is initiated.
3. Some measurement techniques may have been tried previously on your species (or a similar one) with minimal success. Knowing the monitoring history may help you avoid repeating mistakes made earlier.
4. This information will be used in developing the ecological model and setting objectives.

5. A compilation of existing information will identify parties that should be included in the development of the monitoring project. For example, an assessment of the distribution of a species across its range might identify the need to coordinate monitoring of populations on adjacent Federal lands managed by another office or agency. An assessment of threats may identify commodity groups who should be involved in the development of a monitoring project, since their resource use may affect or be affected by the results.

Sources of information are varied and are rarely in an accessible published form. Much of the knowledge about a species resides in the experience of individuals and may be difficult to extract.

Natural Heritage Programs and Conservation Database Centers associated with a state agency or The Nature Conservancy maintain databases on the location and condition of rare species populations. They also provide access to that information in adjacent states. Native plant societies and conservation groups active in a region may have information on a species and may also be able to put you in contact with amateur and professional botanists and wildlife ecologists who know about the species. Academic experts who have worked with the species or related species may sometimes be found at universities or colleges. Experts associated with private consulting firms and with regional agencies and those with federal governments (for example, in the United States, the U.S. Fish and Wildlife Service) may also provide advice.

Herbaria may be a source of information on additional plant populations. Records of collections from national and state museums have often been an important source of information for locating rare animals. Specimen labels often contain habitat notes, and some herbaria and museums have computerized these to facilitate searching and summarizing. Many Heritage Programs/Databases have completed searches for rare species and may have the information in an accessible form. Be cautious, however, about using collection records. Specimens may not be accurately identified (those that have been annotated as part of a recent study are the most reliable), may be misfiled, or may be poor representations of a species. Place names may provide only general location information or may even be incorrect. These problems often increase with the age of the specimen.

Published information on rare species is most often found in symposium proceedings, technical reports, and project reports. This information can be difficult to locate through conventional computerized searches and is often best found through contact with reputable sources. Often Natural Heritage Programs maintain extensive collections of unpublished and published literature on sensitive species.

All existing information should be documented and stored in a single place (you should duplicate and archive one copy to protect from loss). A summary of the information that should be included is given in Box 3.2. For many species, little is known, and many of the information items must be filled in with hypotheses. Avoid simply leaving the information out. Your hypotheses are likely better than nothing, and, by forcing yourself to try to describe all of the species' biology and threats, you will identify those information items that are critical to your ecological model and to the monitoring design. These may require additional study before initiating monitoring.

The sources of the information in your summary should be documented. Cite published sources and personal communication, and comment about the reliability of the information. Hypotheses and your observations should be clearly identified.

These summaries are time-consuming, but they have benefits in addition to improving the quality of monitoring projects. The summaries can be referenced or included in biological evaluations and assessments. They can be helpful in training technicians or other specialists. They also communicate your observations and knowledge of the species to your successors. Once completed, the summaries are easily updated, incorporating new information as it becomes available.

Box 3.2. *COMPONENTS OF INFORMATION THAT MAY BE USEFUL IN A REVIEW OF A SPECIES.*

Summarizing all that is known or hypothesized about each of these components is not only helpful in setting priorities among species and populations for monitoring, it is also critical for developing ecological models, designing studies, and ensuring that anecdotal information about a species is not lost during changes in personnel.

SPECIES BIOLOGY

Plants

Life history
Life expectancy (long or short-lived)
Reproductive ecology
  Pollinators
  Flowering period
  Annual variability in flowering
  Seed maturation period
  Seed production
  Seed viability
Seedling ecology
  Regularity of establishment
  Germination requirements
  Establishment requirements

Animals

Food habits
Predators
Competitors
Breeding season
Age at maturity
Number of offspring produced annually
Frequency of breeding
Home range

POPULATION

Population size (range, average)
Annual variation
Number and distribution of populations

Productivity of different populations

Migratory patterns

HABITAT

Soil
Elevation
Aspect
Slope
Moisture
Community

(Continued)

Box 3.2. COMPONENTS OF INFORMATION THAT MAY BE USEFUL IN A REVIEW OF A SPECIES. (Continued)

Vegetation structure
Competition
Large-scale disturbance (e.g., fire, floods)
Small-scale disturbance (e.g., animal burrows or trails, windthrow)
Landscape connectivity

THREATS
Natural

Disease
Predators
Succession
Weed invasion

Anthropogenic

On-site (grazing, logging, poaching, etc)
Off-site (changes in hydrology, pollinators, habitat fragmentation)

TREND

CAUSES OF TREND

MANAGEMENT OPTIONS

ASSESS AVAILABLE AND REQUIRED RESOURCES

Managers must be committed to the monitoring project and willing to expend the resources required for a successful project. Priorities and allocation of time and dollars are the responsibility of managers. Managers are also the ones who will make decisions based on the monitoring. Be wary of your inclination to do self-driven monitoring, where you choose to devote what resources you can toward your favorite monitoring project. Although the monitoring may be implemented as long as you are there to do it, if you leave, your pet project will probably die. A monitoring project needs other advocates besides the specialist(s), preferably in management.

Once management is supportive, you should consider three limiting factors when designing a monitoring project: 1) the skill level of those planning and implementing the project; 2) the equipment available; and 3) the time and money available for field work and analysis.

The project may require special skills at the planning level. Depending on the complexity of the project and your knowledge, you may need a statistician or someone with expertise in sampling design. State or regional offices may have people who can help. You may be able to solicit or contract advice from specialists associated with universities, private consulting firms, and conservation groups. Experts associated with state agencies and federal agencies also can provide technical expertise. Use as many resource people as possible for review. Special skills may also be needed at the implementation level. Field work that will be completed mostly by summer technicians may need to be designed differently than that done by experienced specialists.

Most monitoring projects require inexpensive equipment, such as measuring tapes, pin flags, and a camera (a list of standard field equipment is provided in Chapter 5). Some projects

may require specialized equipment, such as Global Positioning Systems (GPS), survey equipment, and video equipment. These are becoming more commonly available at agency and organization offices. Other specialists in other disciplines may have ideas about useful equipment that will reduce your field time. Many of these people also have experience in sampling plants and animals and can provide ideas and help sharpen your thinking through discussion.

Finally, the time required must be compared with the time allocated for a monitoring project. Most specialists are fairly good at estimating field time for gathering data. Estimating the office time required is more difficult. For simple projects, estimate at least one work week to develop and document the objectives and to design the monitoring. Complex projects requiring consideration of various points of view and extensive review will take much longer. We estimate that the time needed to collate, digitize, and analyze monitoring data usually exceeds that required to initially collect it in the field. To estimate analysis and reporting time, multiply field time by two to five times, depending on the complexity of the data gathered. Qualitative data will take less time to analyze and report than a detailed, data-intensive method that requires statistical analysis.

It is important that the time required for monitoring be estimated liberally. Many field datasets have not been analyzed because time needed for analysis was not included in the budget. Managers must know and support the total time required for completion of the monitoring project.

PRIORITIES, RESOURCES, SCALE, AND INTENSITY

Allocating your management and monitoring resources among your priorities is not an easy task because of the interplay among a number of priorities, the scale and intensity of management and monitoring, and the available resources. Given limited monitoring resources, scale and intensity are inversely related. You can choose to monitor many populations (large scale) with low intensity or devote all your monitoring resources toward monitoring a single population intensely. If you have many high-priority species, limited monitoring resources may allow you to monitor only a single population of each species at a low intensity. This explicit consideration of the interplay of priorities, resources, scale, and intensity is critical to the effective allocation of monitoring resources.

Priorities, resources, scale, and intensity are interrelated. Change one, and you will change one or all of the others.

Selecting Scale

Monitoring scales can vary from a single, small, local population of a few individuals (local scale) to many, large populations and the range of the species (landscape scale). The scale should be decided explicitly, because scale has important implications for monitoring design (see Chapter 8). The selection of scale will be guided by management considerations and priorities and will be limited by resources available for monitoring. Decisions about scale have to be made at both the landscape and local levels.

Landscape scale can be defined in a number of ways:

- All known populations of the species
- Populations on federal and state lands
- Populations within an administrative unit
- Populations within a watershed
- Populations within a vegetation type
- Populations within a management unit (e.g., an allotment, a wilderness area)
- Populations within a treatment area
- Populations with a specific management treatment (e.g., populations in logged areas)

Establishing a system of monitoring populations of a species across its entire range provides the most accurate measure of the overall trend and condition of a species. Because of the extent of required coordination efforts for species that cross administrative boundaries, however, such rangewide approaches are unfortunately rarely attempted. If you share a species of limited distribution with only one or two other organizations or agencies, consider trying to coordinate monitoring efforts. For species that cross several administrative boundaries, efforts at interagency regional planning and ecosystem management provide hope that future coordination of rangewide monitoring of species may become easier.

Once you have identified the landscape scale and the pool of populations that you will consider, you need to decide if all populations at that scale will be monitored or only a portion of them (perhaps because of limited monitoring resources). If monitoring only a portion, you must decide if you want to either draw a sample of populations from all those that occur at that scale or select specific populations. If you wish to draw conclusions about all of the other populations at that scale from the portion monitored, you will need to draw a random sample of monitored populations from the entire set of populations. For example, if you monitor only populations that are easily accessible along roads, your sample would be biased (not random) and only represent roadside populations. You will be unable to draw any conclusions about nonroadside populations. You may, however, decide that you will select only roadside populations because those are the ones about which you have conservation concerns. This is a perfectly valid approach, as long as you recognize that you are limited to conclusions only about roadside populations.

In statistical terms, when you identify the set of all populations that are of interest, you define a "sampling universe" from which you will randomly draw "sampling units" (in this example, individual populations). You must carefully consider both the sampling universe and sampling units if you want to be able to draw conclusions about several populations. These concepts are described in more detail in Chapter 8 and also apply to consideration of scale at the single population and macroplot level (below).

A sampling universe is the collection of potential sampling units from which a random sample of units will be selected and measured, and to which the results of the sample can be applied.

Note that randomly selecting a single population to monitor does not mean you can draw conclusions about all of the populations. Common sense and biological experience suggest that a single population cannot possibly represent the range of conditions and trends occurring at other populations. You need a sample of populations (i.e., several of them) to have enough information to draw more extensive conclusions. You may, however, be able to use qualitative monitoring at other populations to support conclusions that trends or conditions at these are similar to the site you are monitoring quantitatively.

At a local scale of a single population you will face the same sample-versus-selection issue previously described for populations unless your population is small enough to be completely monitored or sampled. If you wish to draw conclusions about the entire population you are monitoring, you must randomly sample from the entire population. Sometimes this is not possible. For example, a population comprised of individuals dispersed over a very large area may be difficult and time-consuming to sample randomly. Some portions of the population may be physically inaccessible; if you exclude them from your sampling universe, you cannot extrapolate information taken from accessible sampling units to the inaccessible portion of the population. Again, this makes biological sense. You know that the portion of a population on a cliff face is not going to perform the same as the portion found on the accessible slopes below the cliff.

One option is to select a portion of the population as a key area or macroplot, monitor only within that area, and agree among interested parties that the results will be applied to management of the entire population (see Chapter 15). The drawback is that you must assume that the key area functions as an indicator for the entire population. Inferences cannot be made to the entire population based on data. Changes measured on the macroplot may or may not represent those occurring outside of the macroplot. This problem can be partially addressed by supplementing the quantitative studies within a macroplot with qualitative studies dispersed through-

out the population. While you will still be unable to conclusively state that the changes observed within the macroplot represent those outside the macroplot, the supporting evidence may be sufficiently strong for management decisions.

Situating a macroplot requires some decisions. Will the plot(s) be located in the area most likely to be affected by adverse management? Will you attempt to locate the plot(s) in a representative area of the population, and if so, how will you define what is representative? Will your main criteria be ease of access? Chapter 8 discusses these issues in more detail.

Selecting Intensity

Intensity of monitoring can be defined as the complexity of methods used to collect information. Monitoring intensity roughly equates to time, but also relates to the skills required to collect information. Monitoring can be generally classified into qualitative and quantitative techniques. Within each class, levels of intensity also vary.

Qualitative techniques are usually less intensive than quantitative, but can be effective for many situations. Low-intensity monitoring may be designed as a warning system that triggers more intensive monitoring or research if a problem appears. In other situations, low-intensity techniques may provide adequate data for making decisions. Because changes monitored by these low-intensity or qualitative techniques must be fairly large or obvious before they are detected, it is often appropriate to take immediate management action based on these measures. Implementing a high-intensity study to quantify an obvious problem only delays remedial action.

Some examples of qualitative approaches follow. These are covered in more detail in Chapter 4. They are approximately ordered in increasing intensity, although the actual order depends on how each technique would be implemented in a particular situation.

- *Presence or absence.* Noting whether the species of interest is still at the site may be an effective way to monitor a species with many roadside populations. Populations located along roads can be noted by a "windshield check" (i.e., viewed through the windshield of an automobile without leaving the vehicle) by other specialists in the course of their work.
- *Site-condition assessment.* Site-condition assessments provide a repeated evaluation of the quality of the habitat. The monitoring is designed to detect obvious and dramatic changes that can be recorded photographically, with video, or in written descriptions aided by a standard form.
- *Estimates of population size.* Visual evaluation of population size, often in classes (such as 0, 1–10, 11–100, 101–1000, 1001+), provides more information than simply noting presence or absence.
- *Estimation of demographic distribution.* A population's demographic distribution is the percentage of the population or number of individuals within classes such as juvenile, nonreproductive adult, reproductive, and senescent.
- *Assessment of population condition.* In this approach the observer evaluates the condition of the population by noting occurrence and extent of utilization, disease, predation, and other factors.
- *Photopoints.* Photopoints are pictures that are retaken from the same position of the same frame at each observation (see Chapter 4).
- *Photoplots.* Photoplots straddle the division between qualitative and quantitative monitoring. These are usually close-up photographs of a bird's-eye view of a plot within the frame. Plot size varies with camera height and lens type, but commonly ranges from 50cm × 30cm to 1m × 1m. Photoplots can provide a qualitative record of a small portion of the population, or they can be used as a plot to measure cover and/or density (see Chapter 4).

- *Boundary mapping.* Mapping the perimeter of a local population monitors change in the area occupied by the population.

Quantitative monitoring requires the measure or count of some attribute. Three basic types of quantitative approaches can be described (in increasing intensity): census, sample, and demographic.

A census of the population counts or measures every individual. The main advantage of this approach is that the measure is a count and not an estimate based on sampling. No statistics are required to characterize the current status or changes over time. The changes measured from year to year are real, and the only significance of concern is biological. A sample measures only a portion of the population. No sample is an identical representation of the population as a whole. It is an estimate of that population; thus, some error is associated with the sample (the difference between the sample estimate and the real value of the population). Statistics is the tool used to assess that error (see Chapters 7, 8, 9, and 10). A sample of quantitative data should only be taken if the results are to be analyzed statistically because the error associated with that sample can be quite large and statistical analyses are needed to correctly interpret observed changes.

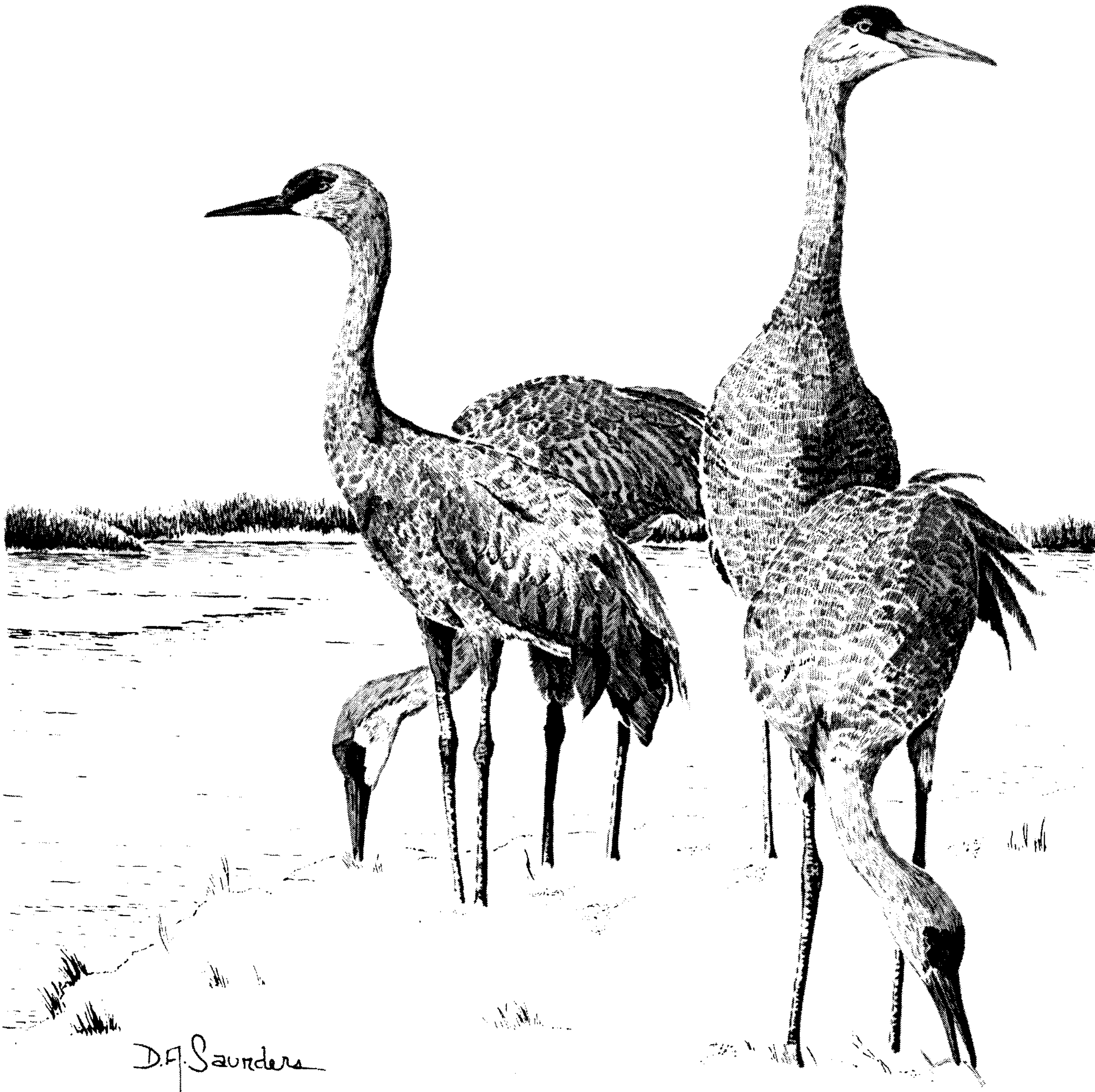
Some monitoring designs avoid statistics by doing a complete census or full counts in a small portion of the population in a representative plot. For example, height may be used to measure plant vigor annually. Rather than sampling, a single representative plot is established in the middle of the population, and the height of all individuals within that plot is measured. No statistics are necessary because you know the true average height of all the plants in the plot. If the decision has been made to base management changes on the changes within the plot, this is an acceptable approach, but be aware that the average height of the plants in the plot is not an estimate of the average height of the plants in the population. (Your sampling universe in this example is the plot.)

Demographic monitoring involves marking and monitoring the fate of individuals through time. It is extremely labor-intensive and represents the most intense level of monitoring that can be used. We do not discuss demographic techniques in this handbook. Good introductions to demographic monitoring for plants can be found in Elzinga et al. (1998), Menges et al. (1986), Menges et al. (1990), and Pavlik (1993). Good introductions to these methods for monitoring animal populations can be found in Burgman et al. (1993) and Crouse et al. (1987).

MANAGEMENT IMPLICATIONS

Allocating monitoring resources is a critical initial stage in the development of a monitoring project. Ranking priorities and selecting scale and intensity are not trivial activities, but are fundamental to the effective design of good monitoring. Using teams and soliciting review will help focus decisions about allocation and will help avoid selecting monitoring methods that fail to provide the information needed for management or that provide unnecessarily expensive information.

CHAPTER 4
Qualitative Techniques for Monitoring



Grus canadensis
Sandhill crane
Artist: D. Andrew Saunders

Qualitative monitoring techniques were introduced in Chapter 3 as a less-intensive alternative to quantitative monitoring techniques. Several chapters in this book are devoted to providing the tools to effectively design and analyze quantitative monitoring projects. Do not let the relative length of the discussion of qualitative techniques mislead you into thinking that they are always less effective than quantitative sampling techniques. In many monitoring situations, they are completely adequate to provide the information needed to make decisions (which is, after all, the entire point of monitoring) and have several advantages over quantitative monitoring using sampling. Qualitative techniques are usually simpler and require less time to plan and implement, may cover a larger area or a greater percentage of the population, are easier to evaluate, and are often easier to communicate to managers and stakeholders. You should always consider a qualitative technique first and determine if it will meet your needs before you begin considering a quantitative monitoring project.

QUALITATIVE ASSESSMENT OF POPULATIONS AND HABITAT

Presence/Absence

Presence/absence techniques note whether the species still occurs at a site. The key advantages are that no special skills are required (anyone who can recognize the species can do the monitoring) and that the monitoring requires very little time. The main disadvantage is that presence/absence observations provide no information on trend, except when the population disappears.

A presence/absence approach may be useful for large or showy plants that grow along roads and are visible during a drive-by visit. Animal species that produce recognizable signs of occupancy such as an active raptor nest or rookery or a prairie dog town may be monitored by presence/absence approaches. Simple check-list surveys have been shown to provide a powerful and simple means of tracking bird population trends over large areas. These surveys simply involve noting which species were present at a site on any given visit and, if repeated sufficiently often, will usually capture important trends in abundance and range shifts (Droege et al. 1998).

You can enlist specialists from other disciplines or volunteers to monitor the presence or absence of the species while they are performing other work. The technique can effectively monitor occurrences across the landscape and is especially appropriate for species with many small but obvious populations.

You can improve the consistency and usefulness of observations with a short form to report population visits. Fields to include are observer, date, and time spent at site. You might also add a field for noting whether the survey was a drive-by or walk-through, a comment field for specific threats or problems, and a field for listing photographs. You will make it easier for others to make observations in the course of other work (and more likely that they will do it) by putting together a packet of maps and data sheets for them to carry with them in their vehicles. We recommended including a map of the entire area showing population areas marked in red and the outlines and names of all overlying topographic quadrangles. This large-scale map should be accompanied by a packet of photocopies of portions of topographic maps, each clearly labeled (e.g., "lower right of Cobalt Quad") and with the population locations shown. Make it easy to flip through (use 8.5" × 11" sheets in a binder) and easy to locate things (e.g., alphabetic tabs for the photocopied topographic maps).

Estimates of Population Size

Estimates of population size require only a small amount of additional time and effort over that needed for presence/absence. The advantage of estimates is that they provide a gross index of population trend. The key disadvantage is that, because of variability among observer estimates, only large changes can be reliably detected.

Establishing some guidelines will improve the repeatability of estimates. You will need to decide, for example, if all individuals will be included or only large or reproductive ones. Estimates that include small, cryptic individuals can be especially variable among observers. Conversely, estimates that include only reproductive individuals may vary year to year because of the variability of reproduction in response to annual weather patterns and other factors. The best choice of which types of individuals to include in a visual estimate of population size will depend on the ecology of the species and the situation, but you must ensure that the counting units are specifically identified and can be consistently applied by all observers.

If the population is very large or spread over a large area, consider using smaller, defined subunits. For plants, you may wish to mark several macroplots in which the number of plants is estimated. These should be small enough that an observer can view the entire macroplot from a single vantage point. For animals, "walk-throughs" of an area, looking for individuals or signs of their presence (e.g., tracks, scat, active burrows, nests, pellets) may suffice, as long as the area of interest is thoroughly covered (usually with multiple passes); specialized habitats of the species are known, identified, and inspected; and the inspections are done during the season and time of day when the species is present and active.

Another option in estimating population size is to use classes rather than to require the observer to provide a number. In most situations where this approach is used, class boundaries are closer at the low end (e.g., 1–3, 4–10, 11–30, 31–60, 61–100, 101–200, 201–500, 501–1000, 1001–5000, and so on). A logarithmic series (1, 2, 4, 8, 16, 32, etc.) has also been used (Muir and Moseley 1994). An alternative logarithmic series sometimes used is 1–10, 11–100, 101–1000, etc. Note that at low numbers you could simply count individuals rather than estimate.

Estimation of Population Condition

You can develop standard field observation sheets to aid observers in making consistent notes about population condition. The types of data fields included will vary by species, habitat, and situation. An example of an observation sheet for plant populations is shown in Box 4.1. Examples of potential fields for animal populations include the following:

- Number of individuals
- Number of juveniles, immatures, and adults
- Number of females and males
- Number of breeders versus nonbreeders
- Number of young associated with adults
- Evidence of activity (scat, browse, pellets, burrows, nests, tracks)
- Evidence of limiting agents (scarcity of food or water, predators, competitors, exotic species, poaching activity)
- Habitats associated with juveniles, immatures, and adults

Site Condition Assessment

This technique evaluates the condition of the habitat through repeated subjective observations. Assessments can focus on a single activity, potential disturbances, or site characteristics.

Existing conditions may have to change dramatically before verbal descriptions clearly show that a change has occurred. Training of observers and the use of photographs illustrating condition categories may reduce between-observer differences. Because of variability of visual estimates among observers, site condition assessments are often more effective at capturing the appearance of a new disturbance than at estimating changes in an existing disturbance. Observers

Box 4.1. *RARE PLANT—QUALITATIVE OBSERVATIONS*

Species: ______________ Date of Observation: ____ Observer(s) ______________

New Record? ❑ Yes ❑ Unknown ❑ No

Element Occurrence Record Number (Heritage or other State System Identifier):

Legal: T R S 3 3 3 T R S 3 3 3 County:

Driving and walking description:

Topo quad name(s): ____________ Attach a topographic map (7.5 min) of location

Landowner: ____.

PHOTOGRAPHS

Roll/Frame	From	Toward	Description

POPULATION DESCRIPTION

Population area (estimate size and draw on topographic map) ❑ 1 m^2 ❑ 10 m^2 ❑ 100 m^2 ❑ 1000 m^2 ❑ 1 ha

Population Size ❑ Count or ❑ Estimate? Give counts or estimates by age class:

___juveniles (defined as: ______________________________________)

___nonreproductive adults (defined as: __________________________)

___reproductive (__flower, ___fruit, or ___both)

___senescent (defined as: ______________________________________)

Herbivory: Percentage of individuals affected: ____________________

Average amount of each individual affected: ____________________

Domestic or native grazer? Sign? ______________________________

Insect impacts? ____

Disease? ____

Collections: (give collection numbers):

Other notes:

may, however, miss new conditions for several visits until they become obvious. A careful observer, for example, may note an exotic weedy plant species when there are only a few plants, but many observers will miss an infestation until it becomes quite large.

Site condition assessments are most effective when observers articulate their qualitative assessment quantitatively. For example, requiring an observer to estimate the size or areal extent of a weed population, even using broad size classes, provides a better measure of the situation than general descriptive terms such as "common."

Site condition assessments should be done with a standard field sheet used each time the study area is visited. Standard fields and questions should prompt the observer to look for certain conditions and to assess conditions in as quantitative a manner as possible.

The types of observations are specific to the habitat, species, and issues; thus, a specifically tailored field sheet must be developed for each situation. Examples of data fields for site condition assessment of plant and animal populations include the following:

- Associated vegetation (successional changes)
- Exotic species
- Fire
- Flooding
- Slope movement
- Animal disturbances (burrowing, trampling, herbivory)
- Mining (exploration, material removal, other)
- Logging
- Domestic livestock grazing
- Off-highway vehicles
- Recreation
- Road construction or maintenance
- Weed control
- Condition of fences
- Signing
- Condition of road barriers

Boundary Mapping

Boundary mapping involves measuring or monumenting the boundaries of the population and tracking changes in spatial location or size. Highly accurate maps illustrating boundaries and features of populations can be generated by computer-aided drawing and design programs (CADD) and standard survey equipment such as a theodolite or transit with an electronic distance measurer (EDM) or a global positioning system (GPS) (see Chapter 5).

For some species, mapping the locations of population areas on a low-level aerial photograph may be adequate. For example the plant, *Primula alcalina*, an eastern Idaho endemic, is found on low terraces associated with spring-fed streams. These habitat areas are fairly small (ranging from 10 to 200m^2), but can be easily distinguished and located on a 1:4000 scale aerial photograph. To monitor this species, all population areas or clusters within a 250-hectare meadow were mapped. The presence and size of each cluster will be monitored by periodic remapping (Elzinga 1997).

PHOTOPLOTS AND PHOTOPOINTS

Photoplots

Photoplots are photographs of a defined, small area (a plot), usually the size of the photograph frame or slightly smaller, taken from above at a specified height to provide a "bird's-eye view" of the plot. Photoplots can be an effective qualitative record of condition within the limited area of the plot from year to year. Their key value is in forming a permanent visual record of the past, allowing factors and changes to be evaluated that might not have been considered when the monitoring was initiated. Photoplots can be used to evaluate invasion by exotic or weedy plant species, successional changes, soil disturbance, trampling, and other changes in habitat.

Photoplots are usually defined on the ground with a standard-sized frame. Typical ones are shown in Figure 4.1. A permanent monument in two corners of the frame ensures that the same area is rephotographed every year.

If you can identify and count individuals within the plot on the photo, photoplots can be used to measure density. Photographic density plots are advantageous in situations where field time is limited, perhaps because of shortness of season or difficult accessibility. The required field time is reduced, but recognize that by deferring counting of individuals to the "slow" time of year, the total time, including office time, will be much longer for each unit using this approach compared with completing counts while in the field (Bonham 1989).

When using photoplots to measure density, test the method on the target species before using it extensively. Serious problems often appear unexpectedly. Individuals are usually less obvious on a photo than they are in real life (enlarging the photo or projecting it as a slide onto a screen can sometimes help). Counts will likely also be underestimates of total density because individuals hidden under taller plants will not be counted.

You can also use photoplots as permanent measurement units for cover. Again, the total time required is much greater using photoplots for cover than measuring or visually estimating cover in the field because of the increased processing time in the office, but the approach may be useful where field time is limited.

Cover can be measured on a photo in two ways. One is to lay a grid over the photo with a known number of intersections, and note the number of "hits" on the target species. The drawbacks of this approach are that species with low cover may be missed completely (Foster et al. 1991; Meese and Tomich 1992) and that it may be difficult to identify small individuals (Leonard and Clark 1993).

Another method is to define canopy polygons on the photo and to planimeter the area encompassed by the polygons. The drawbacks of this approach are that plants with lacy canopies are usually overestimated and that boundaries may be difficult to delineate for irregularly shaped species (Winkworth et al. 1962). If the overestimation is consistent from year to year, it will not affect the monitoring value of the method (because trend is what is of interest), but observers will probably draw polygons around lacy or open canopies differently.

The scale of the photograph will affect the estimate of cover. If, for example, the photographic scale was 1:100, the ground area covered on the photograph by even a small-diameter pin or crosshair would be very large, thus dramatically overestimating cover. The smaller the relative surface covered by a pin or crosshairs, the closer the measure will be to the true cover. If cover is measured on the projected image of a slide, the pin or crosshair bias will vary depending on the projected scale. The ratio of pin area to ground surface area should be minimized, and the scale used in the photographs or projection should be kept constant throughout the monitoring project.

Using photoplots to estimate cover or density in a sampling study by using each photoplot as a sampling unit is unlikely to be successful. In practice, the small square or rectangular plot forced by the field of view of a camera is usually an inefficient design for measuring cover or density because of the spatial variability of biological populations (this is covered extensively in Chapter 8). You would likely have to take many pictures to achieve an adequate sample size.

Figure 4.1. Examples of photoplots that have been used in rangeland studies by the U.S. Bureau of Land Management. Frame size and shape will depend on vegetation characteristics, objectives, and camera lens size.

Several photoplot methods have been published. Schwegman (1986) describes a frame made of PVC pipe. A camera with a 28mm lens is suspended on the frame 1.4m above the ground. The camera frame is attached to a $1m^2$, gridded frame that rests on the ground surface. Frames can also be constructed to suspend the camera over an offset plot, so that the observer can remain a few meters away and not trample the area near the plot (Windas 1986).

Stereo pairs can be made of photoplots with a stereo adapter for the lens or by taking two frames. This will enable you to view a three-dimensional perspective of your photoplot. Wimbush et al. (1967) used two cameras with 28mm lenses, placed 76mm apart, for a stereo pair of a 125cm × 80cm plot from a height of 120cm. Ratliff and Westfall (1973) placed a camera

with a stereo adapter about 130cm above the ground surface to photograph a stereo pair of a square-foot frame. This gave about a 1:7 scale on a standard 3.5" × 5" photograph. Wells (1971) used two cameras, each with a 25mm, wide-angle lens, mounted 15cm apart, to make stereo pairs. The frame supporting the cameras was 132cm above the ground, resulting in a stereo frame of a quadrat 1m × 1.5m. Pierce and Eddleman (1970) created stereo pairs of a $1m^2$ plot by taking two frames, 18cm apart, with a camera with a 55mm lens, suspended 152cm above the ground.

Photoplots taken with a telephoto lens may be especially effective for large plants growing on steep, erosive slopes that cannot be physically accessed by an observer. In a study by the U.S. Bureau of Land Management near Salmon, Idaho, a series of photographs were taken with a 400mm telephoto lens. Each series formed a long, linear, photographed plot from the bottom to the top of the slope. Several of these permanent photographed plots were established. The target species, *Physaria didymocarpa* var. *lyrata*, was clearly visible on the sparsely vegetated slope, especially when in bloom. Individuals could be relocated from year to year, and the total number of individuals counted within each plot. Although this was a good idea in theory and one that worked well most years in practice, in some years poor retakes (either because of poor-quality photographs or failure to retake the exact same frames) resulted in complete loss of monitoring data for an entire year (BLM Salmon Field Office, unpublished data).

Photopoints

Photopoints are landscape or feature photographs retaken each time from the same spot and filling the same frame so that differences between years can be compared. Use photopoints abundantly as a standard part of monitoring for documenting the following:

1. Location of study site. Consider taking photos at the parking spot and along the walking path to the study site. At the monitoring site, photographs taken from the boundary of the population or study site facing both toward and away from the site can help relocate boundaries if other monuments are lost.
2. Transects and macroplots. Photographs taken at each end of a transect or at the four corners of a macroplot can help to relocate the transect or plot and provide a visual record of general conditions.
3. Habitat conditions. Photographs of the site can help assess changes in plant cover, weed invasion, and disturbances.
4. Population conditions. Plant height, flowering effort, plant size, and levels of herbivory are some of the conditions that can be illustrated with photopoints.

Todd (1982), Rogers et al. (1984), and Brewer and Berrier (1984) provide overviews and suggestions for establishing and using photopoints. Two examples of the use of photopoints for monitoring long-term change are Sharp et al. (1990) and Turner (1990). Hart and Laycock (1996) provide an annotated bibliography of 175 publications that use repeat photography, giving the number of repeat photographs, dates, habitat type, and location of each study.

Hints for Monitoring with Photopoints and Photoplots

1. A good, 35mm camera is essential for quality monitoring photographs. A camera that allows control of both shutter speed and aperture is best. Disposable cameras are convenient, but should only be used for recording images where detail and sharpness are not critical such as photographs of the parking area and the route to a monitoring site. Digital cameras may become a more widely used option as they become more affordable, but digital images may be suspect because of the potential for alteration.

2. Lenses should be chosen with care. Generally, lens sizes of 35 to 75mm are appropriate for photoplots, and lenses from 50 to 105mm for photopoints. For photoplots, a wide lens is best. These open up to f1.6 or f1.8, allowing you to take quality photographs in low-light conditions. The wide diameter of the glass allows maximum light to pass through the lens and can dramatically improve the quality of the photograph and the depth of field (see #4, below). These lenses are more expensive, but may be worth the investment if photoplot monitoring will be extensive or if quality is critical. Avoid fish-eye lenses (<35mm) because of the distortion. Also avoid telephoto lenses unless they are specifically required for a given project; these generally do not give as sharp an image as smaller lens sizes and do not function well under low-light conditions.
3. Some cameras come with lenses that zoom in and out with the touch of a button, but the actual focal length is unknown. It is difficult to retake the exact same frame with this kind of camera. Even with a manually operated zoom lens, it can be difficult to get the exact focal length unless you are at an end of the zoom scale. Standard lenses, rather than zoom ones, provide better repeatability.
4. Use the smallest aperture (the largest f-stop) possible, given the light conditions and restrictions on shutter speed. Small apertures increase the depth of field of the photograph, meaning that a larger range of distances from the camera is in focus. This can be especially important for monitoring photoplots.
5. Use the slowest shutter speed possible to maximize the depth of field. Shutter speed should probably be no slower than 1/60th of a second unless the camera is supported by a tripod and the air is very still (no moving vegetation). If even a slight breeze is blowing, increase the shutter speed to reduce blurring caused by moving vegetation.
6. A tripod improves the quality of photographs in nearly all situations and is especially critical for low-light conditions (such as dark woods). A tripod can also help to maintain a standard camera height, if this is recorded. This reduces the different camera angles caused by varying heights of different photographers.
7. Take three or four frames of the same picture, each at a slightly different exposure. Multiple frames are cheaper than return trips to retake photographs because the first ones are all overexposed or underexposed.
8. Most professional photographers prefer slide film to print since both high-quality prints and slides can be made from slide film. Slower films (ASA 25 to 100) give better clarity and less graininess, but faster film (ASA 200 to 400) may be needed for shady areas.
9. Use the first frame of a series as a record frame (a picture of a clipboard with date, time of day, location, and subject). This will save many hours of trying to match boxes of slides with field notes. Use chalkboard or beige paper with the information written in heavy black marker. Avoid reflective, white, dry-erase board or bright white paper. These are often unreadable in a photograph because of glare.
10. Use record frames whenever changing subjects, locations, or film. The first frame of a film should always be a record frame.
11. If photographs will not be curated immediately, include a record board in a bottom corner of each frame.
12. When taking general landscape photographs, include enough horizon in the picture to aid relocation.
13. If the photopoint picture does not include any horizon (e.g., pointed at the ground) use pairs of photographs—the first from the photopoint containing something recognizable or the horizon, and the second of the desired frame.

14. Map photopoint locations on an aerial photograph or a topographic map. Use a symbol that illustrates both the photopoint and the direction of the photograph.
15. Some studies have used permanent monuments such as rebar or T-posts to mark photopoint locations. These are recommended for situations that lack visual indicators to use in relocating photopoints. A riparian area, for example, probably contains diverse enough habitat features to allow you to relocate photopoints by using the previous photographs. In other types of sites such as a large meadow, a dense forest, or a sagebrush grassland site with little topography, photopoints may be difficult to relocate from the photograph. In these situations, a permanent monument can save much time. Geographic Positioning Systems (GPS) can also be used to map and relocate photopoints, but if accurate units are not readily available to all field-going personnel, a physical marker is probably better.
16. Keep a photo log in your field notes in which you write the film number, description, and location of each photograph or series of photographs.
17. Invest in a camera that records the date on every photograph.
18. Curate photographs immediately after developing. Write identifying information in pencil on a label placed on the back of each photo. You can write directly on the slide frame. You can also purchase special pens from photo supply companies designed for writing on the back of photographs. Do not use pen or marker; these may bleed through the photograph, or smear onto another photograph if photos are stacked. Invest in archival-quality plastic sleeves for photographs, slides, and negatives, and store them in labeled, three-ring binders. Photographs kept in boxes or envelopes are seldom looked at again.

VIDEO PHOTOGRAPHY

The most common use of video photography is as a visual record of the site, similar to the use of still cameras and photopoints. Video can provide a good visual overview of the site (and verbal as well, if commentary accompanies the film), providing a better sense of features and conditions than still photographs. The disadvantage is that video footage is difficult to retake since a video cannot easily be played in the field to compare with current conditions. This drawback can be overcome, however, with the creation of stills from the portions of the video that best represent the features that are being monitored. These stills can then be retaken with a regular, 35mm camera equipped with a zoom lens to match focal length of the video image.

Plant cover in quadrats or transects can be recorded by video photography. This application has been most widely used in marine studies (Whorff and Griffing 1992; Leonard and Clark 1993), primarily because diving costs associated with underwater sampling are expensive. Video requires extensive laboratory time for processing and analyzing the images, but only a fraction of the field time that most other sampling techniques require (Leonard and Clark 1993). The drawback of using video for sampling vegetation is that the resolution of the image may make species identification difficult and may limit the detection of small species (Leonard and Clark 1993).

REMOTE SENSING TECHNIQUES

Remote sensing encompasses a range of techniques involving the collection of spectral data from a platform that does not touch the object of interest. This definition is somewhat vague because of the variety of remote sensing techniques, from taking a photoplot with a camera suspended from a hot-air balloon to satellite-based imagery.

Satellite imagery includes several types of spectral data and several platforms. In general, the resolution of satellite imagery does not lend itself to the site-specific monitoring situations addressed in this handbook. Although there may be some use for satellite imagery to identify community types known to contain the target species, gap analysis tests have shown that the level of resolution is often inadequate to identify small habitat islands (Stine et al. 1996). Thus, the use of satellite imagery to identify or stratify habitat should proceed with caution. This may, however, be changing as satellite imagery technology improves. The IKONOS satellite launched in September 1999 is capable of $1m^2$, black-and-white resolution and $4m^2$, multispectral resolution. This is adequate for detecting disturbances such as OHV tracks and identifying distinctive habitat islands, but its use may be limited by its expense. For some examples of the use of satellite imagery in landscape-level monitoring, as well as excellent overviews of applications of satellite imagery in natural resource management, see Frohn (1998), Holmgren and Thuresson (1998), Lillesand and Kiefer (1999), Luque et al. (1994), Sample (1994), Lyon and McCarthy (1995), Verbyla (1995), Wilke and Finn (1996), and Wynne and Carter (1997).

Aerial videography can be used for mapping and monitoring landscape features such as plant communities. Most of the systems described in the literature involve the use of video cameras associated with low-level, fixed-wing flights (Bartz et al. 1993; Nowling and Tueller 1993; Redd et al. 1993). The cameras record spectral data (not necessarily from the visual range), which are immediately processed into digital information associated with a certain pixel size. Pixel size is determined by flight altitude. These pixels are then classed based on their spectral signature. Pixel size ranges from 3m × 3m to 50cm × 50cm, with the cost increasing as the pixel size decreases.

Aerial photography captures visual spectral data (sometimes infrared), generally from a fixed-wing aircraft or helicopter. Most agency offices have access to recent air photo coverage of their entire administrative unit at 1:12,000 to 1:24,000 scales. Photo series are usually in stereo-pairs, which, with some practice, can be viewed three-dimensionally through a stereoscope (see Chapter 5). The National High Altitude Photography Program (NAPP) provides complete coverage of the United States at 1:40,000 scales (with object resolution of 1 to $2m^2$). NAPP repeats these photos about every 5 years. USGS uses these to produce Digital Orthophoto quads, also at $1m^2$ resolution.

These photographs can be extremely valuable for identifying community and population boundaries, for stratifying sites, and for documenting study locations. Aerial photographs can also help identify features and disturbances that are not apparent from the ground. In some offices, older photo series may be available to compare with newer ones. These comparisons can provide an historic perspective on changes in disturbances and human use, ground cover, and even species composition.

Low-level aerial photography (scales of 1:500 to 1:6000) may be commissioned for a specific project. Although expensive, if low-level aerial photographs can be used for monitoring in place of ground measurements, the savings in personnel time may make aerial photography competitive with more conventional monitoring techniques. For plant monitoring, low-level photography is especially applicable to woody species, very large herbaceous perennials, and overall community cover and habitat condition assessments. It obviously will not work well for small species or for a species obscured by a taller canopy. For animal species, low-level photography can be used to monitor habitat, as well as activity for some species (such as prairie dogs). Similarly, it has been used to estimate nesting populations of herons and seabirds, aggregations of waterfowl and large mammals, and to track locations of vernal pools important to breeding amphibians. For examples of low-level photography successfully used to monitor a community, see Knapp et al. (1990) and Jensen et al. (1993).

We also recommend Avery and Berlin's (1992) *Fundamentals of Remote Sensing and Airphoto Interpretation* as a guide to photo interpretation. This book has over 440 black-and-white photographs as examples, including 160 stereo-pairs and 50 color photographs. The layout and fascinating photographs make this book attractive and extremely readable. Of primary interest

for vegetation monitoring is a chapter entitled "Forestry Applications," which also addresses community cover mapping from aerial photographs in nonforested types. The authors describe how individual range plants and grassland types can be identified at scales of 1:500 to 1:2500 and individual trees and large shrubs at 1:2500 to 1:10,000. Typical diagnostic features used to identify different species are plant height, shadow, crown margin, crown shape, foliage pattern, texture, and color. For some forested areas, diagnostic keys to species identification on aerial photographs have been developed; the authors include references to these guides.

OTHER REMOTE MONITORING SYSTEMS

For some animal species, remote monitoring systems have become increasingly useful for detecting simple presence or absence at a site. These systems include weatherproof cameras and infrared-beam counters triggered by animals passing nearby, and audio recording systems that activate regularly for short periods and record ambient sounds. Remote camera systems have been useful for documenting and sometimes estimating (through mark-recapture based on distinct body markings) populations of large and elusive carnivores and forest ungulates (Foresman and Pearson 1998). Remote counter systems also have been evaluated (see Garner et al. 1995). Automated sound-recording systems have been useful in some locations for monitoring frog populations via call-based indices of frog abundance. These systems can be preprogrammed to record at intervals throughout the nocturnal calling period, thereby greatly reducing counting effort and disturbance. Note, however, that with both remote camera systems and automated sound-recorders, considerable effort often is expended to process data (photos, tapes) out of the field.

MANAGEMENT IMPLICATIONS

Qualitative techniques can be very effective monitoring tools and can provide adequate data for making management decisions in many situations. Adequate qualitative methods are almost always preferred over quantitative because of their low cost. Qualitative techniques include estimates of species or site conditions, photographic records, and remote sensing techniques.

CHAPTER 5
General Field Techniques

Atragalus convallarius var. finitimus
Pine Valley Mikvetch
Western Utah and eastern Nevada
Artist: Jeanne R. Janish

A suite of field techniques are useful to employ for all monitoring projects. These include adequate monumentation of study areas and study sampling units, using available tools, and making good collections of plant and animal material for later reference.

BASIC FIELD EQUIPMENT

Box 5.1 lists relatively inexpensive field equipment that is useful for monitoring. We recommend storing all or most of the equipment in a box that can be taken to the field. This precludes forgetting an item and allows trial of other methods if the planned method fails during the pilot study.

Box 5.1. *STANDARD USEFUL FIELD EQUIPMENT*

100-m tape (two to four)
10-m tape (two to four)
30-cm ruler
Aluminum tags and nails
Ball of string, brightly colored
Binder clips, large (can be used to hold tapes to intermediate monuments)
Binder clips, small (to hold field sheets together and to your clipboard)
Binoculars (7×–10× magnification)
Calculator with statistical functions and random-number generator
Camera and three rolls of film
Chaining pins (for holding tape ends)
Clinometer
Clipboard
Compass
Data sheets
Diameter tape (if you work in wooded systems)
Ensolite pad or knee pads (for cushioning)
Field notebook
Flagging (at least four rolls, two each of an unusual color or pattern)
Graph paper (for creating impromptu field sheets)
Hammer
Hand lens
Mechanical pencils (three)
Meter stick
Nails (large enough to serve as markers or to hold down tapes)
Newspaper for plant press
Permanent markers, two each of two colors in both thick- and thin-tipped (eight)
Photo-ID clipboard sheets
Pin flags
Plant press
Plastic bags (several ziploc and a few garbage bags for storing collections)
Plumb bob
PVC frame (adjustable into square plots ranging from 10cm × 10cm to 1m × 1m)
Rebar of varying lengths, 20cm to 50cm long (at least 30 pieces)
Screwdrivers (two) for securing temporary tapes
Small hatchet or machete (for work in forested areas)
Spray marking paint (two colors)
T-posts (several)

Measuring Tapes

Tapes come in a variety of lengths, increments, materials, and cases. The type of tape you wish to purchase is largely a matter of personal preference. Some considerations are as follows:

Length

Short tapes are less expensive, lighter, and easier to use than long tapes, but you will usually need some long tapes for monitoring. At least one tape that is a minimum of 100m long is recommended.

Increments

Tapes can be purchased in both English and metric units. We recommend all monitoring studies use metric units, to allow for potential publication and exchange of information. Metric is the unit of choice for scientific studies. Government agencies in the United States, especially in the forestry and range shops, have long used English units, but those conventions are changing. Tapes may be purchased with English units on one side and metric units on the other.

Increments can be in millimeters, centimeters, and decimeters and with marks at the meter and half-meter points. The most versatile tape for field use has centimeters marked and numbered, decimeters identified with heavier marks and numbered, and meters numbered and marked with the heaviest marks and/or alternative colors. Tapes that are not numbered at every increment take longer to read and increase the risk of error.

Materials

Tapes come in steel, fiberglass, and cloth. Steel is most accurate over the life of the tape. Stretching is virtually nonexistent. The tape length will change depending on temperature, but this amount of error is insignificant for most monitoring work. Steel tapes are the tape of choice for work requiring extreme accuracy. They may be appropriate for permanent transects, where repositioning the transect in exactly the same place is important. Steel tapes are expensive, heavy, and difficult to use because of kinking. Some steel tapes come with a nylon coating that reduces the tendency to kink.

A universally useful metal tape is the steel loggers tape, which has a hook or a ring at the end and a retractable case. These tapes are useful if a number of measurements will be taken and time spent unwinding and winding a tape would be burdensome. Hooks can be set into a tree or a stake and released with a flick of the wrist, retracting the tape automatically. These features may make the relatively expensive loggers tape worth its initial high cost. We recommend wrapping the first 10cm of the tape with electrical tape or some other protective tape to prevent the end of the tape from wearing as it snaps back into the case. Another useful tip is to replace the standard hook at the end of the tape with a bent horseshoe nail. These press into a tree more easily and provide a more controlled release than a standard hook.

Fiberglass tapes will stretch over the life of the tape and when under tension. The amount of stretch is related to material, age, use, and tension. Some manufacturers offer fiberglass tapes with as little as 0.01% stretch per pound of tension over 4.5 pounds, a standard similar to steel tapes. These tapes are advertised to retain accuracy over the life of the tape. Fiberglass tapes are light, durable, and easy to handle. Cloth tapes are also light and easy to use, but are less durable than fiberglass and are prone to substantial stretching.

Cases

Most tapes come in open reel cases, allowing for rapid pick-up. Some tapes come in enclosed metal cases, which provide better protection, but, if the tape twists inside the case and binds up, make it difficult to repair. Surveyor's rope, or rope chain, lacks any case because it is designed to be pulled from site to site rather than rolled up. These usually come in 50m to 100m lengths. Their advantages are that they are designed to withstand dragging and can save roll-up time when sites are close together. They can be coiled rather than reeled up and are generally lighter

than reel tapes. The disadvantage is that surveyor's rope is usually marked in increments of 5cm or 10cm, which is not useful for making measurements to the nearest centimeter.

Paint and Flagging

Paint specially designed for outdoor marking use is available from forestry supply companies. These paints are longer lasting than regular paint, although more expensive. You may wish to choose an unusual color (rather than the standard orange or red). Florescent colors are recommended; they are easily spotted and can even be seen by color-blind workers. Yellow paint can provide an intermediate color choice for stakes that need to be relocated, but should not be glaringly attractive to vandals. Avoid blue paint in forested areas because it is a standard color for marking cut trees on logging units (at least in the United States).

Flagging comes in a wide variety of colors and patterns; a stock of an unusual type can be useful for unique markings in a project area full of orange and pink flagging. Biodegradable flagging is available, but it is more likely to be eaten by animals and often becomes brittle, breaking in cold weather.

Compass

A compass that will be used primarily for route finding should have the following characteristics: 1) a mechanism for adjusting declination; 2) a moveable, transparent housing with vertical lines to aid in map work; 3) azimuths by degrees, 0 to 360; and 4) a folding mirror to increase accuracy of sightings. Most compasses used by resource specialists are of this type.

A compass that will be used primarily to make sightings on objects should be an optical bearing type. These are similar in construction to clinometers, with a liquid-filled compass dial encased in a block of plastic or aluminum with a viewing hole in the otherwise featureless housing. To use these, you sight on an object while reading the azimuth through the viewing hole. While this type of compass provides very accurate azimuth sighting, it is difficult to use for map work because it lacks the built-in protractor and movable housing of the folding mirror type of compass, and declination is not adjustable (0° [360°] always reads magnetic north).

Binoculars

Binoculars are required for many animal studies, such as those that involve identifying birds or butterflies from a distance. Binoculars are also often surprisingly useful in plant studies, and may be necessary for identifying tree species when the leaves are high above the ground and for locating epiphytes in the canopy. Binoculars should be at least 7×–10× power, preferably close-focusing.

Field Notebooks and Data Sheets

While data from most monitoring studies will be collected on preprinted data sheets (see Chapter 6), field notebooks are still needed to record general observations and notes. It is strongly recommended that you keep a field notebook as a log of daily field activities. Field notebooks are also necessary for recording information on collections (see section on Collecting and Pressing Plants) and photographs.

Field notebook systems vary from biologist to biologist. Standard, bound, field notebooks and binder-system notebooks are available from forestry supply companies. Bound notebooks are more durable than binder systems and keep all data in a single notebook, allowing ready reference to earlier field notes. The potential exists, however, for loss of an entire field season's worth of notes. At a minimum, photocopy field notebooks daily if possible, and store photocopies in a safe place.

Binder systems are less durable than bound notebooks, and pages can tear out and be lost. This disadvantage may be offset by three advantages over bound notebooks: 1) Sheets can be removed after each day's field work and stored in a safe place. This can be especially important if field work is done in a remote place where daily photocopying is not possible. 2) Binder sheets

can be used in a laser printer to prepare preprinted sheets. You may wish, for example, to have a section in your field notebook just for tracking collections or photographs. A preprinted entry sheet can save time for these standard types of field notes. 3) Binder systems are less expensive and can be used over several field seasons by purchasing more filler paper.

Both bound and binder-system notebooks are available with waterproof paper. This paper is recommended even in arid climates where a field day in the rain is rare, since the paper will not be destroyed by being dropped in a creek or soaked in a backpack by a leaky water bottle. Waterproof paper can also be used in a laser printer or photocopy machine. The laser and photocopy ink will not smear, fade, or run on this kind of paper.

Waterproof paper is best matched with special pens, available from forestry supply companies for $5 to $10. These pens will not smudge, rub, or wash out. Standard ink pens should never be used; they can bleed and will wash out if the field notes are soaked. Pens are recommended as standard field and scientific practice. Incorrect entries and notes should be struck with a single slanted line (for one character) or a single horizontal line for several letters. Also suitable for field notes are hard pencils (number 4 or 5 leads), which will actually make an impression on the paper. Soft pencils (standard number 2 leads) are not suitable for any type of field notebooks or data sheets because the lead can fade and smudge to the point of illegibility and will become unreadable if wetted.

Handy Tools

Clipboards

Look for a clipboard that contains an area for storage of additional data sheets and a metal cover that can be quickly flipped over the data sheet in the event of inclement weather.

PVC Frames

You can make frames from PVC pipe and connectors, materials available in most hardware stores. Cut pipe into lengths that can be put together with the connectors into various frame sizes (e.g., 10cm × 10cm, 50cm × 50cm, 50cm × 1m, 1m × 1m). Construct the frame so the inside of the frame is the correct size.

Aluminum Tags

You can purchase tags of soft aluminum wrapped around cardboard or wood on which a physical impression can be made with an ink pen or other sharp object. You can also purchase prenumbered tags.

Pocket Stereoscope

This tool enables you to look at stereo pairs of aerial photographs in stereo (three dimensions). It can be handy in the field for locating study plots on aerial photographs when landmarks are scarce.

Rock Picks

This geologist's tool can be very handy for digging up plants to collect as herbarium specimens.

Hip Chain

A hip chain is used primarily by foresters and surveyors. It is a box, worn on your belt, that measures the amount of fine string that is fed out as you walk, thus enabling you to measure long distances without using a tape or counting paces. Measured distances are not accurate enough for fine measurements, but the hip chain can be an excellent tool to measure distance from a known landmark to the study site or to provide rough measurements of population boundaries. The main advantage is that measurement does not require using your hands or keeping track of paces.

Clinometers

These look similar to an optical bearing–type compass. They measure slope, heights, and vertical angles.

Plane Table and Alidade

This is an old surveying tool, but may still be very useful in places where good maps are rare and GPS systems are unavailable. It can also be used to draw an accurate map of a small study site. A plane table is a level mapping surface, about a half a meter square, attached to a tripod. The alidade is basically a telescope mounted on a straight edge. The plane table is set up and leveled at a central location. Direction to a point is measured using the alidade. Distance can be measured with the alidade and a stadia rod based on the principle of similar triangles. Straight edges on the mapping table are used to generate a hand-drawn map to scale.

Reinhardt Redy-Mapper™

With this tool you can quickly and easily map population boundaries in the field to scale. It is essentially a pocket plane table that hangs around your neck. It consists of a 25cm × 25cm sheet of hard plastic with a translucent disk attached at the center. The disk accepts pencil lead and is the drawing surface for the project (it can be cleaned after the map is transferred to paper). Angles and distances are determined by compass and tape or by pacing; the mapping tool facilitates translating those angles and distances to scale. The Redy-Mapper can be used while traversing the boundaries of a population or habitat area, or it can be used to map from a central location. Although this tool has been partially replaced by electronic tools (Electronic Distance Measurers, GPS units, and computer-generated maps), it is still useful. It is much less expensive, quicker, and more accurate than many of the available GPS units. With this tool, you can map a population boundary almost as fast as you can walk it.

Electronic Distance Measurer (EDM)

An EDM is a survey tool that reads distance and direction between stations and records the values electronically. The information from an EDM can be downloaded into standard survey software or drawing programs to generate maps. The instrument is fast and precise, but requires user training both in field techniques and software applications. An EDM can be used to map population boundaries, permanent plot locations, individual plants, and site features.

Global Positioning System (GPS)

These electronic systems interface with satellites to enable the user to locate or relocate a spot on the earth's surface. Their accuracy depends on the system and access to satellites. Expensive, survey-grade instruments can be accurate to within millimeters, but most units owned by resource management agencies and available to biologists range in accuracy from 0.5m to 50m. Factors such as availability of satellites, controls, terrain and tree canopy cover can diminish their accuracy.

The accuracy of GPS data can typically be substantially improved through a postprocessing procedure that removes bias introduced by the atmosphere and other sources. This postprocessing is a differential GPS (DGPS) technique that corrects bias errors at one location (the area you are monitoring) with measured bias errors at a known position (a base station or reference receiver).

While they are likely not accurate enough to map individuals of many types of populations, the more accurate units may be useful for mapping large, widely spaced, long-lived plants such as trees or cacti. GPS can also be used for mapping population boundaries and locations and for locating sampling units within a sampled area. Many GPS systems can download information electronically into a mapping program to produce accurate figures for monitoring plans and reports.

The technology in this field is changing rapidly, and accuracy of relatively inexpensive instruments is improving quickly. GPS units will likely become more widely available and commonly used in the next few years.

Pocket Electronic Distance Measures (PEDMs)

PEDMs are manufactured primarily for construction use, but have been used for outdoor resource work as well. They are now available from forestry supply companies. Most units rely on sound waves combined with an invisible light beam to measure distance. They can measure distances up to about 80m, less in heavy brush or timber. For outdoor use, best results are achieved with both a transmitter and a reflector (indoors, the transmitter can measure the reflected sound waves from a solid object, like a wall). Accuracy of these units can be as good as 2mm. Accuracy, however, can be negatively affected by other sounds; they work poorly in the rain, along water courses, and in noisy urban areas. The units are also difficult to aim; sighting along the edge of the device can help. To ensure accuracy, three consistent measures should be taken for each distance.

Applications for these instruments are many. Distances from a baseline can be measured without a tape. PEDMs could be used for line intercepts, using the electronic distance from a reflector set up at the end of the transect line. They can also be used to determine limiting distance (plant in or out of plot). Almost any distance currently measured with a tape could be measured with a pocket EDM. The amount of time saved is potentially tremendous, especially for long distances or in dense vegetation where pulling a straight tape is difficult. The price for these units is under $200.

GENERAL HINTS FOR STUDY MONUMENTS

Monuments that permanently mark plots, macroplots, transects, or population boundaries are a critical part of the success of a monitoring project. This section contains some hints to secure monuments.

Assess Potential for Loss

No monument is completely safe, but some are more at risk than others. Visible markers such as brightly painted stakes will always be removed in areas frequented by people. Some markers such as pin flags are attractive to animals and may be pulled up from the ground by deer nibbling on the flagging. Flagging, especially the biodegradable type, is attractive to animals and rarely lasts more than a few weeks in areas with grazers and browsers. Wooden lathe is easily broken and rarely lasts more than a few weeks in areas with domestic or native grazers. Stakes made of PVC pipe may last only a few years because they are photodegradable in bright sunlight and subject to breakage in the winter when the cold makes the material brittle.

Natural catastrophes should be considered. Fire is possible in almost any habitat. Use only metal monuments for studies that are needed for more than 1 year. Do not depend on trees for monuments, although they can be used for a backup.

Stakes and T-Posts

Monuments such as T-posts or fence posts are often stolen. Tall stakes are also attractive as animal rubs and raptor roosts, inviting damage to the monument, as well as increasing animal use within the study area. Using shorter markers (such as rebar no greater than 0.5m high) will at least partially resolve the attractant problem. Cutting the top 12 inches from a T-post reduces its value and lessens the chance that it will be stolen. T-posts should be sunk as deeply as possible in the ground to make it difficult for a casual vandal to pull them out.

Monuments should never be a hazard to people or animals. Safety should be a primary concern in the selection and placement of monuments.

Inexpensive stakes can be made from angle iron, rebar, or aluminum conduit. Aluminum conduit is much lighter weight than rebar or angle iron, so it may be preferable for remote locations where material is packed in. It is also easier to cut in the field than steel rebar. Lengths of 60 to 70cm are good for most soils. Lengths should be shorter if you plan to use them on more shallow soils and longer on sites with deep, loose, or saturated soils. At least one-half to two-thirds of the total length should be below ground. Always be considerate of safety issues. Short stakes are a hazard to livestock, wildlife, horseback riders, and off-highway vehicle (OHV) operators. You can minimize the risk to animals by forming a loop of the upper third of the stake. This can easily be done with a large box end wrench once the stake has been pounded into the ground. If the ground is soft, bend the stakes before pushing them into the ground.

Short stakes can be easily pulled out, or knocked out by livestock, if not sunk deeply enough into the ground. Where vandalism is a problem, special stakes are available that can be used for monuments or anchoring lines. Duck-bill tree anchors have hinged winged plates that are closed as the stake is driven downward, but open up as the stake is pulled up. The lower half of screw or auger stakes look like ice augers. The effort of turning these stakes out of the ground deters all but the most determined vandal. Both stakes are commonly available from forestry supply companies.

Marking Trees

Paint can be used for marking trees if the study is short-lived (<2 years) or if study sites will be visited annually. Use only marking paint designed for outdoor use. Use bright unnatural colors if this does not conflict with aesthetics or attract vandals. More subdued colors can be used; however, because these can be difficult to see, additional travel and monumenting information will be needed. Paint a concentrated spot on both sides of the tree and a ring all the way around the tree below the spots. For longer-term marking, supplement paint with another marker (such as a blaze or a tag), since paint can fade or be sloughed off with bark. Blazes are preferred for marking trees where damage to the tree and visible impacts are acceptable, because they are easily spotted from a distance and will not fall off the tree or be pulled out by vandals. Blazes can last decades on some trees.

You can use tags to supplement blazes and paint with information such as study or sampling unit number. Tags can also be used to mark trees, but they occasionally fall from the tree and are difficult to spot from any distance. Numbered metal tags are commercially available from forestry supply companies. Tags should be consistently affixed at eye level or breast height (4.5 feet). Aluminum nails should always be used because they pose minimal hazard to sawyers or mill operators. Even in areas where future logging is unlikely, remember trees last far beyond the life of many studies and protective designations, and there is no justification for using a nail that might potentially endanger a person in the future. Aluminum nails are readily available from forestry supply companies.

The heads of tag nails should be slanted downward with about an inch protruding to allow for tree growth. This allows the tag to slide to the head of the nail, and reduces the chance that it will be enveloped by bark.

You should identify and map marked trees in the field notes and the methodology section of the Monitoring Plan (see Chapter 15). By including information on species and diameter for each marked tree, it is easier to relocate them later. Tree diameters should be measured at the forestry standard of 4.5 feet from the ground (diameter breast height).

Landmark References

All monuments should be supplemented with references to visible permanent landmarks. Obvious landmarks within or adjacent to a study site, such as a rock outcrop, can be used to identify the location of monuments. Directions from the landmark to the monument should include both

measured distance and compass direction (note whether declination or magnetic). A photograph of the landmark from the monument that includes the monument in the foreground helps in relocation.

On sites lacking nearby landmarks, triangulation can be used to identify the location of a monument in relation to distant landmarks. This involves measuring the compass direction toward landmarks such as mountain tops and permanent (you hope) man-made objects such as water towers or microwave towers. By measuring the direction to two objects, your location on the ground is fixed by the angle formed by those objects and your location. The site can then be relocated in the future. Triangulation is most accurate when the angle formed by the two triangulation points is approximately 90°.

Adding "Insurance"

No monuments that are required for continuation of the study (e.g., permanent quadrat corners or transect ends) should be without insurance, in case the primary monuments are lost. One option is to bury physical markers such as large nails or stakes. A single buried nail next to a monument may be disturbed or dug up when the monument is disturbed. Better insurance is to use four buried nails, each exactly 1 m from the primary monument on the four compass directions. A metal detector can then be used to locate the nails if the primary monument is removed.

A second option is to survey the monument using survey or forestry-grade survey instruments. You can survey the monument from a permanent known point or from two or more inconspicuous secondary monuments.

Photographs of the monument also provide a means to reposition new monuments if the original is destroyed. Take photographs from all four directions. These photographs are also useful as visual documentation of the condition of the site.

MONUMENTING STUDY AREAS

Critical to the success of a monitoring project is relocating study areas. Often study areas are not documented in project notes because the initial investigator assumes s/he will be returning the following year to do the measurements. Many studies have not been continued because study areas could not be relocated after the originator leaves.

Description of the location of study areas should include the following:

1. Driving instructions from a well-known landmark, including direction and mileage to the nearest 0.1 mile. A hand-drawn map is helpful if a number of roads exist in the area, especially if topographic maps are outdated.
2. Walking directions, including compass direction and distance (paced), to study area. Again a map is helpful.
3. Study-site location marked on a topographic map (such as a United States Geological Survey 7.5-minute quadrangle) and on a recent aerial photograph.
4. Compass direction from the study site toward at least three prominent, permanent landmarks such as mountain tops. If dense forest vegetation requires that trees be used, at least six trees should be included, and these trees should be monumented.
5. Photos should be taken as needed, such as at the parking spot, along the path to the site, or in several directions at the study site. Each photo should include compass readings to describe direction of the photo and should be marked for location on the topo map or aerial photograph.
6. If available, a GPS unit may be useful for recording the study location.

MONUMENTS FOR PERMANENT MEASUREMENT UNITS

Often, monitoring will take place within permanent units. Examples include plots that are rephotographed (photoplots), larger plots in which all individuals are counted or mapped, and transects used for counting or estimating cover. Sometimes these permanent units are sampling units used in a study incorporating sampling (see Chapter 7).

Chapters 8 and 9 describe the benefits of permanent sampling units. If the correlation between sampling periods is high (e.g., a quadrat with many plants will have many plants the following year and a quadrat with few plants will have few plants the following year) designs using permanent sampling units require far fewer sampling units compared with designs using temporary sampling units.[1] The increased efficiency depends on the degree of correlation, but for many plant and animal species, the benefits of a permanent design in terms of reduced sampling units can be quite dramatic.

A second benefit of permanent sampling units and the primary benefit for permanent units in nonsampling studies is the increased interpretation of change that results from knowledge of spatial habitat features. For example, if most of a plant population declines except for that portion near a creek, one might suspect drought as a cause. With a permanent design, you know which units are near the creek. With a temporary design you have less opportunity for these types of interpretations.

These benefits must be weighed against the increased cost of establishing and monumenting permanent units. Permanent markers (such as t-posts and rebar) can be expensive and awkward to pack long distances. Placing permanent markers is time-consuming; however, if markers are easily relocated, time may be saved in subsequent years by using permanent sampling units rather than relocating temporary plots each time the monitoring study is measured. In some types of habitats, however, permanent markers may be difficult to find on subsequent years, increasing the cost over temporary plots every year. For example, locating a short stake in a large grassland may be extremely difficult. Metal detectors and global positioning units can help you find such markers, but these add cost and field time. In some habitats such as drifting sand dunes permanent markers are not even feasible and you have no choice but to use a temporary design.

Obviously, if you have designed your monitoring study to use permanent units, you must ensure that the units are actually permanent. Guarding against vandalism and loss may require creativity. You must also use enough monuments to ensure that the same measurement unit is measured each year.

A fence post or stake is usually adequate to mark points such as a photo point or the center of a large measurement unit such as point counts for birds. In habitats with structural variability (such as a streamside area), these points may also be accurately relocated from the photograph itself. In a less-diverse area (such as a large meadow or sagebrush grassland) or a densely vegetated area (such as a shrub community or forest) stakes are usually necessary.

Permanent transects should be marked at each end, noting from which side of the stake the transect originates (especially important for thick markers such as fence posts). Long transects should also include intermediate markers approximately every 10 to 20m. You can weld flat pieces of metal to these intermediate markers and secure the measuring tape to them using binder clips. Gutter spikes provide a wide flat head suitable for securing transects that are close to the ground. Permanent transects too long for measuring tapes (such as animal survey routes) can use fenceposts placed within sight distance of each other to guide an observer on the correct route.

It is probably adequate to monument long, narrow quadrats along the long edge as you would a transect. You can simply reposition the tape along this one boundary and measure the

---

[1] *In a temporary design, new sampling units are located each time the population is monitored.*

plot using a meter stick to determine if individuals are in or out of the far boundary of the quadrat. Quadrats wider than a meter should be monumented along both long sides.

Small permanent plots must be repositioned more precisely than large ones. The reason for this is intuitively obvious if you think of the smallest plot possible, a single point. Even if you intend that 50 point intercepts[2] along a transect be permanent sampling units, a small move of the transect could eliminate all correlation between years for each point. While monumenting individual points in a point-intercept study is not practical (see more on this in Chapter 12), the benefits of using small, permanent quadrats[3] may outweigh their cost. Circular quadrats can be marked with a single marker in the center, whereas square or rectangular quadrats should have at least two corners marked with large nails to facilitate their precise relocation.

COLLECTING AND PRESSING PLANTS

Plastic Field Mounts

Plants that are not succulent or wet from precipitation can be preserved as field mounts (Burleson 1975). Plants are cleaned of dirt and dead leaves and arranged on the adhesive side of a sheet of plastic acetate. A second sheet of acetate is carefully and firmly rolled onto the first. With some practice, good clean mounts can be made using this method. The advantage of this method is that you can mount plants permanently in the field, providing a quick and inexpensive record of species encountered. Leaves and flowers also retain much of their original color. Plants rarely mold (Burleson 1975). A disadvantage is that collections cannot be manipulated (flowers teased open, etc.) for later identification. Only the features that you expose at the time you bind the sheets of acetate together will be visible.

Collections preserved in acetate sheets are not suitable for herbarium deposition because these collections tend to degrade over time (10 to 20 years). Rare species should be preserved following standard collection and curation methods (as described in the next section).

Pressed and Dried Collections

Any collection impacts a population, although impacts of collections made in large populations are insignificant. Rare plants may be especially prone to collection damage because of their small populations and the propensity of botanists to collect rare species. This destructiveness must be weighed against the information gained. Plants should not be collected in populations of less than 50 individuals. In smaller populations, small portions of plants may be collected if absolutely necessary.

It is especially important when collecting rare species to make each collection of herbarium quality. It is senseless to destroy an individual of a rare species if the collection is so poor that it is not worth storing for future use.

Within any population, plants will vary significantly in size and reproductive status. In general, choose individuals that are of moderate size. If you chose an exceptionally large or small individual, note it. Try to choose an individual with both fruit and flowers.

If the plant is small (can reasonably be fitted onto a 11½" × 16½" sheet of herbarium paper), the rule is to collect the entire plant, root and all. An incomplete plant is of little use as a herbarium specimen since any feature may be evaluated in a taxonomic study.

If the plant is larger, collect the entire plant press in portions, for example, lower third, middle third, and upper third. If the plant size is completely unmanageable, for example, a

---

[2]*These are mostly used for monitoring plants and are described in detail in Chapter 12. At each point intercept, you note whether the species you are monitoring occurs under that point or not.*

[3]*Small quadrats are often used in frequency studies, as well as for estimation of cover or biomass. These are described in more detail in Chapter 12.*

woody species, collect branches that contain leaves, fruit, or flowering structures and first-year and second-year bark. The height of the individual from which the collection was made should be noted.

Press plants immediately after they have been collected; plants are much easier to arrange and press if they are not wilted. Collections carried around for a day tend to become pretty bedraggled. If immediate pressing is not an option, place the plants in a vasculum (an airtight metal container designed to hold collected plants). A large tin can or pickle bucket will also work, and, in a pinch, so will plastic bags. The hard-sided container is preferable because it keeps plants from getting smashed, but most botanists use plastic bags simply because they take so little room. Small plastic bags are useful for keeping plants separate by collection site, habitat, or species. Information written in pencil (not pen) will survive several days in a plastic bag with moist plants. Plants placed in the refrigerator or a field ice chest will keep much better than those that get warm. Plants that have wilted in the field can be partially revived by placing them in a plastic bag with some soaked paper towels and storing them in a refrigerator overnight.

Standard plant presses are a sandwich made of two pieces of wood lattice (to allow moisture to escape), with blotters and pieces of cardboard layered inside. The blotters help to absorb moisture, and the cardboard helps to flatten the plants, as well as to let moisture vent. Two compression straps hold the sandwich together very tightly. You can purchase plant presses for $30 to $40 or make them. The standard size is 18" × 12".

The plants are usually pressed within newspapers forming multiple layers of cardboard, blotter, newspaper, plant, newspaper, blotter, cardboard. The keys to a good pressing job are patience, practice, and a penchant for neatness. Roots should be carefully but completely cleaned. Leaves must be individually smoothed flat. Plants with small, compound leaves or many leaves require special care. Leaves should be pressed in a way that represents both upper and lower leaf surfaces. Arrange flowers in various positions showing all of the features. Some flowers should be gently torn open before pressing; once the plant is dried, it will be very difficult to open flowers and peer at stamens or other small parts. Think about all of the characteristics you may want to see later for identification and descriptive purposes, and remember the collection you are pressing may later be glued to a sheet of herbarium paper.

It is not easy to flatten a three-dimensional plant onto a two-dimensional sheet of paper. Thick roots may need to be longitudinally sliced to press. Bushy plants are also problematic. You may need to do some judicious pruning of your collection, but you must be careful not to change the aspect of the plant. Leave clues that this plant once had more branches or flowers than it does now, or prune severely one inflorescence, or one portion of the plant to show features clearly, and leave the rest intact to show the general form of the plant. Long stems can be bent into a "V" or an "N" shape to fit.

Wetland plants are often succulent or saturated and may mold in the plant press before they dry. If the plant seems to hold large amounts of internal water, allow it to dry slightly before pressing, although be cautious of wilting.

Aquatic plants also require special handling (Prescott 1980). Because these plants are often limp and drooping, with highly dissected, tangled leaves, they are difficult to arrange before pressing. It is easiest to place the plant directly on herbarium paper under water, and arrange the floating specimen in a life-like position. The herbarium paper, supported by a metal or plastic sheet, can then be carefully lifted from the water and placed in a tilted position to drain. The drained herbarium sheet can be placed in a standard drier between blotters and ventilators. A sheet of muslin over the specimen aids drying and reduces sticking.

Once in the press, accelerate drying of wetland and aquatic plant collections by placing the press outside, in full sun, in a breeze. In high humidity, you may need to rig a plant dryer, which can be as simple as a light bulb placed beneath the press. Presses can also be placed overnight in gas ovens, using the heat of the pilot light.

Extremely succulent plants such as cacti are especially problematic. Tissue must first be killed; a recommended method is blanching in boiling water. They must then be dried rapidly.

Some plants can be hollowed out to allow rapid drying and flat pressing. Fosberg and Sachet (1965) give further ideas for dealing with these difficult species.

To keep track of collections, and the field notes corresponding to each collection, you should keep collection numbers. These are sequential numbers assigned to each collection you make. Write the number in the field notebook and on the newspaper sheet in the press.

Collections without supporting information are nearly useless; thus, most collectors carry a field notebook and take careful notes. At minimum, each collection should have the following information:

1. Collection number
2. Date collected
3. Location, described in enough detail to map and relocate the collection site in the future
4. Habitat information, including slope, aspect, substrate, elevation, shade, and moisture regime
5. Associated species and vegetation type
6. Abundance of the species and approximate size of population area
7. Notes on flower color, plant size, and variability
8. Ownership of site

Other information that may be useful includes keying notes, if field keyed; threats noted; and observed ecological information such as herbivory, pollination, and insects. Some botanists prepare preprinted forms and carry them in a binder field notebook.

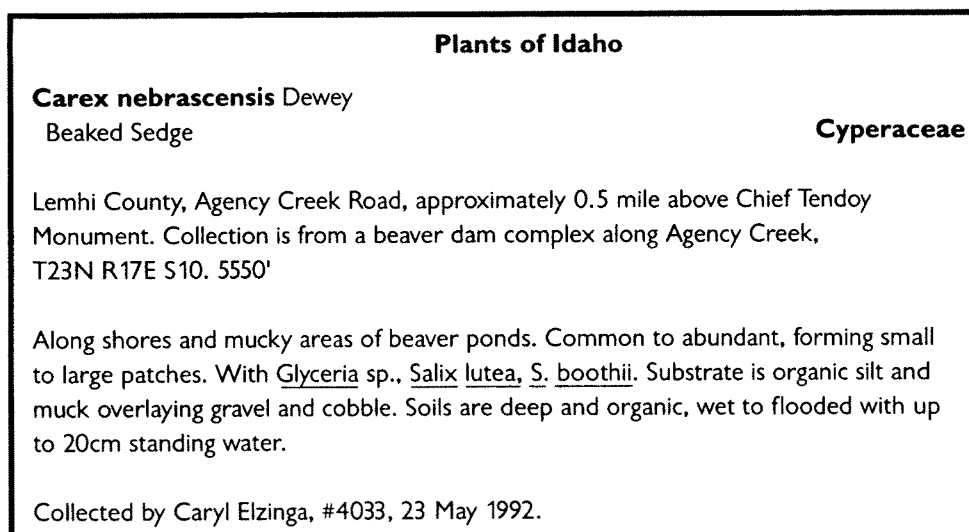
This information helps in identification and later study; it is also needed for the herbarium label that will accompany your collection. Labels will be required if you intend to send your collection to a specialist at a herbarium for verification, but even in an agency or personal herbarium, labels serve important functions. Labels summarize habitat and distributional information for the plant, which provides information about the plant and its ecology. The locational information on a label allows the site where the plant was collected to be relocated. Historic and, occasionally, extant populations of rare plants have been located by looking at the locational information on the labels of herbarium specimens.

An example of a label is shown in Figure 5.1. Labels are usually 3" × 5". Most biological supply companies sell pregummed labels and labels that can be fed through a computer printer. Database systems for herbarium collections are available (or can be created fairly easily), allowing you electronically to store field notes and print standard labels.

Mounting Collections

Mounted collections are easier to handle and examine and are more likely to be used than plants residing in a newspaper. Identification, however, is usually easier with unmounted material, which is why most herbaria prefer to receive specimens unmounted. An unmounted specimen can be examined on both sides and is more easily maneuvered under a dissecting scope. Occasionally, small portions of the plant such as a single flower will be rehydrated to aid certain identification.

Standard paper size for mounting is 11½" × 16½". Mounting paper and herbarium paste are available from biological or herbarium supply houses. Herbarium stock is recommended because the special paper content will not deteriorate nor damage specimens over the long term. Herbarium glue is designed to be unattractive to insects, which can do major damage to herbarium specimens. Normal household glue is an attractant and also deteriorates over time. Plants can also be secured with narrow strips of cloth tape. Well-mounted herbarium specimens could last for a hundred years or more. Larger herbaria may have collections dating from the 1700s.


Plants of Idaho

Carex nebrascensis Dewey
Beaked Sedge **Cyperaceae**

Lemhi County, Agency Creek Road, approximately 0.5 mile above Chief Tendoy Monument. Collection is from a beaver dam complex along Agency Creek, T23N R17E S10. 5550'

Along shores and mucky areas of beaver ponds. Common to abundant, forming small to large patches. With Glyceria sp., Salix lutea, S. boothii. Substrate is organic silt and muck overlaying gravel and cobble. Soils are deep and organic, wet to flooded with up to 20cm standing water.

Collected by Caryl Elzinga, #4033, 23 May 1992.

Figure 5.1. Typical herbarium label.

Loose material such as seeds or small flowers should be placed in an envelope attached to the herbarium sheet. Normal letter envelopes are not recommended, since the seams are not glued all the way to the edges and small seeds can escape from the corners. The paper will also deteriorate over time. Supply companies sell special envelopes that avoid these problems.

COLLECTING ANIMALS

Invertebrates

Taxonomically problematic butterflies can usually be readily identified through close-up photographs of both the uppersides and undersides of wings of captured animals. Patient stalking of alighted individuals also can provide opportunities to secure these photographs. Most other insects must be identified by experts who will dictate preservation methods. These methods might include pinned specimens, specimens in alcohol (labels must be written with India ink), or specimens in glassine envelopes with labels inside. All labels should be on acid-free, 70- to 80-pound, cotton rag to prevent deterioration and should include location, date, and collector's name. Again, the preservation method is best worked out before collecting with the expert responsible for the identifications.

Fish

Fish can usually be adequately documented with photographs. If specimens are sought, they can be frozen for extended periods or fixed in 10% formalin solution. Fish to be preserved in solution should be inserted head-first into jars containing fixative, with large specimens first injected directly into the body cavity and their stomachs incised to prevent rotting. Wide-mouthed glass jars with polypropylene lids and polyfoam liners should be used. Labels should be waterproof and include one specimen label attached to the jaw or inserted into the mouth or opercular area of each specimen and a data label affixed to the outside of each jar. Specimen labels should be written in pencil and should contain specimen number and species name. Data labels on the jar should include a site identifier, gear information, number of specimens, watershed/waterbody identifier, collection date, and the collector's name. Species should be fixed for several days to a week and then transferred to 50% isopropyl alcohol for long-term preservation.

Amphibians and Reptiles

Collecting usually is not required for amphibians or reptiles as diagnostic features are evident on most photographs of most species, particularly photographs taken with a close-up lens and framed to show features important for identification. If specimens must be collected, fix them in 10% buffered formalin with field tags tied above the knee on the right rear leg of frogs and large salamanders and around the neck of small salamanders. Most larvae and small amphibians are penetrated quickly by formalin, but large frogs should have the body cavity opened or the formalin solution injected. To fix reptile specimens, note that injections must be used because the skin is relatively impermeable. For lizards, 10% buffered formalin should be injected under the skin in each leg segment and at the base of the tail. Snakes should be injected at three points between the snout and vent, as well as in the tail, with the hemipenes everted. Turtles should be injected in the neck, limbs, and tail and deep into the body cavity. For both groups, specimens should initially be preserved on "hardening trays" in desired positions for several hours before being transferred to fluid-filled containers for further hardening, if necessary, and then to the preservative solution.

Birds

Birds are well-known taxonomically, and specimens are rarely needed. Documentation is usually sufficient with visual or song identification by qualified observers, photographs, and sound recordings. Note that most birds are migratory, and strict protections apply to such species in most countries such that permits are needed to handle birds and even bird parts. For any voucher, record observer, species, subspecies, sex, age, exact date and locality, and, if collecting specimens in the field, habitat. Use a plastic bag with ice packs to keep specimens cool, and freeze immediately once out of the field.

Mammals

Mammals, like birds, are well known, and vouchers are needed for identification purposes usually only for the more obscure rodents. Simple hair samples can be quite useful (e.g., gathered from barbed wire, thorns, or bark) and can be stored in envelopes marked with the sample number and collection information and later dried and stored in a cool, dry place or a freezer. Because of accidental trap mortality, specimen preparation may nevertheless sometimes be desired. Specimens should be sealed individually in ziplock bags accompanied by ice packs with a label that includes date, location, collector, and field reference number, and they should be put in a freezer as quickly as possible. Note that some species (e.g., bats) can be health hazards and should be handled appropriately. Fluid preservation for mammals is not desirable as it impairs discerning diagnostic pelage patterns and colors.

FIELD HINTS

This list of ideas for making field work easier and more comfortable is certainly not exhaustive, but may be useful.

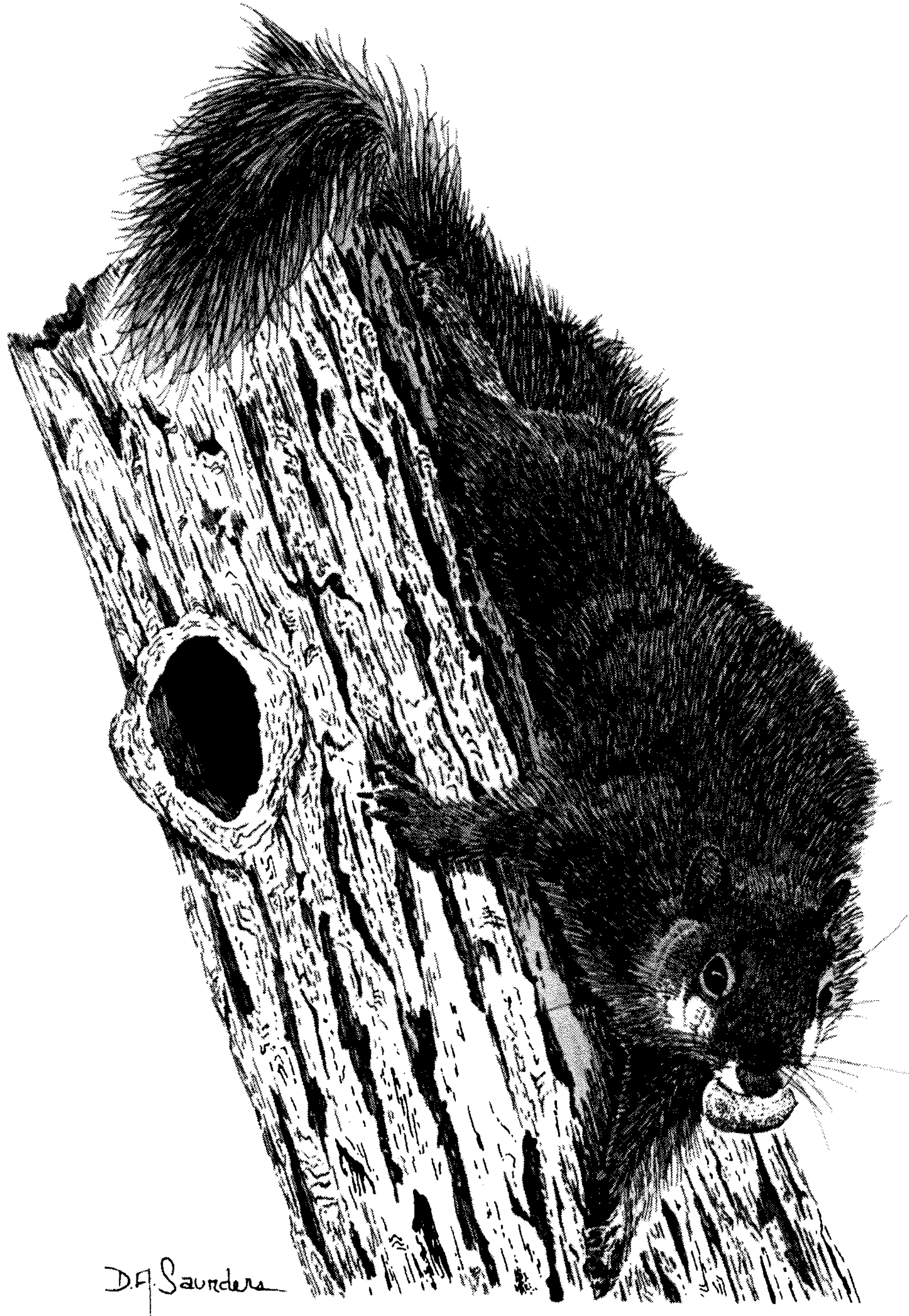
- For comfort while you are measuring tedious plots, take an ensolite pad to kneel on or a small gardening stool to sit on. This will also reduce trampling damage.
- Always paint stakes and monuments just before you leave the site. You lessen the amount of paint on your clothing, equipment, and data sheets.
- Paint stakes and monuments every year that you monitor.
- Paint a stake that marks the corner of a permanent plot carefully to avoid spraying any plants.

- You may want to paint the handles of your field equipment bright orange so they can be easily found if dropped.
- Pin flags have all sorts of uses. They can be temporary markers of your population boundary. They can mark clusters of plants in the field so that you can get a better visual picture of the distribution of individuals in a population. They can mark plot boundaries. Pin flags are often preferable to flagging because they are quickly placed, easily moved, and easily picked up when no longer needed, although flagging hung in trees or shrubs is more visible.
- Screwdrivers are another multipurpose tool. Use them to dig up plant specimens, hold down tapes, and secure temporary frames.
- To keep track of location when counting dense quadrats, use two sticks the width of the plot to mark the temporary counting subplots as you work the length of the plot.
- Two people can usually lay a tape in dense brush much more easily and accurately than one. The first person uses a compass to sight on an object past the end of the transect and then guides the second person with the tape over and around objects.
- Photocopy field notes after each day and store in a safe place (off-site). This will eliminate the chance of losing an entire field season's worth of data and observations.
- Photocopy monitoring data sheets and store off-site from the originals. This will reduce the chance of losing data to a catastrophe such as fire.
- Field vests or cruising vests can be purchased that come with a myriad of different-sized pockets. Compass, pens, pencils, field notebook, clipboard, camera, film, etcetera can all be kept at your fingertips. Most are colored bright orange for visibility.
- If you work in either a wet climate or a very hot, arid one, there will be times while you are sampling vegetation when some form of shade or rain protection would be welcomed. If you are doing a field project that requires fairly long periods in a spot (perhaps mapping small plants in a small quadrat), a moveable gazebo may make you more comfortable. These are like a tent fly without the tent. Look for ones that are free-standing, requiring no stakes, so you can pick it up and move it to the next plot when you move on. You can also rig a large umbrella to a lightweight frame.

MANAGEMENT IMPLICATIONS

Good monitoring depends on conducting field work efficiently and accurately. Equipment should be carefully selected for the task. Study sites and measurement units must be well documented and well monumented. Collections should be curated carefully to ensure that they can function as a reference in the future.

CHAPTER 6
Data Collection and Data Management



Sciurus niger
Eastern fox squirrel
Artist: D. Andrew Saunders

This chapter covers the different methods of recording and managing actual field monitoring data. These methods, if carried out properly, lead to the orderly and efficient processing of information, smoothing the way for data summary, data analysis, and report completion. Otherwise, the data managers are faced with processing messy and confusing datasets. Poorly gathered and poorly managed monitoring data usually stem from a lack of understanding that an enormous amount of time (days, weeks, or even months) can be saved by following some of the guidelines presented in this chapter.

Successful data collection and data management start with the planning of a monitoring study and continue for as long as datasets are archived on computers or hard files. Good data collection methods lead to efficiency in the field and in the office. Detailed documentation of field methods and descriptions of codes or abbreviations help to ensure the integrity of data from the field to the final interpretation of monitoring results.

RECORDING DATA IN THE FIELD

Good data management begins with field data collection methods that minimize errors in recording and that maximize both field and office efficiencies. Data will be captured in the field either electronically (tape recorders, data recorders, computers) or by hand on data forms or in field notebooks. Field data forms remain the most common method of recording field data.

Tape Recorders

Portable tape recorders can reduce the amount of time spent in the field, especially when a person is working alone and needs to record a large amount of data. Most portable tape recorders are relatively inexpensive and light to carry. Voice-activated recorders reduce the amount of button-pushing and shorten transcription time by eliminating the quiet time between data points. Detailed site descriptions and other field observations can be verbally recorded in less time than it takes to write them in a field notebook.

Transcribing the audio recording will require more office time than transcribing from a hard-copy data form. It is difficult to scan the recorded data to look for any patterns or problems or to verify which sample areas have been sampled or which types of data have been gathered. By following a blank field data sheet or a checklist as a guide to consistently gather all categories of information in the same sequence, then the transcription time and confusion can be significantly reduced. Tapes are inexpensive, so record only information from a particular day or a single plot per tape to reduce searching time of the recording when transcribing data. If you have a number of plots on a single tape, periodically note the plot number as you are collecting the data for that plot to facilitate searching later. You can also note the location on the tape of the data for a particular plot in your field notes if your recorder has a counter.

At the beginning of each tape, record the day and the location. Clearly label each tape. Like other electronic devices, tape recorders will occasionally fail, and data could be lost if the tape is damaged. Recording in high winds or with loud background noise may make the data unintelligible. Periodically, rewind and check the recording to ensure that the recorder is working properly and that the data can be heard on the recording. We also strongly advise that transcription of the tapes occur within hours of recording or, at most, within a few days. This will allow you to return and recapture lost data if need be. Transcriptions that occur months later allow no such option. Always carry plenty of spare batteries and use fresh tapes.

Portable Computers or Data Loggers

Recording data directly into a portable computer or electronic data logger can be the most efficient means of collecting field data. This method eliminates the time-intensive, data-entry and data-proofing steps that go along with data recorded on field data forms. Field data can be en-

tered in a predesigned format that will facilitate summary and analysis steps. Some portable computers support the use of Windows-based software (e.g., spreadsheet programs), making data exchanges with desktop computers easy.

Data-loggers designed for field use are quite expensive, but will withstand rigorous field conditions (e.g., dustproof, waterproof). Some electronic data loggers use nonstandardized computer programs that can make the transfer of data to a DOS-based computer difficult.

Laptop, notebook, and palmtop computers are not designed for field use, but with some care can function adequately. Palmtop-size computers can be used in the rain when placed in a gallon-sized zip-locked plastic bag. Data entry and screen viewing can be done through the plastic. Carry plenty of spare bags and periodically inspect for leaks.

Most portable computers will cost several hundred dollars and will require either many batteries or a battery charger with rechargeable batteries. The viewing screen on most portable computers is quite small, which can make it difficult to scroll around a large, data-entry template. Most portable computers are heavy and awkward to use in the field, although some of the palmtop computers are quite light. Data can be lost because of hardware or software problems or if a computer's batteries run dead. Be sure to make some kind of backup of the data at least every day by transferring the field data to another computer, to a floppy diskette, to a flash-memory card, or by printing a hard copy of the data.

Field Data Forms

Field data forms or field notebooks are inexpensive and lightweight, but have several drawbacks. If data are to be summarized or analyzed with a computer then data will need to be transcribed from the field data forms. The data entry and data proofing steps can consume more time than the field data collection. Because wet field data sheets lead to writing smears or streaks and because the pages may become stuck together, print field data forms on waterproof paper. Several paper suppliers sell waterproof paper that can be used in standard printers and photocopiers.

Field data forms should be designed to promote efficiency in field collection and computer data entry. The time required to complete data entry and data proofing steps is profoundly influenced by the design of the field data form. Transcribing data from a poorly designed, sloppily written field data form can take more than 10 times longer than transcribing data from a well-designed, clearly legible data form. An example of a well-designed form is shown in Figure 6.1.

Each set of data should have a cover sheet that stays with the field data at all times. The cover sheet should provide the following information for the data collected: what, why, where, who, how, and when. Detailed information should be provided on the location of study plots, the species or community being studied, the personnel involved, the types of management treatments that have occurred or are being planned, a description of any codes that are used, and a thorough description of the field methodology. See Box 6.1 for a list of the types of information that should be included on the cover sheet. In addition, each field data form should have a complete "header" section that links the form to the project described on the cover sheet. The header should be completely filled out on every page. The header should include at least the following items:

1. Date
2. Location (general area and specific sampling location)
3. Title/project description
4. Species or community name
5. Treatment category (if applicable)
6. Observer (person(s) doing the sampling)
7. Transect or macroplot number (if this information applies to entire data sheet)
8. Page number _____ of _____ total pages
9. Room for additional comments

Density data sheet—3 categories of the counting unit—rectangular quadrats divided into segments

Title/Description: Plant density monitoring in permanent quadrats	Date: 5/15/95—5/25/95	Page 1 of 4
Location: Agate Desert Preserve (SW Oregon)	Species: Lomatium cookii	
Treatment: To burn in 1996	Field personnel: Darren Borgias	Quadrat dimensions Width: 0.25m Length: 90m Segment length: 1m

Notes:

Mplot	Quad	Seg	Seedling	Veg	Flwring	Mplot	Quad	Seg	Seedling	Veg	Flwring
1	1	1	13	2	0	1	2	9	0	3	0
1	1	2	1	0	0	1	2	10	0	2	0
1	1	3	1	6	4	1	2	11	0	1	0
1	1	4	0	8	0	1	2	12	0	1	0
1	1	5	1	7	0	1	2	13	0	5	0
1	1	7	0	3	0	1	2	15	0	2	0
1	1	9	0	4	1	1	2	16	0	3	0
1	1	14	0	5	2	1	2	17	1	3	0
1	1	15	0	1	0	1	2	18	1	1	0
1	1	16	3	6	2	1	2	21	0	2	0
1	1	17	0	1	0	1	2	22	0	4	0
1	1	18	0	1	0	1	2	23	0	5	0
1	1	19	0	1	0	1	2	24	0	1	0
1	1	22	0	2	0	1	2	54	0	1	0
1	1	23	0	2	0	1	2	55	1	2	0
1	1	25	0	0	1	1	2	56	0	3	0
1	1	26	0	1	0						
1	1	27	0	2	0						
1	1	56	0	1	0						
1	1	58	0	1	0						
1	1	60	0	0	1						
1	1	61	0	3	0						
1	1	62	0	1	0						
1	1	63	1	0	0						
1	1	67	0	3	0						
1	1	77	1	2	0						
1	1	81	0	3	0						
1	1	82	5	8	7						
1	1	84	1	0	0						
1	1	87	0	1	0						
1	1	88	1	2	0						
1	2	1	2	0	0						
1	2	2	14	1	1						
1	2	3	16	2	0						

Figure 6.1. Example of a well-organized form for collecting density data.

Box 6.1. *FIELD MONITORING COVER SHEET*

1. *Include header from the field data form. This header should include the following:*
   a. *Title or project description name*
   b. *Location*
   c. *Species or community name*
   d. *Type of study (density, cover, frequency, etc.)*
   e. *Personnel*
   f. *Date(s)*
   g. *Treatment (if applicable)*
   h. *Macroplot or transect, or other location identifier if this information applies to the entire data sheet*

2. *Management objective (see Chapter 14)*

3. *Sampling objective (see Chapter 14)*

4. *Location and layout of study area. Sketch location, including access. Denote key area, macroplot, or transect locations with macroplot numbers, names, and treatments, as applicable, and the approximate bounds of the population being studied. If sampling units are placed along transect lines, show how they are placed. Provide approximate scale.*

5. *Detailed description of data collection methods. This should include sufficient detail so that someone unfamiliar with the project can understand how the data were collected. Consider the following issues:*
   a. *What are the bounds of the sampled population?*
   b. *If you are sampling within macroplots, what is the size and shape of the macroplots and how were they positioned?*
   c. *What is the sampling unit (e.g., quadrats, transects, individuals)?*
   d. *What is the size and shape of the individual sampling units?*
   e. *How are sampling units positioned in the population of interest?*
   f. *How many sampling units were sampled? How was sample size determined?*
   g. *Describe any boundary rules for counts or measurements that occur along the edge of sampling units.*
   h. *For density measurements, define the counting unit (e.g., genet, ramet, stem, flowering stem for plants) and any rules that are used to discriminate among adjacent counting units.*
   i. *For cover measurements, define whether basal or canopy cover is measured and define gap rules. If ocular estimates of cover are made in cover classes, define those classes. For point-intercept cover measurements, describe the point diameter and type of tool used.*
   j. *Include a full description of any codes used on the field data sheets, including species acronyms.*

Preprint as much information as possible. Time can be saved in both the field collection and data entry phases by preprinting as much reference information as possible on the field data form. This eliminates the need for a lot of repetitive writing and reduces mistakes. When communities are being sampled and a large number of species codes are being used, include the full genus and species name on the field data sheet along with a species code. The code will be used during data entry; having the full name listed with the code eliminates serious, data summary problems, such as two species being inappropriately grouped together or the data for a single species being split between two or more categories. If a list of species known to occur in a particular community is available, or if only a subset of the species are being tracked, then preprint the

species codes, and the genus and species names, on the field data form. Preprinting species codes and names saves a lot of writing time in the field, minimizes data transcription errors, and greatly speeds up data entry because the sequence of species stays the same from page to page. The sequence of species can either be alphabetic, by taxonomic or growth-form groupings (e.g., all grasses together, all forbs together, etc.), by relative abundance, or through some combination of these methods (e.g., list the four most common species first with the remainder of the list sorted alphabetically).

Species codes frequently consist of four letters, the first two letters of the genus and the first two letters of the species (e.g., ARHE = *Ardea herodias* or LIOC = *Lilium occidentale*). To avoid using duplicate codes for different species that share the same four-letter acronym consult a national database. For plants, refer to national databases such as the PLANTS National Database, maintained by the National Resources Conservation Service in the United States or the Australia Plant Name Index. For birds of North America consult the AOU (1998) and for birds of the world, see Sibley and Monroe (1990, 1993). For mammals of the world, Wilson and Reeder (1993) is the standard reference. For fishes, consult Eschmeyer (1998). For reptiles and amphibians of North America, the standard nomenclatural source is Collins (1997), whereas global lists for amphibians include Frost (1999) and for reptiles and amphibians include Frank and Ramus (1995). Many of these databases are currently available over the internet and can be found by searching for their titles (see citations).

If a plant or animal is only identified to genus, and some master list of codes is not available, avoid the use of "SP" as an abbreviation for "SPECIES" (e.g., *Bromus* species = BRSP). Instead, adopt some convention such as "ZZ" or "Z1" (e.g., BRZ1) to use whenever a specimen is only identified to genus. This will reduce the number of duplicate codes (many species names actually start with the letters "sp"), and more clearly indicate when species identity is unknown. Six-digit species codes (composed of the first three letters of the genus and species) reduce the number of duplicate codes.

It is important to define any numeric or character codes that are used on the field data sheet. These codes should always be defined in the field data cover sheet, and, when possible, they should appear on the field data sheet itself. For example, if counts are being made in randomly positioned quadrats, and if the particular habitat type of each quadrat is being recorded (e.g., mound, intermound, pool), use a numeric code to define the habitat type rather than writing the full habitat type for each quadrat recorded. Placing the code descriptions near the top of the data sheet ensures that habitat type information will be recorded and summarized properly.

Recording unanticipated information on the data sheet is one of the most common reasons for messy and confusing data sheets. Not all data form needs can be anticipated. Unexpected observations can lead to the need to incorporate additional information onto a field data form. For example, a subset of plants being counted in quadrats may have some peculiar attribute such as yellowish, dried leaves or evidence of flower head herbivory. There is sometimes a tendency to incorporate many detailed comments onto field data forms, taking advantage of any available blank space. Sometimes the same characteristic is described in different ways (e.g., "some flower heads eaten," "inflorescence damaged," "three seed heads with signs of herbivory"). This could create confusion during the data-entry process. Which comments are important and should be entered into the computer with the regular monitoring data? Which comments are insignificant and should be ignored? Which different comments mean the same thing? How should the additional data be used during the data summary process?

The best way to incorporate additional information is to consider how the inclusion of this type of information will impact data summary and data analysis steps. Will it be useful to have a tally of all plants showing some characteristic (such as evidence of flower head herbivory) separate from plants that do not show the characteristic? If so, create a "Notes" column along the margin of a field data form and create a numeric code to assign to any observation exhibiting the characteristic. The code should be described at the top of the field data sheet (e.g., 1 = flower head herbivory noted). These additional data are then easily incorporated into the dataset during

data entry, the observations can be sorted by this additional field, and separate summary statistics can be generated very easily.

Try to design field data forms so that nearly all data entry will be numeric. Data entry is most efficient when data can be entered from the 10-key numeric keypad portion of a computer keyboard. Using a combination of character and numeric data slows down data entry.

Take adequate time to make sure that all handwriting is clearly legible. You should not assume that you will be the only one reading the completed field data sheets. Poor handwriting can significantly slow down the data entry process and can introduce errors into the datasets.

ENTRY AND STORAGE OF DATA IN THE OFFICE

If the quantity of data gathered is small, sometimes the data can be efficiently summarized right off the field data form using a hand calculator. These calculations should be repeated to ensure that no mistakes were made in entering and summarizing the data. Often, however, monitoring data will need to be input into a computer system for summary and analysis. A framework for data entry and storage will reduce loss of valuable data.

Selecting a Computer Software Program

Four categories of software applications for entering and storing monitoring data are available: 1) word processors, 2) relational databases, 3) spreadsheets, and 4) statistical software programs. Word processors used to be the worst place to enter or store ecological monitoring data. Most word processors did not distinguish data files from regular text files containing memos or reports. Data summarization procedures were not available, or they were extremely limited in most word processors. In recent years, however, data table formats have been added to many word processors and some of these support limited spreadsheet type operations. Check carefully to make sure that data can be easily exported to other software applications prior to entering field data into a word processor.

Programs such as Dbase, Paradox, Oracle, and Microsoft Access are examples of relational databases. Relational databases are designed to organize and manage large amounts of information. Custom data entry screens can be created where the user enters data into blank highlighted fields. Most relational databases include some basic data summary procedures (e.g., calculating totals or averages). Entering and storing monitoring data in a relational database may be a logical alternative if data from individual observations (i.e., height of an individual plant or the number of plants in a certain permanent quadrat) are frequently referenced or reported. Relational databases usually have sophisticated tabular reporting features but limited graphical reporting features. Most relational databases can, however, import and export data easily with other software programs. Data gathered as part of a large-scale monitoring network should be stored in a relational database to facilitate data management and data processing (Stafford 1993). You should identify the data standards for the network early in the monitoring design to ensure consistency and completeness.

Programs such as Lotus 1-2-3, Quattro Pro, and Excel are examples of spreadsheet programs. The data entry screen in a spreadsheet is a rectangular matrix of labeled columns and rows. Spreadsheets contain many time-saving data entry procedures. For example, if data were gathered from plots numbered 1 to 100, a few keystrokes can generate a list of plot numbers from 1 to 100 so that 100 individual plot numbers do not have to be entered. Large sections of data can easily be copied or moved within a spreadsheet. Most spreadsheets include at least basic data summary procedures, and some include relatively advanced summary and analysis routines. Descriptive reference information (species, location, dates, treatments, definition of codes, etc.) can be placed in the spreadsheet above the actual data matrix. Spreadsheet programs usually offer sophisticated tabular and graphic reporting features. They also can import and export data

files in many different formats. Data entry onto spreadsheets may be the most efficient means of transferring data from field data forms into a computer file. Even if data are going to be stored in a relational database, it may be more efficient to enter the data in a spreadsheet and then transfer the data to the relational database.

Statistical programs (discussed in Chapter 9) offer powerful data summary, data analysis, and graphing procedures. They all have some kind of data entry mode, usually a screen resembling a spreadsheet. Statistical programs are often not the best choice for data entry. Data entry routines found in statistical packages are often more limited than those found in real spreadsheets; their spreadsheet-like format does not usually allow you to add the type of descriptive reference information that you can enter onto spreadsheets. Finally, the data may not be easily transferred from one statistical package to another. Compare the features in your statistical program with your spreadsheet program before entering large datasets directly into the statistical program. Because most statistical programs readily import data from spreadsheets and relational databases, you are usually better off using spreadsheets for data entry and storage.

Storing Data Files—Filenames and Subdirectories

The naming and storing of files does not seem like a problem when there are only a few data files to input. At first, a data manager may decide to place all data files in a single computer directory called something like "DATA." He or she may name individual files with whatever seems like a logical name at the time the file is created, without adopting any standard conventions for naming files. Confusion starts to increase as more and more data files are created. Soon it becomes difficult to find a particular file, and numerous files may need to be opened until the right one is located. Some files may be accidentally deleted because the data manager thought another file contained data superseding the deleted file. Creating an efficient, standardized system of naming and storing computer data files early in the development of a monitoring program will save a data manager many hours, days, weeks, or months of frustrating data management. Consider a naming protocol that includes the type of data, the species of interest, and the years of data collection.

One efficient method of storing monitoring data is to create separate subdirectories or folders for different sites. This could be done either by establishing a DATA directory with different subdirectories for each site (e.g., all data files from the Middle Fork of the John Day Preserve are stored in C:\DATA\MFJD*.*) or by creating a DATA subdirectory under a site directory (e.g., C:\MFJD\DATA*.*).

Avoid creating many separate files for related monitoring data. Keep related information from different sampling areas or from the same sampling area over different years in the same file. The data will need to be brought together for summary and analysis purposes, and having the data in a single file all along can reduce data management headaches. Figure 6.2 shows a sample format for recording data from multiple macroplots and multiple years in a single file.

Keep data files separate from summary and analysis files. Do not change data files except to update them with a new year of data or to correct mistakes. This will avoid accidentally deleting or changing your original data.

Adequately Documenting Data Files

Each data file should include reference information about the data in that file (Stafford 1993). This information should detail the how, when, what, where, and who information included in the field data cover sheet and in header sections of the field data forms. This kind of information should be included in a file header that appears in the computer file above the rows of actual monitoring data. Any codes contained in the dataset should be listed and described in the file header. A detailed description of the methods used to gather the data should be included in the file header or a reference to another source for this information should be provided. See Figure 6.2 for an example of a completed data file header.

Proofing Datasets

If data were entered into a computer file from field data forms, then the data must be checked for any keystroke errors introduced during data entry. Having someone read off the data from the original data form while another person follows along either at the computer file or on a printed hard copy is one method that works well. Any corrections are noted on the computer printout.

Using a dual entry procedure is an alternative quality control option for catching keystroke errors (Stafford 1993). Two different keystroke operators enter the same data. Any mismatches between the two entered copies are noted, and the original field data sheets are checked to determine which copy is in error.

Making Backups of Entered Data

It is essential that backup copies are made of all computer data files. Hardware, software, and user failures occur on an unpredictable schedule and large amounts of grief can be saved if a regular backup schedule is maintained. Daily backups can easily be made to some external medium, such as a floppy disk. Weekly backups of all files made to a tape drive, CD-ROM, or other high-capacity medium can make data recovery much easier following a hard disk failure. You should "leap frog" backups so that you are not copying to the only backup copy of the data. It is a good idea to keep one copy of the backup at another location (in case a catastrophic fire consumes the office copy).

FILENAME: C:\DATA\AGAT\FLOCO891.WK1
PRESERVE NAME: Agate Desert
ELEMENT NAME: Rogue Valley Mounded Prairie
DATE OF OBSERVATION: 25 May 1989, 14-16 May 1990, 15-21 May 1991
SITE DESCRIPTION: Research Plots LOCO Burn
MACROPLOT NUMBERS: 2, 3, 5-8
TYPE OF MEASUREMENT: Nested Frequency
NUMBER OF QUADRATS: 50
QUAD SIZE/CODED VALUE: 1 = 0.01 m^2, 2 = 0.1 m^2.
DATA CONTACT: Darren Borgias
COMMENTS:
HABIT = Habitat codes; 1 = mound, 2 = flank, 3 = intermound, 4 = pool
THAT = thatch measured in cm
Grouped species codes:
TRNA = Trifolium native (T. variegatum)
TREX = Trifolium exotic (T. subterraneum, T. arvense, T. dubium)
TRSP = Trifolium species (unidentified)
UNK1 = unknown composite
Thatch information can be found to the right of the spreadsheet

For a full list of species codes and a detailed description of field methodology see: Borgias, D. 1993. Fire effects on the Rogue Valley Mounded Prairie on the Agate Desert, Jackson CO.

YEAR	MPLOT	QUAD	HABIT	POSC	TACA	BRSP	VUSP	POBU	HOGE	DEDA	AICA
89	2	1	2	0	1	2	2	0	1	0	0
89	2	2	2	0	1	1	0	0	1	0	0
89	2	3	1	0	1	1	0	0	0	0	0
89	2	4	3	0	1	1	0	0	2	0	0
89	2	5	2	0	1	1	1	0	0	0	0
89	2	6	2	0	1	1	0	0	1	0	0
89	2	7	2	0	1	1	0	0	0	0	0
89	2	8	2	0	1	1	1	0	0	0	0
89	2	9	1	0	1	2	0	0	0	0	0
89	2	10	1	0	1	1	1	1	0	0	0

Continued for the rest of 90 and 91

Continued for the rest of the macroplots

Continued for the rest of the quadrats

Figure 6.2. Example of a spreadsheet file showing the reference information provided in the file header.

MANAGEMENT IMPLICATIONS

Data are expensive to collect and valuable for application to management decisions. Good data management can avoid loss of information, can increase the accuracy of translating data from field to analysis, and can save many hours of frustration and wasted time.

CHAPTER 7
Basic Principles of Sampling

Physaria didymocarpa var. lyrata
Salmon Twin Bladderpod
Endemic to Lemhi Country, Idaho
Artist: Glenn A. Elzinga

Sampling is the act or process of selecting a part of something with the intent of showing the quality, style, or nature of the whole. Monitoring does not always involve sampling techniques. Sometimes, you can count or measure all individuals within a population of interest in a complete census. Other times, you may select qualitative techniques that are not intended to show the quality, style, or nature of the whole population (e.g., subjectively positioned photographed plots).

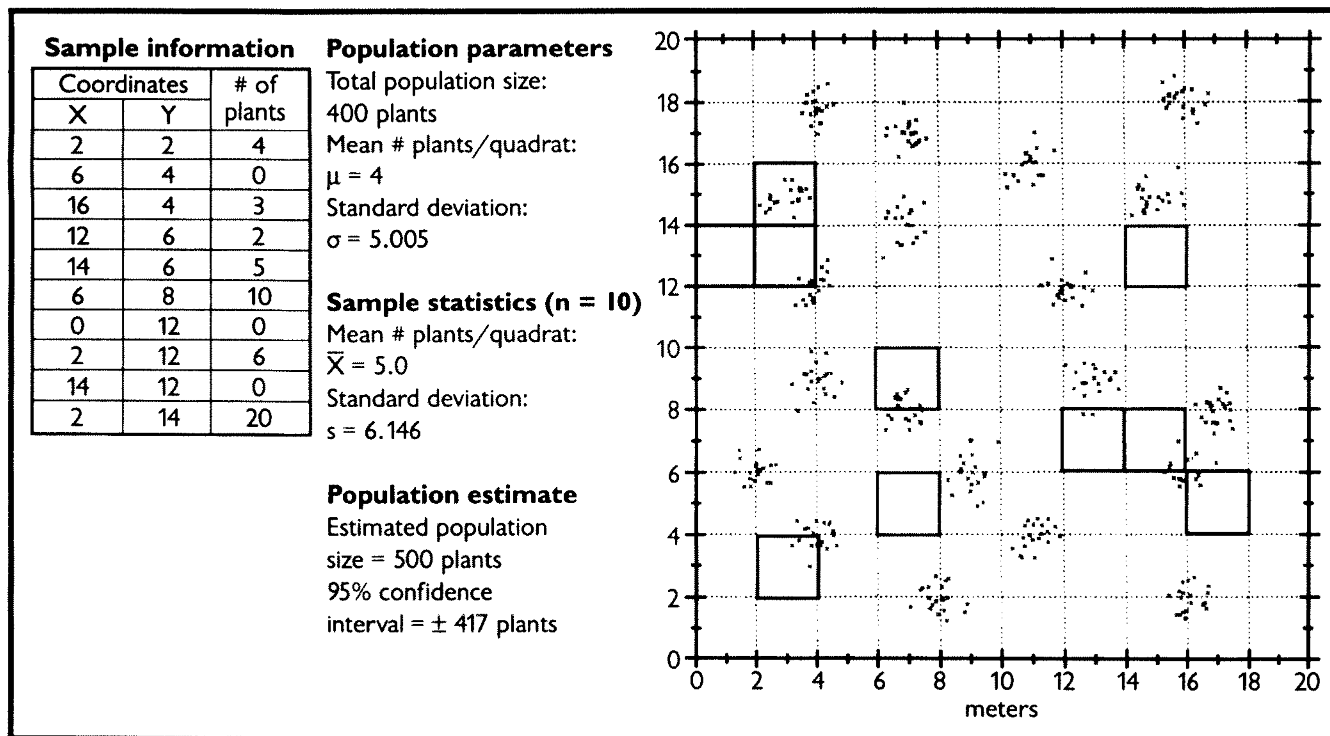
What about those situations where you have an interest in learning something about the entire population, but where counting or measuring all individuals is not practical? This situation calls for sampling. The role of sampling is to provide information about the population in such a way that inferences about the total population can be made. This inference is the process of generalizing to the population from the sample, usually with the inclusion of some measure of the "goodness" of the generalization (McCall 1982).

Sampling will not only reduce the amount of work and cost associated with characterizing a population, but sampling can also increase the accuracy of the data gathered. Some kinds of errors are inherent in all data collection procedures, and, by focusing on a smaller fraction of the population, more attention can be directed toward improving the accuracy of the data collected.

This chapter includes information on basic principles of sampling. Commonly used sampling terminology is defined, and the principal concepts of sampling are described and illustrated. Even though the examples used in this chapter are based on counts of plants in quadrats (density measurements), most of the concepts apply to all kinds of sampling for both plants and animals.

POPULATIONS AND SAMPLES

The term "population" has both a biological definition and a statistical definition. In this chapter and in Chapters 8 and 9, we will be using the term "population" to refer to the statistical population or the "sampling universe" in which monitoring takes place. This sampled population will sometimes include the entire biological population and, at other times, some portion of the biologi-



Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Figure 7.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m × 2m quadrats, along with sample statistics and true population parameters.

cal population. Sometimes, the sampled population will not be comprised of individual organisms in the way we think of biological populations, because the population consists of the complete set of individual objects about which you want to make inferences. These may be individual organisms, or they may be quadrats (plots), points, or transects. We will refer to these individual objects as sampling units. A sample is simply part of the population, a subset of the total possible number of sampling units.

Quadrats are square or rectangular (or rarely circular) sampling units in which an attribute is counted or measured.

Macroplots are relatively large areas, with sampling units such as quadrats, lines, or points randomly located within them.

These terms can be clarified in reference to an artificial population of plants shown in Figure 7.1. This population contains a total of 400 plants, distributed in 20 patches of 20 plants each. All the plants are contained within the boundaries of a 20m × 20m macroplot. The collection of plants in this macroplot population will be referred to as the "400-plant population." A random arrangement of ten 2m × 2m quadrats positioned within the 400-plant population is shown in Figure 7.1. We wish to estimate the total number of plants within the 20m × 20m macroplot. Counts of plants are in the individual quadrats. The sampling unit in this case is the 2m × 2m quadrat. The sample shown in Figure 7.1 is a set of 10 randomly selected quadrats. The sampled population in this case is the total collection of all possible 2m × 2m quadrats that could be placed in the macroplot (N = 100).

The population mean is the sum of all the values for each member of the population divided by the number of the population members. For example, if counting plants in quadrats, the mean is the sum of all the counts in all the quadrats divided by the number of quadrats.

$$\text{Population Mean } (\mu) = \frac{\text{Sum of Values}^{a} \text{ for Each Member of the Population}}{\text{Number of Population Members}}$$

Mathematically, this is given by:

$$\mu = \frac{X_1 + X_2 + \ldots X_N}{N}$$

where

X_1 = value of the first member of the population.
X_2 = value of the second member of the population.
X_N = value of the last member of the population.

or more concisely by:

$$\mu = \frac{\Sigma X}{N}$$

POPULATION PARAMETERS VERSUS SAMPLE STATISTICS

Population parameters are descriptive measures that characterize the population and are assumed to be fixed but unknown quantities that change only if the population changes. Greek letters such as μ and σ are often used to denote parameters. If we count all the plants in all the quadrats that make up the 400-plant population shown in Figure 7.1 (400 plants) and divide by the total number of possible 2m × 2m quadrat locations in the macroplot (100 quadrats), we obtain the true average number of plants per quadrat (4 plants/quadrat). This, assuming we have made no errors, is the **true population mean** (μ). If we know how much each individual quadrat differs from the true population mean, we can calculate another important population parameter, the **true population standard deviation** (σ). The standard deviation is a measure of how similar each individual observation is to the overall mean and is the most common measure of variability used in statistics. Populations with a large amount of variation among possible sampling units will have a larger

The sample mean is the estimate of the population mean from the sample.

$$\text{Sample Mean } (\bar{X}) = \frac{\text{Sum of Values, e.g., Heights, of Each Observation in Sample}}{\text{Number of Observations in Sample}}$$

The equivalent mathematical statement is:

$$\bar{X} = \frac{\Sigma X}{n}$$

[a] *These values can be heights, counts, cover values, etc.*

The population standard deviation is the square root of the population variance (denoted σ^2).

Population Variance (σ^2) = Sum of (Value Associated with Member of Population − Population Mean)^2 / Number of Population Members

Mathematically, this is given by:

$$\sigma^2 = \frac{(X_1-\mu)^2 + (X_2-\mu)^2 + \ldots + (X_N-\mu)^2}{N}$$

or more concisely by:

$$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N}$$

Population Standard Deviation (σ) $= \sqrt{\text{Population Variance}}$

Mathematically, this is given by: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$

standard deviation than populations with sampling units that are more similar to one another.

Sample statistics are descriptive measures derived from a sample (e.g., 10 of the 100 possible 2m × 2m quadrats). Sample statistics provide estimates of population parameters. Sample statistics will vary from sample to sample, in addition to changing whenever the underlying population changes. Roman letters such as $\bar{X}$ for the **sample mean** and s for the **sample standard deviation** are usually used for sample statistics. Consider the following simple example where a sample of three sampling units yields values of 9, 10, and 14 plants/quadrat:

The sample mean ($\bar{X}$) = (9+10+14)/3 = 11 plants/quadrat

The sample standard deviation s is an estimate of the population standard deviation. It is equivalent to the population standard deviation except that μ is replaced by its estimator $\bar{X}$ and N in the denominator is replaced by n − 1.

Mathematically, this is given by:

$$s = \sqrt{\frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \ldots + (X_n-\bar{X})^2}{n-1}}$$

Or more concisely by:

$$s = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}}$$

We could also calculate from this sample a sample standard deviation (s). The sample standard deviation describes how similar each individual observation is to the sample mean. The standard deviation is easily calculated with a simple hand calculator using the "s" or "s_{n-1}" key. The standard deviation (s) for the simple example above is 2.65 plants/quadrat. Consider another simple example with sampling unit values of 2, 10, and 21 plants/quadrat.

The mean ($\bar{X}$) = (2+10+21)/3 = 11 plants/quadrat

The standard deviation (s) for this example is 9.54 plants/quadrat.

Thus, both examples have a sample mean of 11 plants/quadrat, but the second one has a higher standard deviation (9.54 plants/quadrat) than the first (2.65 plants/quadrat), because the individual quadrat values differ more from one another in the second example.

In the example shown in Figure 7.1, the true population mean is 4.00 plants/quadrat, whereas the sample mean is 5.00 plants/quadrat. The true population standard deviation is 5.005 plants/quadrat, whereas the sample standard deviation is 6.146 plants/quadrat.

ACCURACY VERSUS PRECISION

Accuracy is the closeness of a measured or computed value to its true value. Precision is the closeness of repeated measurements of the same quantity. A simple example will help illustrate the difference between these two terms. Two quartz-based clocks, equally capable of tracking time, are sitting side-by-side on a table. Someone comes by and advances one of the clocks by 1 hour. Both clocks will be equally "precise" at tracking time, but one of them will not be "accurate."

Efficient sampling designs try to achieve high precision. When we sample to estimate some population parameter, our sample standard deviation gives us a measure of the repeatability, or precision of our sample; it does not allow us to assess the accuracy of our sample. If counts of plants within different quadrats of a sample are similar to one another (e.g., the example above with a mean of 11 and a standard deviation = 2.65), then it is likely that different independent samples from the same population will yield similar sample means and give us high precision. When quadrat counts within a sample are highly variable (e.g., the example above with a mean of 11 and a standard deviation of 9.54), individual sample means from separate independent samples may be very different from one another, giving us low precision. In either case, if the counting process is biased (perhaps certain color morphs or growth forms of individuals are overlooked), results may be inaccurate.

SAMPLING VERSUS NONSAMPLING ERRORS

Sampling errors result from chance; they occur when sample information does not reflect the true population information. These errors are introduced by measuring only a subset of all the sampling units in a population.

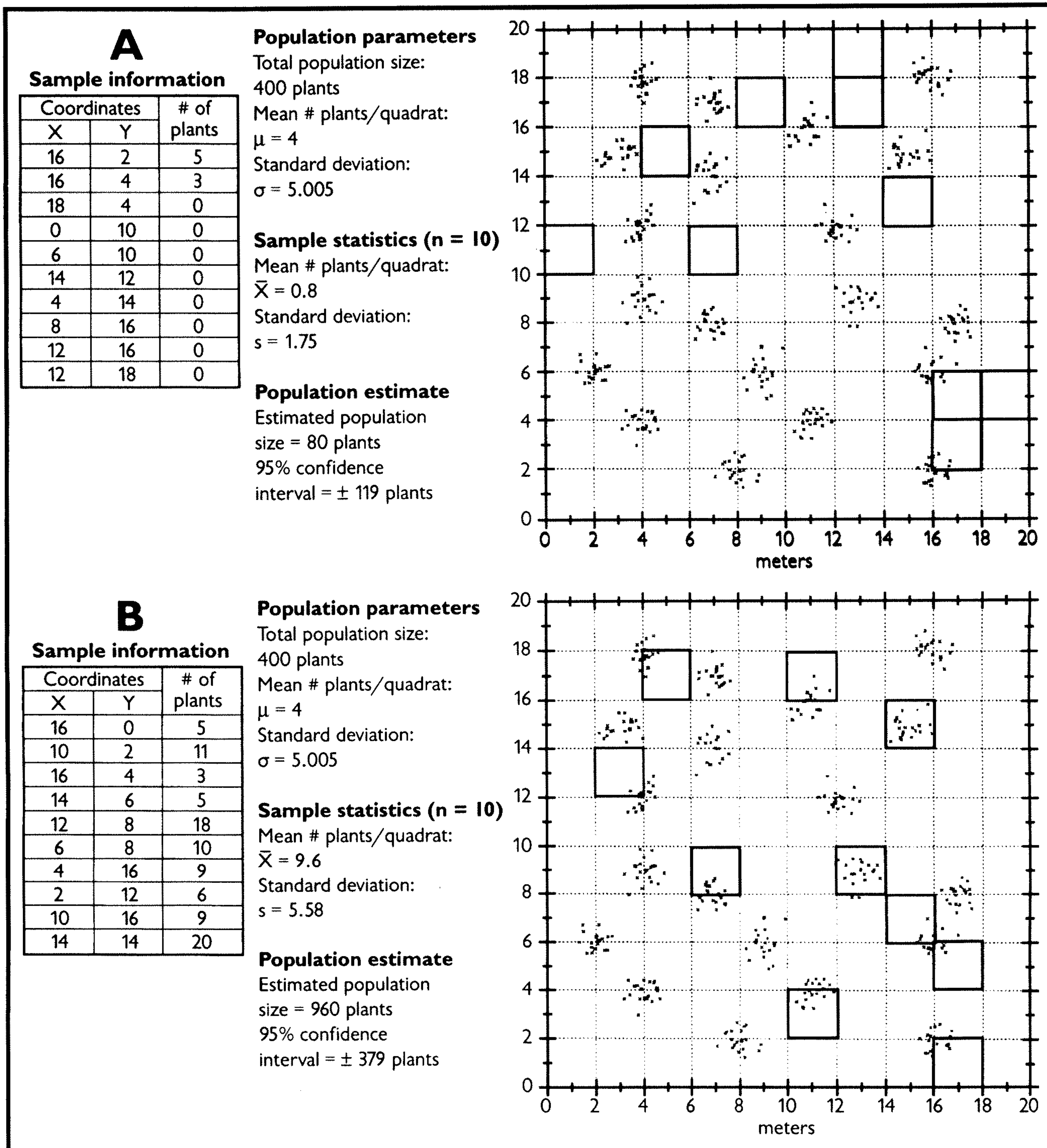
Sampling errors are illustrated in Figure 7.2, in which two separate, completely random samples (A and B) are taken from the 400-plant population shown in Figure 7.1. In each case, ten 2m × 2m quadrats are sampled, and an estimate is made of the total number of plants within the population. The sample shown in Figure 7.2A produces a population estimate of only 80 plants, whereas the sample shown in Figure 7.2B yields an estimate of 960 plants. Both estimates are poor because of sampling error (chance placement of the quadrats resulted in severe underestimates or overestimates of the true population total).

You can imagine the problems that can arise if you monitor the same population 2 years in a row and get sample information that indicates that the population shifted from 960 plants to 80 plants when it really did not change at all. Sampling errors can lead to two kinds of mistakes: 1) missing real changes (missed-change errors) and 2) detecting apparent changes that do not really exist (false-change errors).

The risk of committing sampling errors can be estimated from the sampling data. Some of the basic sampling design tools covered in Chapter 8 enable you to evaluate the effectiveness of your monitoring study by taking a closer look at the sampling data. This can be especially helpful when setting up new projects; an evaluation of pilot sampling data can point out potential sampling error problems, enabling an investigator to fix them at an early stage of the project. Good sampling designs can reduce sampling errors without increasing the cost of sampling.

Nonsampling errors are errors associated with human, rather than chance, mistakes. Examples of nonsampling errors include the following:

- Using biased selection rules such as selecting "representative samples" by subjectively locating sampling units or by substituting sampling units that are "easier" to measure.
- Using sampling units in which attributes cannot be accurately counted or measured. For example, counts of grass stems within a quadrat with counts in the hundreds may lead to numerous counting errors.
- Inconsistent field sampling effort. Nonsampling errors can be introduced if different investigators use different levels of effort (e.g., one investigator makes counts from "eye-level," whereas another counts by kneeling next to the quadrat) or ability (e.g., one investigator can't hear the high-pitched bird calls that another can).
- Transcription and recording errors. Nonsampling errors can be introduced if the data recorder's "7s" look like "1s" to the person entering the data.



A

Sample information

Coordinates X	Coordinates Y	# of plants
16	2	5
16	4	3
18	4	0
0	10	0
6	10	0
14	12	0
4	14	0
8	16	0
12	16	0
12	18	0

B

Sample information

Coordinates X	Coordinates Y	# of plants
16	0	5
10	2	11
16	4	3
14	6	5
12	8	18
6	8	10
4	16	9
2	12	6
10	16	9
14	14	20

Figure 7.2. Examples of sampling errors from sampling the 400-plant population. The population estimates of 80 plants and 960 plants are far from the true population of 400 plants.

- Incorrect or inconsistent species identification. This category also includes biases introduced by missing certain size classes or color morphs.

Because sampling designs and statistical analyses are based on the assumption that nonsampling errors are zero, the number of nonsampling errors needs to be minimized. Ensure that your sampling unit makes sense for the type of measurement technique you have selected. When different personnel are used in the same monitoring study, conduct rigorous training and testing to ensure consistency in counts or measurements. Design field data forms (see Chapter 6) that are easy to use and easy for data transcribers to interpret. Proof all data entered into computer programs to ensure that entered numbers are correct. In contrast to sampling errors, the probability of nonsampling errors occurring cannot be assessed from pilot sample data.

SAMPLING DISTRIBUTIONS

One way of evaluating the risk of obtaining a sample value that is vastly different from the true value (such as population estimates of 80 or 960 plants when the true population is 400 plants) is to sample a population repeatedly and to look at the differences among the repeated population estimates. If almost all the separate, independently derived, population estimates are similar, then you know you have a good sampling design with high precision. If many of the independent population estimates are not similar, then you know your precision is low.

The 400-plant population can be resampled by erasing the 10 quadrats (as shown in either Fig. 7.1 or Fig. 7.2) and placing 10 more in new, random positions. We can keep repeating this procedure, each time writing down the sample mean. Plotting the results of a large number of individual sample means in a simple histogram yields a sampling distribution. A sampling distribution is a distribution of many independently gathered sample statistics (most often a distribution of sample means). Under most circumstances, this distribution of sample means fits a normal or bell-shaped curve.

A distribution of population-size estimates from 10,000 separate random samples using ten 2m × 2m quadrats from the 400 plant population is shown in Figure 7.3A. The x-axis shows the range of different population estimates, and the y-axis shows the relative and actual frequency of the different population estimates. Think of this as the results of 10,000 different people sampling the same population on the same day, each one setting out 10 randomly positioned 2m × 2m quadrats (somehow without negatively impacting the population) and coming up with their own independent population estimate. The highest population estimate out of the 10,000 separate samples was 960 plants, and the lowest population estimate was zero (four of the 10,000 samples yielded a population estimate of zero). The shape of this distribution indicates the magnitude of likely sampling errors. Wide distributions mean that sampling could yield population estimates that are "far" from the true population value. A sampling design that led to the type of sampling distribution depicted in Figure 7.3A would not be useful since few of the estimates approach the true population size of 400 plants. *One of the principal objectives in sampling design is to make the shape of sampling distributions as narrow as possible.*

Fortunately, you do not have to repeatedly sample your population and see how wide your sampling distribution is to determine if you need to change anything. There are some simple statistical tools that provide a convenient shortcut for evaluating the precision of your sampling effort from a single sample. These tools involve calculating standard errors and confidence intervals to estimate sampling precision levels.

Standard Error

A **standard error** is the standard deviation of a large number of independent sample means. It is a measure of precision that you derive from a single sample. To paraphrase the earlier statement regarding an important objective of sampling design, *one of the principal objectives in sampling design is to reduce the size of the standard error.* This formula demonstrates that there are only two

Standard error is the standard deviation of all possible means of samples of size n from a population. The standard error quantifies the certainty with which the mean computed from a random sample estimates the true mean of the population from which the sample was drawn. We estimate the standard error from a random sample taken from the population. The best estimate of the population standard error is

Formula for standard error:

$$SE = \frac{s}{\sqrt{n}}$$

where SE = standard error
s = standard deviation
n = sample size

ways of minimizing standard errors—either 1) increase the sample size (n) or 2) decrease the standard deviation (s):

- Increase sample size. A new sampling distribution of 10,000 separate random samples drawn from our example population is shown in Figure 7.3B. This distribution came from randomly drawing samples of twenty 2m × 2m quadrats instead of the ten quadrats used to create the sampling distribution in Figure 7.3A. This increase in sample size from 10 to 20 provides a 29.3% improvement in precision (as measured by the reduced size of the standard error).
- Decrease sample standard deviation. Another sampling distribution of 10,000 separate random samples drawn from our 400-plant population is shown in Figure 7.3C. The sampling design used to create this distribution of population estimates is similar to the one used to create the sampling distribution in Figure 7.3B. The only difference between the two designs is in quadrat shape. The sampling distribution shown in Figure 7.3B came from using twenty 2m × 2m quadrats; the sampling distribution shown in Figure 7.3C came from using twenty 0.4m × 10m quadrats. This change in quadrat shape reduced the true population standard deviation from 5.005 plants to 3.551 plants. This change in quadrat shape led to a 29.0% improvement in precision over the 2m × 2m design shown in Figure 7.3B (as measured by the reduced size of the standard error). This 29.0% improvement in precision came without changing the sampling unit area ($4m^2$) or the number of quadrats sampled (n = 20); only the quadrat shape (from square to rectangular) changed. When compared with the original sampling design of ten 2m × 2m quadrats, the twenty 0.4m × 10m quadrat design led to a 49.8% improvement in precision. Details of this method and other methods of reducing sample standard deviation are covered in Chapter 8.

How is the standard error most often used to report the precision level of sampling data? Sometimes the standard error is reported directly. You may see tables with standard errors reported or graphs that include error bars that show ± 1 standard error. Often, however, the standard error is multiplied by a coefficient that converts the number into something called a confidence interval.

Confidence Intervals

A **confidence interval** provides an estimate of precision around a sample mean, a sample proportion, or an estimate of total population size that specifies the likelihood that the interval includes the true value.

A confidence interval is the interval within which a true parameter value lies with known probability. It is a measure of the reliability of our sample estimate of the parameter value.

A confidence interval includes two components: 1) the confidence interval width (e.g., ± 340 plants), and 2) the confidence level (e.g., 90%, 95%). The confidence level indicates the probability that the interval includes the true value. Confidence interval width decreases as the confidence level decreases. Three confidence intervals for the design that used a sample of 10 of the 2m × 2m quadrats are shown again in Figure 7.4A, where they are graphed in a format commonly used to report confidence intervals. There is no gain in precision associated with the narrowing of confidence interval width as you go from left to right in Figure 7.4A (i.e., from 95% confidence, to 80% confidence, to 50% confidence); only the probability that the confidence interval includes the true value is altered. Another set of three confidence intervals is shown in Figure 7.4B. Like

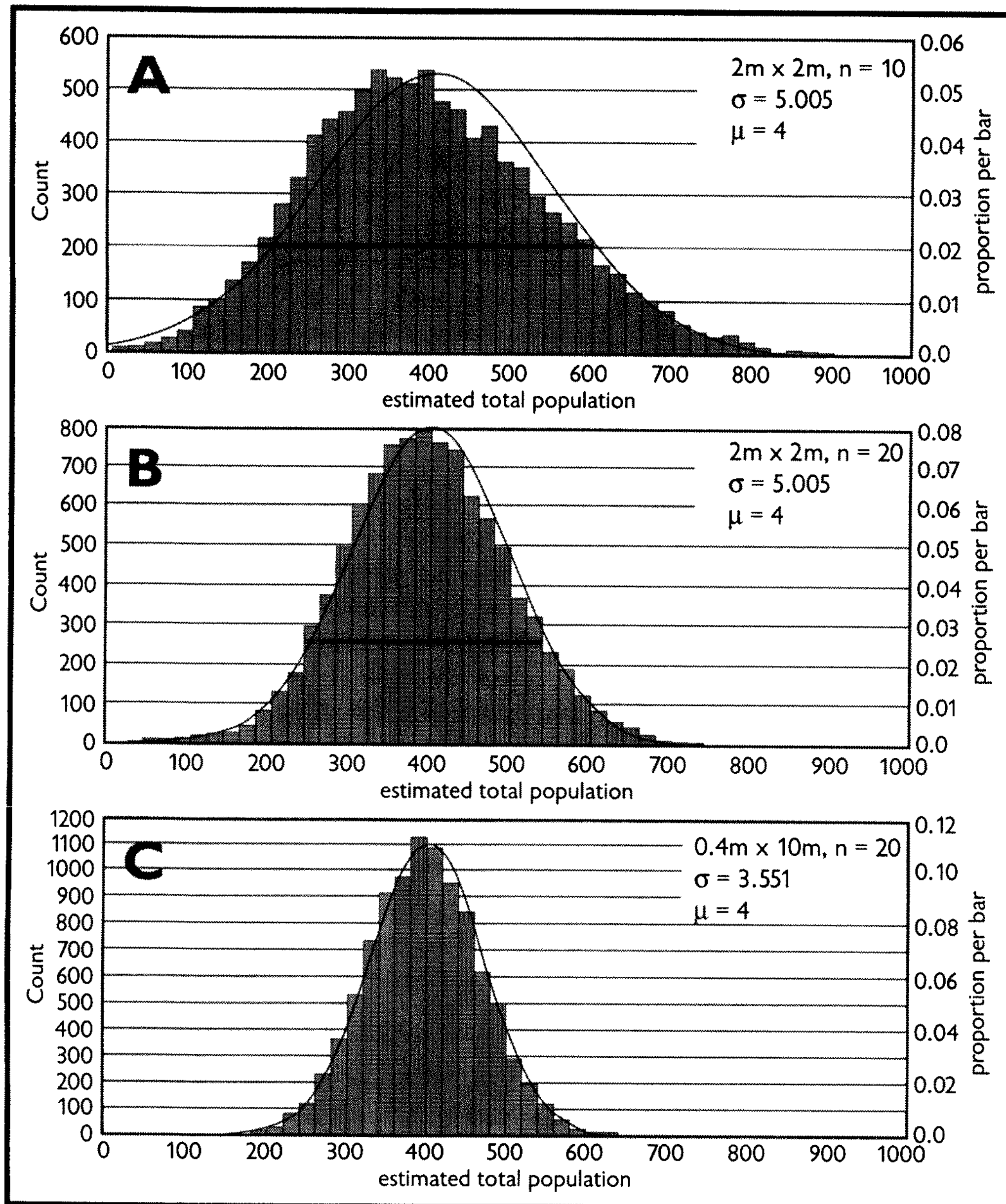


Figure 7.3. Sampling distributions from three separate sampling designs used on the 400-plant population. All distributions were created by sampling the population 10,000 separate times. The smooth lines show a normal bell-shaped curve fit to the data. Figure 3A shows a sampling distribution where ten 2m × 2m quadrats were used. Figure 3B shows a sampling distribution where twenty 2m × 2m quadrats were used. Figure 3C shows a sampling distribution where twenty 0.4m × 10m quadrats were used.

Figure 7.4A, confidence intervals get narrower as we move from left to right in the graph, but this time the confidence level is the same (95%), and the narrower widths came from using different sampling designs. There is a gain in precision associated with the narrowing of confidence interval width as you go from left to right in Figure 7.4B (i.e., from the ten 2m × 2m design to the twenty 2m × 2m design to the twenty 0.4m × 10m design) because we have reduced the uncertainty of our population estimate by tightening the confidence interval width at the same confidence level.

To calculate confidence intervals for sample means, we need two values: 1) the standard error ($SE = s/\sqrt{n}$), and 2) the corresponding value from a table of critical values of the t distribution (see Appendix III for instructions on calculating confidence intervals around proportions).

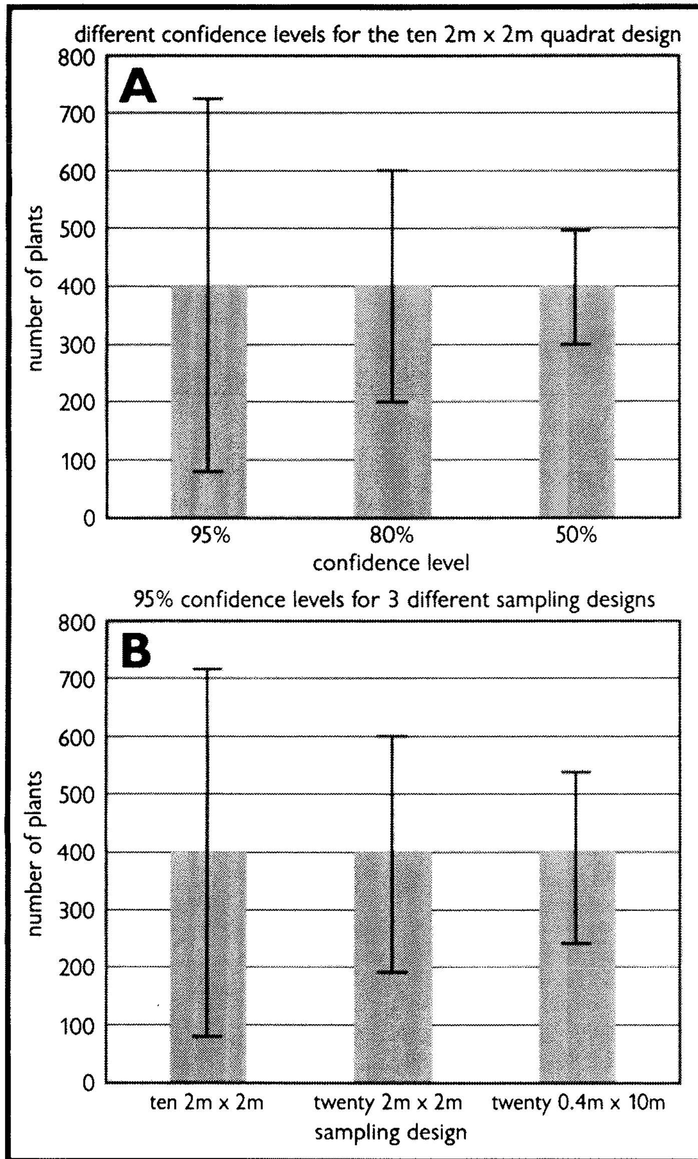


Figure 7.4. Comparison of confidence intervals and confidence levels for different sampling designs from the 400-plant population. Figure A shows three different confidence levels (95%, 80%, and 50%) for the same data set based upon sampling ten 2m × 2m quadrats. Figure B shows 95% confidence intervals for three different sampling designs that differ in the level of precision of the population estimates.

The confidence interval half-width, extending an equal distance on both sides of the mean, is the standard error × the critical *t* value (except when sampling from finite populations; see the next section). The appropriate critical value of *t* depends on the level of confidence desired and the number of sampling units (n) in the sample. Values of the *t* distribution can be found in many statistical texts.[1] To use a *t* table, you must first select the appropriate confidence level column. If you want to be 95% confident that your confidence interval includes the true mean, use the column headed $\alpha(2) = 0.05$. For 90% confidence, use the column headed $\alpha(2) = 0.10$. You use $\alpha(2)$ because you are interested in a confidence interval on both sides of the mean. You then use the row indicating the number of degrees of freedom (v), which is the number of sampling units minus one (n-1).

For example, if we sample 20 quadrats in the macroplot shown in Figure 7.1 and come up with a mean of 5.0 plants and a standard deviation of 4.616, we would calculate a 95% confidence interval around our sample mean:

The standard error ($SE = s/\sqrt{n}$) = 4.616/4.472 = 1.032

The appropriate *t* value from a *t* table for 19 degrees of freedom (v) is 2.093. One-half of our confidence interval width is then

SE × *t*-value = 1.032 × 2.093 = 2.160

Our 95% confidence interval can then be reported as 5.0 ± 2.16 plants/quadrat, or we can report the entire confidence interval width from 2.84 to 7.16 plants/quadrat. This indicates a 95% chance that our interval from 2.84 plants/quadrat to 7.16 plants/quadrat includes the true value.[2]

Another way to think of 95% confidence intervals calculated from sampling data is that the interval specifies a range that should include the true value 95% of the time. If you are calculating 95% confidence intervals and independently randomly sample a population 100 different times, you should see that approximately 95 of the intervals will include the true mean and 5 will miss it (Fig. 7.2A shows a sample that misses the true mean). This relationship is shown in Figure 7.5 where 100 independent population estimates are graphed with 95% confidence intervals from

---

[1] *Links to the on-line tables on the Web can be found on our Web page (see Preface).*

[2] *This is not a very precise estimate, but it would improve with the application of the finite correction factor. In this example, we have sampled 20 of the 100 possible quadrats, or 20% of the population. We would apply the finite correction factor described in the next section.*

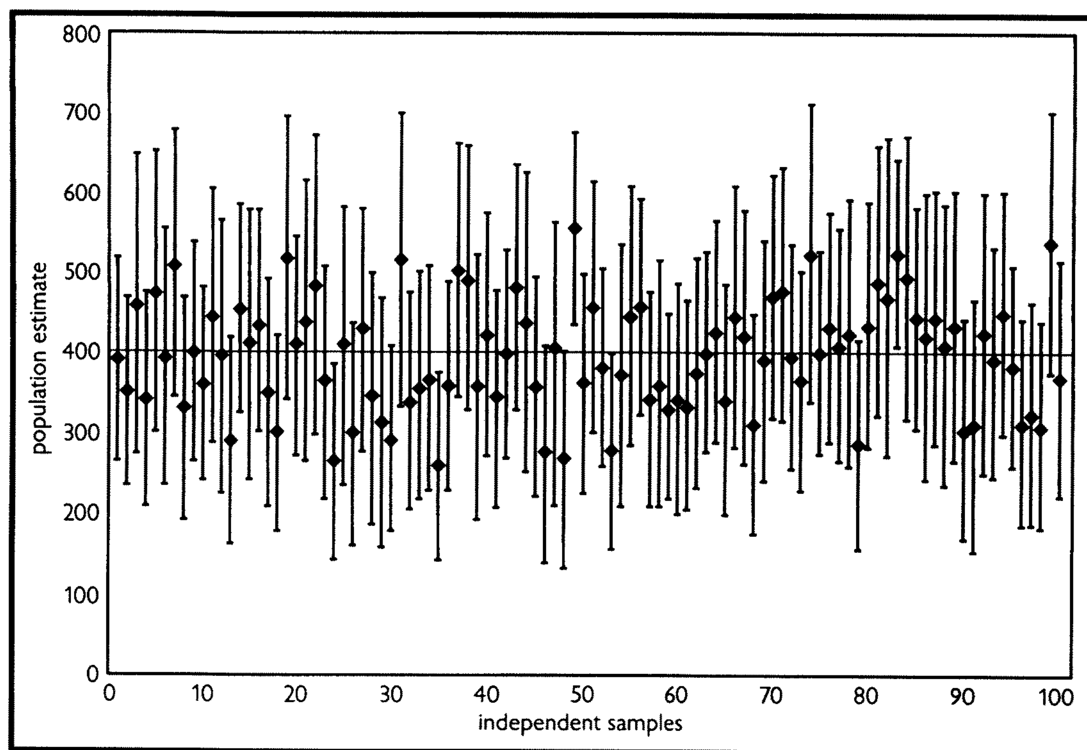


Figure 7.5. Population estimates from 100 separate random samples from the 400-plant population. Each sample represents the population estimate from sampling twenty 0.4m × 10m quadrats. The horizontal line through the graph indicates the true population of 400 plants. Vertical bars represent 95% confidence intervals. Four of the intervals miss the true population size.

the 400-plant populations using samples of twenty 0.4m × 10m quadrats. You will notice that the solid diamonds, used to show each of the 100 population estimates, fluctuate around the true population value of 400 plants. You will also notice that 96 out of 100 confidence intervals shown in Figure 7.5 include the true value. If the confidence level had been set at 80%, then approximately 20 of the intervals would have failed to include the true value. A 99% confidence level would have led to approximately only one interval out of the 100 that did not include the true population size (to capture the true value more often, the individual confidence interval widths for a 99% confidence level are wider than the confidence interval widths for a 95% confidence level).

FINITE VERSUS INFINITE POPULATIONS

If we are sampling with quadrats and no two quadrats may overlap, there is a finite number of quadrats that can be placed in the area to be sampled (this is called sampling without replacement). If the sampled area is large, then the number of quadrats placed in the area may be very large as well, but nonetheless finite. On the other hand, an infinite number of points or lines could be placed in the area to be sampled. This is because points, at least theoretically, are dimensionless, and lines are dimensionless in one direction. This means, at least for all practical purposes, that a population of points or of lines is infinite.

If the area to be sampled is large relative to the area that is actually sampled, the distinction between finite and infinite is of only theoretical interest. When, however, the area sampled makes up a significant portion of the area to be sampled, we can apply **the finite population correction factor,** which reduces the size of the standard error.

The finite population correction factor (FPC) should always be applied if you are sampling more than 5% of the population. It is applied to confidence intervals, as well as statistical tests (see Chapter 9).

Formula for the finite population correction factor:

$$FPC = \sqrt{\frac{N-n}{N}}$$

where N = total number of potential quadrat positions
n = number of quadrats sampled

Here is an example where the FPC is applied to the standard error:

$$SE' = (SE)\sqrt{\frac{N-n}{N}} \qquad SE' = (0.73)\sqrt{\frac{100-30}{100}} = 0.61$$

where SE' = corrected standard error
SE = uncorrected standard error
N = total number of potential quadrat positions
n = number of quadrats sampled

When n is small relative to N, the equation is close to 1, whereas when n is large relative to N, the value approaches zero. The standard error ($s/\sqrt{n}$) is multiplied by the finite population correction factor to yield a corrected standard error for the finite population.

Consider the following example. The density of plant species X is estimated within a 20m × 50m macroplot (total area = 1000m^2). This estimate is obtained by collecting data from randomly selected 1m × 10m quadrats (10m^2). Sampling without replacement, there are 100 possible quadrat positions.

Thus, our population, N, is 100. Let us say we take a random sample, n, of 30 of these quadrats and calculate a mean of eight plants per quadrat and a standard deviation of four plants per quadrat. Our standard error is thus: $s/\sqrt{n} = 4/\sqrt{30} = 0.73$. Although our sample mean is an unbiased estimator of the true population mean and needs no correction, the standard error should be corrected by the finite population correction factor.

Because the standard error is one of the factors used to calculate confidence intervals (the other is the appropriate value of t from a t table), correcting the standard error with the finite population correction factor makes the resulting confidence interval narrower. It does this, however, only if n is sufficiently large relative to N. A rule of thumb is that unless the ratio n/N is greater than 0.05 (i.e., you are sampling more than 5% of the population area), there is little to be gained by applying the finite population correction factor to your standard error.

The finite population correction factor is also important in sample size determination (see Chapter 8) and in adjusting test statistics (see Chapter 9). The finite population correction factor works, however, only with finite populations, which we will have when using quadrats, but will not have when using points or lines.

FALSE-CHANGE ERRORS AND MISSED-CHANGE ERRORS

False-change errors and missed-change errors relate to situations where two or more sample means or proportions are being compared with some statistical test. This comparison may be between two or more places or the same place between two or more periods. These terms are pertinent to both the planning and the interpretation stages of a monitoring study. Consider a simple example where you have sampled a population in two different years and now you want to determine whether a change took place between the two years. You usually start with the assumption, called the null hypothesis, that no change has taken place. You may make two types of decisions when interpreting the results of a monitoring study: 1) you can decide that a change took place, or 2) you can decide that no change took place. In either case, you can be right, or you can be wrong (Fig. 7.6).

The conclusion that a change took place may lead to some kind of action. For example, if a population of a rare fish is thought to have declined, a change in management may be needed. If a change was detected but did not actually occur, this constitutes a false-change error, a sort of false alarm. Controlling this type of error is important because taking action unnecessarily can be ex-

pensive (e.g., a range permittee is not going to want to reduce grazing intensity along a stream bank if a decline in a rare fish population really did not take place). There will be a certain probability of concluding that a change took place even if no difference actually occurred. The probability of this occurring is usually labeled the *P* value, which is one of the types of information that comes out of a statistical analysis of the data (see Chapter 9). The *P* value reports the likelihood that the observed difference was the result of a false-change error. For example, if a statistical test comparing two sample means yields a *P* value of 0.24, this indicates that there is a 24% chance of obtaining the observed result even if no true difference exists between the two sample means.

monitoring for change — possible errors

	no change has taken place	there has been a real change
monitoring system detects a change	false-change error (Type I) α	no error (Power) $1 - \beta$
monitoring system detects no change	no error $(1 - \alpha)$	missed-change error (Type II) β

Figure 7.6. Four possible outcomes for a statistical test of some null hypothesis, depending on the true state of nature.

Some threshold value for this false-change error rate should be set in advance so that the *P* value from a statistical test can be evaluated relative to the threshold. *P* values from a statistical test that are smaller than or equal to the threshold are considered statistically "significant," whereas *P* values that are larger than the threshold are considered statistically "nonsignificant." Statistically significant differences may or may not be ecologically significant, depending on the magnitude of difference between the two values. The most commonly cited threshold for false-change errors is the 0.05 level, but there is no reason to arbitrarily adopt the 0.05 level as the appropriate threshold. The decision of what false-change error threshold to set depends on the relative costs of making this type of mistake and the impact of this error level on the other type of mistake, a missed-change error.

When monitoring a rare species, we are usually most concerned about declines. The conclusion that no change took place usually does not lead to changes in management practices. Failing to detect a true change constitutes a missed-change error. Controlling this type of error is important because failing to take action when a true change actually occurred may lead to the serious decline of a population.

Statistical power is the complement of the missed-change error rate (e.g., a missed-change error rate of 0.25 gives you a power of 0.75; a missed-change error rate of 0.05 gives you a power of 0.95). High power (a value close to 1) is desirable and corresponds to a low risk of a missed-change error. Low power (a value close to 0) is not desirable because it corresponds to a high risk of a missed-change error.

Since power levels are directly related to missed-change error levels, either level can be reported and the other level easily calculated. Power levels are often reported instead of missed-change error levels, because it seems easier to convey this concept in terms of the certainty of detecting real changes. For example, the statement "I want to be at least 90% certain of detecting a real change of 5 plants/quadrat" (power is 0.90) is simpler to understand than "I want the probability of missing a real change of 5 plants/quadrat to be 10% or less" (missed-change error rate is 0.10).

An assessment of statistical power or missed-change errors has been virtually ignored in the field of environmental monitoring. A survey of over 400 papers in fisheries and aquatic sciences through the 1980s found that 98% of the articles that reported nonsignificant results failed to report any power results (Peterman 1990). A separate survey, reviewing toxicology literature, found high power in only 19 out of 668 reports that failed to reject the null hypothesis (Hayes 1987). Similar surveys in other fields such as psychology or education have turned up "depressingly low" levels of power (Brewer 1972; Cohen 1988).

It is not clear why missed-change errors have traditionally been ignored in environmental monitoring. Perhaps researchers have not been sufficiently exposed to the idea of missed-change

errors nor understood how considering power can improve their work. Perhaps people have not realized the potentially high costs associated with making missed-change errors. Most introductory texts and statistics courses deal with the material only briefly. Computer packages for power analysis have only recently become available.

The situation has improved in recent years. A literature review in the 1980s would not have turned up many articles dealing with statistical power issues. A literature review today would turn up dozens of articles in many disciplines from journals all over the world (see Peterman [1990] and Fairweather [1991] for good review papers on statistical power). In the 1990s, ecologists and conservation biologists began paying more attention to power concerns (Andren 1996; Gibbs et al. 1998; Green and Young 1993; Osenberg et al. 1994). A number of recent wildlife biology papers discuss power issues and monitoring wildlife population trends (Beier and Cunningham 1996; Hatfield et al. 1996; Kendell 1992; Taylor and Gerrodette 1993; Van Strien et al. 1997; Zielinski and Stauffer 1996). A few papers have been published specifically on power analysis and amphibian populations (Hayes and Steidl 1997; Reed and Blaustein 1995).

False-change and missed-changed errors are related (although not directly). Reducing one increases the other (discussed and graphically portrayed below). Balancing these when designing a monitoring study requires consideration of which error is more costly in terms of management and natural resources. Most commonly in the management of rare species, we are concerned about a decline; committing a missed-change error (missing a true decline) may be very costly in terms of the viability of the species because we may fail to implement management action until the decline becomes very obvious. In other situations, a conclusion that no change took place may trigger a management action. For example, if you were trying to control weeds, and if the monitoring suggested no changes were resulting from your current management, you would likely institute alternative or more intensive management. Similarly, if your management was attempting to increase a rare species, and if your monitoring suggested no change, you might change the type of management being implemented. In both of these situations a missed-change error would result in increased management activity that may not be necessary (i.e., your current management may actually be effective at reducing the weed population, or increasing the rare species, but your monitoring does not detect it), but committing such an error and changing management would probably not be detrimental to the resource you are trying to manage. A false-change error, however, may make you believe that your management is effective at decreasing the weed or increasing the rare species when, in fact, your management is ineffective and neither has actually changed.

MINIMUM DETECTABLE CHANGE

Another sampling design concept that is directly related to statistical power and false-change error rates is the size of the change that you want to be able to detect. This will be referred to as the minimum detectable change or MDC.

The MDC is the size of the change you identify in the management objective (see Chapter 14). Setting MDCs requires considering both the biological implications and the monitoring costs. If power and the false-change error rate remain the same, detecting a small change will require more intensive monitoring (usually more sampling units) than detecting a large change. With a large enough sample size, statistically significant changes can be detected for changes that have no biological significance (Johnson 1999).

How large a change should be considered biologically meaningful? Should a 30% change in the mean density of a rare plant population be cause for alarm? Should a population decline of 20% of a rare animal be of concern? What about a 15% change or a 10% change? If, for example, an intensive monitoring design leads to the conclusion that the mean density of a plant population increased from 10.0 plants/m^2 to 10.1 plants/m^2, does this represent some biologically

meaningful change in population density? Probably not. Further, a design that detected such a small change wasted limited monitoring resources.

Setting a reasonable MDC can be difficult when little is known about the natural history of a particular species (see Chapter 14 for general suggestions). The initial MDC, set during the design of a new monitoring study as part of the objectives, can be modified once monitoring information demonstrates the size and rate of population fluctuations.

HOW TO ACHIEVE HIGH STATISTICAL POWER

Statistical power is related to four, separate, sampling design components by the following function equation:

Power = a function of (α, MDC, n, and s)

where

α = false-change error rate

MDC = minimum detectable change

n = number of sampling units

s = standard deviation

Power can be increased in the following four ways:

1. Increasing the acceptable level of false-change errors (α).
2. Increasing the MDC.
3. Increasing the number of sampling units sampled. This method of increasing power is straightforward, but keep in mind that increasing n has less of an effect than decreasing s because the square root of sample size is used in the standard error equation ($SE = s/\sqrt{n}$).
4. Reducing standard deviation. This means altering the sampling design to reduce the amount of variation among sampling units (see Chapter 8).

Note that the first two ways of increasing power are related to making changes in the sampling objective, whereas the other two ways are related to making changes in the sampling design (see Chapter 14).

POWER AND TRADEOFFS—A GRAPHIC COMPARISON

In this section we take a graphic look at how altering these factors changes power. The comparisons in this section are based on sampling a fictitious plant population where we are interested in assessing plant density relative to an established threshold value of 25 plants/m^2. Any true population densities less than 25 plants/m^2 will trigger management action. We are only concerned with the question of whether the density is lower than 25 plants/m^2 and not whether the density is higher. In this example, our null hypothesis (H_0) is that the population density equals 25 plants/m^2, and our alternative hypothesis is that density is less than 25 plants/m^2. The density value of 25 plants/m^2 is the most critical single density value since it defines the lower limit of acceptable plant density.

The figures in this section are all based on sampling distributions where we happen to know the true plant density. Recall that a sampling distribution is a bell-shaped curve that depicts the distribution of a large number of independently gathered sample statistics. A sampling distribution defines the range and relative probability of any possible sample mean. You are more likely

to obtain sample means near the middle of the distribution than you are to obtain sample means near either tail of the distribution.

A sampling distribution based on sampling our fictitious population with a true mean density of 25 plants/m^2 is shown in Figure 7.7A. This distribution is based on a sampling design using thirty 1m × 1m quadrats where the true standard deviation is ± 20 plants/quadrat. If 1000 different people randomly sample and calculate a sample mean based on their 30 quadrat values, approximately half the individually drawn sample means will be less than 25 plants/m^2, and half will be greater than 25 plants/m^2. Approximately 40% of the samples will yield sample means less than or equal to 24 plants/m^2. A few of our 1000 individuals will obtain estimates of the

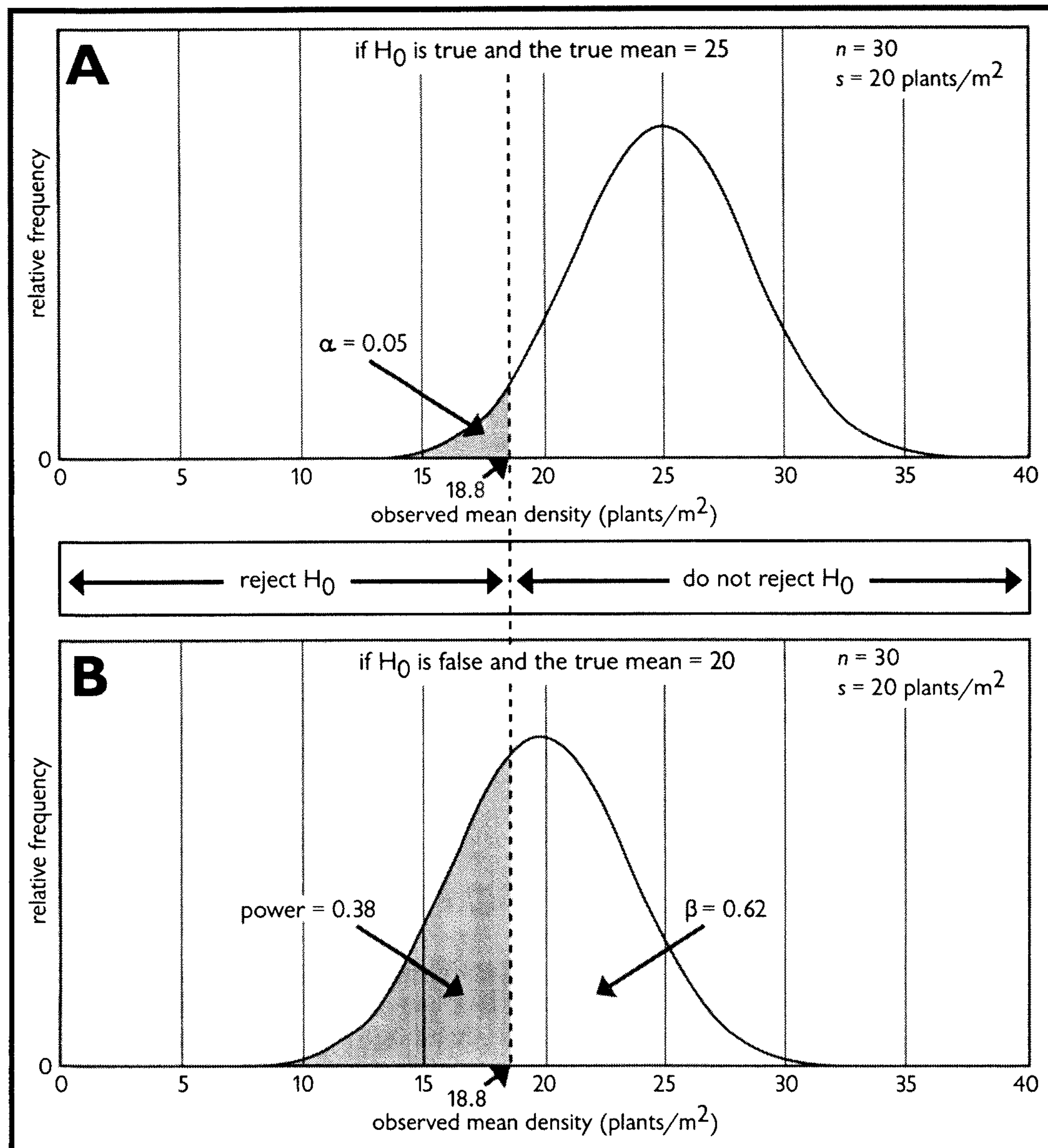


Figure 7.7. Example of sampling distributions for mean plant density in samples of 30 quadrats where the among-quadrat standard deviation is 20 plants/m^2. Part A is the sampling distribution for the case in which the null hypothesis, H_0, is true and the true population mean density is 25 plants/m^2. The shaded area in part A is the critical region for $\alpha = 0.05$ and the vertical dashed line is at the critical sample mean value, 18.8. Part B is the sampling distribution for the case in which H_0 is false and the true mean is 20 plants/m^2. In both distributions, a sample mean to the left of the vertical dashed line would reject H_0, and to the right of it, would not reject H_0. Power and β values in part B, in which H_0 is false and the true mean = 20, are the proportion of sample means that would occur in the region in which H_0 was rejected or not rejected, respectively.

mean density that deviate from the true value by a large margin. One of the individuals will likely stand up and say, "my estimate of the mean density is 13 plants/m^2," even though the true density is actually 25 plants/m^2. As interpreters of the monitoring information, we would conclude that, since 999 of the 1000 people obtained estimates of the density that were greater than 13, the true density is probably not 13. Our best estimate of the true mean density will be the average of the 1000 separate estimates (this average is likely to be extremely close to the actual true value).

Now that we have the benefit of 1000 independent estimates of the true mean density, we can return to the population at a later time; take a single, random sample of thirty 1m × 1m quadrats; calculate the sample mean; and then ask the question, "what is the probability of obtaining our sample mean value if the true population is still 25 plants/m^2?" If our sample mean density turns out to be 24 plants/m^2, would this lead to the conclusion that the population has crossed our threshold value? Seeing that our sample mean is lower than our target value might raise some concerns, but we have no objective basis to conclude that the true population is not, in fact, still actually 25 plants/m^2. We learned in the previous paragraph that a full 40% of possible samples are likely to yield mean densities of 24 plants/m^2 or less if the true mean is 25 plants/m^2. Thus, the probability of obtaining a single sample mean of 24 plants/m^2 or less when the true density is actually 25 plants/m^2 is approximately 0.40. Obtaining a sample mean of 24 plants/m^2 is consistent with the hypothesis that the true population density is actually 25 plants/m^2.

How small a sample mean do we need to obtain to feel confident that the population has indeed dropped below 25 plants/m^2? What will our interpretation be if we obtained a sample mean of 22 plants/m^2? Based on our sampling distribution from the 1000 people, the probability of obtaining an estimate of 22 plants/m^2 or less is around 20%, which represents a one-in-five chance that the true mean is still actually 25 plants/m^2. Based on the sampling distribution from our 1000 separate samplers, we can look at the likelihood of obtaining other different sample means. The probability of obtaining a sample of 20 plants/m^2 is 8.5%, and the probability of obtaining a sample of 18 plants/m^2 is 2.9% if the true mean density is 25 plants/m^2.

Since in most circumstances we will only have the results from a single sample (and not the benefit of 1000 independently gathered sample means), another technique must be used to determine whether the population density has dropped below 25 plants/m^2. One method is to run a statistical test that compares our sample mean to our density threshold value (25 plants/m^2). The statistical test will yield a *P* value that defines the probability of obtaining our sample mean if the true population density is actually 25 plants/m^2. As interpreters of our monitoring information, we will need to set some probability threshold *P* value to guide our interpretation of the results from the statistical test. This *P* value threshold defines our acceptable false-change error rate. If we run a statistical test that compares our sample mean to our density threshold value (25 plants/m^2), and if the *P* value from the test is lower than our threshold value, then we conclude that the population density has, in fact, declined below 25 plants/m^2. Thus, if we set our *P* value threshold to 0.05 and the statistical test yields a *P* value of 0.40, then we fail to reject the null hypothesis that the true population density is 25 plants/m^2. If, however, the statistical test yields a *P* value of 0.022, this is lower than our threshold *P* value of 0.05, and we would reject the null hypothesis that the population is 25 plants/m^2 in favor of our alternative hypothesis that the density is lower than 25 plants/m^2.

The relationship between the *P* value threshold of 0.05 and our sampling distribution based on sampling thirty 1m × 1m quadrats is shown in Figure 7.7A. The threshold density value corresponding to our *P* value threshold of 0.05 is 18.8 plants/m^2, which is indicated on the sampling distribution by the dashed vertical line. Thus, if we obtain a mean density of 18 plants/m^2, which is to the left of the vertical line, we reject the null hypothesis that the population density is 25 plants/m^2 in favor of an alternative hypothesis that density is lower than 25 plants/m^2. If we obtain a mean density of 21 plants/m^2, which is to the right of the vertical line, then we fail to reject the null hypothesis that the population density is really 25 plants/m^2.

So far, we have been discussing the situation where the true population density is right at the threshold density of 25 plants/m^2. Let us look now at a situation where we know the true density has declined to 20 plants/m^2. What is the likelihood of our detecting this true, density difference of 5 plants/m^2? Figure 7.7B shows a new sampling distribution based on the true density of 20 plants/m^2 (standard deviation is still ± 20 plants/m^2). We know from our previous discussion that sample means to the right of the vertical line in Figure 7.7A lead to the conclusion that we cannot reject the null hypothesis that our density is 25 plants/m^2. If our new sample mean turns out to exactly match the new true population mean (i.e., 20 plants/m^2), will we reject the idea that the sample actually came from a population with a true mean of 25 plants/m^2? No, at least not at our stated *P* value (false-change error) threshold of 0.05. A sample mean value of 20 plants/m^2 falls to the right of our dashed threshold line in the "do not reject H_0" portion of the graph, and we would have failed to detect the true difference that actually occurred. Thus, we would have committed a missed-change error.

What is the probability of missing the true difference of 5 plants/m^2 shown in Figure 7.7B? This probability represents the missed-change error rate (β), and it is defined by the nonshaded area under the sampling distribution in Figure 7.7B, which represents 62% of the possible sample mean values. Recall that the area under the whole curve defines the entire range of possible values that you could obtain by sampling the population with the true mean = 20 plants/m^2. If we bring back our 1000 sampling people and have each of them sample thirty 1m × 1m quadrats in our new population, we will find that approximately 620 of them will obtain estimates of the mean density that are greater than the threshold value of 18.8 plants/m^2 that is shown by the vertical dashed line.

What about the other 380 people? They will obtain population estimates fewer than the critical threshold of 18.8 plants/m^2, and they will reject the null hypothesis that the population equals 25 plants per quadrat. This proportion of 0.38 (380 people out of 1000 people sampling) represents the statistical power of our sampling design, and it is represented by the shaded area under the curve in Figure 7.7B. If the true population mean is indeed 20 plants/m^2 instead of 25 plants/m^2, then we can be 38% sure (power = 0.38) that we will detect this true difference of 5 plants/m^2. With this particular sampling design (thirty 1m × 1m quadrats) and a false-change error rate of $\alpha = 0.05$, we run a 62% chance ($\beta = 0.62$) that we will commit a missed-change error (i.e., fail to detect the true difference of 5 plants/m^2). If the difference of 5 plants/m^2 is biologically important, a power of only 0.38 would not be satisfactory.

We can improve the low-power situation in four different ways: 1) increase the acceptable false-change error rate, 2) increase the acceptable MDC, 3) increase sample size, or 4) decrease the standard deviation. New, paired, sampling distributions illustrate the influence of making each of these changes.

Increasing the Acceptable False-Change Error Rate

In Figure 7.7B, a false-change error rate of $\alpha = 0.05$ resulted in a missed-change error rate of $\beta = 0.62$ to detect a difference of 5 plants/m^2. Given these error rates, we are more than 12 times more likely to commit a missed-change error than we are to commit a false-change error. What happens to our missed-change error rate if we specify a new, higher, false-change error rate? Shifting our false-change error rate from $\alpha = 0.05$ to $\alpha = 0.10$ is illustrated in Figure 7.8 for the same sampling distributions shown in Figure 7.7. Our critical density threshold at the $P = 0.10$ level is now 20.21 plants/m^2, and our missed-change error rate has dropped from $\beta = 0.62$ down to $\beta = 0.47$ (i.e., the power to detect a true difference of 5 plants/m^2 increased from 0.38 to 0.53). A sample mean of 20 plants/m^2 will now lead to the correct conclusion that a difference of 5 plants/m^2 between the populations does exist. Of course, the penalty we pay for increasing our false-change error rate is that we are now twice as likely to conclude that a difference exists in situations when there is no true difference and our population mean is actually 25 plants/m^2. Changing the false-change error rate even more, to $\alpha = 0.20$ (Fig. 7.9), reduces the probability of

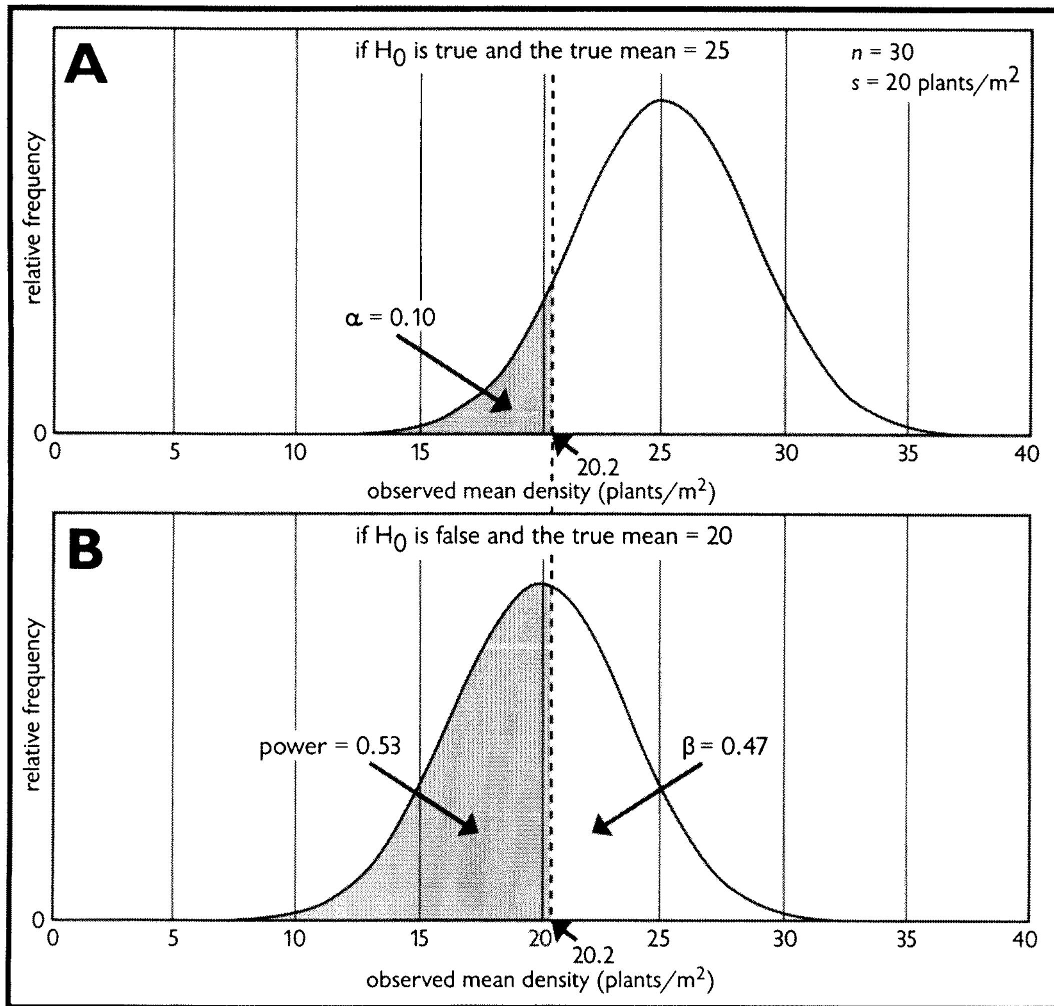


Figure 7.8. The critical region for the false-change error in the sampling distributions from Figure 7.7 has been increased from α = 0.05 to α = 0.10. Part B, in which the H_0 is false and the true mean = 20, shows that power is larger for α = 0.10 than for Figure 7.7 where α = 0.05.

making a missed-change error down to β = 0.29 (i.e., giving us a power of 0.71 to detect a true difference of 5 plants/m^2).

Increasing the Acceptable Minimum Detectable Change

Any sampling design is more likely to detect a true, large difference than a true, small difference. As the magnitude of the difference increases, we will see an increase in the power to detect the difference. This relationship is shown in Figure 7.10B, where we see a sampling distribution with a true mean density of 15 plants/m^2, which is 10 plants/m^2 below our threshold density of 25 plants/m^2. The false-change error rate is set at α = 0.05 in this example. This figure shows that the statistical power to detect this larger difference of 10 plants/m^2 (25 plants/m^2 to 15 plants/m^2) is 0.85 compared with the original power value of 0.38 to detect the difference of 5 plants/m^2 (25 plants/m^2 to 20 plants/m^2). Thus, with a false-change error rate of 0.05, we can be 85% certain of detecting a difference of 10 plants/m^2 or greater from our threshold of 25 plants/m^2. If we raised our false-change error from α = 0.05 to α = 0.10 (not shown in Figure 7.10), our power value would rise to 0.92, which creates a sampling situation where our two error rates are nearly equal (α = 0.10, β = 0.08).

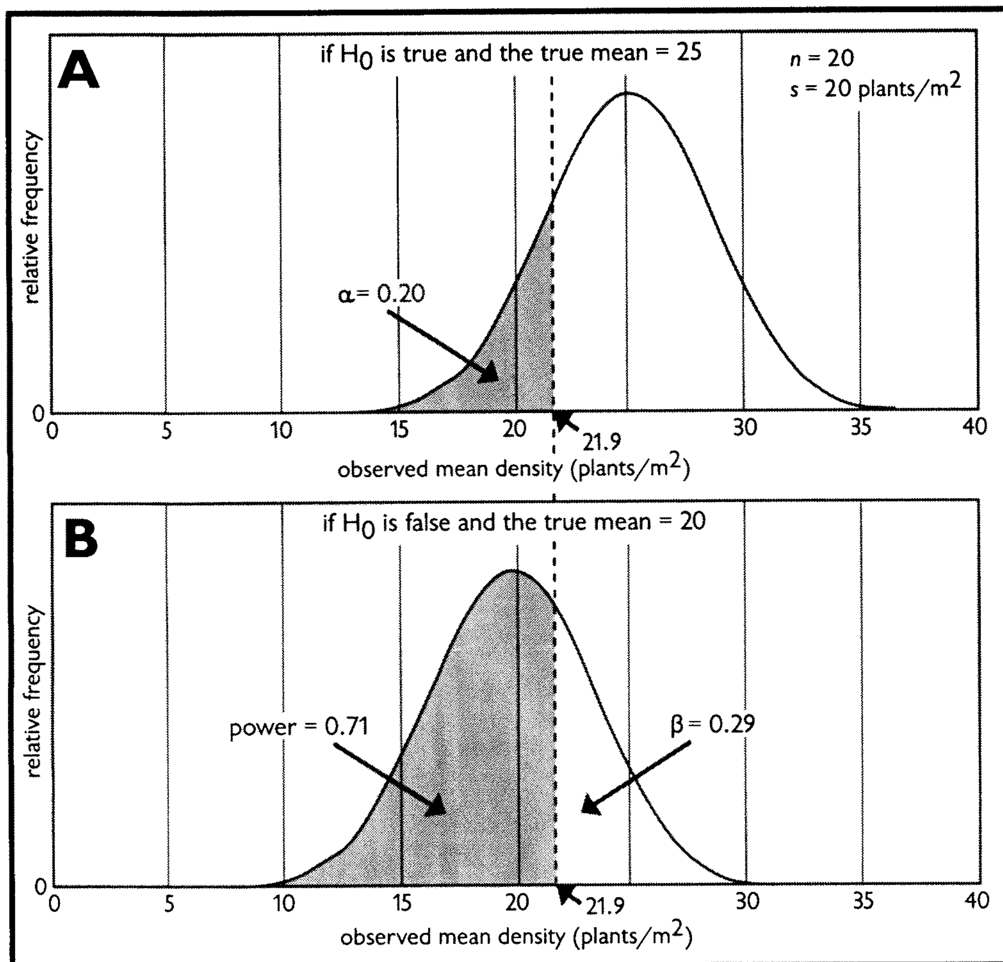


Figure 7.9. The critical region for the false-change error in the sampling distributions from Figure 7.7 has been increased from α = 0.05 to α = 0.20. Part B, in which the H_0 is false and the true mean = 20, shows that power is larger for α = 0.20 than for Figure 7.7 where α = 0.05 or Figure 7.8 where α = 0.10. Again, a sample mean to the left of the vertical dashed line would reject H_0, while one to the right of it would not reject H_0.

Increasing the Sample Size

The sampling distributions shown in Figures 7.7 to 7.10 were all created by sampling the populations with n = thirty 1m × 1m quadrats. Any increase in sample size will lead to a subsequent increase in power to detect some specified minimum detectable difference. This increase in power results from the sampling distributions becoming narrower. Sampling distributions based on samples of n = 50 are shown in Figure 7.11, where the true difference between the two populations is once again 5 plants/m^2 with a false-change error rate threshold of α = 0.05. The increase in sample size led to an increase in power from power = 0.38 with n = 30, to power = 0.54 with n = 50. Note that the critical threshold density associated with an α = 0.05 is now 20.3 plants/m^2 as compared with the threshold of 18.8 plants/m^2 when n = 30.

Decreasing the Standard Deviation

The sampling distributions shown in Figures 7.7 to 7.11 all are based on sampling distributions with a standard deviation of ±20 plants/m^2. The quadrat size used in the sampling was a square 1m × 1m quadrat. If individuals in the plant population are clumped in distribution, then it is likely that a rectangular shaped quadrat will result in a lower standard deviation (see Chapter 8 for a detailed description of the relationship between standard deviation and sampling unit size

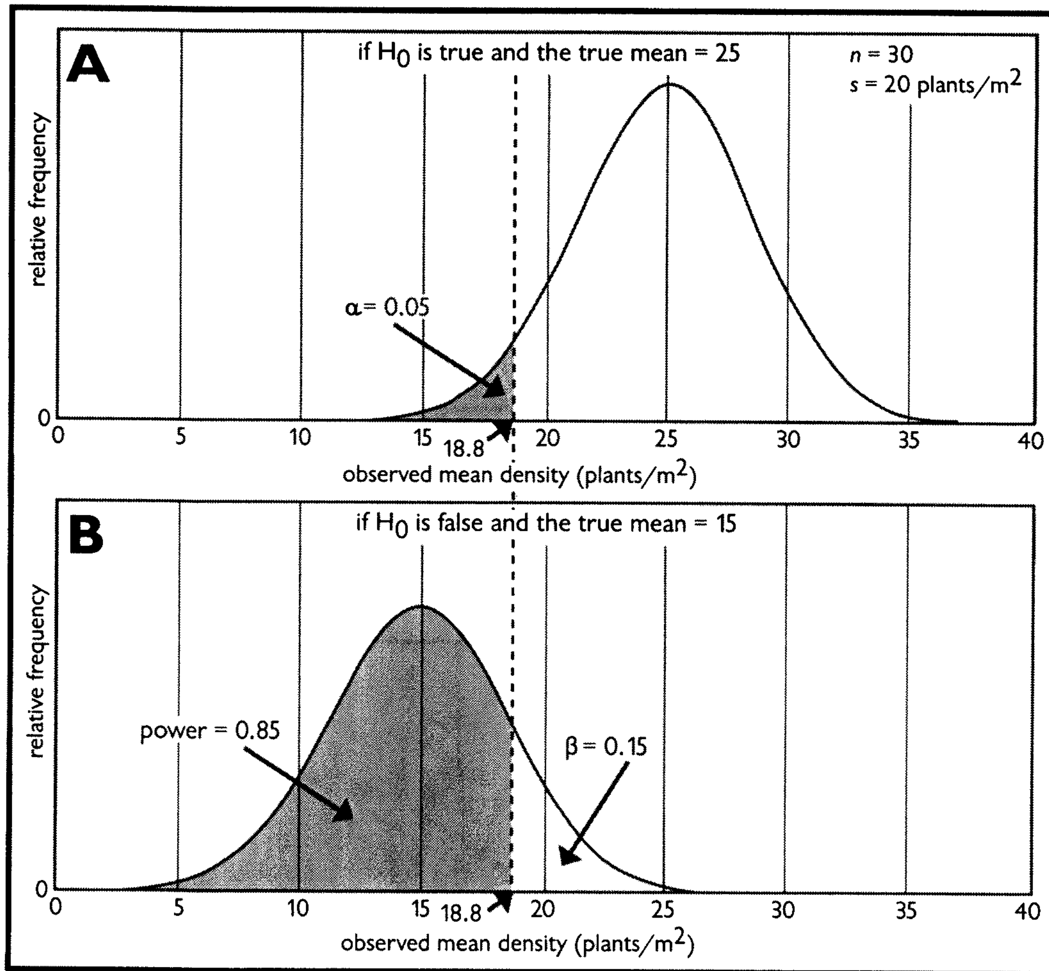


Figure 7.10. Part A is the same as Figure 7.7; in part B, the true population mean is 15 plants/m^2 instead of the 20 plants/m^2 shown in Figure 7.7. Note that power increases (and β decreases) when the new true population mean gets further from the original true mean of 25 plants/m^2. Again, a sample mean to the left of the vertical dashed line would reject H_0, while one to the right of it would not reject H_0.

and shape). Figure 7.12 shows sampling distributions where the true population standard deviation was reduced from ±20 plants/m^2 to ±10 plants/m^2. Note that the critical threshold density associated with an α of 0.05 is now 21.9 plants/m^2 compared with a threshold of 18.8 plants/m^2 when the standard deviation was ±20 plants/m^2. This reduction in the true standard deviation came from a change in quadrat shape from the 1m × 1m square shape to a 0.2m × 5m rectangular shape. Note that quadrat area (1m^2) stayed the same, so that the mean densities are consistent with the previous sampling distributions shown in Figures 7.7 through 7.11. This reduction in standard deviation led to a dramatic improvement in power, from 0.38 (with s = 20 plants/m^2) to 0.85 (with s = 10 plants/m^2). Reducing the standard deviation has a more direct impact on increasing power than increasing sample size, because the sample size is reduced by taking its square root in the standard error equation ($SE = s/\sqrt{n}$). Recall that the standard error provides an estimate of sampling precision from a single sample without having to enlist the support of 1000 people who gather 1000 independent sample means.

POWER CURVES

The relationship between power and the different sampling design components that influence power can also be displayed in power curve graphs. These graphs typically show power values on

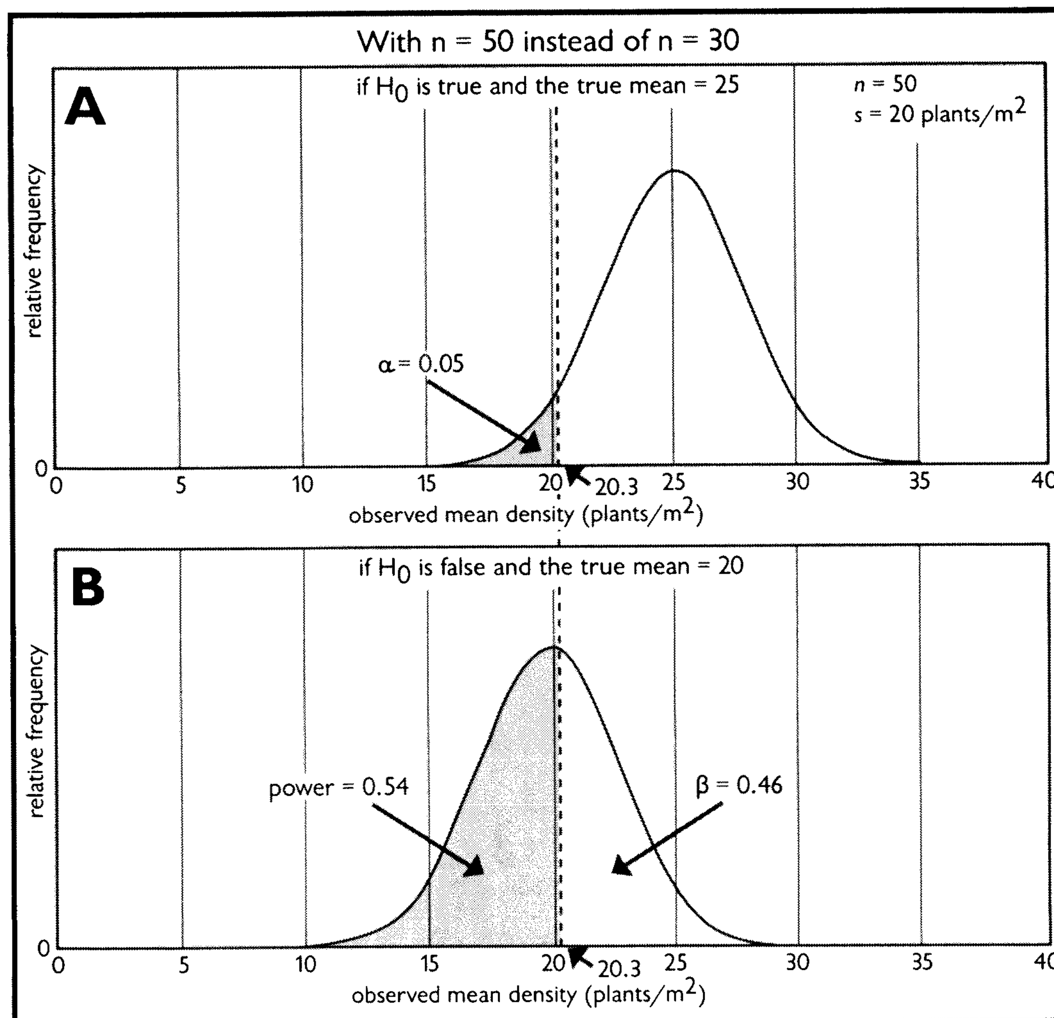


Figure 7.11. The sample size was increased to n = 50 quadrats from the n =30 quadrats shown in Figure 7.7. Note that power increases (and β decreases) at larger sample sizes. Again, a sample mean to the left of the vertical dashed line would reject H_0, while one to the right of it would not reject H_0.

the y-axis and either sample size, MDC, or standard deviation values on the x-axis. Figure 7.13A shows statistical power graphed against different magnitudes of change for the same hypothetical dataset described above and shown in Figures 7.7 to 7.10. Four different power curve lines are shown, one for each of the following four different false-change (α) error rates: 0.01, 0.05, 0.10, and 0.20. The power curves are based on sampling with a sample size of 30 quadrats and a standard deviation of 20 plants/m^2. For any particular false-change error rate, power increases as the magnitude of the minimum detectable change increases. When $\alpha = 0.05$, the power to detect small changes is very low. For example, we have only a 13% chance of detecting a difference of 2 plants/m^2 (i.e., a density of 23 plants/m^2, which is 2 plants/m^2 below our threshold value of 25 plants/m^2). In contrast, we can be 90% sure of detecting a minimum difference of 11 plants/m^2. We can also attain higher power by increasing the false-change error rate. The power to detect a change of 8 plants/m^2 is only 0.41 when $\alpha = 0.01$, but it increases to 0.69 at $\alpha = 0.05$, to 0.81 at $\alpha = 0.10$, and to 0.91 at $\alpha = 0.20$.

A different set of power curves are shown in Figure 7.13B, where the sample size is $n = 50$ instead of the $n = 30$ shown in Figure 7.13A. This larger sample size shifts all of the power curves to the left, making it more likely that smaller changes will be detected. For example, with a false change error rate of $\alpha = 0.10$, the power to detect a difference of 7 plants/m^2 is 0.88 with a sample size of $n = 50$ quadrats compared with the power of 0.73 with a sample size of $n = 30$ quadrats.

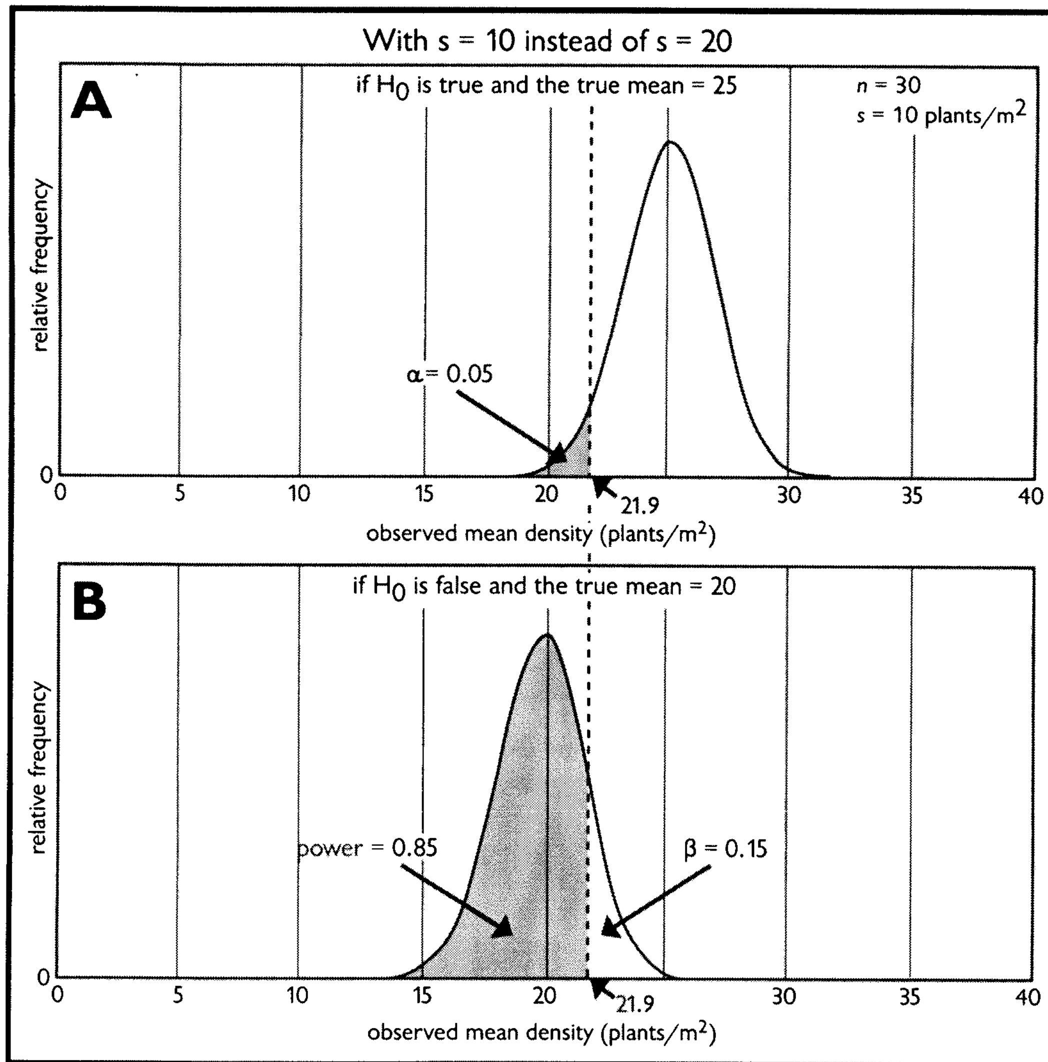


Figure 7.12. The standard deviation (s) of 20 plants/m^2 shown in Figure 7.7 is reduced to ten plants/m^2. Note that power increases (and β decreases), as the standard deviation decreases. Again, a sample mean to the left of the vertical dashed line would reject H_0, while one to the right of it would not reject H_0.

Figure 7.13C illustrates the effect of reducing the standard deviation from 20 plants/m^2 to 10 plants/m^2. The smaller standard deviation shifts all of the power curves to the left and results in much steeper slopes. The smaller standard deviation leads to substantially higher power levels for any particular MDC value. For example, the power to detect a change of 5 plants/m^2 with a false change error rate of $\alpha = 0.10$ is only 0.53 in Figure 7.13A as compared with the power of 0.92 in Figure 7.13C.

USE OF PRIOR POWER ANALYSIS DURING STUDY DESIGN

Power analysis can be useful during both the design of monitoring studies and in the interpretation of monitoring results. The former is sometimes called "prior power analysis," whereas the latter is sometimes called "post-hoc power analysis" (Fairweather 1991). Post-hoc power analysis is covered in Chapter 9.

The use of power analysis during the design and planning of monitoring studies provides valuable information that can help avoid monitoring failures. Once some preliminary or pilot

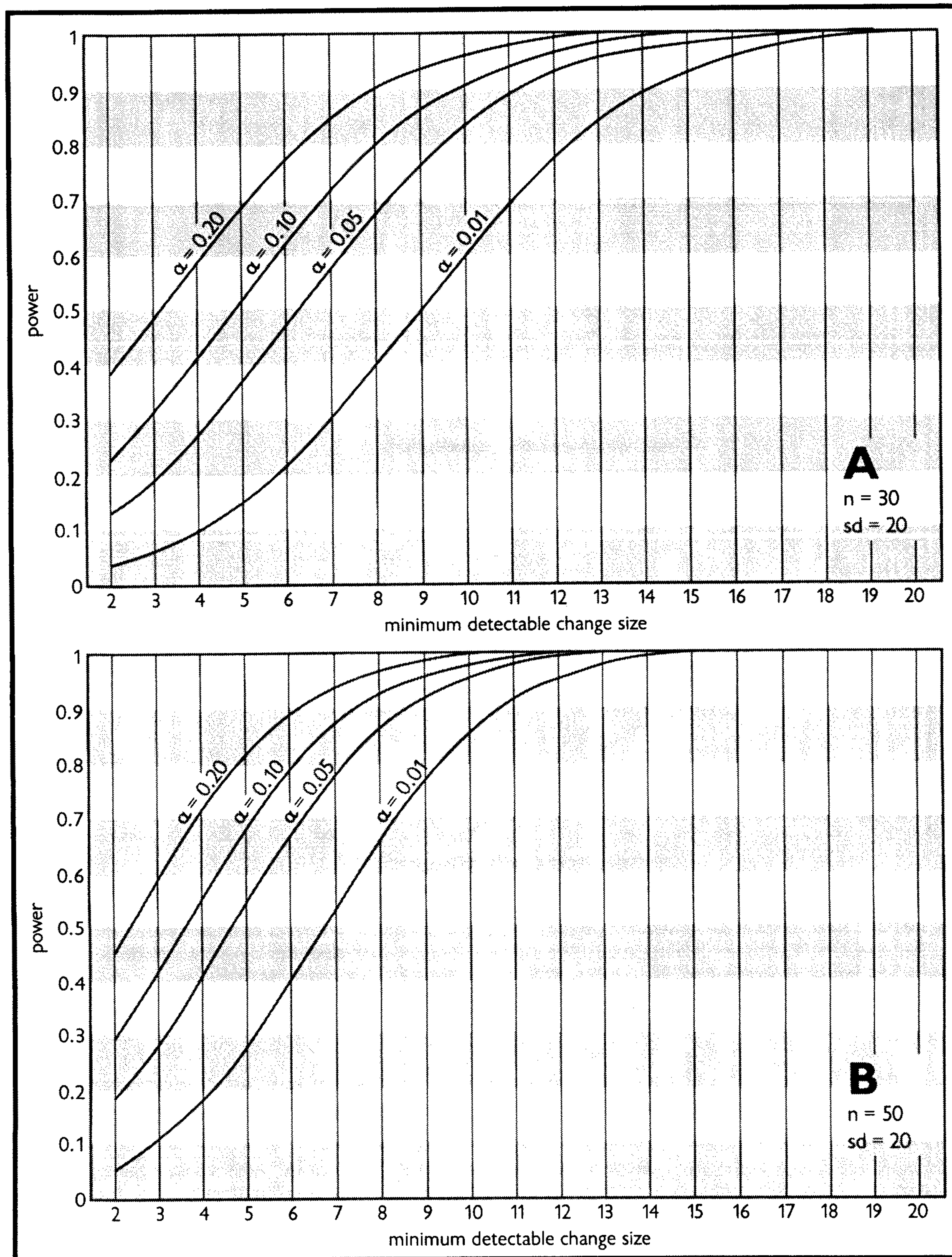


Figure 7.13. Power curves showing power values for various magnitudes of minimum detectable change and false-change error rates when the standard deviation is 20. Part A shows power curves with a sample size of 30. Part B shows power curves with a sample size of 50. Part C shows power curves with a standard deviation of 10 plants/m^2.

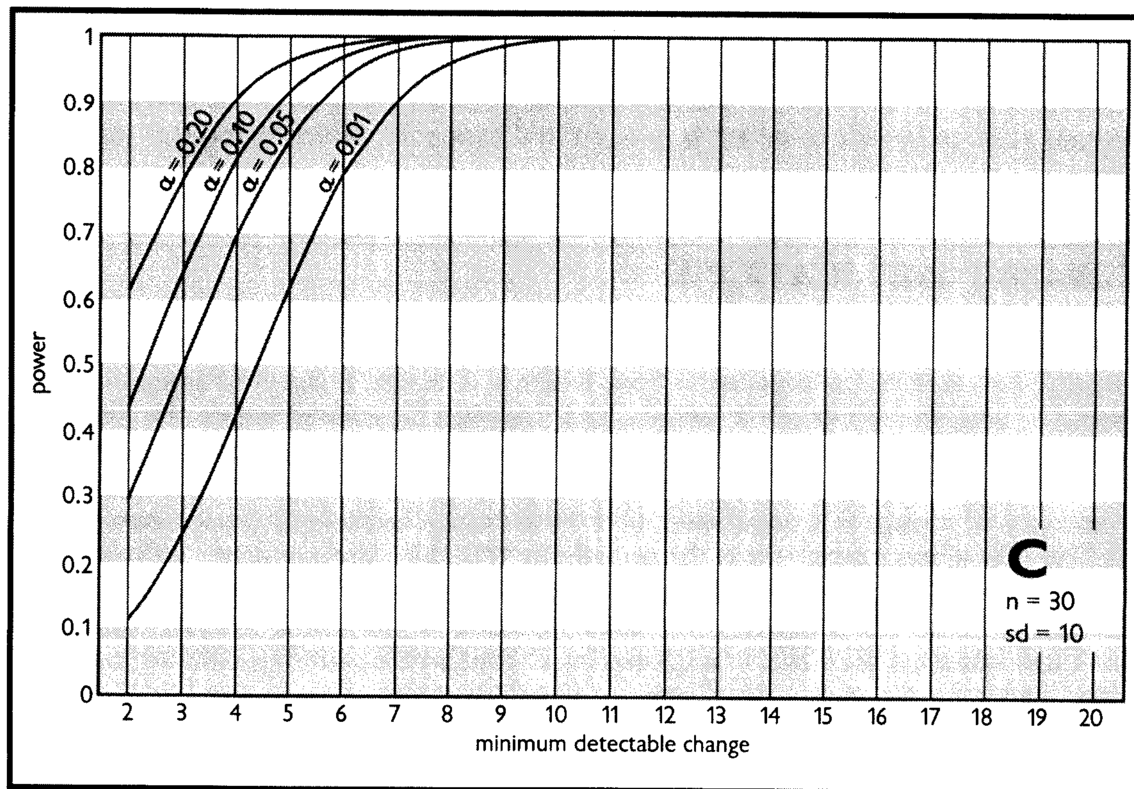


Figure 7.13. *(Continued)*

data have been gathered, or if previous monitoring data are available, power analysis can be used to evaluate the adequacy of the sampling design. Prior power analysis can be done in several different ways. All are based on the power function described earlier:

Power = a function of (α, MDC, n, and s)

The power of a particular sampling design can be evaluated by plugging sample standard deviation, sample size, the desired MDC, and an acceptable false-change error rate into equations or computer programs and then solving for power (Thomas and Krebs 1997).[3] If the power to detect a biologically important change turns out to be quite low (high probability of a missed-change error), then the sampling design can be modified to try to achieve higher power.

Alternatively, a desired power level can be specified and the terms in the power function can be rearranged to solve for sample size. This will give you assurance that your study design will succeed in being able to detect a certain magnitude of change at the specified power and false-change error rate. This is the format for the sample-size equations that are discussed in Chapter 8 and presented in Appendix II.

Still another way to do prior power analysis is to specify a desired power level and a particular sample size and then rearrange the terms in the power function to solve for the MDC (Rotenberry and Wiens 1985; Cohen 1988). If the MDC is unacceptably large, then attempts should be made to improve the sampling design. If these efforts fail, then the decision must be made to either live with the large MDC or to reject the sampling design and perhaps consider an alternative monitoring approach.

The main advantage of prior power analysis is that it allows the adequacy of the sampling design to be evaluated at an early stage in the monitoring process. It is much better to learn that

---

[3] *See our Web page (address in Preface) for links to on-line calculators and programs that calculate power.*

a particular design has a low power at a time when modifications can easily be made than it is to learn of low power after many years of data have already been gathered. The importance of specifying acceptable levels of false-change and missed-change errors along with the magnitude of change that you want to be able to detect is covered in Chapter 14, which introduces sampling objectives.

MANAGEMENT IMPLICATIONS

Sampling involves measuring a part to draw conclusions about the whole. A sample never corresponds perfectly, however, to the population from which it is drawn. Substantial sampling error may be associated with the results of the sample, and this must be assessed before the results of the sample are applied to management of the whole. For estimates of a population characteristic (e.g., total size, average length), confidence intervals are used to assess the precision of the estimate. For estimates of change in a population, both false-change and missed-change errors must be assessed. The false-change error rate is the probability that the sample suggests a change that actually did not occur in the population. The missed-change error rate is the probability that the monitoring study failed to detect a change that actually occurred. Historically, missed-change errors have had less attention than false-change errors, although in monitoring, missing an unacceptable change may be the more critical error. An understanding of these basic principles of sampling is required for design of a useful and efficient monitoring study.

CHAPTER 8
Sampling Design

Danaus plexippus
Monarch butterfly
Artist: D. Andrew Saunders

Design is critical to any sample-based monitoring study. The consequences of poor study design are many: lost time and money, reduced credibility, incorrect (or no) management decisions, and unnecessary resource deterioration, to name just a few. Take your time during this stage to design a study that will meet your management and sampling objectives in the most efficient manner.

Six basic decisions, which are discussed in detail in this chapter, must be made in designing monitoring studies based on sampling:

1. What is the population of interest?
2. What is an appropriate sampling unit?
3. What is an appropriate sampling-unit size and shape?
4. How should sampling units be positioned?
5. Should sampling units be permanent or temporary?
6. How many sampling units should be sampled?

Throughout this handbook we encourage you to initiate your monitoring study with a pilot study. This is essentially a trial run of your monitoring design. A pilot study accomplishes three critical things: 1) it provides estimates of the standard deviation needed to plug into sample size formulas to determine an adequate sample size to meet your sampling objective (Chapter 14); 2) it exposes problems at an early stage; and 3) it demonstrates whether a monitoring design is feasible. Based on the pilot study you perform, you may find that you cannot meet your objectives within the constraints of the time and money available. One solution to this dilemma is to change from sample-based monitoring to monitoring based on a qualitative technique or a complete census. Other solutions include choosing a different attribute to measure or changing your management and sampling objectives to reflect a less precise estimate (in the case of a target/threshold objective) or detection of a larger change (in the case of a change/trend objective).

These decisions must be made based on site-specific information and objectives. There is no "right" sampling-unit size and shape, just as there is no "right" number of sampling units. In most situations, these decisions can be made only through on-site assessment by pilot sampling.

The sampling-design issues discussed in this chapter pertain to monitoring studies in which all of the sampling units or individuals are available for measurement. In animal studies, individuals are often secretive and difficult to count. For these types of animals, most of the sampling-design issues discussed in this chapter are not applicable. Chapter 13 covers these situations.

WHAT IS THE POPULATION OF INTEREST?

As we learned in Chapter 7, the population consists of the complete set of units about which we want to make inferences. We are using "population" in the statistical, rather than the biological, sense. That both biologists and statistitians use the term "population" for different things creates ongoing confusion. To clarify the term, we describe four types of populations: biological populations, target populations, sampled populations, and statistical populations (Box 8.1).

A "biological population" is often difficult to define. A plant species that only occurs within a 100-hectare wetland with no other of this species found for over 100 km would likely be unanimously considered a biological population. A group of animals isolated on a single mountaintop would likely be considered unequivocally a population. Most plant and animal groupings, however, are less obviously isolated from others, creating a problem of identifying boundaries of the biological population. You will need to consider the biological population when assessing population rarity and risk (Chapter 3) and when developing ecological models that include immigration, emigration, and movement within and between biological populations (Chapter 14).

Management activities usually take place within some type of administrative boundary that does not respect the boundaries of the biological population. The portion of the biological population that you manage and are interested in we call the "target population." For example, if we

Box 8.1. *FOUR POPULATIONS*

A rare plant species grows in a 300 hectare wet meadow, isolated by about 40 km from the nearest of the 5 known occurrences of this species. Within the meadow, estimates of the number of individuals of this small perennial species range up to a million or more. A portion of the wet meadow (approximately 100 hectares) is managed by your office. This area was fenced 5 years ago to eliminate livestock grazing. The remainder is privately owned and lightly grazed; the landowner refuses to allow any monitoring on his land. You are limited to spending only 2 days per year monitoring the portion of the population managed by your office. You recognize that you cannot possibly sample the entire 100 hectares in a single day (the other day will be spent on data analysis and report-writing). Travel is difficult across the wet meadow, and you are concerned about disrupting a great heron rookery. You decide to establish a 100m × 100m monitoring site within the 100 hectares. Within this monitoring site, you will annually estimate density. After trials of several sizes and shapes of quadrats, you select a 25m × 0.5m quadrat for sampling, resulting in 800 potential quadrats that can be placed in the 100 × 100m area without overlap.

- *Biological Population: all plants within the 300 hectare wetland. (This is an easy example; most biological population boundaries are much more difficult to draw.)*
- *Target Population: all plants within the 100 hectares managed by your office.*
- *Sampled Population: all plants within the 100 × 100m monitoring site.*
- *Statistical Population: the 800 quadrats that may be potentially sampled.*

are interested in the success of a rare fish population as measured by the average length, our target population may be all the individuals of that species in a spring system of a preserve that has been set aside for that species' protection. Similarly, our target population might be all of the individuals of a rare plant species occurring within a particular wet meadow.

In sampling, the difference between the target population and the population you actually sample (the "sampled population") must be understood. When target populations are small and distributed in some uniform area such as all plants within a fenced pasture, we may be able to position sampling units throughout the entire target population. However, two factors usually lead to defining a new sampled population: 1) irregular target population boundaries, and 2) target populations that cover a very large geographic area.

When the target population is small, but has irregular boundaries, then we might fit some regular-shaped polygon such as a square or rectangle over the bulk of the population (as illustrated in Figure 8.1A). This newly defined area, often referred to as a **macroplot,** becomes our sampled population. The macroplot is usually permanently marked. The use of a macroplot facilitates the positioning of sampling units (see below) and ensures that the same area is sampled each year.

Macroplots are relatively large areas, with sampling units such as quadrats, lines or points randomly located within them.

We can make statistical inferences only to the boundaries of the sampled population (i.e., to the area within the macroplot), not to the entire target population. This approach works well for small target populations; a large population, however, would necessitate a very large macroplot, resulting in long distances between sampling units. The time necessary to travel to each sampling unit would make the design inefficient.

If the target population covers a very large geographic area, constraints of time and money, coupled with the tremendous variability usually encountered when sampling a very large popula-

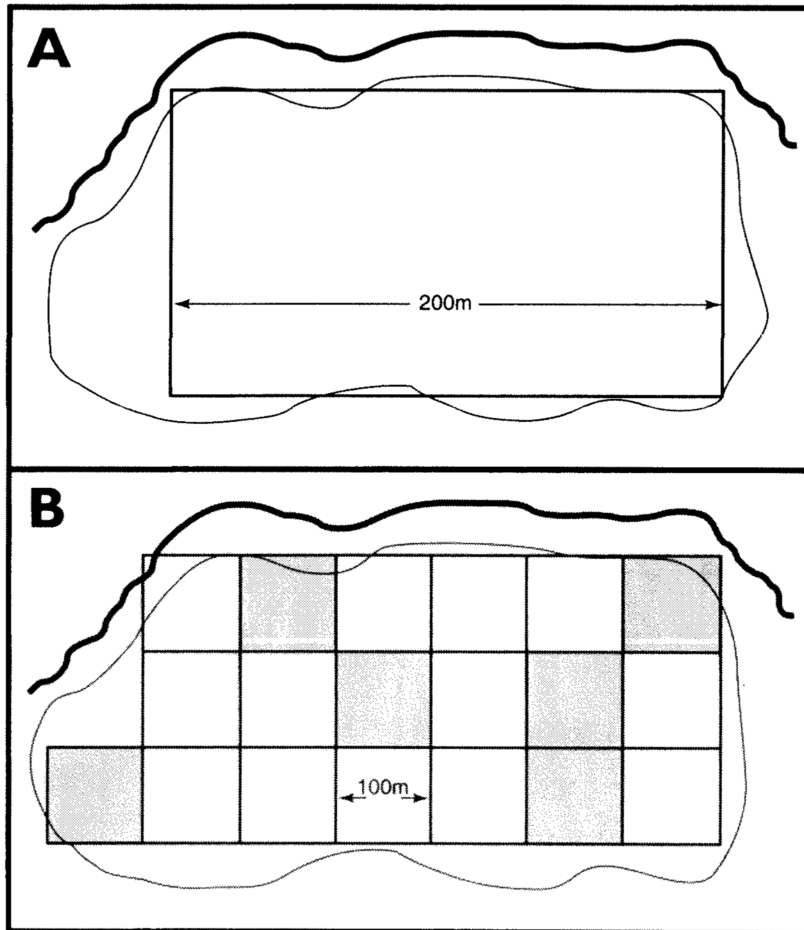


Figure 8.1. Positioning of macroplots (rectangles and squares) within irregularly shaped target populations (thin lines). The thick irregular line denotes a river. Figure 8.1.A. A single 200m × 75m macroplot is placed over the bulk of the target population. Inferences can be made only to the area within the macroplot (i.e., the macroplot is the sampled population). Figure 8.1.B. Target population covers a much larger area (note scale change). Six 100m × 100m macroplots are randomly placed within the target population. Inferences can be made to the entire target population (i.e., the sampled population is the same as the target population). Figure 8.1.C. A single square macroplot is placed in the target population. Inferences can be made only to the area within the macroplot (i.e., the macroplot is the sampled population). Figure 8.1.D. Subjective placement of a macroplot within a "representative" key area (dotted line).

tion, often require further restriction of the sampled population to a smaller geographic area. There are several ways this can be accomplished:

1. A sample of macroplots can be randomly positioned within the target population (Fig. 8.1B). If sampling takes place within each macroplot, then we have something called a two-stage sampling design, described in detail later in this chapter. Statistical inferences can be made to the entire target population, and the sampled population and the target population are the same.
2. A single macroplot can be subjectively positioned within the target population (Fig. 8.1C). The sampled population is the macroplot. No inferences to the target population are possible because there is no way of determining how "representative" this macroplot is of the target population.
3. A few macroplots can be subjectively positioned within the target population. Inferences can be made only to the area encompassed by the macroplots. In other words,

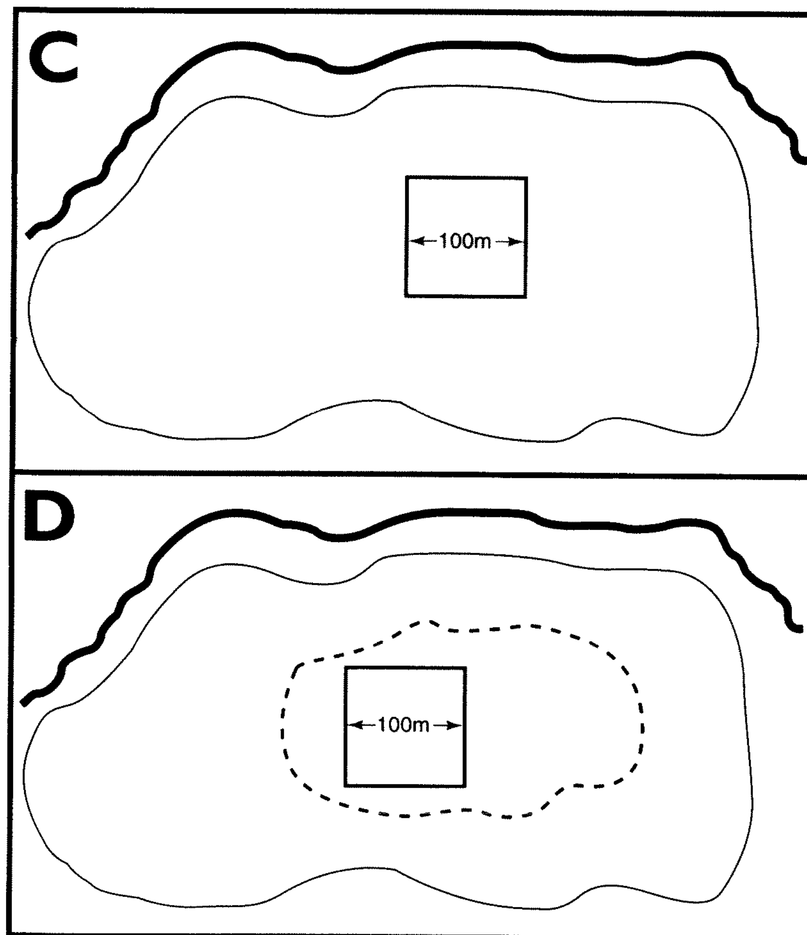


Figure 8.1. *(Continued)*

the sampled population is the area within the macroplots. The value of subjective positioning is that you can place the macroplots in the areas you consider most representative or critical.

When the target population area becomes very large and difficult to sample, we may select one or a few key areas in which we will conduct monitoring (Fig.8.1D). The key area concept is widely used, particularly in rangeland monitoring. Using this approach, key areas are selected (subjectively) that we hope reflect what is happening on a larger area. We may believe that the key area(s) are representative of a larger area (such as a pasture) or are critical or sensitive areas where we are most interested in detecting a problem.

Although we would like to make inferences from our sampling of key areas to the larger areas they are chosen to represent, this cannot be done statistically because the key areas were chosen subjectively. We could, of course, choose to sample the larger areas, but the constraints of time and money coupled with the tremendous variability usually encountered when sampling very large areas often make this impractical. The key area concept represents a compromise.

Careful definition of the sampled and target population remains critical. Remember the monitoring data only represents what is happening in the sampled population. Here are examples:

1. The key area is sampled with randomly placed quadrats. The key area is the sampled population.
2. One macroplot is subjectively positioned within the key area (Fig. 8.1D). You can only make inferences to the area inside the macroplot. Your sampled population is

the macroplot, not the key area. (You could reduce the size of the key area to the macroplot, making them the same.)

3. Several macroplots (a sample of macroplots) are randomly positioned within the key area. This is a two-stage sampling design. Inferences may be drawn about the key area in which the macroplots were randomly placed. The sampled population is the key area.

Because statistical inferences can be made only to the key areas that are actually sampled, it is important to develop objectives that are specific to these key areas. It is equally important to clarify that actions will be taken based on what happens in the key area, even when it cannot be demonstrated statistically that what is happening in the key area is happening in the area it was chosen to represent. It is also important to base objectives and management actions on each key area separately. Values from different key areas should never be averaged, because this gives the impression that key areas are sampling units used to sample a much larger area than is really the case. Key areas are selected with particular intent; they are not randomly selected sampling units. Averaging values from key areas results in a "mean" value for which we can have no measure of precision.

It is important to explicitly recognize the difference between your target population and your sampled population so you know the limitations of your data. You can only draw statistical inferences about your sampled population. You do not know how well the observations in the sampled population compare with the target population, unless you sample the entire target population. In management, it may be acceptable to make decisions for the entire target population based on the results from the sampled population. All stakeholders may have agreed to abide by the results from the sampled population (knowing there exists a risk that results may not represent the target population), or you may decide to collect qualitative or other ancillary data in the target population that supports the results in the sampled population. Consider the following questions:

- How limited are your monitoring resources?
- How difficult will it be to sample the entire target population?
- How comfortable will you (or the decision-maker) be in making management decisions for the entire target population based on the information gathered from a more limited sampled population?
- If the sampled area is located toward the middle of the population, will you miss changes that occur near the edge of the target population?

WHAT IS AN APPROPRIATE SAMPLING UNIT?

The type of sampling unit you select depends on the attribute you are measuring, which should be detailed in a specific management objective (see Chapter 14). Density, cover, frequency, biomass, and size of plant or animal populations are the attributes most commonly monitored. Attributes related to individual measures of performance such as height or number of flowers for plants and length and weight of animals are also often of interest (Box 8.2).

In many cases, simply determining the attribute you are going to measure determines the sampling unit. If you are going to measure density, frequency, or biomass, the sampling unit will be a quadrat. For cover, however, you have several choices. The sampling unit can be a line intercept, a point intercept, or a quadrat (Chapter 12 gives information to help you decide which of these to choose). If you are measuring something about individuals, the sampling unit is the individual (although, as we will see later, you will often incorporate quadrats). Most animals,

Box 8.2. *EXAMPLES OF SAMPLING UNITS*

- *Individual plants. Plants are the sampling units for attributes such as plant height, number of flowers per plant, or cover if the cover measurements are made on individual plants (e.g., tree stem diameters, bunchgrass basal area measurements).*
- *Individual animals. Animals are the sampling units for attributes such as height, length (e.g., snout-vent length in amphibians), condition (e.g., kidney fat index in ungulates), parasite loads, or reproductive rates (e.g., number of yearlings accompanying adult females).*
- *Plant parts. Fruits might be the sampling units if the attribute is the number of seeds per fruit or the percentage of fruits containing some seed herbivore. Or, you may be interested in estimating the number of flowers per inflorescence, in which case the inflorescence is the sampling unit.*
- *Quadrats (plots). Most estimates of plant density, frequency, or biomass require the use of quadrats, which represent the sampling units. Quadrats can also be the sampling units for measurements of vegetation cover if visual estimates of cover are made within quadrats. Most estimates of animal density, frequency, or biomass require the use of quadrats, sometimes called belt transects if greatly elongated, which represent the sampling units.*
- *Lines (transects). When cover is measured using the line-intercept method, the line is the sampling unit. Lines can also serve as sampling units when points (for cover) or quadrats (for cover, density, or frequency) are positioned along lines and the points or quadrats are not far enough apart to be themselves considered the sampling units (because they are not independent of one another). The line-intercept method is occasionally used to estimate animal populations based on the probability of transects intercepting animal tracks (Becker 1991).*
- *Points. When cover is measured with the point-intercept method and the points are randomly positioned, then the points are the sampling units. Points are sometimes used for sampling animals, mainly colonial ones that form large aggregations, such as corals.*
- *Point frames or point quadrats. When plant cover is measured using point frames or point quadrats and these frames or quadrats are randomly positioned, then the point frames or point quadrats are the sampling units. Point frames are not recommended because they are inefficient for measuring cover in most vegetation types (see Chapter 12).*
- *Distance (plotless) methods. There is a class of techniques to estimate density called distance or plotless techniques. The sampling unit with these techniques is usually the individual distance between a randomly selected point and the nearest plant or between a randomly selected plant and its nearest neighbor. Distance measures are inaccurate for most plant populations (see Chapter 12). Distance methods are also used in animal studies, but are different from those used in plant studies in that they attempt to overcome the problem of incomplete detectability of individual animals. These methods are discussed in Chapter 13.*

however, are too secretive or elusive to be directly counted; therefore, specialized sampling techniques (covered in Chapter 13) are needed to estimate most population parameters for animals.

Certain sampling designs incorporate sampling units at more than one level. These are called multistage sampling designs (Krebs 1998). The two-stage sampling design, discussed below, is one example. A random sample of primary sampling units is selected. Then, a subsample is taken from each of the primary sampling units. This subsample is made up of secondary sampling units (these are often called elements to differentiate between the two types of units).

The collection of sampling units from which you draw your sample is the statistical population. For example, a macroplot 20m × 50m will contain 4000 frequency quadrats 50cm × 50cm in size (quadrats do not overlap). The statistical population is the 4000 quadrats. If you were sampling with density quadrats 50cm × 25m in size, the statistical population is the total number of these that could fit into the 20m × 50m macroplot: 80 quadrats. These are finite statistical populations (see Chapter 7), unless so many potential quadrats exist within a large area that the number is essentially infinite. If you were sampling using line intercepts, the statistical population is all the potential line intercepts that could be placed within the 20m × 50m macroplot. Because lines have no width (theoretically, at least) an infinite number could be placed within the macroplot. The statistical population is thus infinite. This concept of infinite or finite populations has important implications for determining sample size and for analysis (see below and also Chapter 9).

WHAT IS AN APPROPRIATE SAMPLING UNIT SIZE AND SHAPE?

Considerations

The most efficient sampling unit size and shape depend on the type of attribute being measured and the morphology and spatial distribution of the species (or the object of your study such as nests, burrows, motorcycle tracks). The most efficient design is usually the one that yields the highest statistical precision (smallest standard error and narrowest confidence interval around the mean) for either a given area sampled or a given total amount of time or money available. Several factors must be considered:

Travel and Setup Time Versus Searching and Measuring Time

As the sampling unit increases in size, the time required to measure the unit increases. For estimating density in quadrats, for example, you must consider whether it is more important to minimize the number of sampling units or the total area (or proportion) of the population sampled. When sampling along transects, you must consider the time required to set up each transect, the travel time between them, and the time needed to measure each transect. Consider the size of the area you are sampling (is it a kilometer between each sampling unit?) and how difficult it is to get from one sampling unit position to another (are you sampling on a cliff face?). Also consider how hard it is to locate and measure the target species within each sampling unit. For large or conspicuous species such as large mammals, trees, or tall herbaceous plants that occur at low densities, having a large sample area or a long transect is not much of a problem because you can see all of the individuals, even from a distance. For small, obscure species that may be hidden under the vegetation canopy or under the leaf litter, you might have to search very carefully; in this case minimizing the total sample area or length may be critical.

Spatial Distribution of Individuals in the Population

Very few biological populations are randomly distributed in the area they occupy. If they were, different configurations of the same sampling unit size or length would perform similarly. Most populations, however, are aggregated or clumped in their distribution. For clumped distributions, sampling units that intersect some clumps of the species will reduce both the number of sampling units with zero counts and the number of sampling units with very high counts. This

decreases the variation among the quadrats and increases the precision of estimates. It is best if the sampling unit length (i.e., the length of the long side of the quadrat or the length of the transect) is longer than the mean distance between clumps.

As an example, consider the species *Primula wilcoxii*, which grows on the shaded side of terraces on a terraced slope in the foothills near Boise, Idaho. The terraces are approximately 1.5 meters apart. In this case, 1m × 1m quadrats to estimate density would be a very poor choice, because many of these would fall between terraces, resulting in many zero values. Some of the 1m × 1m quadrats, however, would fall right on the terraces, and very high counts of this species would be obtained for these quadrats. For this species at this terraced site, quadrats of 0.5m × 2.5m performed well.

Depending on the nature of your population, orientation of sampling units can be very important. For example, you want to orient rectangular quadrats to capture the variability within the quadrats rather than between the quadrats. This results in lower, among-quadrat variance and higher precision. Thus, if there is some gradient such as elevation or moisture to which the population responds differently, you want to make sure your rectangular quadrats follow that gradient to incorporate the variability within the quadrats. In the *Primula wilcoxii* example, the rectangular quadrats were most efficient when placed perpendicular to the terraces. If you were making counts of butterflies along transects,[1] you would orient the transects across changing habitats rather than run parallel to them. Similarly, if you were estimating cover using a line-intercept transect on a site that had a moisture gradient running from the east (wet edge) to the west (dry edge), you would orient your transects from east to west along the gradient.

Edge Effects

Edge effects are an important consideration for quadrat sampling units. The edge of a quadrat is its outer boundary. The more edge a quadrat has, the greater the difficulty in determining whether individuals near the edge are in or out of the quadrat. Rectangular quadrats have more edge per unit area than squares or circles. Although this is an important issue, Chapter 12 discusses ways to minimize the nonsampling error associated with edge bias when sampling plants (stationary animals would follow the same conventions). For most animals, determining whether an individual is in or out of the quadrat may be more difficult because they are moving and because you usually cannot necessarily measure the distance to them. This is, for example, an important issue when counting birds, some of which are often entering and leaving a study plot while a count is being made. Chapter 13 discusses these issues. You must be consistent in applying whichever boundary convention you choose and to make sure, through training and documentation, that others involved in the monitoring during the first and all subsequent years use the same convention.

Abundance of Target Population

If the density is relatively high, you will want to use smaller quadrats because you do not want to be counting hundreds to thousands of occurences in each quadrat. Conversely, if density is relatively low, you will want to use larger quadrats to avoid sampling many quadrats with no individuals in them.

Ease in Sampling

The considerations here are the difficulties in searching the entire sampling unit and keeping track of what portions have already been searched. With large quadrats for measuring density, for example, long, narrow rectangles are easier to search because you can start at one end and keep track of counts at intervals along the quadrat. With large, square quadrats, you will probably have to subdivide the quadrat area to ensure that you do not double-count.

---

[1] *These types of transects in which counts are made are actually very long, narrow quadrats, compared to true transects which have no width (theoretically dimensionless).*

Disturbance Effects

If the sampling unit size/shape is so large that you have to stand in the sampling unit to search through it, you risk impacting the population through your sampling. This is particularly important when sampling permanent sampling units, because the changes you observe over time may simply be the result of your impacts to the sampling units and not reflect the true situation in the sampled population as a whole. It is also a problem when using temporary sampling units, however, especially if you impact areas of the sampling unit before you have searched them.

Computer-Simulated Comparisons of Sampling Designs

The importance of selecting an efficient sampling unit configuration is often ignored when developing a monitoring study. Sampling units of different configurations perform differently, and the efficiencies to be realized from using an appropriate configuration can be substantial. We will use a particular type of sampling unit, quadrats for estimating density, to explore this issue further using computer simulation.

As stated earlier, rectangular quadrats perform better than square or circular quadrats when sampling clumped populations, but two unanswered questions remain: 1) What are the actual trade-offs of changing quadrat size and shape on the number of quadrats to sample or on the total area sampled? 2) As you make quadrats larger, you will presumably have to sample fewer of them—but how many fewer? You can investigate these questions in the field, but you are somewhat limited in the number of different sizes and shapes you can try, and there are potential negative impacts from repeated sampling across the entire area.

Salzer (unpublished data) evaluated these sampling design decisions using computer-simulated sampling. Two populations, each with 4000 plants, were created on the computer. Plants in both populations exhibited a clumped distribution pattern, although they differed in the degree of clumping. One of the populations had plants that were distributed along a gradient. Random samples of these virtual populations were drawn by computer, using density quadrats of different sizes and shapes.

Consider the population of 4000 plants depicted in Figure 8.2. This population was termed the "clumped-gradient population" because the plants were both clumped and distributed along a gradient (note that this gradient follows the x-axis: there are more clumps near the left side of the macroplot than there are near the right side). This population was subjected to 30 different sampling designs that differed in the width and length of the quadrats. The following quadrat widths were used: 0.25m, 0.5m, 1.0m, 2.0m, and 4.0m. The following quadrat lengths were used: 1m, 2m, 5m, 10m, 25m, and 50m. Every combination of quadrat width and quadrat length was used to sample the population (i.e., 0.25m × 1m, 0.25m × 2m . . . 4m × 25m, 4m × 50m). Sampling was conducted so that the long side of the quadrat was oriented so the quadrat included as much of the changing gradient as possible (i.e., the long side was oriented parallel to the x-axis of the population as depicted in Figure 8.2).

For each of the 30 sampling designs, the entire population was sampled (i.e., all the quadrats that fit in the population, without overlapping), so that true, parametric values for the mean density and standard deviation could be calculated for every design. This is desirable for comparing various sampling designs, but is nearly impossible to achieve in a field setting. The true parametric values were entered into a sample size formula to determine how many quadrats would need to be sampled to attain the desired precision. The precision level selected was an estimated mean density with a 95% confidence interval that was no wider than ±30% of the mean value. This brought performance of each sampling design into a common currency—the number of quadrats to sample—so that they could be compared with one another. By knowing the size and number of quadrats being used, the proportion of the population sampled was also calculated.

Figure 8.3 depicts the interaction between quadrat width, quadrat length, number of quadrats, and proportion of the population sampled. A typical quadrat configuration used in

CLUMPED-GRADIENT POPULATION

meters

100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0

0 5 10 15 20 25 30 35 40 45 50

Figure 8.2. The "clumped-gradient population." A population of 4,000 plants aggregated into clumps and responding to a gradient that runs from left to right (along the x-axis). Note the much greater number of clumps near the left side of the population.

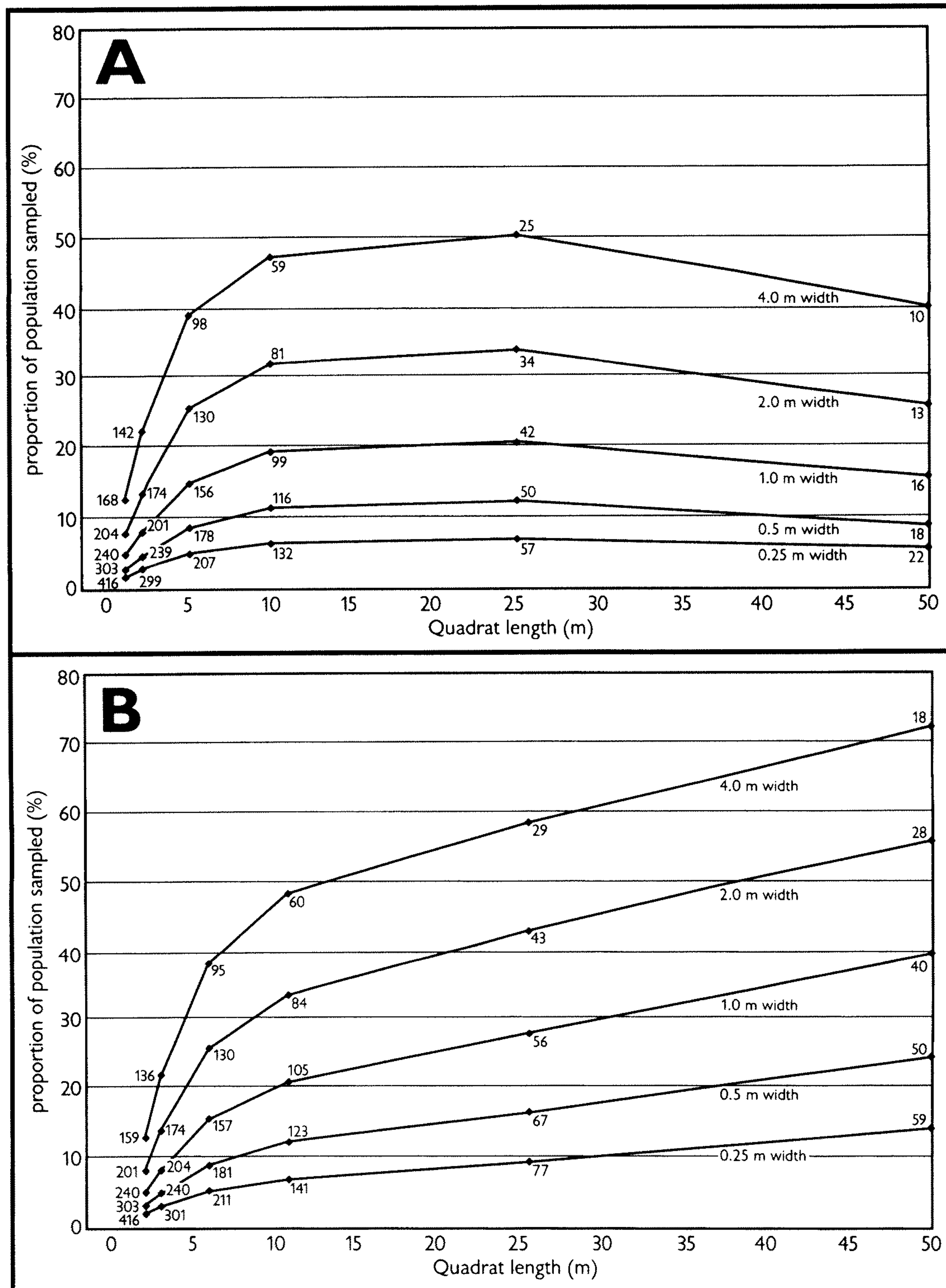


Figure 8.3. Comparison of 30 sampling designs to sample density. Designs used quadrats of different widths (0.25m, 0.5m, 1.0m, 2.0m and 4.0m) and lengths (1m, 2m, 5m, 10m, 25m, 50m) for a total of 30 different quadrat configurations. All designs achieved the same level of precision. Numbers next to data points are the number of quadrats that must be sampled to meet the desired level of precision in the estimate of mean density using that particular quadrat size and shape. Figure 8.3A shows the results when quadrats are oriented along the gradient shown in the population of Figure 8.2 (i.e., the long edge of the quadrat along the x-axis of the population, including as much of the gradient variability as possible). Figure 8.3B shows the results when quadrats are oriented perpendicular to the gradient (i.e., the long edge of the quadrat along the y-axis of the population) shown in Figure 8.2. Figure 8.3C shows the results from sampling a similar population of 4000 plants that lacks a gradient but has much denser clumping (i.e., more unoccupied space between clumps). This densely clumped population is shown in Figure 8.5.

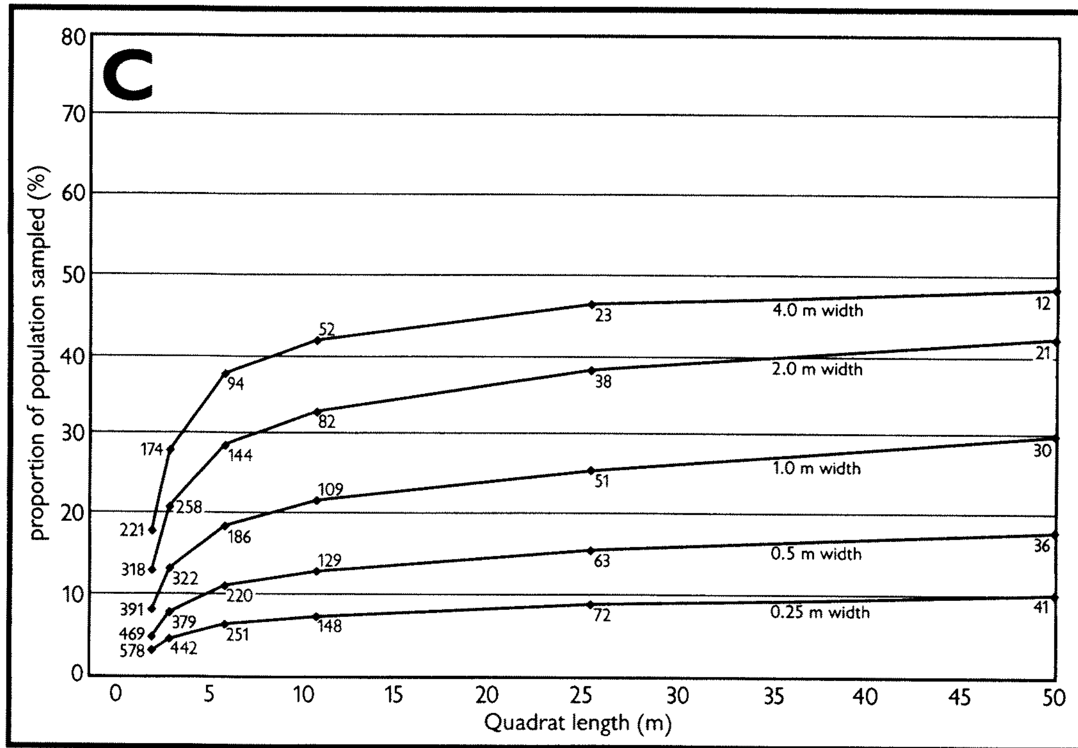


Figure 8.3. *(Continued)*

monitoring is the 1m × 1m quadrat. Note that 240 of these quadrats would need to be sampled to meet the same precision of the estimate as sampling only sixteen 1m × 50m quadrats. The design requiring the fewest quadrats (ten 4.0m × 50m) requires sampling about 40% of the area. Some of these designs offer smaller sample sizes and smaller proportions of the population. For example, compare these two designs: sample twenty-five 4m × 25m quadrats (50% of the entire population), in which case you must count about 2000 plants, or sample twenty-two 0.25m × 50m quadrats (5.5% of the population), in which case you must count only about 220 plants. The design that minimizes both the area to be searched and the number of quadrats to locate is the longest, thin quadrat of 0.25m × 50m. While the 1m × 1m quadrat is typically used in sampling vegetation, it is almost never the most efficient size. Similar considerations exist for square or circular sampling units used for sampling animal populations.

Which of the 30 designs is best? It depends on the difficulty of counting the plant within each quadrat, the time required for placement of quadrats, and the importance of edge effects. In selecting a quadrat size and shape consider:

- How conspicuous is your target plant or animal? Can it be spotted at eye-level or does it take careful searching of every square centimeter of sample area? If large and easily visible from eye level, you might choose a wider quadrat size, leading to a smaller sample size. The larger proportion of the population sampled might not carry much of a penalty (cost) if the portions of the quadrats between clumps can be searched rapidly.
- How quickly can you locate sampling units? If travel between sampling units is difficult because of topography or dense vegetation, sampling fewer larger quadrats would probably save time.
- How big a problem is edge effect? Are plants single-stemmed with small diameter stems clearly arising from a rooted point so that boundary decisions are relatively

rare and quickly made when they do occur or are the target plants bunch grasses with a wide basal area and amorphous shapes, requiring many difficult and time-consuming boundary decisions? Are the animals relatively slow-moving, or is determining whether they are in or out of the sampling unit difficult because they are moving quickly?

If plants are small and inconspicuous with distinct, single-rooted stems, look for a design that has both a small sample size and samples a small proportion of the population. The twenty-two 0.25m × 50m quadrats would be a good choice in this case. Realize, however, that even if minimizing the sample area is critical, you will not want to sample 416 of the 0.25m × 1m quadrats (2.1% of the population area).

Results for the same clumped-gradient population with quadrat orientation reversed (i.e., with the long side parallel to the y-axis) are shown in Figure 8.3B. Rather than looking at the individual sample sizes, concentrate on just the relative proportion of the population that must be sampled. With this quadrat orientation, quadrats located near the left of the macroplot will have high numbers of plants, while quadrats located near the right of the macroplot will have low numbers. This pattern of high and low quadrat counts is undesirable, producing a high standard deviation and wide confidence intervals. With the 4m × 50m quadrat, you need to sample over 70% of the population. You would be better off counting all of the plants in the macroplot (conducting a complete census) than using this quadrat size. Clearly it is better to use a narrower quadrat that is oriented in the opposite direction.

Results from a population of 4000 plants that are more tightly clumped with the clumped centers randomly distributed (without a gradient) are shown in Figure 8.3C. (You can see this population in Figure 8.5). Because of the tighter clumping of plants in the dense-clumped population, sample sizes are even greater for small and square or short and wide quadrats than they were for the clumped-gradient population. This is because quadrats with plants tend to have higher counts and there are more quadrats with zero plants, a situation that increases the standard deviation. It would take, for example, 578 of the 1m × 0.25m quadrats to achieve the desired level of precision in the dense-clumped population as compared with 416 in the clumped-gradient population. With increasing clumping, the advantages of long, narrow quadrats also increase. Conversely, if plants are randomly distributed, quadrat shape has no influence on the number of quadrats to sample. This, however, is seldom the case in nature.

Even though the narrower quadrat sizes perform better statistically, there are practical limitations that must be considered. For example, when sampling the virtual dense-clumped population by computer using different shapes of quadrats with an area of $1m^2$, a 2cm-wide × 50m-long quadrat performed better (n = 98 quadrats) than a 1m × 1m quadrat (n = 394), but the 2cm width would be a ridiculous shape to try to use in the field, because of the tremendous amount of "noise" introduced by edge effect.

In many monitoring situations, especially for herbaceous plants or small, slow-moving animals, a 0.25m or 0.5m quadrat width works well for estimating density. (This width would probably be inappropriate, however, for large or sparsely distributed plants, or for large or fast-moving animals.) Either is a convenient width to search in. Widths larger than 1m or 2m are difficult to search because it is hard to see individuals at the far edge (unless all the individuals are fairly large and there is minimal associated vegetation to obscure your line of sight). The quadrat length should be determined by the size of the area that you are working in and the spatial distribution of the species you are counting. You want to avoid getting many sampling units with zeros, so you want your quadrats to be long enough to incorporate several clumps. You also do not want your quadrats so long that you have to count thousands of individuals—the time involved and the potential measurement error associated with counting that many individuals would be too great. Box 8.3 gives a procedure for comparing the efficiency of different density quadrat sizes and shapes through pilot sampling.

Box 8.3. *A PROCEDURE TO COMPARE THE EFFICIENCY OF DIFFERENT QUADRAT SIZES AND SHAPES USING PILOT SAMPLING*

Select several good candidates of quadrat dimensions that are multiples of the two dimensions of the area you want to sample. For example, in a 50m × 100m macroplot where you want to orient quadrats with the long side along the 50m side of the macroplot, you might select 5m, 10m, 25m and 50m (all factors of 50m) and widths of 0.25cm and 0.50cm. Randomly locate some initial number (e.g., 10) of 0.5m × 50m quadrats in the population of interest. Position the quadrats according to the design you plan to use (this will allow you to use the data from this initial test as part of your actual sample). Attach one end of a 50m tape to a pin or stake, pull it tight and treat one edge of the tape as the center of your quadrat. Count all plants that are within 0.25cm of either side of the tape edge (total width = 0.5m) and record separately, by side, on a field data sheet (Figure 8-B). You should also subdivide the long dimension of the quadrat and record plant counts separately within each segment (e.g., every meter) along your tape. This enables you to look at the performance of quadrats of different lengths.

You can save space by recording the segment number only if you have actual plant counts for that segment. For example, you have laid out your tape and started searching along both sides of the tape. You find your first plants (three of them on the left side and two of them on the right side in the third segment of the tape (between 2m and 3m along the tape). The next plants (two of them on the left side, none on the right) are found in the seventh segment (between 6m and 7m along the tape). The entries on the field data sheet would look like Figure 8-B.

Continue this counting and recording procedure until all your preliminary quadrats have been sampled. Now you can use a hand calculator to calculate means and standard deviations for different size and shape quadrats. To compare quadrats of different sizes you should calculate the coefficient of variation (CV) for each quadrat size. The CV is calculated as follows:

$$CV = s/\bar{x}$$

Where: $\bar{x}$ = The sample mean
s = The sample standard deviation

Unlike the standard deviation, which has a magnitude dependent on the magnitude of the data, the coefficient of variation is a relative measure of variability. Thus, coefficients of variation from different sampling designs can be compared. The smaller the coefficient of variation the better. If two designs have similar coefficients of variation, choose the design that will be easiest to implement.

If, after evaluating the performance of different quadrat sizes, you select a size and shape that was some subcomponent of the larger quadrat sampled, you can still use the data as part of your first year's set of data. To do this you should randomly select the subcomponent from each of your pilot quadrats. Using the previous example, if you elected to use a 0.25m × 50m quadrat, you could randomly select one half of each of the 0.5m × 50m quadrats that you sampled as part of your pilot effort.

		PLANT COUNTS		
PLOT #	SEGMENT #	Left	Right	Total
1	3	3	2	5
1	7	2	0	2

Figure 8-B. Examples of entries on a field data sheet when plants are found in the third and seventh segments of plot number 1.

Other Sampling Units

Your prime design objective when selecting a sampling-unit size and shape is to try to reduce the variability between sampling units while maintaining a size and shape that is practical in the field. Many of the design principles described for density quadrats are applicable to other types of sampling units. Transects should be long enough to intersect clumps of the target species and should be oriented to include as much of the gradient variation as possible. Plots for visually estimating cover or measuring biomass are typically quite small and often square or rectangular, because it is difficult to estimate cover, to clip vegetation, or to estimate biomass in large or long plots. These small quadrats can be arranged, however, along a transect, with the transect, not the quadrats, treated as the sampling unit. This design is really a two-stage sampling design, with the transects serving as the primary sampling units and the quadrats serving as secondary sampling units. We treat this in more detail below.

Chapter 12 describes sampling-unit design considerations for most of the typical methods of measuring plants: density, cover measured by point intercept, line intercept and quadrat estimation, biomass measurements, and frequency measurement. Chapter 13 describes special considerations in sampling-unit design for animal studies.

Determining Sampling-Unit Size and Shape in Real Populations

The best way to determine the appropriate sampling-unit size and shape is to approach every new sampling situation without a preconceived idea of the configuration you will use. Sampling-unit size and shape should be determined during pilot sampling. If possible, wander around the population area and study the spatial distribution of the species you will be sampling (for plants, use pin flags or flagging to improve the visibility of clumps). Attempt to answer the following questions: 1) At what scale(s) can you detect clumping? 2) How large are the clumps, and what are the distances between clumps? 3) How long will sampling units need to be to avoid having many sampling units containing none of the species in them? 4) How narrow will density quadrats need to be to avoid counting hundreds or thousands of the species whenever the quadrat intersects a dense clump? 5) How wide an area can be efficiently searched from one edge of a quadrat? 6) How big a problem will edge effect be?

HOW SHOULD SAMPLING UNITS BE POSITIONED IN THE POPULATION?

There are three requirements that must be met by a monitoring study with respect to positioning sampling units in the population to be sampled: 1) some type of random, unbiased sampling method must be employed; 2) the sampling units must be positioned to achieve good interspersion of sampling units throughout the population; and 3) the sampling units must be independent of each other. Before discussing different methods of random sampling, we will discuss these three characteristics in more detail.

1. Random (unbiased) sampling. Critical to a valid monitoring study design is that the sample has been drawn randomly from the population of interest. Several methods of random sampling can be used, many of which are discussed below. The important point is that all the statistical-analysis techniques available to us are based on knowing the probability of selecting a particular sampling unit. If some type of random selection of sampling units is not incorporated into your study design, you cannot determine the probability of selection, and you cannot make statistical inferences about your population. Preferential sampling, the practice of subjectively selecting sampling units, should be avoided at all costs.

2. Interspersion. One of the most important considerations in sampling is good interspersion of sampling units throughout the area to be sampled (the target population). Although Hurlbert (1984) uses the term "interspersion" to apply to the distribution of experimental units in manipulative experiments, the term can also be applied to sampling units in observational studies. The basic goal is to have sampling units well interspersed throughout the area of the target population. For this reason, the practice of placing all the sampling units, whether quadrats or points, along a single or even a few transects must be avoided. This is true even if the single transect or few transects are randomly located.
3. Independence. Independence means the sampling units are spaced far enough apart so that measurements are not spatially correlated. For example, if quadrats are not spatially correlated, high mortality in Quadrat A does not necessarily mean there will be high mortality in Quadrat B, at least not because of its proximity to Quadrat A. If your design has quadrats located closely along a transect, each quadrat is in close proximity to two others, and changes in each quadrat will probably be correlated with two others (or more). In simple random sampling, there will always be some quadrats located close together simply by chance. The difference is that this correlation only affects some of the quadrats, and the degree of correlation fluctuates randomly with the spatial location of the randomly placed quadrats.

We discuss eight types of random sampling: simple random sampling, stratified random sampling, systematic sampling, restricted random sampling, cluster sampling, two-stage sampling, double sampling, and taking a random sample of individuals. These are summarized in Table 8.1 and are described in more detail below.

Table 8.1. Summary of Random Sampling Types

SAMPLING TYPE	RECOMMENDED USES	ADVANTAGES	DISADVANTAGES
Simple random sampling	Useful in relatively small geographic areas with homogeneous habitat, when the number of sampling units is not likely to be large.	The formulas necessary to analyze data are the simplest of all sampling types.	By chance, some areas within the target population may be left unsampled. The travel time is considerable when the sampling area and/or sample size is large. Restricted random sampling and systematic random sampling outperform simple random sampling when populations have a clumped distribution.
Stratified random sampling	Useful when the attribute of interest responds very differently to some clearly defined habitat features. Since it involves taking a simple random sample within each stratum, each stratum should consist of a relatively small geographic area with homogenous habitat, and the number of sampling units in each stratum should not be too large.	Results in more efficient population estimates than simple random sampling when the attribute measured varies with clearly defined habitat features.	The mathematic formulas required for analysis are more complex than those used for simple random sampling. When the geographic area within any stratum is large and/or the number of sampling units is likely to be large, then one of the other types of sampling listed below will be more efficient. By chance, some areas within each stratum may be left unsampled.
Systematic sampling	Useful for any sampling situation, as long as the first sampling unit is selected randomly and the sampling units are far enough apart to be considered independent. Can also be used as part of cluster and two-stage sampling designs.	When the conditions given in the cell to the left are met, this is the best type of sampling design to use. There is better interpersion of sampling units than with simple random sampling. The data can be gathered much more efficiently than with simple random sampling and still be analyzed using the formulas for simple random sampling.	In the uncommon event that the number of possible samples is limited to fewer than about 25–30 (see text), systematic sampling may lead to questionable results; in this situation you should use restricted random sampling.

(Continued)

Table 8.1. Summary of Random Sampling Types (*Continued*)

SAMPLING TYPE	RECOMMENDED USES	ADVANTAGES	DISADVANTAGES
Restricted random sampling	Although more useful than simple random sampling in most situations, restricted random sampling should be used only when the number of potential samples is fewer than 25–30. Otherwise, systematic sampling is the better choice.	Like systematic sampling, restricted random sampling results in better interspersion of sampling units than with simple random sampling. If the number of potential samples is less than 25–30, restricted random sampling is better than systematic sampling. The data can be analyzed using the formulas for simple random sampling.	The design is not as efficient as systematic sampling when the number of potential samples is greater than 25–30.
Cluster sampling	Cluster sampling is used to select a sample when it is difficult or impossible to take a random sample of the individual elements of interest. A cluster of elements is identified, and a random sample (usually using systematic sampling) is taken of the clusters. Every element within each cluster is then measured. In monitoring, cluster sampling is most often used to estimate something about individuals (e.g., mean height, number of flowers/plant). In this situation, quadrats are the clusters.	It is often less costly to sample a collection of elements in a cluster than to sample an equal number of elements selected at random from the population. Except in rare situations, it is not practical to take a random sample of individuals. Instead, the attribute of interest is measured on every individual in a sample of quadrats (which function as the clusters).	All the elements within each cluster must be measured. If the clusters contain large numbers of the element of interest, two-stage sampling is more efficient. Other disadvantages include the difficulty in determining how many clusters should be sampled versus how large each cluster should be, the more complex calculations required for analysis, and the fact that most statistical software packages do not include these calculations.
Two-stage sampling	Similar to cluster sampling in identifying groups of elements (such as plants) and taking a random sample (usually using systematic sampling) of these groups. In two-stage sampling, however, a second sample of elements is taken within each group. Like cluster sampling, the main use of two-stage sampling is to estimate some value associated with individuals.	Same advantages as cluster sampling. The two types are the only efficient means of estimating some attribute associated with individuals. When the number of individuals in each group (quadrat) is large, two-stage sampling is more efficient than cluster sampling.	There are standard deviations associated with both stages of sampling (unlike cluster sampling, which has no standard deviation associated with the values measured at the second stage). This results in more complicated formulas in arriving at estimates of values and standard errors (although the standard deviation of the secondary sample can be ignored as long as the finite-population correction factor is not applied to the standard error of the primary sample).
Double sampling	Useful when the variable of interest (e.g., actual measurements of biomass) is difficult to measure, but is correlated with an auxiliary variable (e.g., ocular estimates of biomass) that is more easily measurable. The second variable is measured in a large number of sampling units, while the first variable is measured in only a subset of the sampling units. The samples are often taken using systematic sampling.	If the auxiliary variable is relatively quick to be measured and is highly correlated with the variable of interest, double sampling is much more efficient in estimating a variable that is difficult to measure than directly measuring the variable.	The formulas for data analysis and sample-size determination are much more complicated than for simple random sampling, and most statistical software programs do not include the necessary calculations.
Taking a random sample of individuals	This can only be accomplished in rare situations. When the objective is to measure something on individual plants, it is best to use either cluster or two-stage sampling. See text for further information.	In those few situations where it is possible to take a random sample of individuals, the calculations necessary for analysis are simpler than those for either cluster or two-stage sampling.	It is not practical to take a simple random sample of individuals in most monitoring situations.

Simple Random Sampling

A simple random sample is one that meets the following two criteria: 1) each combination of a specified number of sampling units has the same probability of being selected; and 2) the selection of any one sampling unit is in no way linked to the selection of any other (McCall 1982).

One method for selecting random sampling units in a simple-random-sampling design is the **simple random-coordinate method**. While this is probably the most commonly used method, it has serious problems for many sampling units. As shown in Figure 8.4, random coordinates are selected for each of two axes. The point at which these intersect specifies the location of a sampling unit. Coordinates that fall out of the target population boundaries are rejected. This method will work for small sampling units such as plots used to measure frequency,[2] but it will not perform well when the sampling units are lines or long rectangles or when sampling units are points of the center of a large circular sampling unit (e.g., bird counts often sample a circular area with a radius of up to 100m extending from a given sampling point, which often results in a large part of the sampled area falling outside a study site). Two problems with the coordinate method are difficult to overcome:

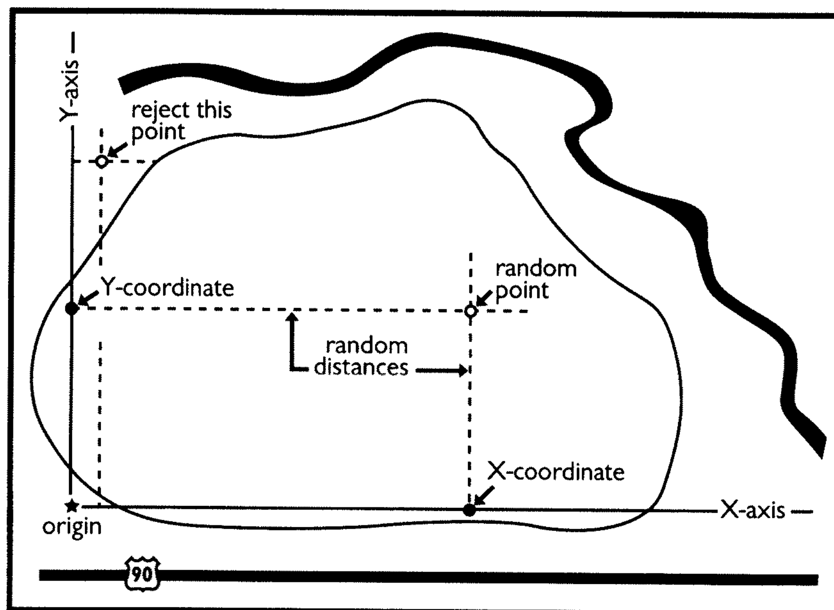


Figure 8.4. Locating points using the simple random coordinate method (adapted from Chambers and Brown 1983). Although this method will work to position points or square quadrats, the grid-cell method is much better for locating long, narrow quadrats or lines.

1. No unbiased method exists to deal with randomly located points that send a portion of the sampling unit out of the target population (a common occurrence with large or long sampling units). If you reject such points, your sample will be biased toward the center of the population (i.e., you will be less likely to sample the edges of the population). If you "reflect" the line or quadrat from the population edge back into the population, you bias your sampling toward the edges of the population.
2. This technique introduces the probability of overlapping sampling units. This is, for example, a major problem with bird surveys, in which some birds can be detected up to 100m away, necessitating that sampling points be separated by twice that distance. For quadrats (either rectangular ones, or circular ones located by their center point) overlap is highly undesirable, because we will not be able to use the finite-population correction factor discussed later in this chapter. For transect line intercepts, you could address overlap by selecting random compass orientations from each randomly located point; lines represent an infinite population regardless of their orientation and so we never use the finite population correction factor. This

[2] *Although such a random selection procedure is justified for sampling with point intercepts, frequency quadrats and biomass and cover estimation quadrats, the time required to position 100 to 200 or more of these small sampling units makes this procedure impractical. Instead, some type of systematic approach is usually used.*

approach, however, eliminates the possibility of orienting lines consistently along the gradient.[3]

A better method for locating random sampling units is the **grid-cell method**. The grid-cell method eliminates the problems associated with the random coordinate method and is one of the most efficient and convenient methods of randomly positioning quadrats. The sampled population area is overlaid with a conceptual grid (there is no need to actually lay out tapes and strings to achieve this), where the grid-cell size is equivalent to the size of each sampling unit. Consider the dense, clumped population example introduced earlier. We have overlaid a grid of 4m × 10m quadrats on this population (Fig. 8.5). If we want to sample ten 4m × 10m quadrats from this population, we would first divide the population into 125 different 4m × 10m cells, as shown on Figure 8.5. Since we are sampling without replacement, 125 possible quadrat positions (5 along the x-axis times 25 along the y-axis) are possible, none of which overlap. Once one is sampled, it will not be sampled again (at least not during the same sampling period). More information on implementing the grid-cell method in the field is given in Chapter 11.

As its name suggests, simple random sampling is the simplest kind of random sampling, and the formulas used to calculate means and standard errors are easier than with many of the more complex types of designs discussed below. But unless you are planning to use permanent quadrats to detect change, simple random sampling should only be used in relatively small geographic areas where a degree of homogeneity is known to exist. If the sampling area is large and/or the sample size is relatively large, as it often is for frequency or point-intercept simple random sampling, the time spent in locating quadrats or points and traveling between locations can be considerable.

Another problem with simple random sampling is that, simply by chance, some areas may be left unsampled. Figure 8.6 shows a simple random sample of a hundred 1m × 1m quadrats positioned within a 50m × 100m macroplot. By chance, some large portions of the macroplot did not receive any sampling units. This can be especially problematic in populations that are clumped. Computer-simulated sampling (Salzer, unpublished data) suggests that both restricted random sampling and systematic sampling designs (described below) result in more precise estimates than simple random sampling when sampling clumped distributions (the most common situation in biologic populations).

Stratified Random Sampling

Stratified random sampling involves dividing the population into two or more subgroups (strata) before sampling. Strata are generally delineated in such a manner that the sampling units within the same stratum are very similar, while the units between strata are very different. Simple random samples are taken within each stratum.

Strata should be defined based on the response (of the attribute that you are estimating) to habitat characteristics that are unlikely to change over time. Examples of characteristics that might be used to delineate strata are soil type, aspect, major vegetation type (e.g., forest or grassland), and soil moisture. You should avoid defining strata based on characteristics related to the attribute you are estimating, since this is likely to change with time, leaving you stuck with strata that are no longer meaningful. For example, if you are interested in estimating the density of species X, and you note that the east half of the target population is much more densely populated than the west half, avoid basing your strata on this fact alone. If there is an obvious habitat feature responsible for this difference such as aspect, then base your strata on this habitat feature. If there is no obvious reason for the difference, you are probably better off using a simple-

---

[3] *You should orient sampling units to include as much of the gradient variation as possible within the sampling unit. This maximizes variability included within the sampling unit and minimizes the variability between them, and can dramatically increase the efficiency of the sampling design. See the computer-simulated sampling design example above.*

Figure 8.5. The dense clumped population overlaid with a grid of 4m × 10m quadrats. There are 125 possible quadrat locations for this size and shape of quadrat. The 4m × 25m quadrat (50 possible quadrat locations) and 4m × 50m quadrat (25 possible quadrat locations) are also shown. The 4m width was used for illustration only. A better quadrat design would be thinner (e.g., 0.25m or 0.5m) but would not show up well on the figure.

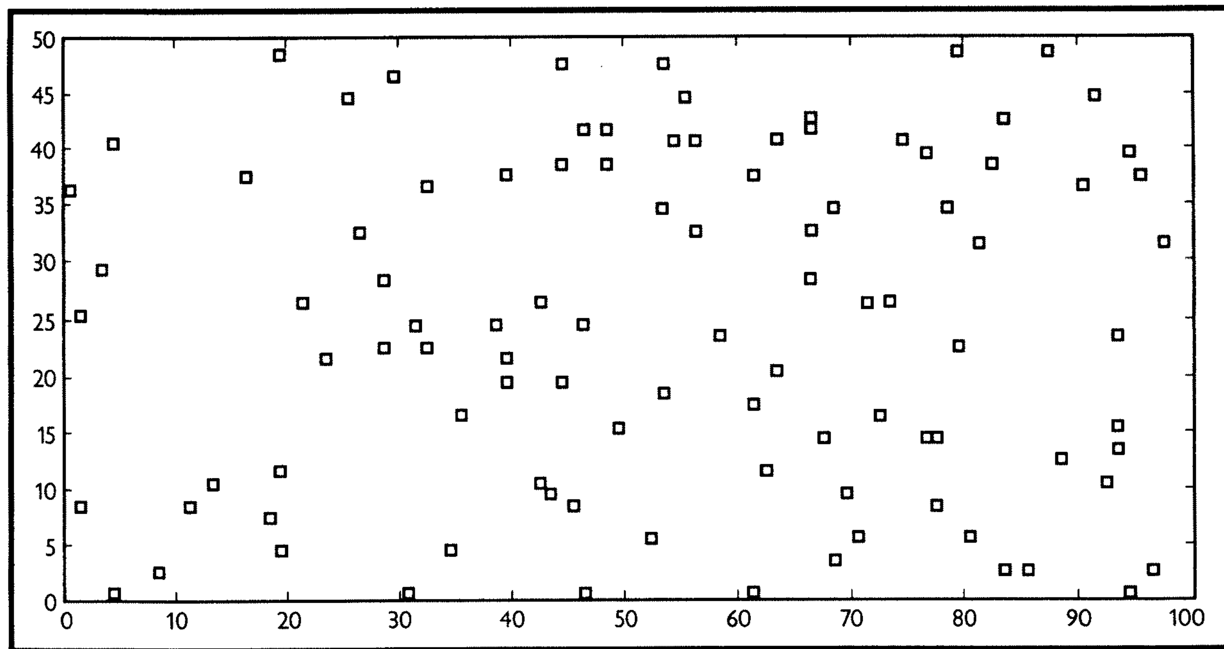


Figure 8.6. A simple random sample of 100 1m × 1m quadrats positioned within a 50m × 100m macroplot. Simply by chance, some large portions of the macroplot did not receive any sampling units.

random-sampling procedure, because you might find that your management will result in more recruitment of species X into the west half of the target population, leaving you with a stratified random sampling procedure that is less efficient than simple random sampling.

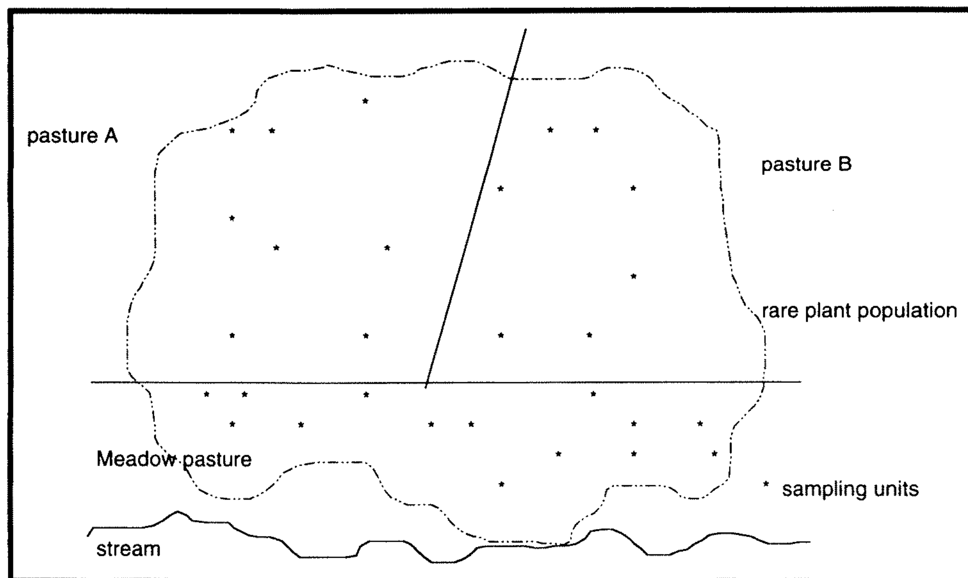


Figure 8.7. A rare plant population grows in a meadow along a stream and up an adjacent slope. The population area is grazed in the spring in Pasture A and in the fall in Pasture B. The meadow has recently been excluded from livestock grazing except for a short duration low intensity graze in the early spring before green-up. The three areas are treated as strata in a stratified random sample.

Figure 8.7 depicts a rare plant population that occurs within three grazing pastures, each with different grazing regimes. We decide to use each pasture as a sampling stratum. Through pilot sampling, we discover that the meadow portion of the population is more variable than the portion growing on the adjacent slope in the upland pasture, and we allocate more sampling units to that stratum.

Sampling units do not have to be allocated in equal numbers to each stratum. In fact, one of the benefits of stratified random sampling is that, when the attribute of interest responds differently to different habitat features, you can increase the efficiency of sampling over simple random sampling by allocating different numbers of sampling units to each stratum. Sam-

pling units can be allocated: 1) equally to each stratum, 2) in proportion to the size of each stratum, 3) in proportion to the number of target individuals in each stratum, or 4) in proportion to the amount of variability in each stratum.

Figure 8.8 illustrates a stratified random sampling scheme used in the U.S. Fish and Wildlife Service's National Wetlands Inventory of the United States (Dahl and Johnson 1991). A sample of many plots, each 4 miles square, was allocated to three strata in the state of North Carolina. Notice how the coastal stratum, because it has more habitat variability and greater suspected wetland density, is sampled more intensively. This differential sampling intensity, with greater effort allocated to strata with higher density and/or greater variability, is a common feature of stratified random sampling.

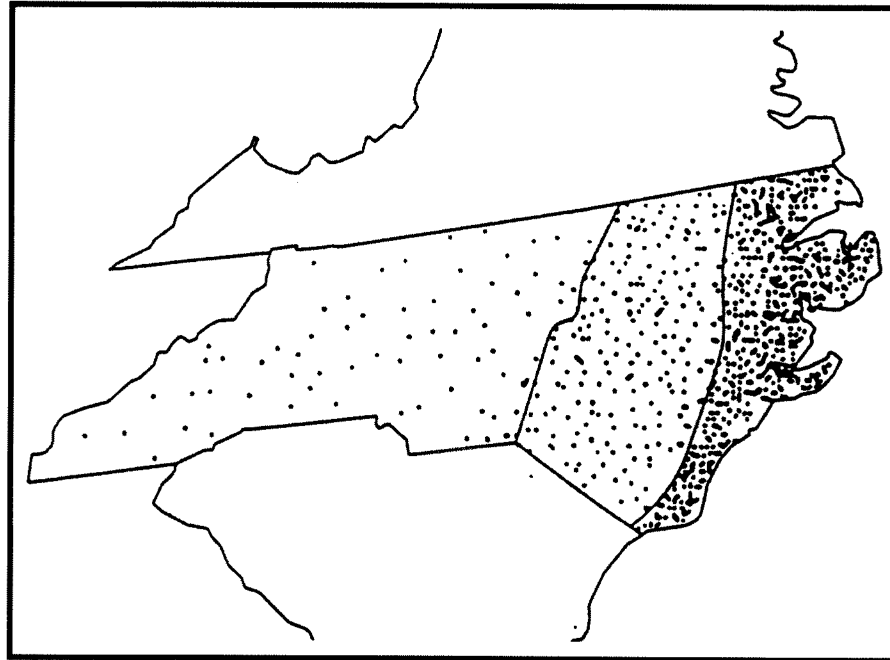


Figure 8.8. A stratified random sampling scheme. This example, from the National Wetlands Inventory (Dahl and Johnson 1991), shows how a sample of many plots, each 4 mi^2, was allocated to three strata in the State of North Carolina.

The major advantage of stratified random sampling is an increase in the efficiency of population estimation over simple random sampling when the attribute of interest responds very differently to some clearly defined habitat features that can be treated as strata. The principal disadvantage is the more complicated formulas that must be used both to determine allocation of sampling units to each stratum and to estimate means and standard errors. Since we are taking a simple random sample within each stratum, the possibility exists that, simply by chance, areas within one or more strata may be left unsampled. Additionally, each stratum should be somewhat homogeneous and cover a relatively small geographic area (for plants and stationary animals); otherwise the method will be less efficient than systematic and restricted random sampling.

Refer to Appendix IV for the formulas necessary to calculate sample sizes when using stratified random sampling and for the formulas to calculate statistics. Other good references include Cochran (1977), Krebs (1998), and Thompson (1992).

Systematic Sampling

A systematic approach is commonly used in sampling plant and animal populations. It is one of the easiest ways to locate sampling units throughout a sampled population because of low setup and travel time between sampling units. It also ensures good interspersion of sampling units. The regular placement of quadrats along a transect is an example of systematic sampling. The starting point for the regular placement is selected randomly. To illustrate, let us say we decide to place ten 1m^2 quadrats at 5m intervals along a 50m transect. The selection of the starting point for systematic sampling must be random. Therefore, we randomly select a number between 0 and 4 to represent the starting point for the first quadrat along the transect and place the remaining nine quadrats at 5m intervals from this starting point. Thus, if we randomly select the 3m mark for the first quadrat, the remaining quadrats will be placed at the 8, 13, 18, 23, 28, 33, 38, 43, and 48m points along the transect. This is illustrated in the transect along the left side of Figure 8.9.

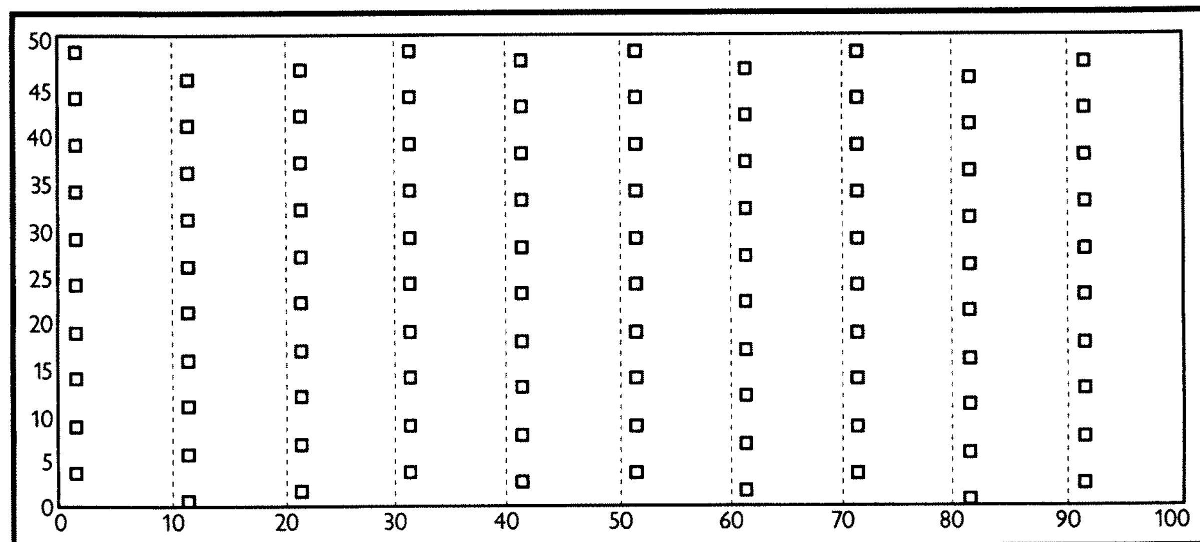


Figure 8.9. A 50m × 100m macroplot, sampled by 100 1m × 1m frequency quadrats. The quadrats are aligned along transects. Both the transects and the quadrats are systematically positioned with a random start. A random starting point is selected for the transects along the baseline, while separate random starting points are selected for the quadrats along each transect.

Systematic sampling with a random starting point is commonly used in animal studies because it permits easily identifying sampling points and because it generally, but not always, yields estimates of comparable accuracy and precision to those provided by purely random sampling. For example, for sampling fishes in small streams, a systematic sampling approach has been recommended (Hankin and Reeves 1988) because it delivers comparable precision, is generally representative, and avoids the work of identifying a complete list of sampling sites required by random sampling. Litter searches of quadrats for amphibians and small lizards are also frequently made in a systematic fashion. Similarly, point counts for birds are almost invariably arrayed in a systematic fashion along counting "routes" or transects.

A common use of systematic sampling in vegetation studies is to facilitate the positioning of quadrats for frequency sampling and of points for cover estimation. Using this approach, a baseline is laid across the sampled population, either through its center or along one side of it. Transects are run perpendicular to the baseline beginning at randomly selected points along the baseline (if the baseline runs through the middle of the population, transects are run in either of two directions; the direction for each one can be randomly determined by tossing a coin). Quadrats or points are then systematically positioned along each transect. The starting point for the first quadrat or point along each transect is selected randomly.

Systematic samples, if well designed, can safely be analyzed as a simple random sample. Milne (1959) analyzed data taken from random and systematic samples of 50 totally enumerated biologic populations and found that there was no error introduced by assuming that a centric systematic sample is a simple random sample and by using all the appropriate formulas from random sampling theory (Krebs 1998:228). Milne's (1959) conclusion was that "with proper caution, one will not go very far wrong, if wrong at all, in treating the centric systematic-area sample as if it were random." Note, however, that Milne compared random samples to centric systematic samples, illustrated in Figure 8.10. The units of a centric systematic sample lie on equidistant parallel lines (these can be thought of as transects) arranged in a manner such that, in effect, the area is divided into equal squares (see dotted lines) and a sampling unit taken from each square. Thus, the sampling units are spaced a considerable distance apart with maximum interspersion of sampling units throughout the sampled population.

The design shown in Figure 8.9 ensures good interspersion of sampling units throughout the sampled population. Here, a 50m × 100m macroplot was sampled by a hundred $1m^2$ frequency

quadrats, with a 100m baseline along the southern edge. The quadrats were aligned along transects. In this example both the transects and the quadrats were systematically positioned with a random start. In the case of the transects, a random number between 0 and 9 was selected. That number was 1. The first transect therefore began at the 1m mark along the baseline, with subsequent transects beginning at 11m, 21m, up to 91m. In the case of the quadrats, a random number between 0 and 4 was chosen for each transect, the first quadrat positioned at that point, and subsequent quadrats placed at increments of 5m from the first quadrat. Thus, for transect number 1 the first quadrat was located at the 3m mark, with subsequent quadrats located at the 8m, 13m . . . 48m marks.[4]

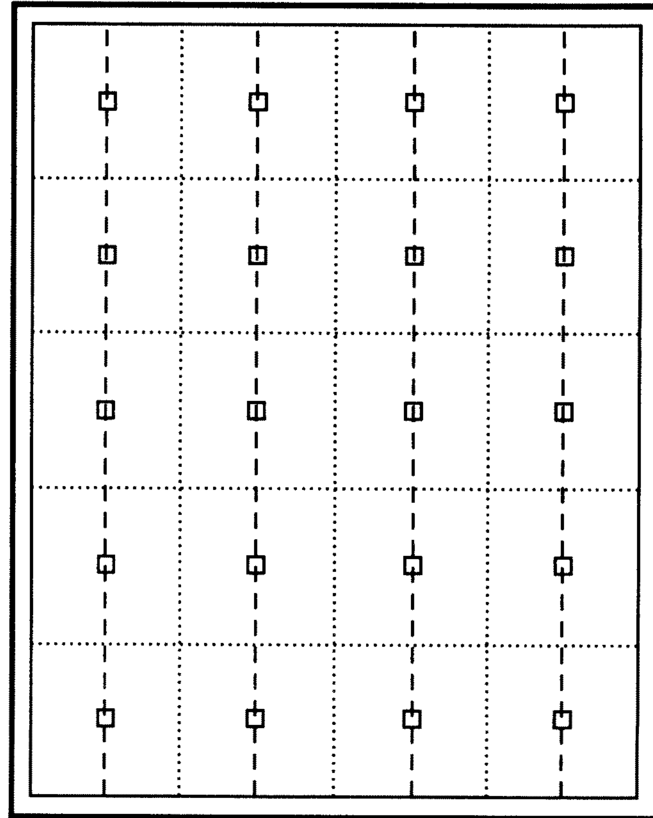


Figure 8.10. A centric systematic sample (adapted from Milne 1959). Small squares are sampling units, dashed lines are transects, and dotted lines show how the sampling units fall in the center of each subunit of area.

Good interspersion of sampling units throughout the sampled population is one of the principal advantages of systematic sampling. Strictly speaking, however, systematic sampling is analogous to simple random sampling only when the population being sampled is in random order (see, for example, Williams 1978). Populations in random order are rare in biology; most natural populations of both plants and animals exhibit a clumped spatial distribution pattern. This means that nearby units tend to be similar to (correlated with) each other. If, in a systematic sample, the sampling units are spaced far enough apart to reduce this correlation, the systematic sample will tend to furnish a better mean and smaller standard error than is the case with a random sample, because with a random sample one is more likely to end up with at least some sampling units close together (see Milne 1959; discussion of sampling an ordered population in Schaeffer et al. 1979). Computer simulation has validated this conclusion. For example, for density estimation, Salzer (unpublished data) found through Monte Carlo simulations that systematic designs outperform simple random sampling in terms of precision when sampling clumped populations.

On a cautionary note, systematic sampling for density estimation can lead to questionable results if the sampling design creates a situation where there are only a small number of potential samples. For example, consider the macroplot shown in Figure 8.11. Ten 1m × 50m quadrats are systematically positioned in the macroplot with a random starting point at the 2m position on the x-axis, and the quadrats spaced at 10m intervals after that. In this case, since the position of all quadrats is fixed once the first quadrat is positioned, there are only 10 possible samples to draw from, depending on which of the 10 possible starting points is randomly selected in the first 10m segment of the population (0, 1, 2, 3, 4, 5, 6, 7, 8, or 9). The sampling distribution (distribution of all possible sample mean values) for this sampling design might resemble a uniform (flat) distribution instead of the smooth, bell-shaped curve of the normal distribution, because

---

[4]*What is the sampling unit in Figure 8.9? You have two options: You can treat the sample as if the quadrats had been selected as a simple random sample or you can calculate separate frequency values for each transect and treat the transect as the sampling unit. The implications of each option will be clearer once you have been introduced to cluster sampling and two-stage sampling, discussed below.*

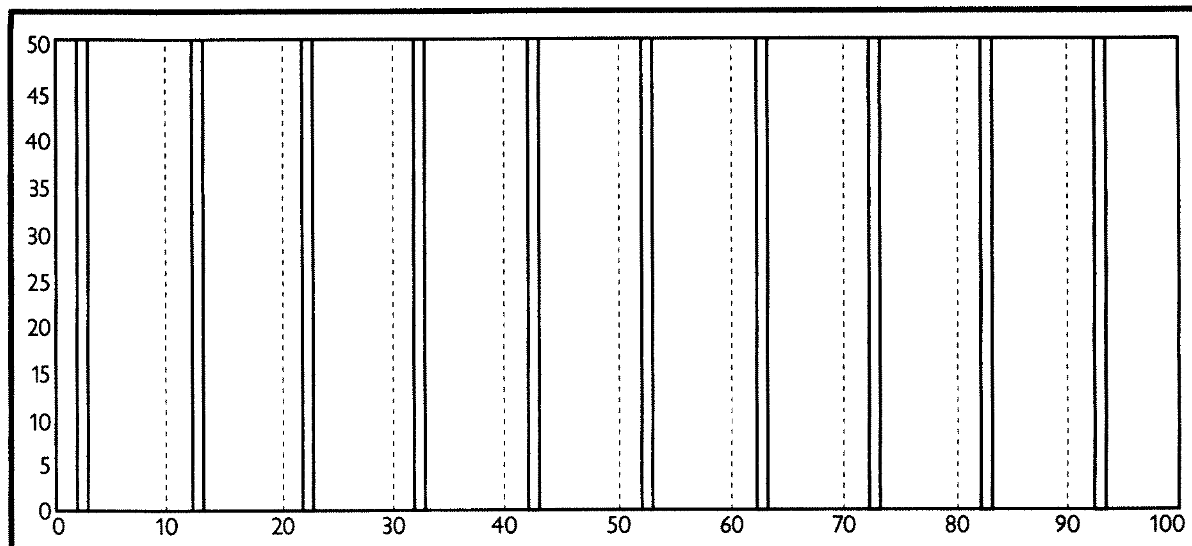


Figure 8.11. A systematic sample of 10 1m × 50m quadrats in a 50m × 100m macroplot. Note that there are only 10 possible samples, corresponding to which of the 10 possible starting points in the first 10m segment of the baseline (x-axis). In this case, the sample started at the 2m mark.

there are only 10 different sample means possible. Treating such a sample as if it were a simple random sample could result in poor estimates of the sample standard error. The next type of sampling design, restricted random sampling, solves this problem. Except for this somewhat uncommon situation, however, systematic sampling is preferred over restricted random sampling. If more than 30 possible systematic samples may be drawn, systematic sampling is acceptable.

Another caution is that situations do arise in which systematic sampling can seriously bias estimates if the pattern of the sampling units intersects some pattern in the environment (e.g., dune ridges and slacks; Goldsmith et al. 1986). One example is estimating food abundance for wildlife in croplands planted in a regularly repeated fashion. Systematic sampling, depending on how it was applied, might consistently locate sampling units between or on top of crop rows and thereby yield substantially different estimates. This has occurred, for example, in studies of availability of waste corn for waterfowl.

If some periodic pattern does exist, the data analysis will not reveal this, and your estimates, particularly of standard errors, will be wrong. Although this type of periodic pattern is rare in nature, you should be alert to the possibility.

Restricted Random Sampling

In restricted random sampling, you determine the number of sampling units, n, you will need to meet your monitoring objective (sample size determination is discussed below), then divide your population into n equal-sized segments. Within each of these segments, a single sampling unit is randomly positioned. The sample of n sampling units is then analyzed as if it were a simple random sample.

Figure 8.12 is an example of a restricted random sampling procedure. This is the same 50m × 100m macroplot as we used in our discussion of systematic sampling. In this case, however, we divide the x-axis into ten 10m segments. Within each of these segments we randomly select a single quadrat location. This gives us 10 possible random locations within every 10m segment of the x-axis. Every quadrat location in the macroplot still has an equal probability of selection. The same technique can also be applied to the y-axis if there is more than one possible quadrat position along that axis.

The restricted-random-sampling procedure can also be used when the sampling unit is a transect instead of a quadrat. Divide the population into equal-sized segments and allocate a single transect to each segment. If you are locating sampling units such as quadrats or point intercepts along transects (similar to Figure 8.9), you may want to use a combination of the restricted

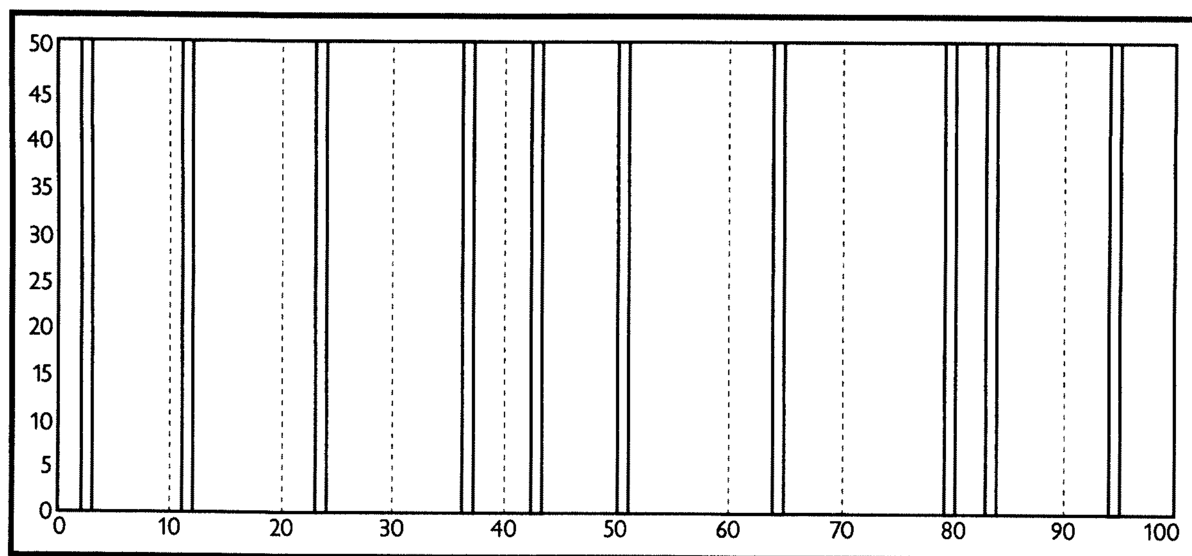


Figure 8.12. A restricted random sample of 10 1m × 50m quadrats in a 50m × 100m macroplot. One quadrat is randomly positioned within each 10m segment of the baseline (x-axis).

and systematic designs. If, for example, you decide to run 10 transects, each with 50 point intercepts, perpendicular in one direction from a baseline, you could divide the baseline into 10 equal segments, randomly locate beginning points for each transect within each of these 10 segments, and then systematically space the point intercepts along each transect (as in Figure 8.11, except with points systematically positioned along one edge of each quadrat).

Restricted random sampling is similar to both stratified random and systematic sampling. It is similar to stratified random sampling in that we have effectively stratified our macroplot into 10 strata. However, unlike stratified random sampling, the strata are arbitrary, and we take only one sampling unit in each stratum. As with systematic sampling, we divide our population into equal-sized segments. With systematic sampling, however, only the first sampling unit is randomly determined; all the others are spaced at equal intervals from the first.

Similar to systematic sampling, restricted random sampling results in very good interspersion of sampling units throughout the target population. Furthermore, Salzer (unpublished data) has shown through simulation studies that restricted random sampling results in more precise estimates of density than simple random sampling. He has also demonstrated the procedure to be more robust than systematic sampling when the number of possible systematic samples are few, because with restricted random sampling designs you do not constrain the number of potential samples from which you can draw. The principal disadvantage of restricted random sampling is that you can, purely by chance, end up with sampling units positioned side-by-side. This can leave larger portions of the sample area unsampled than is the case with a systematic design. When the number of potential systematic samples is large enough (more than 25 to 30), you are probably better off choosing a systematic sample. Otherwise, use the restricted random design.

Cluster Sampling

Cluster sampling[5] is a method of selecting a sample when it is difficult or impossible to take a random sample of the individual elements of interest. With cluster sampling, we identify groups or clusters of elements and take a random sample of these clusters. We then measure every element within each of the randomly selected clusters.

---

[5]*Cluster sampling should not be confused with cluster analysis, a technique used in classification and taxonomy.*

In monitoring, cluster sampling is most often used when the objective is to estimate something about individuals such as parasite loads in animals or the mean number of flowers per plant. For example, you may want to track the average height of plant X in population Y. There are too many plants in the population to feasibly measure all of them. Five quadrats are randomly placed in the population, and the heights of all plants within these quadrats are measured (Fig. 8.13).

Cluster sampling and two-stage sampling are the only two efficient designs that can be used to sample individual plant and animal characteristics. Examples for application to plant monitoring include estimating number of seeds produced per plant, biomass per plant, and average height or size per plant. In these examples, a quadrat is employed as the cluster and each plant is an element. Examples of the application of two-stage and cluster sampling to animal studies include estimating the size of birds' eggs (nests are the cluster and eggs are the elements), the number of eggs per nest (nests may be located using trees or a quadrat as the cluster and nests as the element), food habits of fish (e.g., a seine catch as the cluster and each fish stomach as an element), and the size of beetles (the trap is the cluster and each beetle the element). Sometimes,

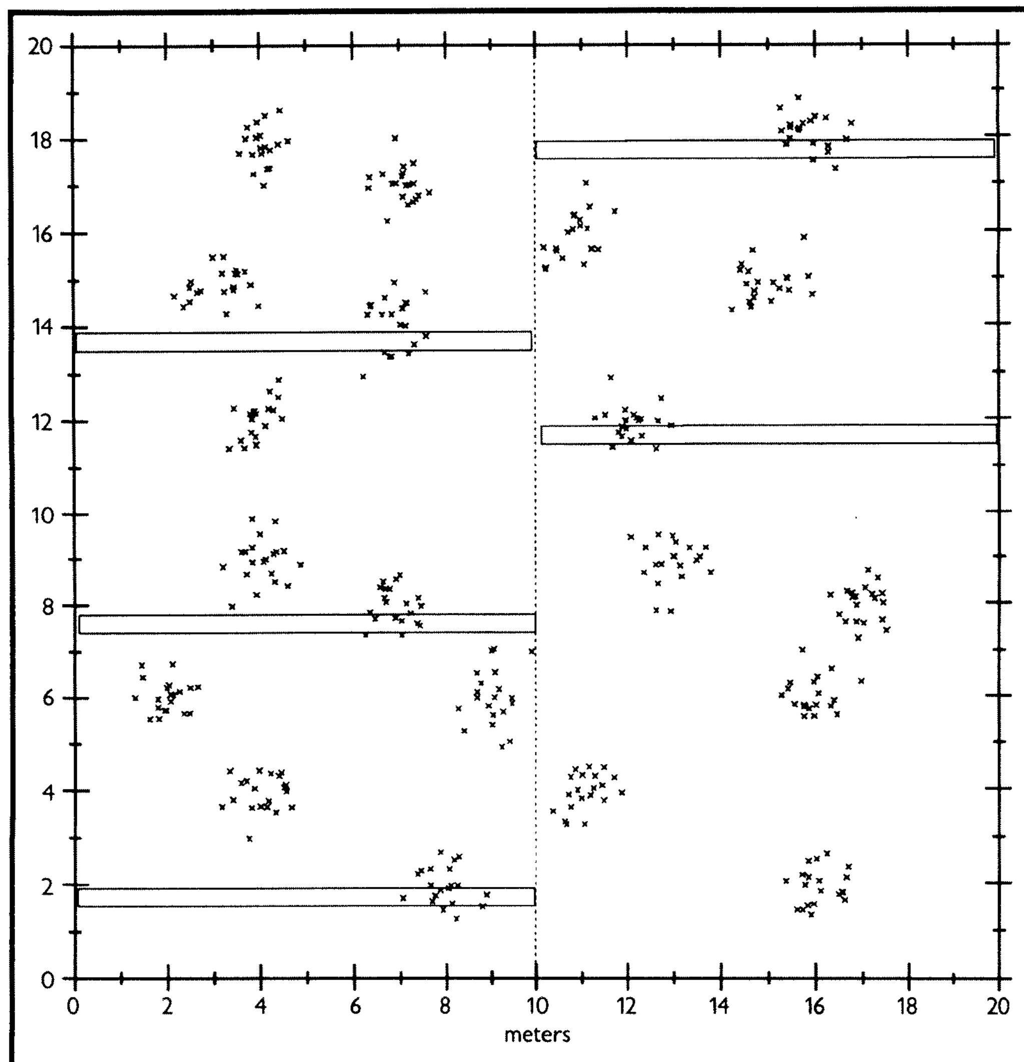


Figure 8.13. An example of cluster sampling to estimate the mean height of plants in a population. Five quadrats are randomly placed in the population and the heights of all plants within these quadrats are measured.

the elements are erroneously treated as independent sampling units. Careful articulation of the method of positioning sampling units should help avoid this problem.

With animals, cluster sampling is also sometimes of a temporal rather than a spatial nature, such that repeated counts are made during randomly determined visits to a site instead of making the single counts at randomly determined times, thereby greatly saving on time needed to reach sites to make counts.

The advantage of cluster sampling is that it is often less costly to sample a collection of elements in a cluster than to sample an equal number of elements selected at random from the population (Thompson 1992). It is most efficient when different clusters are similar to each other and incorporate much variability within. Because individuals near each other tend to be similar, this condition will not be realized with square clusters (Thompson 1992). Therefore, just as with simple random sampling for density estimation, cluster sampling using long, narrow quadrats to delineate clusters will be more efficient than using square quadrats.

Cluster sampling has several disadvantages. First, all elements within each cluster must be measured. If the clusters contain large numbers of the element of interest, two-stage sampling, described below, will be more efficient. Second, it is often difficult to figure out how many clusters should be sampled versus how large each cluster should be. Third, more complex calculations are required. Most statistical software packages do not include these calculations. A worked example is provided in Appendix IV.

Two-Stage Sampling

Two-stage sampling is similar to cluster sampling in that we identify groups of elements about which we wish to make inferences. We then take a random sample of these groups. However, instead of measuring every element in each group as we would if doing cluster sampling, we take a second sample of elements within each group. The groups sampled are called primary sampling units, while the elements sampled are called secondary sampling units. The secondary sampling units can be either a simple random sample of elements or a systematic sample of elements. Figure 8.14 shows a two-stage sampling design. Like cluster sampling, the main use of two-stage sampling is to estimate some value associated with individual plants. It has also been used to increase the precision of counts of large mammals, for example, deer (Freddy and Bowden 1983) and wildebeest (Norton-Griffiths 1973).

An example of two-stage sampling is its use in estimating the number of flowers per plant produced by species X. We might randomly locate a sample of quadrats in the target population. Within each quadrat we then take a random sample of plants and count the number of flowers on each plant selected. The quadrats are the primary sampling units and the plants are the secondary sampling units.

Two-stage sampling may also involve macroplots and quadrats. For example, you are interested in the mean density per quadrat of a salamander species, and you want to be able to make statistical inferences to a large area. The area is relatively homogeneous, with no logical basis of stratification. Seven 50m × 100m macroplots (primary sampling units) are randomly distributed throughout the population, and fifteen 0.20m × 25m quadrats (secondary sampling units) are randomly sampled within each macroplot.

Both these examples involve simple random sampling at both stages. Either or both of the stages may involve different types of sampling. A common type of two-stage sampling involves simple random sampling at the primary stage and systematic sampling at the second stage. We have already seen examples of this: when quadrats or points (secondary sampling units) are systematically located with a random start along transects, and the transects (primary sampling units) are run from randomly selected points along a baseline. Of course, the transects could be positioned using another type of design such as restricted random sampling or systematic sampling. The point is that the two stages can involve different sampling designs.

Figure 8.14. Two-stage sampling to estimate the number of flowers per plant on a particular species of plant. Five 4m × 50m quadrats (primary sampling units) are randomly located in the sampled population and three 1m × 25m quadrats (secondary sampling units) are randomly located within each of the five larger quadrats. The number of flowers per plant is counted within all of the selected 1m × 25m quadrats.

The practical advantage of two-stage sampling, compared with a simple random sample of the same number of secondary units, is that it is often easier or less expensive to observe many secondary units in a group than to observe the same number of secondary units randomly spread over the population (Thompson 1992). Travel costs are therefore reduced with two-stage sampling. Two-stage and cluster sampling designs are the only two efficient designs that can be used to sample individual plant and animal characteristics.

Because sampling occurs at both stages, there are standard deviations associated with estimates of the values at both stages (unlike cluster sampling, which has no standard deviation associated with the values measured at the second stage). This results in more complicated formulas for estimating values and standard errors. A less complicated alternative is to follow Cochran (1977), ignoring the standard deviation of the secondary sample as long as the finite population correction is not applied to the standard error of the primary sample. For example, if we had a sample of quadrats arranged along a transect, we could simply use the mean of each transect's collection of quadrats as our unbiased estimate of the transect value. We then treat the collection of transect values as a simple random sample. This allows us to use standard statistical computer programs to perform our analysis.

Platts et al. (1987) provides good worked examples of calculating means and standard errors from two-stage sampling when you wish to consider the standard deviation of the secondary sample (Appendix IV). More detailed discussions can be found in Cochran (1977:279), Krebs (1998), and Thompson (1992).

Comparison of Sampling Designs: the Sampling Unit Revisited

Often we will arrange small sampling units (quadrats for measuring frequency, visual estimates of percent cover and biomass or point intercepts for measuring cover) along a transect. Should these be considered a random sample of the smaller units (using the transects only for locating these units) or should the transect itself be considered the sampling unit?

Technically, when we use transects as the sampling units, whether for frequency quadrats, cover point estimates, biomass quadrats, or visual cover estimation quadrats, we are really conducting two-stage sampling. The transects are the primary sampling units, and the quadrats or points are the secondary sampling units. Standard deviations are associated with both the primary sample of transects and the secondary sample of quadrats or points. Two-stage designs take into account both sets of standard deviations. The result is a much more complex set of equations that standard statistical programs will not calculate. Although we could subject these data to the more complex formulas of two-stage sampling, there is no need to do so. Cochran (1977:279) points out that we can ignore the standard deviation of the secondary sample as long as we do not use the finite-population correction factor in our analysis. We can simply use the mean of each transect's collection of quadrats or points as our unbiased estimate of the transect value.

For small quadrats that are used to visually estimate percent cover or estimate biomass or estimate density of small animals, it is generally best to group those along a transect and consider the transect the sampling unit. This allows us to use small quadrats of practical size in the field while taking advantage of the benefits of elongated sampling units (the transects) that cross the variability inherent in the population. By treating the transects as the sampling units, we get the best of both worlds.

For frequency or point intercept cover data you should usually treat the quadrats or point intercepts as the sampling units rather than the transects along which these are located. Estimates will be more precise and significance tests more powerful because of the larger sample sizes realized by using quadrats or point intercepts rather than transects as the sampling units. There are at least two situations, however, in which you might want to treat the transects as the sampling units. The first of these is when the transects are permanent (see discussion on permanent vs. temporary sampling units below). If you have reason to believe that the average values

per transect are more correlated between years than are the quadrat or point values, you may choose to analyze the transects rather than the quadrats or points as the sampling units.[6]

The second situation in which you might want to treat the transects as the sampling units when systematically sampling frequency quadrats or cover point intercepts is when the quadrats or points are not far enough apart to be considered independent. This is more likely to be a problem in already established studies, where quadrats or points were placed contiguously or a very short distance apart. Hopefully, you will design new studies in such a manner that the quadrats or points are spaced far enough apart to achieve independence.

Independence means that the sampling units are not spatially correlated, that the response of the species in Quadrat A is not related to the response of the species in Quadrat B because of their proximity to one another. If the quadrats or points are far enough apart that they can be considered independent, we have the benefit of increasing our sample size dramatically (because the point or plot is the sampling unit instead of the transect) while keeping the field efficiency of locating sampling units rapidly along a transect. Conversely, the contiguous placement of quadrats along a transect or the separation of such quadrats by small distances (e.g., one "pace"), practically ensures that adjacent sampling units will be correlated. This will result in an underestimation of the standard error and questionable results

Determining how far apart to place sampling units along a transect for them to be considered independent can be difficult. Chapter 12 discusses this issue in more detail for plants. This is a particular problem for animals as well, especially those that are detected at long distances by their calls and hence easily double-counted. These issues are discussed in Chapter 13. Probably the best way to determine spacing of sampling units along transects is to consider the degree of interspersion of your design. The goal is to have sampling units interspersed as well as possible throughout the area of the target population (see previous discussion on interspersion). Once you have delineated the area you intend to sample, strive for a design in which the spacing between transects is about the same as the spacing between sampling units. If you do this, it is likely that the issue of independence will take care of itself.

Double Sampling

Double sampling, sometimes called two-phase sampling, involves the estimation of two variables. Because one of these variables, the variable of interest, is difficult and expensive to measure, it is measured in only a relatively small number of sampling units. To improve the rather poor precision of the estimate that normally results from a small sample, an auxiliary variable that is much easier to measure is estimated in a much larger number of sampling units. The variable of interest is measured in a subsample of the sample of units in which the auxiliary variable is measured.

The idea of double sampling will become clearer with examples. The technique is often used in estimating aboveground biomass in rangelands. Because it is slow and expensive to clip, dry, and weigh biomass in many sampling units, observers train themselves to visually estimate biomass. Once trained, the observers randomly locate quadrats within a target population and visually estimate the biomass in all the quadrats. For example, 100 quadrats are so estimated. Then, in a subsample of these quadrats, say 10, the visual estimates are made as in the other quadrats, but after these estimates are recorded, the aboveground biomass is clipped, dried, and weighed. Thus, for these 10 quadrats we have two estimates of biomass, one from the visual estimate, the other from the actual weighing of the clipped biomass.

---

[6]*It is highly unlikely that you will be able to accurately reposition point intercepts as permanent sampling units, but a transect of point intercepts may be highly correlated from year to year and thus be suitable for consideration as a permanent sampling unit.*

Double sampling is also used in forest surveys to estimate the volume of trees in a stand. Trained observers make a visual estimate of volume for a large sample of standing trees, while accurate volume measurements that require felling are limited to a small subsample of trees (Thompson 1992).

In wildlife management, double sampling may be used for surveys of breeding waterfowl. Aerial surveys estimate abundance over an extensive area, but a subsample of the survey areas are subjected to more thorough ground surveys. The ground surveys are used to adjust the bias inherent in aerial surveys (Routledge 1999).

In all these cases, the subsample on which the variable of interest is actually measured is more accurate, but the precision of the estimate can be greatly improved by considering the measurements on the auxiliary variable. The improvement in precision depends on how well the auxiliary variable correlates with the variable of interest. In the examples given above, this relates to how well the trained observers actually estimate biomass, tree volume, or abundance of waterfowl.

If the auxiliary variable is relatively quick to be measured and is highly correlated with the variable of interest, double sampling is much more efficient in estimating variables that are difficult to measure compared to directly measuring the variable in all sampling units. A disadvantage is that the formulas for data analysis and sample-size determination are much more complicated than formulas for simple random sampling. Refer to Cochran (1977:327-358) or Thompson (1992:139-147) for the formulas needed to analyze double-sampling data.

Taking a Simple Random Sample of Individual Plants or Animals

Let us say that we want to estimate something about a population of individual plants such as their mean height or the mean number of flowers per plant, and that the population is too large to measure this variable on every single plant in the population. Easy, you say; we will just take a simple random sample of plants, measure the variable on the sample, and calculate the mean and standard error for the sample. We can then construct a confidence interval around the estimate at whatever confidence level we choose (e.g., a 95% confidence interval). Although it might seem logical to take a simple random sample of plants, for most plant populations this is not feasible.

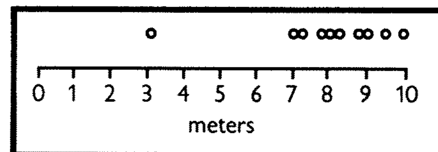


Figure 8.15. Distribution of individuals of plant species X along a 10-meter transect. A randomly positioned point on the transect will be far more likely to be closest to the individual at the 3m mark than to any of the other plants.

One way that is often—and incorrectly—used is to select a random sample of points in the population and to take the nearest individual to each of these points. Unfortunately, this works only if the population of plants or animals is randomly distributed, a condition rarely met by natural populations. If, as is typically the case, the individuals are spatially distributed in patches, this technique most decidedly will not result in a simple random sample of individuals. Consider Figure 8.15, which shows the distribution of individuals of a hypothetical plant species along a 10m transect. Note that nine of the ten individuals are clumped in the last 3 meters of the transect, while a single individual occurs at the 3m mark. A randomly positioned point along this transect would have about a 50% probability of being closest to this isolated individual and about a 20% chance of being closest to the individual at the 7m mark. The probability of the point lying closest to any of the other eight individuals is much less.

Thus, in a clumped population of plants, a "random" sample of individuals chosen by taking the individuals closest to randomly located points will be biased toward those individuals that are isolated from the majority of the population. These individuals may either be much larger than the majority of plants in the population because of reduced intraspecific competition or much smaller because they occupy suboptimal habitat. Let us say we are interested in estimating the mean height of such a population. By biasing our estimate toward the isolated plants in the population we may greatly underestimate or overestimate the mean height of such

a population. The same is true for any other attribute associated with individual plants that we may wish to estimate such as number of fruits per plant. Obviously, for populations of plants that follow a clumped or patchy distribution—which is by far the majority of populations—such a sample of individuals cannot be used to adequately characterize the population.

How, then, can you take a random sample of individuals? One way is to completely enumerate every individual in the population by, for example, mapping every individual and numbering each one from 1 to *n*. A simple random sample could then be taken by drawing random numbers between 1 and *n*. This, of course, would be extremely time-consuming except for small populations, in which case you might be able to measure the attribute on every individual in less time. For example, if you are interested in mean height of plants, you could simply measure the height of every plant in the population and not sample at all. If, however, you need to estimate the mean number of flowers per plant and each plant has several hundred flowers, selecting individual plants randomly from a complete list might be a reasonable approach, although for most practical purposes it is far too time-consuming.

Another possibility is to take a systematic random sample of individuals. With this method you gather information from every *nth* individual in the population. This method will work if you are planning to conduct a complete census of the population, but you are also interested in estimating some attribute from a subset of the individuals (e.g., number of flowers/plant). Before you start you need an estimate of the following two types of information: 1) the approximate size of the population, and 2) the approximate number of individual plants you will need to sample (calculated as a proportion of the total population size). For example, if you estimated a total population size of 1000 plants and your sample-size calculations from pilot sampling identified a sample size of 100 plants, you would count the number of flowers on every 10th plant encountered. You choose a random number between 1 and 10. Say the number is 4. Then, starting at one edge of your population you systematically count the plants. You place a pin flag next to plant number 4, another next to plant number 14, and so on until you have counted all the plants. You can then come back and count the flowers on the flagged plants. This sample can properly be analyzed as a simple random sample.

The most practical approach to estimating attributes of individual plants and animals usually employs cluster sampling or two-stage sampling designs using quadrats as primary sampling units. In a cluster sample you would measure the attribute on all the plants in the primary sampling unit (the quadrat). If individuals are still too numerous within the quadrat to measure all of them, you could employ a two-stage sampling design by positioning smaller quadrats (secondary sampling units) within each large quadrat (primary sampling units).

SHOULD SAMPLING UNITS BE PERMANENT OR TEMPORARY?

A critical decision in sampling designs for monitoring is whether to make your sampling units temporary or permanent. When sampling units are temporary, the random sampling procedure is carried out independently at each sampling period. For example, your sampling objective involves detecting change in density over time of a plant species in a 50m × 100m macroplot. In the first year of sampling you place twenty-five 0.5m × 25m quadrats within the macroplot by randomly selecting 25 unique sets of coordinates and counting the number of the species in each quadrat. In the second year of sampling, you place another twenty-five 0.5m × 25m quadrats by randomly selecting a new set of coordinates and counting the number of the species in each quadrat. The sampling units (quadrats) in this example are temporary, and the two samples are independent of each other.

Using the same sampling objective, you could decide to use permanent quadrats. In the first year of sampling you randomly place the 25 quadrats as described above and count the number of individuals in each quadrat. This time, however, you permanently mark the locations of the 25 quadrats. In the second year of sampling, you count the number of individuals in the same quadrats. In this example the sampling units are permanent, and the two samples are dependent.

The principal advantage of using permanent instead of temporary sampling units is that for many species the statistical tests for detecting change from one period to the next in permanent sampling units are much more powerful than the tests used on temporary sampling units. This advantage translates into a reduction in the number of sampling units that must be sampled to detect a certain magnitude of change.

To see why this is so, let us consider the process used in comparing the samples between two periods when using permanent quadrats. If we were using temporary quadrats, we would calculate separate means and standard errors for the two samples and compare these using a statistical test (such as a *t*-test) for independent samples (see Chapter 9). With permanent quadrats, however, we calculate only one mean and one standard error. This requires some explanation. Each quadrat at time one is paired with the same quadrat at time two. The data from which we calculate the mean and standard error consists of the set of differences between each of the quadrats at time one and its corresponding quadrat at time two. For example, we randomly positioned five permanent quadrats in a population and counted the number of plants in each quadrat in 1993 and again in 1994. Data from these permanent quadrats yielded the values in Table 8.2.

Quadrat Number	Number of Plants in 1993	Number of Plants in 1994	Difference Between 1993 and 1994
1	5	5	0
2	5	5	0
3	5	5	0
4	6	6	0
5	3	3	0
			mean difference 0 standard error 0

Table 8.2. Density Data Taken From Five Permanent Quadrats in 1993 and 1994.

Note that the permanent quadrats are extremely effective at detecting the lack of change from year to year. Because in our example the difference between 1993 and 1994 was zero in every quadrat, there is no variation between sampling units, and the standard error is actually 0. Had temporary quadrats been used in both years, it is quite likely that the estimates for each year would have been different just because of chance. For this reason more temporary sampling units (perhaps many more) would have been required to reach the same conclusion that no change had occurred.

Because we are interested only in the change that takes place within each permanent sampling unit between two periods, the difference between sampling units at either period is not nearly as important as it is when using temporary quadrats. Consider the following example. To detect change in cover of species X between two periods, 10 transects were randomly positioned in the target population in 1990. The beginning, middle, and end points of each transect were permanently marked. Fifty points were systematically positioned (with a random start) along each transect and "hits" recorded on canopy cover of species X. The estimate of cover along each transect is then this number of hits divided by the total number of possible hits, 50. Thus, a transect with 34 hits would have a cover estimate of 68% or 0.68. The data from these two years are shown in Table 8.3. (This example is also displayed graphically in Figs. 9.10 and 9.11 of Chapter 9.)

Even though the cover estimates are highly variable between transects for both 1990 and 1994 (for example the mean cover for 1990 is 0.44 with a 95% confidence interval of 0.27 to 0.62), the standard error of the mean difference is relatively small. A 95% confidence interval around this mean difference is –0.02 to –0.12. In fact, in lieu of doing a paired statistical test (such as a paired *t*-test), you could simply look at the 95% confidence interval around the mean difference to see if it includes 0. If not, then you can declare the change significant at a *P* value of 0.05 (*P* values are explained in Chapter 9).

If you had collected these data using temporary transects (i.e., independent samples at both sampling periods), you would have concluded that no change took place. In fact, with the large

Transect Number	Cover in 1990	Cover in 1994	Difference Between 1990 and 1994
1	0.22	0.20	-0.02
2	0.32	0.26	-0.06
3	0.06	0.06	0.00
4	0.86	0.80	-0.06
5	0.62	0.58	-0.04
6	0.54	0.50	-0.04
7	0.50	0.32	-0.18
8	0.28	0.24	-0.04
9	0.36	0.18	-0.18
10	0.68	0.64	-0.04
			Mean difference -0.07
			Standard error 0.02

Table 8.3. Cover Values Taken Along 10 Permanent Transects of 50 Points Each in 1990 and 1994.

degree of variability between transects, you would have needed unreasonably large numbers of transects to detect the change that only 10 permanent transects were able to detect.

This simple comparison suggests that permanent sampling units would always be advantageous, but their value must be balanced against their disadvantages. Time and equipment costs associated with permanent sampling units are higher than temporary ones. Sampling units must be marked well with permanent markers. These can be costly and time-consuming to install during the first year and difficult to find on subsequent years. Permanent markers may not be feasible in some situations because of the nature of the habitat or for safety reasons (see Chapter 5).

Another disadvantage of a design using permanent sampling units is that you usually need 2 years of data to determine adequate sample size. The only exception to this is when you have some basis to estimate the degree of correlation (the correlation coefficient) of sampling units between years when estimating means (e.g., density sampling) or a model of how the population is likely to change when estimating proportions (e.g., frequency sampling). We will discuss this at more length in the next section.

Impacts either from investigators or from animals may bias your results. By going back to the same sampling unit locations each year, you might negatively impact the habitat in or near the permanent sampling units. In addition, permanent markers may also attract wildlife, domestic livestock, wild horses, or burros. This might lead to differential impacts to the vegetation in or near the sampling units. If markers are too high (e.g., t-posts or other fence posts), livestock may use the markers for scratching posts and differently impact the sampling units. Wildlife impacts may also occur. Raptors, for example, might use the markers as perches; this could result in fewer herbivores in the sampling units than elsewhere in the target population, with resulting differences in the attribute being measured. Songbirds also might use the perches, defecating seeds and changing the plant community.

The advantage of permanent sampling units varies depending on degree of correlation between two measurements. Permanent sampling units will be the most advantageous when there is a high degree of correlation between sampling-unit values between two periods. This condition often occurs with long-lived plants (e.g., trees, shrubs, large cacti, or other long-lived perennial plants and long-lived and relatively sedentary animals). If, however, there is low correlation between sampling units between two periods, then the advantage of permanent quadrats is diminished. This could occur, for example, with annual plants, if their occurrence in quadrats one year does not greatly depend on their occurrence in the previous year. Small mammals and many insects with highly mobile populations provide another example. Even for these species, however, permanent quadrats may still outperform temporary quadrats if recruitment most often takes place near parents.

Permanent Sampling Units to Estimate Density

Let us examine two very different situations involving permanent density quadrats. Figure 8.16 compares sample sizes needed to detect different levels of change in density in a clumped population of 4000 plants using permanent and temporary quadrats. All sampling was done with

0.25m × 50m quadrats. In this example, there was no recruitment of new plants; all change between year 1 and year 2 was the result of plant mortality. This created a strong correlation between quadrat counts between the two periods for the low mortality changes. The x-axis shows the percent change in mean plant density (equivalent to percent mortality in this example). The y-axis shows the number of quadrats that needed to be sampled to detect the true population change with false-change and missed-change error rates both set at 0.10. When the change in mean plant density between the first and second sampling periods was less than 50%, permanent quadrats were much more effective than temporary quadrats at tracking the change. For example, for detecting a 5% change, 22 permanent quadrats performed as well as 338 temporary ones!

The advantage of permanent quadrats occurs when counts between two periods correlate with one another. This is true in the situation depicted in Figure 8.16 because no new plants show up in new locations. The opposite extreme, illustrated by Figure 8.17, shows population changes caused by 100% mortality of the original population combined with various levels of recruitment from plants in completely new positions. Permanent quadrats no longer provide any advantage over temporary ones, and the disadvantages of permanent quadrats would lead you to a temporary quadrat design.

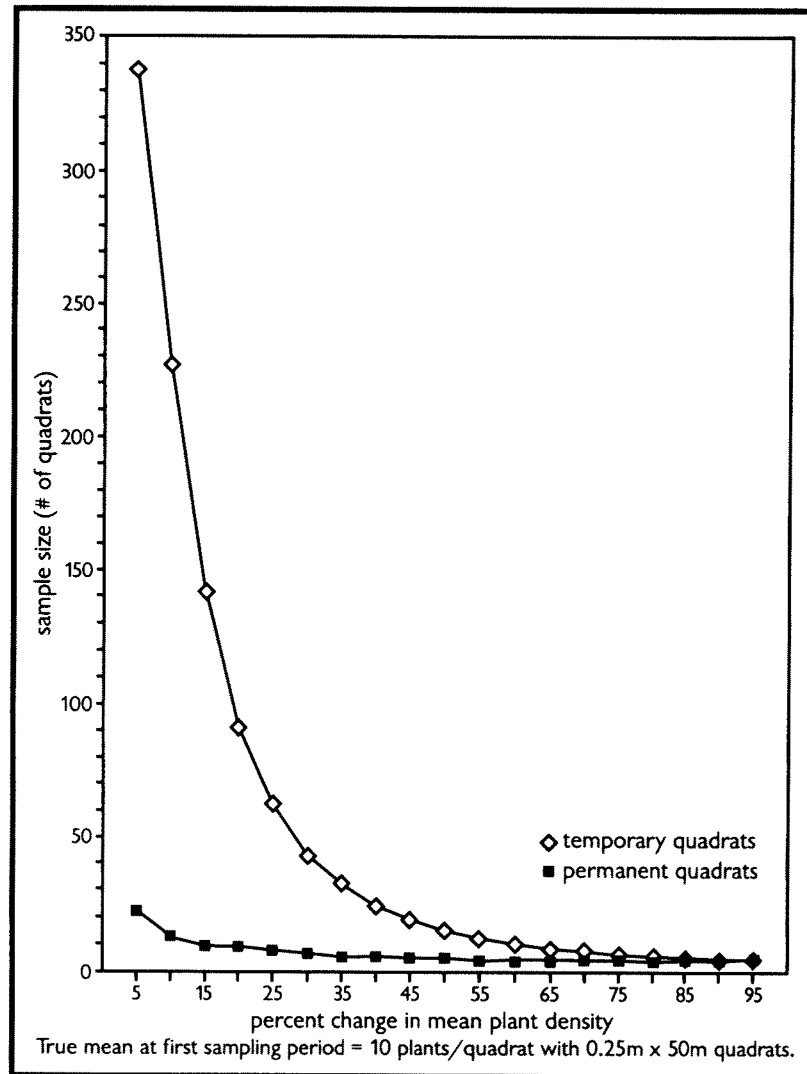


Figure 8.16. Sample sizes needed to detect different degrees of population decline from an artificial clumped population of 4,000 plants using temporary vs. permanent quadrats. All changes are due to mortality of the original population without any recruitment of new plants. Note the much better performance of permanent quadrats in detecting changes below 50%.

Most populations will show a combination of mortality and recruitment, as opposed to the extreme situations shown in Figures 8.16 and 8.17. For most species, permanent quadrats will provide greater precision with the same number of quadrats or equivalent precision at smaller sample sizes, because the locations of new individuals will likely be correlated with the location of old individuals given typical patterns of reproduction. You must balance the magnitude of this increase in precision (or reduction in sample size) against the disadvantages of using permanent sampling units.

Permanent Frequency Quadrats and Points

The discussion so far has centered on the use of paired quadrats for estimating density. This type of sampling is analyzed by means of a paired *t*-test (this will be covered in Chapter 9). The paired *t*-test would also be used to analyze changes in paired quadrats used to estimate cover and to analyze changes in permanent transects such as those used for line intercept sampling or for

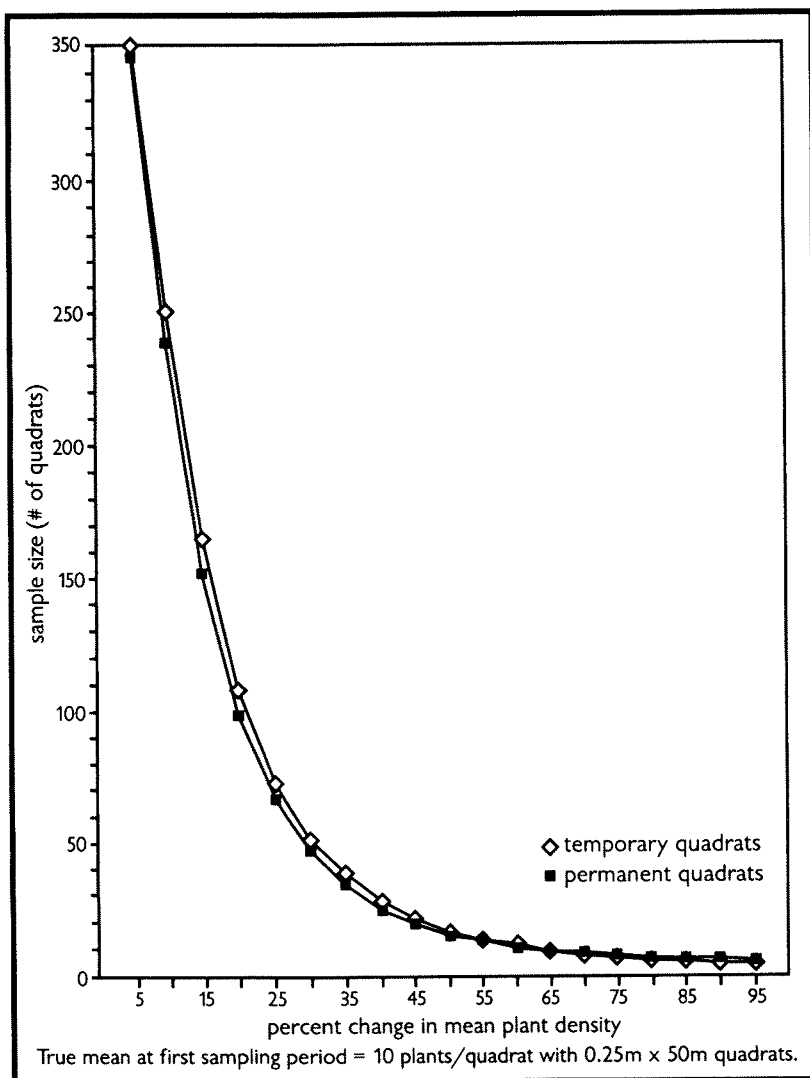


Figure 8.17. Sample sizes needed to detect different degrees of population decline from an artificial clumped population of 4,000 plants using temporary vs. permanent quadrats. All changes result from 100% mortality of the original population with various levels of random recruitment. Temporary and permanent quadrats perform about the same in this situation.

point or quadrat sampling in systematic sampling designs (when the transects, as opposed to the quadrats or points, are treated as the sampling units).

When frequency quadrats or points are treated as the sampling units, a different set of tests is used to determine if a statistically significant change has taken place. The chi-square test is used when these types of sampling units are temporary (i.e., randomly located in each year of measurement), while McNemar's test is used when the quadrats or points are permanently located in the first year of measurement. These tests are discussed in Chapter 9, but it is important here to point out that—just as for permanent designs that use transects or quadrats for estimating density or cover—it is sometimes much more efficient to make use of permanent frequency quadrats or points.

Salzer (unpublished data) concludes that under certain population-change scenarios, permanent frequency quadrats offer large reductions in sample size over those required for temporary quadrats. In the most extreme example, 87 permanent quadrats perform as well as 652 temporary quadrats in detecting a 5.5% decline in frequency (with the false-change and missed-change error rates both set at 0.10). In other situations, little difference exists between permanent quadrat designs and temporary quadrat designs.

The sample-size differences between temporary and permanent frequency designs depend on the particular nature of population changes. For this reason, the determination of whether to use permanent or temporary frequency quadrats must be evaluated on a case-by-case basis, taking into account the life history of the species, the sample size advantages of using the permanent design, and the disadvantages associated with designs using permanent quadrats.

Appendix II contains more information on the use of permanent frequency designs and should help you decide when to use one.

HOW MANY SAMPLING UNITS SHOULD BE SAMPLED?

An adequate sample is vital to the success of any successful monitoring effort. Adequacy relates to the ability of the observer to evaluate whether the management objective has been achieved. It makes little sense, for example, to set a management objective of increasing the density of a

rare plant species by 20% when the monitoring design and sample size will not likely detect changes in density of less than 50%.

Deciding on the number of sampling units to sample (which we refer to as "sample size") should be based on the following considerations:

1. Sample size should be driven by specific objectives. If you are targeting point-in-time estimates (parameter estimation), you need to specify how precise you want your estimates to be. If you are trying to detect changes in some average value, you need to specify the magnitude of the change you wish to detect and the acceptable false-change and missed-change error rates (refer to Chapters 7 and 14 for further guidance).
2. Sample size should be based on the amount of variability in actual measurements. You should assess this variability during pilot sampling. Once you have tried various sampling-unit sizes and shapes and have decided on a particular one, start randomly positioning the sampling units in the population. After you have sampled some initial number of sampling units, stop and do some simple number-crunching with a hand calculator and evaluate the variation in the data. This is called sequential sampling and is discussed below. You can enter standard deviations into sample-size equations or computer programs, and the output will inform you whether you have sampled enough. If you have not sampled enough, sample size equations (or a computer program) will calculate the number of sampling units you need to sample to meet your objective.
3. Sample-size formulas and computer programs make two assumptions. The first is that sampling units are positioned in some random manner. This is discussed above. The formulas and programs also assume a distribution of sample means (a sampling distribution) from your population fits approximately a normal distribution. If your population is highly skewed, this latter assumption will not be true for small sample sizes. We discuss this issue in more detail in Chapter 9.
4. Sample sizes required differ between infinite versus finite populations. We introduced this concept in Chapter 7. Most computer programs and standard, sample-size equations assume that the population you are sampling from is infinite. This will always be the case if you are estimating cover using either points or lines, because these are considered dimensionless. If, however, you are sampling a relatively small area, and you are making density, frequency, cover, or biomass assessments in quadrats, then you should account for the fact that you are sampling from a finite population. This means there is some finite number of quadrats that can be placed in the area to be sampled. The sample-size formulas provided in Appendix II include a correction factor called the Finite Population Correction (FPC). If you are sampling more than 5% of a population, applying the FPC "rewards" you by reducing the necessary sample size. Appendix II describes how to apply the FPC to sample-size determination.[7]
5. Precision increases with sample size but not proportionately. This is illustrated in Figure 8.18 in an example where the statistical benefits of increasing sample size diminish once you reach about $n = 30$. You should seek to increase statistical precision and power not by simply increasing sample size, but by reducing the standard deviation to as small a value as possible through good design.
6. Problems may occur in the use of many published sample-size formulas. Most formulas that are designed to determine sample sizes for "point-in-time" estimates

---

[7] *Chapter 9 describes how to apply the finite correction factor to results of significance tests. Appendix III describes how to apply the finite correction factor to analysis of confidence intervals.*

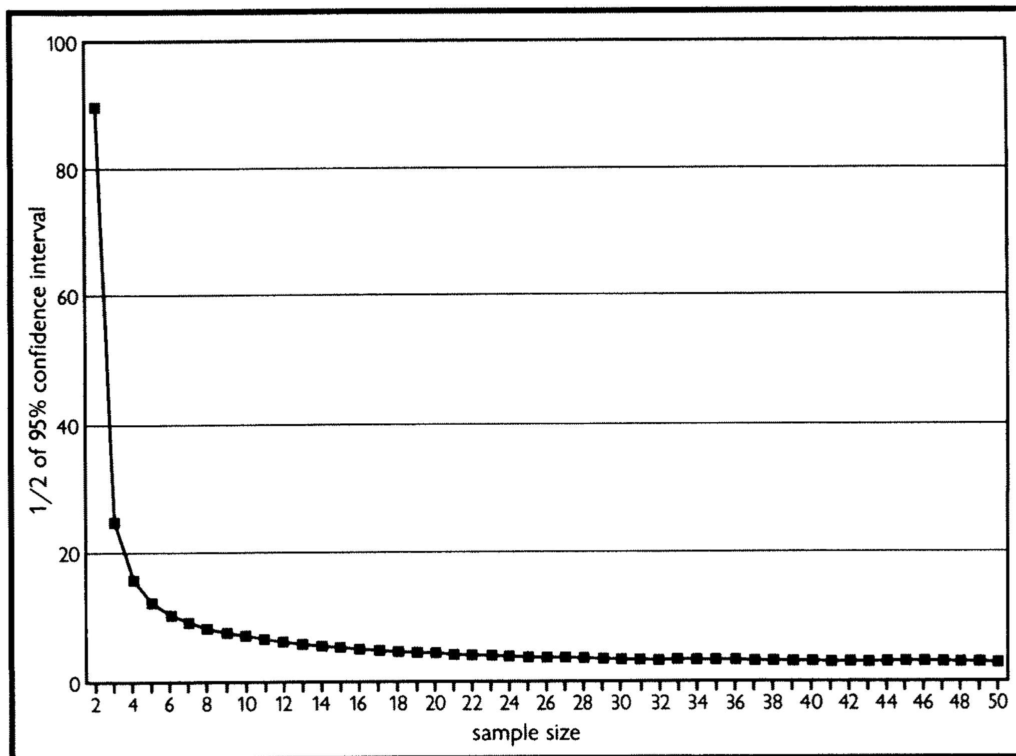


Figure 8.18. Influence of sample size on level of precision. Sample sizes necessary to achieve different levels of precision at a constant standard deviation of 10. Note that there is no effective improvement in precision after about n = 30.

(parameter estimation) with specified levels of precision do not account for the random nature of sample variances. They do not include a "level of assurance" (also known as a tolerance probability) that you will actually achieve the conditions specified in the sample-size equations and obtain a confidence interval of a specified width. Blackwood (1991) provides a layperson's discussion of this topic and reports the results of a simulation that illustrates the concept. Kupper and Hafner (1989) provide a correction table to use with standard, sample-size equations for estimates of single population means or population totals. A modified version of this table and instructions on how to use it are included in Appendix II.

Information Required for Calculating Sample Size

Appendix II gives equations for calculating sample sizes for the following sampling objectives:

1. Estimating means and totals
2. Detecting differences between two means when using temporary sampling units
3. Detecting differences between two means when using permanent sampling units
4. Estimating a proportion
5. Detecting differences between two proportions when using temporary sampling units
6. Detecting differences between two proportions when using permanent sampling units

Equations for calculating sample size for cluster samples, two-stage samples, and stratified random samples are given in Appendix IV. Computer programs are also available that implement

these equations to calculate sample size.[8] Both the formulas and programs require you to insert some of the following values from your objectives (see Chapter 14) and your pilot sampling:

- Estimating means and totals. You must specify the precision desired (confidence interval width), the confidence level, and an estimate of the standard deviation among sampling units.
- Detecting differences between two means when using temporary sampling units. You must specify the false-change error rate, the power of the test, the magnitude of the smallest change you wish to detect, and an estimate of the standard deviation (the standard deviation among sampling units is usually assumed to be the same for both periods).
- Detecting differences between two means when using permanent sampling units. You must specify the false-change error rate, the power of the test, the magnitude of the smallest change you wish to detect, and an estimate of the standard deviation (this is the standard deviation of the differences between the paired sampling units, not the standard deviation of the population being sampled in the first year).
- Estimating a proportion. You must specify the precision desired (confidence interval width), the confidence level, and a preliminary estimate of the proportion to be estimated (if you do not have any idea of what proportion is to be expected, you can conservatively estimate the sample size by assuming the proportion to be 0.50).
- Detecting differences between two proportions when using temporary sampling units. You must specify the false-change error rate, the power of the test, the magnitude of the smallest change you wish to detect, and a preliminary estimate of the proportion in the first year of measurement (using a value of 0.50 will conservatively estimate the sample size).
- Detecting differences between two proportions when using permanent sampling units. You must specify the false-change error rate, the power of the test, the magnitude of the smallest change you wish to detect, and an estimate of the sampling-unit transitions that took place between the 2 years. This last estimate is specific only to this design and is discussed separately in Appendix II.

Your management and sampling objectives already include most of the information required to calculate sample size using either the equations of Appendix II or the computer programs. What is missing is 1) an estimate of the standard deviation, for those situations where you wish to estimate a mean value or detect change between two mean values; and 2) a preliminary estimate of the population proportion, when estimating a proportion or detecting change between two proportions using temporary sampling units. For proportions you have the flexibility of simply entering 0.50 as your preliminary estimate of the population proportion (this provides a conservative estimate of sample size). Alternatively, you can use an estimate derived from pilot sampling. When dealing with mean values, however, you must have an estimate of the standard deviation. This is the subject of the next section.

Sequential Sampling to Obtain a Stable Estimate of the Mean and Standard Deviation

In several places in this chapter we have stressed the need for pilot sampling. The principal purposes of pilot sampling are to assess the efficiency of a particular sampling design and, once a particular design has been chosen, to generate the values needed for calculating the sample size

---

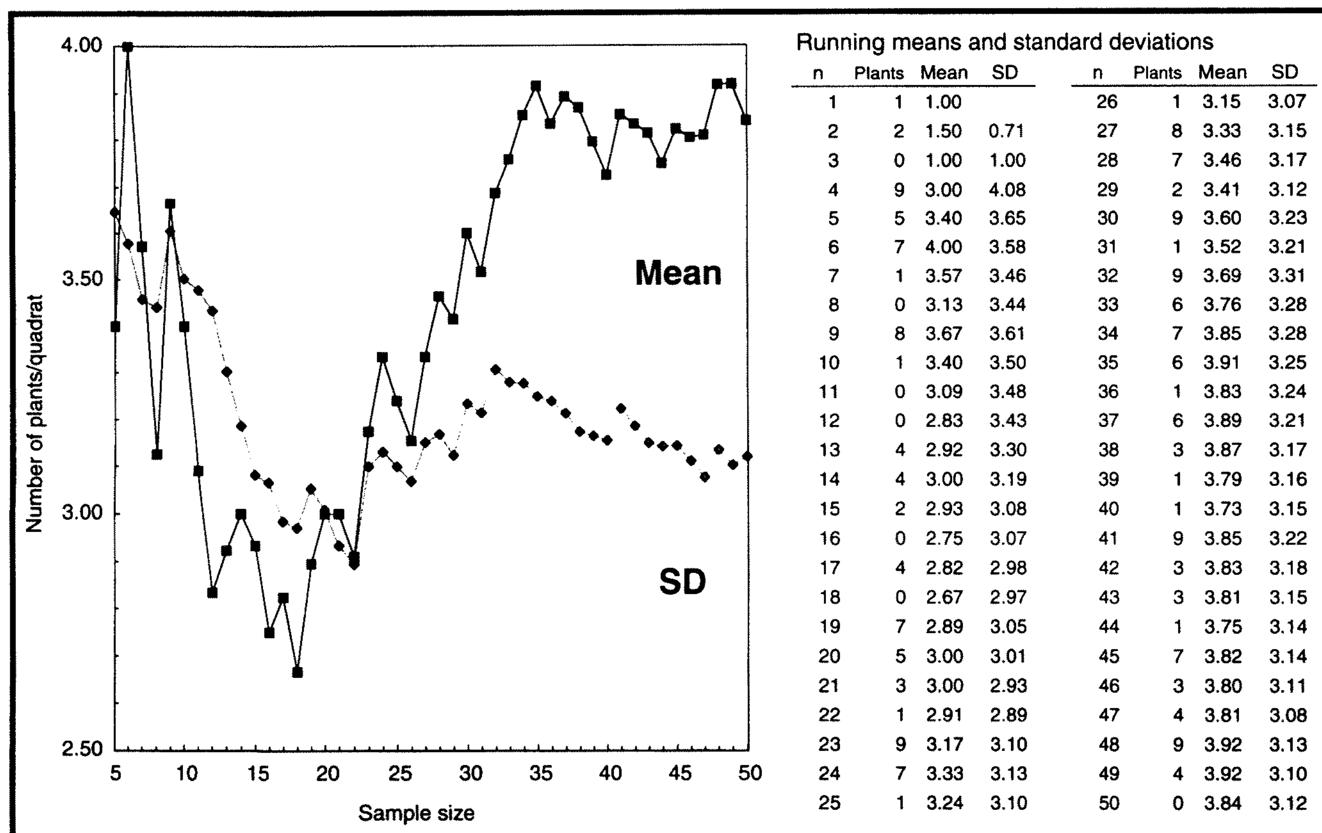
[8] *Surprisingly, many of the general statistical programs, despite their expense, do not include routines for calculating sample size. Thomas and Krebs (1997) reviewed 29 computer programs for calculating sample size; a link to an online copy can be found on our web page. Several freeware or shareware programs are available. Links to these can also be found on our web page (see Preface).*

required to meet the sampling objective. Pilot sampling enables us to obtain stable estimates of the population mean and the population standard deviation and to calculate the **coefficient of variation.** The estimate of the standard deviation derived through pilot sampling is one of the values we use to calculate sample size, whether we use the formulas of Appendix II or a related computer program. Sequential sampling is the process we use to determine whether we have taken a large enough pilot sample to properly evaluate different sampling designs and/or to use the standard deviation from the pilot sample to calculate sample size.

The coefficient of variation (CV) is calculated by dividing the sample standard deviation by the sample mean. The coefficient of variation is useful because, as a measure of variability, it does not depend upon the magnitude and units of measurements of the data. This allows direct comparison of CV's from different studies and of different sampling designs. It also enables us to derive estimates of sample size when we do not have data from pilot studies but do have an idea of the magnitude of CV from similar studies and sites.

We begin by gathering pilot sampling data using some selected sample size. The selection of this initial sample size will depend upon the relative amount of variation in the data—if many of the sampling units yield numbers similar to one another, then you may want to perform the first sequential sampling procedure after n = 8 or 10. If you see high variation among the sampling units, then you may want to start with a larger number (e.g., $n \geq 15$) or, perhaps preferably, consider altering the size and/or shape of your sampling unit before doing the first iteration of the sequential sampling procedure.

Calculate the mean and standard deviation for the first two sampling units, calculate it again after putting in the next sampling unit, and then repeat this procedure for all of the sampling units sampled so far. This will generate a running mean and standard deviation. Look at the four columns of numbers on the right of Figure 8.19 for an example of how to carry out this proce-



Running means and standard deviations

n	Plants	Mean	SD	n	Plants	Mean	SD
1	1	1.00		26	1	3.15	3.07
2	2	1.50	0.71	27	8	3.33	3.15
3	0	1.00	1.00	28	7	3.46	3.17
4	9	3.00	4.08	29	2	3.41	3.12
5	5	3.40	3.65	30	9	3.60	3.23
6	7	4.00	3.58	31	1	3.52	3.21
7	1	3.57	3.46	32	9	3.69	3.31
8	0	3.13	3.44	33	6	3.76	3.28
9	8	3.67	3.61	34	7	3.85	3.28
10	1	3.40	3.50	35	6	3.91	3.25
11	0	3.09	3.48	36	1	3.83	3.24
12	0	2.83	3.43	37	6	3.89	3.21
13	4	2.92	3.30	38	3	3.87	3.17
14	4	3.00	3.19	39	1	3.79	3.16
15	2	2.93	3.08	40	1	3.73	3.15
16	0	2.75	3.07	41	9	3.85	3.22
17	4	2.82	2.98	42	3	3.83	3.18
18	0	2.67	2.97	43	3	3.81	3.15
19	7	2.89	3.05	44	1	3.75	3.14
20	5	3.00	3.01	45	7	3.82	3.14
21	3	3.00	2.93	46	3	3.80	3.11
22	1	2.91	2.89	47	4	3.81	3.08
23	9	3.17	3.10	48	9	3.92	3.13
24	7	3.33	3.13	49	4	3.92	3.10
25	1	3.24	3.10	50	0	3.84	3.12

Figure 8.19. A sequential sampling graph. Running means and standard deviations are plotted for increasing sample sizes. Note how the curves smooth out after n = 35.

dure. Most hand calculators enable you to add additional values after you have calculated the mean and standard deviation, so you do not have to input the previous values again.

Use graph paper (or preferably a computer spreadsheet program) to plot the mean and standard deviation against sample size (Fig. 8.19). We suggest starting your graph at $n = 3$ or $n = 5$ for reasons that will become clear later. You are looking for a smoothing of the graph, suggesting the mean and standard deviation have stabilized.

A laptop or field computer is extremely valuable for creating sequential sampling graphs during a pilot sampling program. Spreadsheet programs allow you to enter your data in a form that can later be analyzed (saving time on later data entry) and at the same time create a sequential sampling graph of the running mean and standard deviation. You can also reorder the data (as though you had measured the sampling units in a different order) and replot the sequential sampling graph.

Now, let us apply these concepts. Examine Figure 8.20. The graph shows two sampling sequences of the same population using the same sampling units. The difference between them is that, simply by chance, in the first sequence several plots with large numbers of plants were the first to be sampled, while in the second sequence the first plots to be sampled had only a few plants (or none). Where would you stop sampling in either of these sequences (consider the curve "smoothed")? One strategy would be to reorder the sampling units and evaluate alternative sequential sampling graphs. A better strategy would be to re-evaluate the sampling design. Look again at Figure 7.2 in Chapter 7. Do you think the 0.4m × 10m quadrat is an efficient sampling-unit design? If, after sampling 20% to 30% of the possible sampling units, your sequential sampling graph has not stabilized, you should definitely reconsider your design. Figure 8.21 shows a good sampling design with the curve smoothing at about $n = 12$. The samplers could have saved a substantial amount of effort by stopping long before they did.

Figure 8.22 illustrates the problems that may arise from plotting your graph beginning with the first data point. Here, a large initial value of six plants and the scale of the y-axis in Figure 8.22A give an illusion of smoothing. Figure 8.22B shows a graph of the same data reordered, with the first quadrat containing only two plants. Even more important, in this example the sampler should have recognized that a problem existed in the sampling design by the time they had sampled $n = 20$ (or earlier). A sequential sampling graph showing a repeated pattern of spikes followed by gradual declines is indicative of a poor design. The spikes are quadrats containing many of the species. The gradual declines are caused by encountering several quadrats sequentially with zero observations in them. In practice, if more than a quarter of your initial 10 quadrats contain none of the species, stop immediately and reevaluate your sampling-unit size and shape. Another hint that a problem exists in this sampling design is that the standard deviation is consistently greater than the mean. Compare this with the good design illustrated in Figure 8.20 and the acceptable design of Figure 8.19.

Once the mean and standard deviation have apparently stabilized, use those values in the appropriate sample-size equation or computer program to generate the actual sample size required. We do not recommend using sequential sampling graphs alone to determine sample size.

Alternatives to Sequential Sampling to Obtain an Estimate of the Standard Deviation

Pilot sampling, using the sequential sampling procedure described above, is by far the best means of deriving an estimate of the standard deviation to enter into a sample-size equation or computer program. Two less effective methods will be briefly discussed.

The first method is to use data from similar studies to estimate the standard deviation. Although not as reliable as a pilot study, you may have conducted a study using the same study design and measuring the same attribute in the same vegetation type. The standard deviation of the sample from this study can be used as an estimate of the standard deviation of the population that is the focus of the current study.

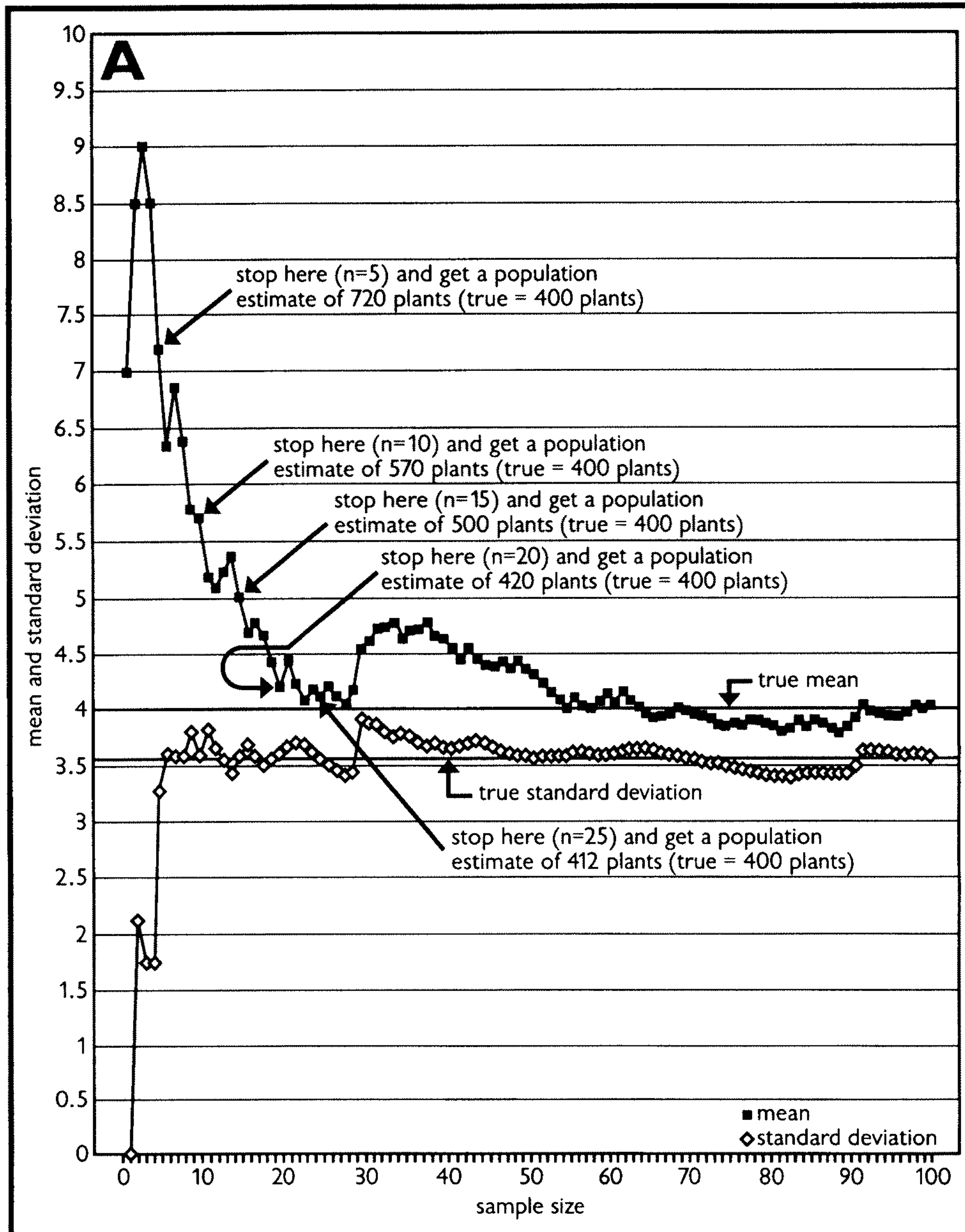


Figure 8.20. Sequential sampling graphs of the 20m × 20m "400 plant population" introduced in Chapter 7. The population was sampled using a 0.4m × 10m quadrat. The entire population consists of 100 quadrats. Notice how far estimates are from the true mean if they are made before the curves smoothing out. In Part A many of the quadrats sampled at the beginning had large values. Note how we would have overestimated the population if we had stopped too soon.

The second method relies on professional judgment. As pointed out by Krebs (1998), an experienced person may have some knowledge of the amount of variability in a particular attribute. Using this information you can determine a range of measurements to be expected (maximum value – minimum value) and can use this range to estimate the standard deviation of a measure. Table 8.4, adapted from the table in Dixon and Massey (1983), gives the appropriate conversion factor to be multiplied by the range value to come up with an estimate of the population standard deviation.

Figure 8.20B. This figure is the same sample as Part A, but with the data randomly reordered. If we'd used the initial values shown in this graph (prior to the curves leveling off), we would have seriously underestimated the true mean value, as opposed to overestimating as in Part A.

To illustrate how to use this table, let us assume we know from working with a particular species that in a sample of size 30 we could expect a range of 0 individuals per quadrat to 100 individuals per quadrat (this process assumes a normal distribution so we should not have too many quadrats with zeros in them). The range in this case is 100 – 0 = 100 individuals. The conversion factor for a sample of size 30 is 0.245. Our estimate of the population standard deviation is, therefore, 100 individuals × 0.245 or 24.5 individuals per quadrat.

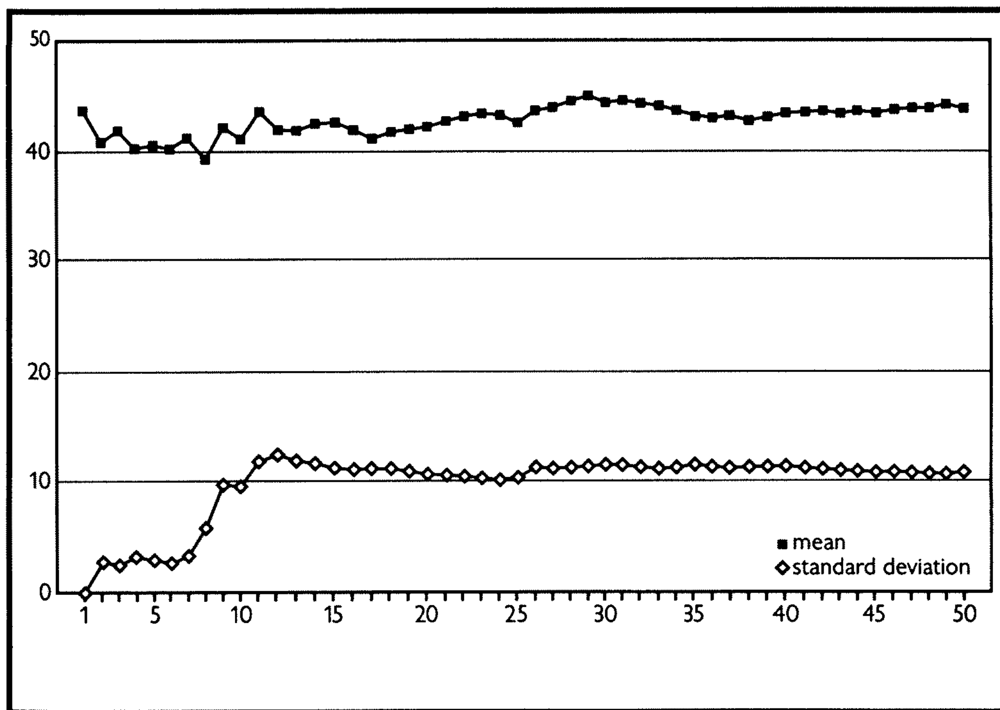


Figure 8.21. Sequential sampling graph of vegetation height measurements. Note how the graphs have flattened out long before the sampling ended.

Although this method can be used, it should be emphasized again that data from a pilot study are more reliable and are preferable to this method.

Estimating the Standard Deviation When Using Permanent Sampling Units

Estimating the standard deviation for a design that uses permanent sampling units is difficult because it is the standard deviation of the difference between the sampling units between the two years that must be entered into the sample size equation or computer program, and this is a value that you will not have until you have collected data in the second year. Thus, your pilot study must span 2 years before you can accurately estimate the sample size required to meet your sampling objective. You would like, however, to make a reasonable estimate from the first year's data of the standard deviation of the difference. This will give you a good chance of having used a large enough sample size the first year, with the result that you will not have to add more sampling units the second year and will be able to use the first year's data in your analysis. Following are some methods you can use for this purpose.

You can estimate the standard deviation using the alternative methods discussed in the section above. Remember, however, that it is the standard deviation of the difference that must be estimated, so if you use data from previous studies, they must be studies that used permanent sampling units. If you use the expected range to estimate the standard deviation, it must be the range of the differences, not the range of the data for any one year.

There is another way you can calculate the necessary sample size by having only the first year's pilot data. This method requires that you have some knowledge of the degree of correlation (correlation coefficient) expected between the permanent sampling units between years. Appendix II provides a formula by which you can estimate the standard deviation of the differ-

Figure 8.22. Sequential sampling graphs for *Astragalus applegatei* at the Euwana Flat Preserve. Part A shows what can happen when the y-axis is set at too large a range, because of initial large values. This can make it appear that the running mean and standard deviation has smoothed out when in fact they haven't. Part B illustrates the real situation: neither statistic has smoothed out even by n = 100. This is a poor sampling design. See text for further elaboration.

Sample Size	Conversion Factor	Sample Size	Conversion Factor
2	0.886	19	0.271
3	0.591	20	0.268
4	0.486	25	0.254
5	0.430	30	0.245
6	0.395	40	0.231
7	0.370	50	0.222
8	0.351	60	0.216
9	0.337	70	0.210
10	0.325	80	0.206
11	0.315	90	0.202
12	0.307	100	0.199
13	0.300	150	0.189
14	0.294	200	0.182
15	0.288	300	0.174
16	0.283	500	0.165
17	0.279	1000	0.154
18	0.275		

Table 8.4. Conversion Factors Used to Estimate the Population Standard Deviation. To estimate the standard deviation of a variable from knowledge of the range for samples of various sizes, multiply the observed range (maximum – minimum value) by the table values to obtain an unbiased estimate of the standard deviation. This procedure assumes a normal distribution. From Dixon and Massey (1983) and reproduced in Krebs (1998).

ence between years by using the standard deviation of the first year's sample and the correlation coefficient. This is something you might have from similar studies on the same species (although in that case you would probably already have an estimate of the standard deviation of the difference between years that you could use). Based on your knowledge of the life history of the species you are dealing with, you might make an initial estimate of correlation. For example, if you are monitoring a long-lived perennial and do not anticipate a lot of seedling recruitment (or if you expect seedling recruitment to be very close to parent plants), you might estimate that the correlation coefficient between years is relatively high, say about 0.80 or 0.90. You then plug this coefficient into the formula, along with your estimate of the standard deviation of the first year's data.

Whichever method you use to estimate the standard deviation of the difference, once you have collected the second year's data, you will still need to enter the actual observed standard deviation of the difference into an equation or a computer program to calculate actual sample size. You can then modify your initial estimate of sample size accordingly.

MANAGEMENT IMPLICATIONS

Good sampling design can dramatically increase the precision of the estimates of population characteristics while reducing field costs. While good design may be time-consuming at the planning stage of a monitoring study, the investment pays well throughout the life of the monitoring and in the application of data to management decisions. Six design features must be addressed when planning a monitoring study using sampling:

1. What is the population of interest?
2. What is an appropriate sampling unit?
3. What is an appropriate sampling-unit size and shape?
4. How should sampling units be positioned?
5. Should sampling units be permanent or temporary?
6. How many sampling units should be sampled?

CHAPTER 9
Statistical Analysis

Arctomecon californica
Golden Bear Poppy
Near Lake Mead in Arizona and southwest Nevada
Artist: Jeanne R. Janish

With two exceptions, quantitative data collected through monitoring must be subjected to some type of statistical analysis. The two exceptions involve the following two types of data: 1) data gathered from a complete census, and 2) data gathered by sampling techniques that do not incorporate some type of random selection process (see Chapter 8). A census provides you with complete information about the target population. The means, totals, or proportions resulting from a complete census are the actual population values (assuming no measurement error such as errors in counting or in identifying plants). If there is no sampling error, no statistical analysis is necessary. Any changes in these population values between years are real. All that remains is to determine whether the changes have any biological significance.

While census data are highly valuable and applicable, quantitative data gathered without using some type of random sampling procedure may be useless. The fact that statistics cannot be applied to nonrandom sampling procedures makes proper analysis and interpretation of the data virtually impossible; this should reinforce the need to use a random sampling procedure in designing and implementing monitoring.

Statistics are extremely important to sample-based monitoring. They enable us to make management decisions even when we have access to only part of the information. For example, you might like to know the true number of individuals in a given area. Because the area is large, however, and the individuals far too numerous to count, the best you can do is take a random sample of quadrats within this area and estimate the total number of individuals from this sample. The use of statistics enables you to derive an unbiased estimate of this total and, more important, assess how good this estimate is.

No doubt, you will use calculators with statistical functions or computer software programs to analyze your data. For that reason, this chapter emphasizes principles and concepts and contains a minimum of mathematical formulas.

USING GRAPHS TO EXPLORE THE NATURE OF YOUR DATA

Several types of graphs can be used to examine your data before analysis: normal probability plots, density plots, box plots, and combinations of these. These are particularly important in the initial stages of designing your study. Graphs of pilot study data can, for example, help show whether you are using an efficient sampling-unit size and shape, or whether your data meet the assumptions of **parametric statistics.** (Parametric statistics are those statistics used to estimate **population parameters** such as means and totals; we discuss below the assumptions you must make when using them).

A statistic is an estimate of a population parameter derived from a sample.

A parameter is a quantity that describes or characterizes a population. Examples of parameters are the population mean, population variance, population standard deviation, and population coefficient of variation.

Graphing your monitoring data is likely to reveal patterns in your data that will not be apparent if all you do is calculate standard summary statistics like the mean and standard deviation. Figure 9.1 shows four samples, each of which has a mean of 100 and a standard deviation of 10. Without graphing the individual data points, we would probably assume that these four samples had the same or very similar distributions. The graph in Figure 9.1 shows how wrong we would be.

An excellent and concise discussion of using graphs for exploratory data analysis can be found in Ellison (1993). Following are some of the most valuable of these graphs, along with examples of each.

Normal Probability Plot

A normal probability plot is a good way to inspect your data to determine if they approximate a **normal distribution.** Most statistical packages produce these plots. The observed values are plotted against the values that would be expected if the data came from a normal distribution. If the

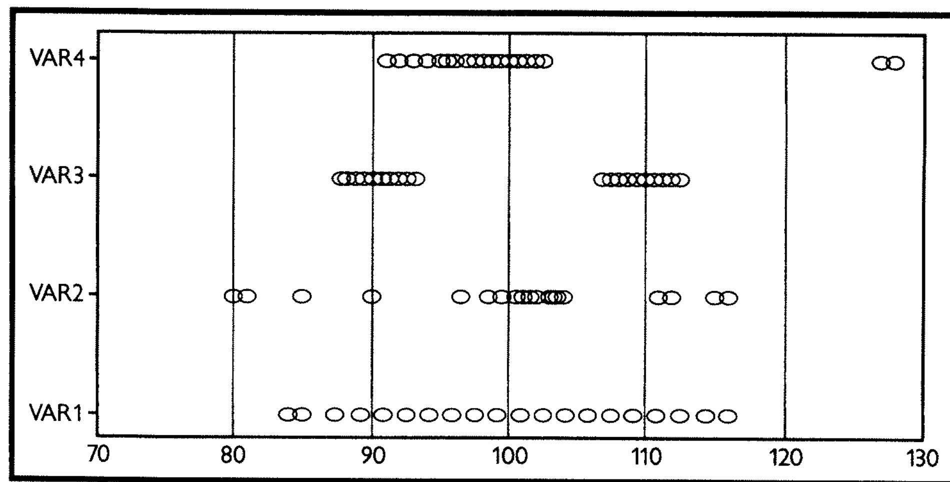


Figure 9.1. Four different samples of size 20, each of which has a mean of 100 and a standard deviation of 10. For these samples, the two summary statistics alone (mean and standard deviation) are insufficient to fully characterize the population. The differences in data distributions become apparent only after the individual data points are plotted.

data come from a normal distribution, the plotted values fall along a straight line extending from the lower left corner toward the upper right corner (Fig. 9.2A).

When a normal probability plot forms a pattern like that shown in Figure 9.2B, we know they do not conform to a normal distribution. This sample of plant heights contains many very small values (short plants) and a few large values (tall plants), a distribution that is common in biology. If you were to take the logarithms of the data, the resulting values would more closely approximate a normal distribution. For this reason, the distribution is called a lognormal distribution. This plot has alerted us that we need to be careful when applying parametric statistics to this data set (see below).

A normal distribution (also called the Gaussian distribution) is a family of distributions that form familiar bell-shaped curves. They are symmetric, with observations more concentrated in the middle than in the tails. The shape of the curve varies based on the underlying population characteristics, but is described by two parameters: the population mean and the population standard deviation.

Density Plots

A histogram is a type of density plot. Each bar in a histogram illustrates the density of data values found between the lower and upper bounds of the bar. Figure 9.3, A and B, illustrates examples of histograms. Histograms are familiar and interpretable by a wide audience and are commonly used. Three disadvantages of histograms limit their use for exploratory data analysis (Ellison 1993):

1. The raw data are hidden within each bar. Consider the histogram of cover data presented in Figure 9.3A. Each of the 10 bars (the second and tenth bars have no values in them) contains cover values within a range of 0.1. The third bar contains 11 values between 0.2 and 0.3, but we do not know if this represents 10 values of 0.21, 10 values of 0.29, or any of the other possible combinations of values between 0.2 and 0.3.
2. The number and width of bars is arbitrary. Changing these alters the shape of the histogram and adds some additional information, but the complaint in #1 remains. Figure 9.3B is a histogram of the same data, but with 20 bars instead of 10.
3. Summary statistics (for example, means and medians) cannot be computed from the data illustrated in a histogram.

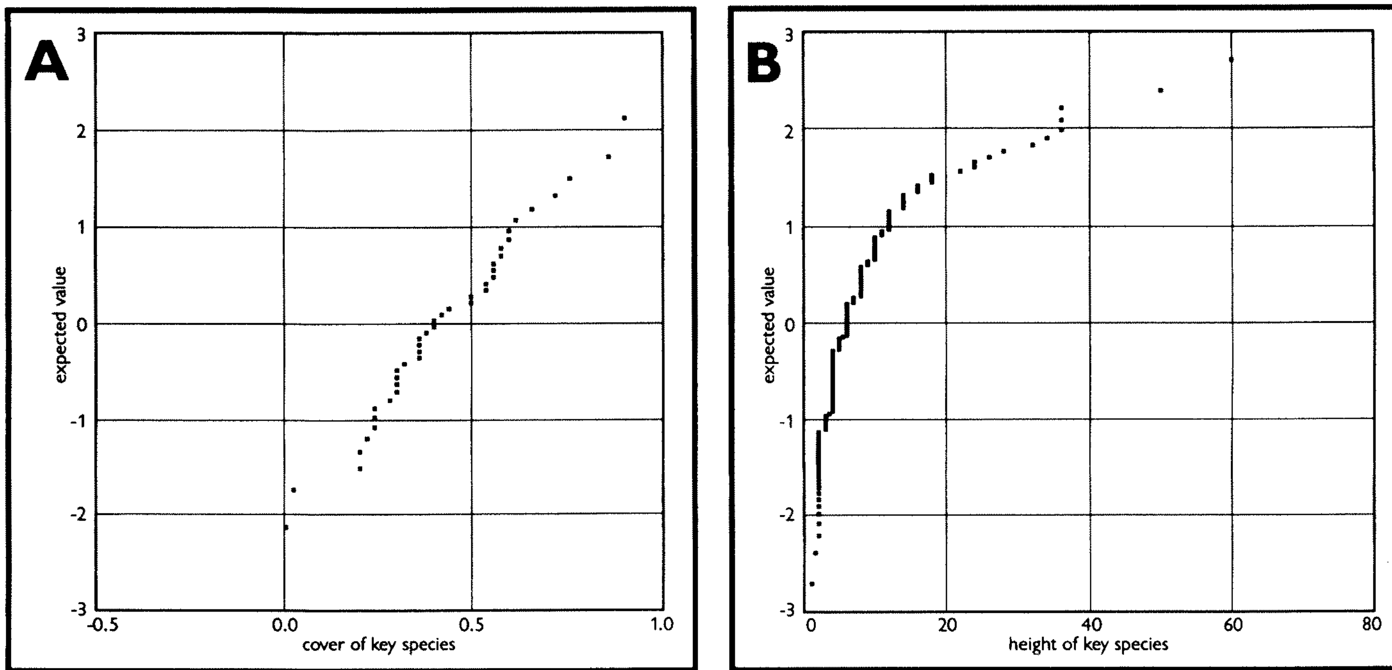


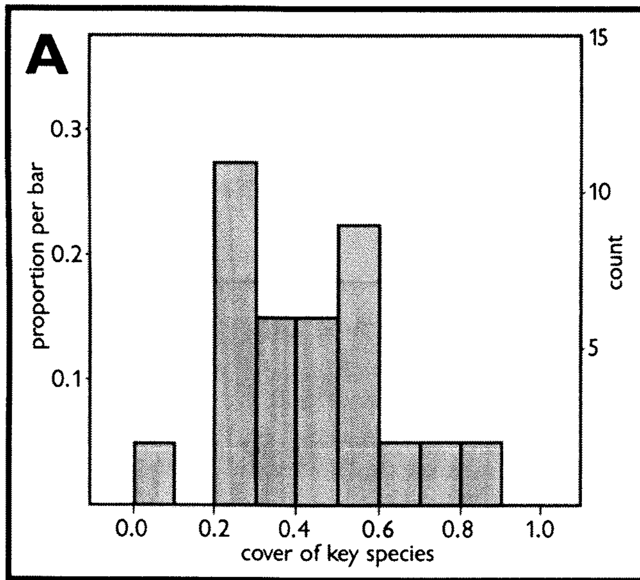
Figure 9.2. Part A is a normal probability plot of cover data from a sample of 40 randomly placed transects, each of which had 50 point-intercept cover measurements (the transects are the sampling units). These data approximate a normal distribution. Part B is a normal probability plot of plant heights. These data are not from a normal distribution.

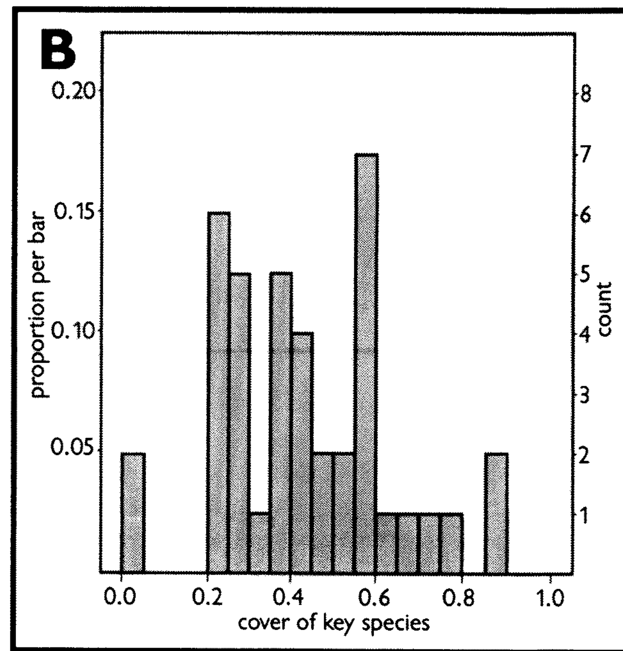
A dit plot, as illustrated in Figure 9.3C, is a better type of density plot, because all the data points are presented, the underlying data structure is maintained, and the graph is easy to understand (Ellison 1993).

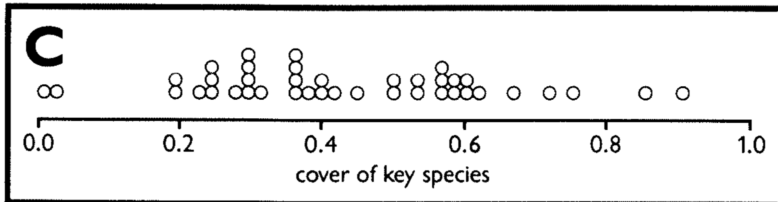
A box plot, also called a box-and-whisker plot (Tukey 1977), is another good way to explore your data. Although less familiar than a histogram, a box plot is a clear and efficient way to convey more information. Figure 9.3D is a box plot of the same cover data. The vertical line in the center of the box indicates the sample **median.** The left and right vertical sides of the box indicate the location of the 25th and 75th percentiles, respectively, of the data. This means that 25% of the data points lie to the left of the left vertical side of the box and 75% to the left of the right vertical side of the box. These 25th and 75th percentiles are often called lower and upper quartiles or hinges. The absolute value of the distance between the hinges (obtained by subtracting the value of the lower hinge from the value of the upper hinge) is the hspread. The whiskers on each side of the box extend to the last data point between each hinge and its inner fence, a distance 1.5 hspreads from the hinge (description after Ellison 1993 and Wilkinson 1991).

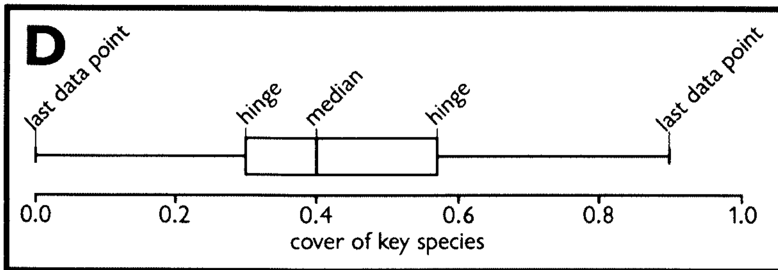
The median is the value that has an equal number of observations on either side, after the observations have been placed in order from smallest to largest.

Box plots also illustrate outliers, which are data points lying farther from the rest of the data than one would usually expect (particularly if one were assuming the data came from an approximately normal distribution). The cover data set contained no outliers. The plant height data (see Fig. 9.2B), graphed in a box plot of plant heights (Fig. 9.4), illustrates the two kinds of possible outliers. Points occurring between 1.5 hspreads and 3 hspreads (the outer fence) are indicated by an asterisk. Points occurring beyond the outer fence (far outliers) are indicated by open circles. These data are measurements of the heights of a plant after grazing. There is, therefore, a preponderance of short plants (note the position of the median; half of all the plants measured are less than about 6cm high), but some individual plants were ungrazed or not grazed as heavily, accounting for the outliers and far outliers shown in the box plot. These data follow a lognormal distribution and, as we learned when we examined a normal probability plot of these same data, we have to be careful when we use parametric statistics with such data sets (see below).









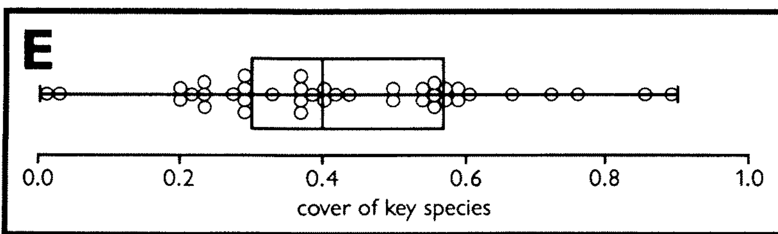


Figure 9.3. Cover data shown in Figure 9.2A explored through a series of graphs. Histograms group the cover observations from the 50 transects into classes. The top left histogram (A) is formed with 10 classes or bars chosen (two of the classes contain no observations), while the top right histogram (B) is with 20 bars (six of which contain no observations). A histograph provides no information about how data points are arranged within each class. In contrast, the dit plot (C) presents each data point. The box plot (D) shows the median, the hinges, and the last data points on either side within the inner fence. In this example, the median is 0.4, with 19 observations of higher cover and 19 observations of lower cover and 2 observations equal to the median value. Combining a box plot over a symmetric dit plot (E) produces a graph with much information in an efficient way.

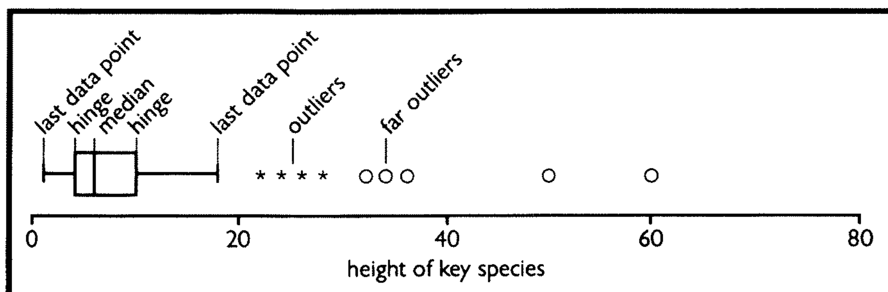


Figure 9.4. Plant height data shown in Figure 9.2B graphed as box plot. This plot illustrates two types of possible outliers. Outliers and far outliers in box plots also alert you that the data likely do not follow a normal distribution.

Sometimes it is helpful to overlay different types of plots. Figure 9.3E overlays a symmetric dit plot onto a box plot. In addition to the information conveyed by the box plot you can see how the individual data points are arrayed.

These examples were constructed using the statistical package SYSTAT (Wilkinson 1991). You should be aware that other statistical packages may use different symbols to indicate outliers and far outliers. They may also define hspreads differently.

SELECTING THE APPROPRIATE STATISTICAL ANALYSIS

The type of statistical analysis to which you intend to subject your data should be determined during the initial stages of your study. Two basic types of analysis can be identified based on the nature of the management and sampling objectives: parameter estimation (with confidence intervals) for target/threshold objectives and significance tests for change/trend objectives.

Confidence Intervals

If you are estimating a quantity based on a single independent sample (i.e., you are not trying to compare the sample to another year or another site), then calculating the precision of your estimate using **confidence intervals** is the correct approach. Confidence intervals can be calculated for a population mean, proportion, or total population size (the term "population" here is referring to your sampled population as described in Chapter 8). Examples include total number of caribou within the sampled area, mean number of purple loosestrife stems per unit area, the proportion of quadrats occupied by salamanders, the mean height or weight of red spruce within your sampled population, the proportion of occupied nesting boxes, the mean number of motorcycle tracks per unit area, and the mean number of frog eggs per vernal pool. Confidence intervals are discussed in more detail below.

A confidence interval is the interval within which a true parameter value lies with known probability. It is a measure of the reliability of our sample estimate of the parameter value.

Calculation of confidence intervals requires meeting several assumptions. You must consider these before using confidence intervals to assess your data. See below.

Significance Tests—An Introduction

If your management objective requires detecting change from one time period to another in some average value (such as a mean or proportion), then statistical analysis consists of a significance test, also called a hypothesis test. This situation often occurs in monitoring and involves analysis of two or more samples from the same monitoring site at different times (usually in different years).[1] The major question asked is whether there has been change in the parameter of interest over a particular period. This parameter is often the mean, but we will also look at situations where the parameter is a proportion. If a change has occurred, the direction of change is a question usually (but not always) of equal importance. Significance tests are used to assess the probability that an observed difference represents a real change in the population or simply results from the random variation that comes from taking different samples to estimate the parameter of interest.

The null hypothesis is usually the no-change hypothesis and is represented by H_o, while the alternative hypothesis of a change (or a change in one direction for one tailed tests) is represented by H_a.

A hypothesis is a prerequisite to the use of any significance test. In monitoring, this hypothesis is usually that no change has occurred in the parameter of interest. This hypothesis of no change is called the **null hypothesis.** If, through

[1] *While we focus in this discussion on the difference between samples over time, which is typical of monitoring studies, these tests can also be used for other comparisons such as different sites, different seasons, different treatments, etc.*

our significance test, we conclude that an observed change in a parameter between 2 or more years is not likely the result of random variation, we reject the null hypothesis in favor of an alternative hypothesis: that there has been a change in the parameter of interest.

The process can be illustrated by example. Let us say we have estimated the density of a rare salamander species in a macroplot in 2 separate years. Each year we have taken a new random sample of forty 0.25m × 5.0m quadrats and counted the number of salamanders in each quadrat. The first year we obtain a mean of 6 salamanders per quadrat, and the second year we obtain a mean of 4 salamanders per quadrat. We wish to determine whether this change is statistically significant or simply the result of random variation inherent in the population of all possible quadrats.

We start with the hypothesis that no real difference exists between the mean of 6 salamanders and the mean of 4 salamanders. What we are really saying is that the true population mean (unknown to us because we are sampling) has not changed, that these two sample means could have been selected simply by chance from the same population.

To test this null hypothesis we must first quantify the difference between these two sample means with a test statistic (Glantz 1997). When the test statistic is sufficiently large, we reject the null hypothesis of no difference between population means and conclude there is in fact a difference. However, we must specify in advance how large this test statistic must be for us to reject the null hypothesis. We do this by specifying a critical or threshold significance level, or ***P* value.** In this example we have specified a threshold *P* value of 20% or 0.20. This threshold *P* value is also called the α level.

The P value is the probability of obtaining a value of the test statistic as large as or larger than the one computed from the data when in reality there is no difference between the two populations (Glantz 1997).

We now enter our data into the computer and conduct a significance test using a statistical software program (do not worry about which test at this point—we will cover this in the following sections), which gives us a calculated *P* value of 0.125. Because this is smaller than the $P =$ 0.20 (selected as our threshold level for determining significance), we conclude that the true population mean has changed. Our calculated *P* value of 0.125 tells us there is a 12.5% chance we are wrong, that there has been no real change at all, or in other words, a 12.5% chance we have committed a false-change error.

If our analysis resulted in a calculated *P* value of 0.85, we would conclude the true population mean has not changed, because the calculated value is greater than our threshold *P* value of 0.20. In this case, we cannot have committed a false-change error (since our conclusion is that no change has taken place), but we may have committed a missed-change error. The probability of a missed-change error (or its complement, power) must also be considered in analysis (see Chapter 7 and below).

Many scientific papers do not report actual *P* values. Instead, they report that an observed difference between samples "was not significant ($P > 0.05$)" or that the difference "was significant ($P < 0.05$)." This practice should be avoided. Actual *P* values calculated from your data should be reported, to enable the readers, who may have different thresholds of significance than you, to make up their own minds. In our example, a *P* value greater than 0.20 would indicate to us that no significant change occurred. But if the actual *P* value were 0.21, we would be more concerned that we may have failed to detect a true change than would be the case if the actual *P* value were 0.85.

Types of Significance Tests

Data can be classified into three basic types: frequency, ranked, and measurement. The appropriate test differs for different types of data (Table 9.1). Selection of the appropriate test also depends on whether your data are paired and the number of years (or comparisons) you want to make. In monitoring, paired data are usually permanent sampling units. See below for discussion of the benefits of paired sampling units and descriptions of the appropriate tests.

Table 9.1. Summary of statistical tests available to analyze typical monitoring data. Note that no parametric test is available for frequency data. Selection between parametric and nonparametric methods is discussed in the text.

PURPOSE OF TEXT	PARAMETRIC TEST	NONPARAMETRIC TEST
Testing for change between 2 years; samples independent; not frequency data	Independent sample t-test	Mann-Whitney U test
Testing for change between 2 years; samples paired (permanent sampling units); not frequency data	Paired t-test	Wilcoxon's signed rank test
Testing for change between 2 years; samples independent; frequency data		Chi-square test (2 x 2 contingency table)
Testing for change between 2 years: samples paired (permanent sampling units); frequency data		McNemar's test
Testing for change between 3* or more years; samples independent; not frequency data	Analysis of Variance; Independent-sample t-tests with Bonferroni correction	Kruskal-Wallis test; Mann Whitney U test with Bonferroni correction
Testing for change between 3* or more years; permanent sampling units; not frequency data	Repeated Measures Analysis of Variance; paired *t*-tests with Bonferroni correction	Friedman's test; Wilcoxin signed rank test with Bonferroni correction
Testing for change between 3* or more years; samples independent; frequency data		Chi-square test (2 x ≥ 3 contingency table)

**Information concerning analysis of changes over several years can also be found in Chapter 10.*

Frequency data (also called "attribute" data or "nominal" data) are class data, with mutually exclusive classes. In monitoring studies, the number of classes is most commonly two (e.g., present or absent). Examples of frequency data include presence/absence in a frequency study using quadrats, point estimates of cover when the point is the sampling unit (because a point either does or does not intersect the species), or responses or nonresponses to baits or broadcast calls in wildlife monitoring. Significance tests available to test changes in frequency are the Chi-square test or McNemar's Test (when the data are paired such as in permanent plots).

Measurement data (also called interval data) are counts or measurements (density, height, weight, number, population size). For measurement data, you usually have an option between a parametric and nonparametric test (Table 9.1). The selection between parametric and nonparametric tests involves an evaluation of your data and consideration of whether your data meet the assumptions of parametric tests. This is discussed in the next section.

Ranked data (also called ordinal data) provide some sense of relative size, allowing ranking from smallest to largest. Examples of this type of data are ocular estimates of cover arranged in 5 classes (dominant, abundant, frequent, occasional, rare), or a height measured in 6 classes (very short, short, medium, tall, very tall, extremely tall). These data are sometimes analyzed with nonparametric tests (Table 9.1). Where ranks are approximately equal in size they can be considered measurement data and analyzed as such using parametric tests (Snedecor and Cochran 1989). For example, estimates of percentage cover in classes can function as measurement data if the classes are the same size (e.g., class 1 = 1% to 10%, class 2 = 11% to 20%, etc.) or if the midpoints of classes are used in analysis. Similarly, the example of plant heights could be treated as measurement data with values of 1 to 5 corresponding to plant height classes if the classes were approximately equal in size. Conversely, if plant height classes were very different in size (e.g., class 1 = 1cm to 5cm in height, class 2 = 6cm to 50cm, class 3 = 51cm to 5m, etc.), the data would be analyzed using nonparametric tests.

Parametric or Nonparametric Statistics?

The use of parametric statistics requires that several assumptions be met, at least approximately (no monitoring data will meet these assumptions exactly):

1. That the population being sampled follows a normal distribution. A normal distribution is the familiar bell-shaped curve illustrated in Figure 9.5. This assumption holds both for the calculation of confidence intervals and for the use of *t*-tests and analyses of variance. (For paired *t*-tests, the *differences* between sampling units should come from a population that follows a normal distribution.)
2. That the sampling units are drawn from populations in which the **variances** are the same even if the means change from the first year of measurement to the next. This assumption, called homogeneity of variances, applies to significance tests to detect changes in means.
3. That the sampling units are drawn in some random manner from the population. This assumption applies for the calculation of confidence intervals and for both parametric and nonparametric significance tests.

Population variance is the sum of (value associated with a given member of the population – population mean)² divided by the number of population members. Mathematically this is expressed

$$\sigma^2 = \frac{(X_1-\mu)^2 + (X_2-\mu)^2+...+(X_N-\mu)^2}{N}$$

The population variance is usually denoted with the Greek δ^2, while the sample variance is usually denoted s^2.

Before data analysis you need to determine if your monitoring data meet these assumptions. Although some tests can be used to assess whether your sample data are normally distributed, it is often most effective to look at graphical analyses of your data. The use of probability plots, dit plots, and box plots to explore your data for normality was discussed at the beginning of this chapter. Fortunately, both *t*-tests and analyses of variance are robust to moderate departures from either normality or homogeneity of variances (Zar 1999).

Several tests are available to determine if the variances of two or more samples are equal, but none is very reliable. The most well-known, Bartlett's test, is not recommended because it is unduly sensitive to departures from normality (Sokal and Rohlf 1994). In fact, Zar (1999:204) recommends no test be used to assess whether the assumption of homogeneity of variances holds, because the analysis of variance is robust to departures from this assumption if sample sizes are equal each year of measurement. The *t*-test is similarly robust.

What happens if your data do not meet the assumptions of normality and homogeneity of variances? Data on biological entities will rarely meet either of these assumptions perfectly. As pointed out by Koch and Link (1970), few if any real data come from a population that is normal, or even quasi-normal. They go on to point out the only consequences of failure to meet the assumption of normality are some distortion of the theoretic risk levels and a reduction in the efficiency of estimation. These problems, however, are far less serious than the failure to meet the assumption of randomness (Koch and Link 1970). For severe departures from these assumptions there are several possible solutions:

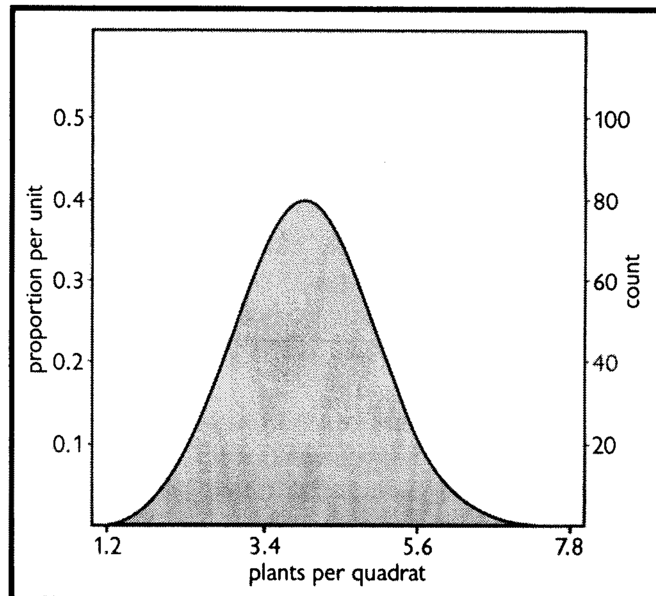


Figure 9.5. A normal distribution. The population is a set of quadrats; the variable measured is number of plants per quadrat. A distribution this close to normal is rare in practice.

1. Increase your sample size. According to Mattson (1984), a sample size of at least 100 sampling units will ensure against problems resulting from severe departures from normality (very skewed distributions). This is conservative; less severe departures from normality will not require as large a sample. We talk more about this below.
2. Transform your data. Transformations, whereby data in the original units are converted to another scale prior to analysis, are often applied to data prior to performing significance tests to make the data conform more closely to the assumptions of normality and homogeneity of variance. The use of transformations is covered in many statistics texts (Fowler and Cohen 1998; Hoaglin et al. 1983; Zar 1999). They will not be covered further here, except to say that their utility for monitoring is limited because of several problems: 1) estimated quantities such as means, variances, and confidence intervals in the transformed scale are typically biased when the data are transformed back into the original scale; 2) the results of statistical analyses expressed in the transformed scale are often difficult to understand or apply; 3) more calculations are required (Gilbert 1987). Li (1964) suggests that the most common transformations are seldom helpful in practice.
3. Use nonparametric statistics. If you are greatly concerned whether your data meet the assumptions of normality and homogeneity of variance, you can use nonparametric statistics, which do not require meeting these assumptions. Note, however, that nonparametric statistics, just like parametric statistics, require that data be collected in a random manner.
4. Use statistical analyses based on resampling. With the advent of personal computers in the 1980s, statisticians began developing new theory and methods based on the power of electronic computation (Efron and Tibshirani 1991). These resampling methods (also called computer-intensive methods) are becoming more popular with ecologists and other scientists and can be used to calculate confidence intervals and to conduct significance testing. Two of the most commonly used methods are bootstrapping (which involves repeated resampling of the original data set with replacement) and randomization (also called permutation) testing (which involves sampling the original data set without replacement). The only drawbacks to the use of these methods are their lack of familiarity and the fact that some of the theories behind them are relatively new and therefore little tested. The advantages of resampling methods, however, are many, including the fact that few of the assumptions required for parametric statistics are needed (except, of course, for the assumption of random sampling) and they are apparently just as powerful (Manly 1997). These methods thus allow for estimation of parameters (including construction of confidence intervals) that would be difficult or virtually impossible to estimate for some datasets using conventional statistics. References to these methods, which are not covered in this handbook, are found in Chernick (1999), Davidson and Hinkley (1997), Good (1999), Good (2000), Lunneborg (2000), and Manly (1997).

Why not use nonparametric statistics all the time? Given that nonparametric statistics require fewer assumptions than parametric statistics, you can ask the question, why bother with parametric statistics at all? The answer lies in the fact that, when the necessary assumptions are at least approximated, parametric statistical tests are more powerful than their nonparametric analogues. If, however, the populations from which you sample are highly skewed, your sample size is small, and—in the case of significance tests—your sample sizes are very different at each time of measurement, you may want to use nonparametric methods or resampling techniques. Otherwise, you are perfectly justified in using parametric statistics.

When should you worry about using parametric statistics? Glantz (1997) offers the following rules of thumb for deciding whether to use parametric statistics in significance testing. If the

variances are within a factor of 2 to 3 of each other, then you can assume that your data meet the assumption of homogeneity of variances. If a density plot of the observations reveals that they are not heavily skewed and contain no more than one peak, then you can assume that the data are close enough to a normal distribution to use parametric statistics. Another "test" of normality is to compare the size of the mean with the **standard deviation.** When the standard deviation is about the same size or larger than the mean and the variable being measured can take on only positive values (which is true for most monitoring data), this is an indication that the distribution is heavily skewed (Glantz 1997).

Population standard deviation is the square root of the population variance.

Hahn and Meeker (1991) point out that confidence intervals designed to include the population mean (as opposed to some other population parameter such as the variance) are relatively insensitive to the assumption of normality. Willoughby (unpublished data) evaluated a population that followed a highly skewed exponential distribution (Fig. 9.6). Parametric confidence intervals were similar (within the expected probabilities) in comparison to confidence intervals calculated by resampling methods (described above) when sample sizes were 150 or more. The parametric confidence intervals were very close to the resampling confidence interval with a sample size 100 and reasonably close with a sample size 50. In other words, the probability that the true mean lay within the calculated parametric confidence interval was not affected by the skewed distribution of the population once the sample size was large enough.

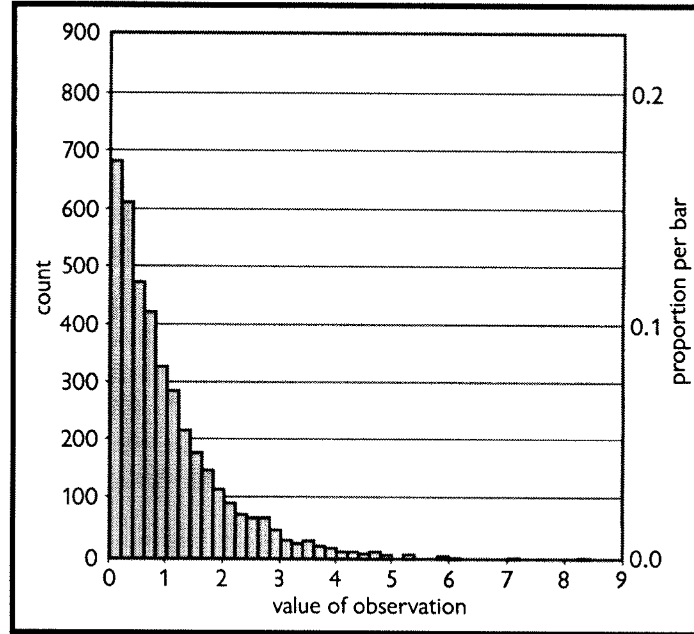


Figure 9.6. Histogram of a simulated population of 4000 observations. The population follows the exponential distribution. The mean of this population is 0.995, and the standard deviation is 0.962. Note the large number of small values and the very long tail to the right caused by a few very large values (though hard to see, there is a single value larger than 8, which is more than eight times the standard deviation).

Cochran (1977:42–43) offers what he terms a "crude rule" for determining how large the sample size must be to use the normal approximation in computing confidence intervals. This rule makes use of Fisher's measure of skewness, often designated G_1. Most statistical programs routinely calculate this measure, although many of them simply use the term "skewness" instead of G_1. The rule is designed so that a 95% confidence probability statement will be wrong no more than 6% of the time. The rule controls only for the total error rate and ignores the direction of the error of the estimate. Cochran's formula is as follows:

$$n > 25G_1^2$$

where n = Sample size.
G_1 = Fisher's measure of skewness.

To illustrate how this formula works, let us use it on the simulated population of 4000 observations that follow an exponential distribution described above (see Fig. 9.6). The computer program, STATMOST, gives a G_1 value (which it labels simply as "skewness") of 1.778.[2]

---

[2]*This skewness value, G_1, is the true value for the entire population. The value based on a sample of this population would be somewhat different. To ensure that we have enough values to obtain a stable estimate of G_1, we could construct a sequential sampling graph of the mean and standard deviation, as shown in Chapter 8. When these values stabilize, we know we have a large enough sample. We then calculate G_1 for this sample and plug it into Cochran's formula.*

Table 9.2. Sample of shrub heights.

HEIGHT (M)	RANK
0.35	1
0.40	2
0.50	3
0.55	4
0.75	5
0.90	6
1.00	7
1.10	8
1.30	9
4.50	10
5.10	11

Entering this value into the above formula yields the value $(25)(1.778)^2 = 79$. We therefore know that we must have at least 79 sampling units to obtain a confidence interval that includes the true mean 94% or more of the time (this is consistent with the empirical results found by Willoughby). It remains necessary, of course, to calculate the sample size required to obtain the level of precision specified in your sampling objective (see Chapter 8).

When should I use nonparametric statistics? You would use nonparametric statistics in the following situations:

- The data are ranked with the categories of unequal magnitude.[3]
- The data are frequency data.
- The data are measured data with a very skewed distribution, and sample sizes are relatively small (e.g., <50).
- The data are measured data with a very skewed distribution, and samples being compared are very different in size (for significance tests).

Nonparametric statistics usually involve ordering (ranking) the data from the smallest value to the largest and using the ranks rather than the values themselves. For example, we might have a sample of 11 shrub heights shown in Table 9.2. Note how this ranking reduces the effect of the two large values, 4.5 m and 5.1 m, on the data set. Since the analysis is based on the ranks, not the actual values, the difference between ranks 9 and 10 is only one unit, rather than 3.2 units in the original units.

For distributions with a large positive skew, however, there are times that you may want to use the median instead of the mean as a measure of the central tendency in the distribution. The median is that value of the variable (after the values have been ranked) that has an equal number of values on either side of it. It divides the frequency distribution in half. Thus, the median for our sample of 11 shrub heights is 0.90, because there are five numbers above and five numbers below the value of 0.90. Notice the difference between this median and the mean of 1.50 for the same data set. The two large values of 4.50 and 5.10 have greatly affected the mean. (In a completely normal distribution the mean and median are equal.)

Distributions with large positive skews are rather common in biology. For example, we may measure the heights of a shrub species several years following a wildfire. A few plants of this shrub species, having survived the fire, might be very tall, while the rest of the plants, being new recruits, might be relatively short. Figure 9.7 shows how this distribution of heights might look. Similarly, populations of long-lived animals are comprised of many small juveniles and relatively few large reproductive adults.

Depending on how severe this skew is, you might want to use nonparametric statistics in analyzing the data, unless your sample size is large enough to use parametric statistics. If you are interested in inferring something about the characteristics of a skewed population, you may wish to estimate the median of the population. If so, you can calculate a confidence interval within which the true population median lies (with some confidence level). Methods can be found in most statistics texts (Hahn and Meeker 1991; Zar 1999). Our web page (see Preface) carries links to sites that calculate confidence intervals around a median, and many statistics packages will calculate confidence intervals around a median.

---

[3] *As stated earlier, ranked data may be analyzed using parametric methods if the ranks are of similar size.*

There are nonparametric analogues to all parametric significance tests discussed in this handbook (Table 9.1). We do not cover most of these in this handbook.[4] We believe that most monitoring data collected from well-designed studies will be suitable for parametric analysis. To analyze studies using ranked data or datasets of measurement data not meeting assumptions of parametric tests, refer to statistical texts such as Hollander and Wolfe (1999), Daniel (1990) and Zar (1999) for descriptions of the tests. Most statistical computer packages perform these tests.

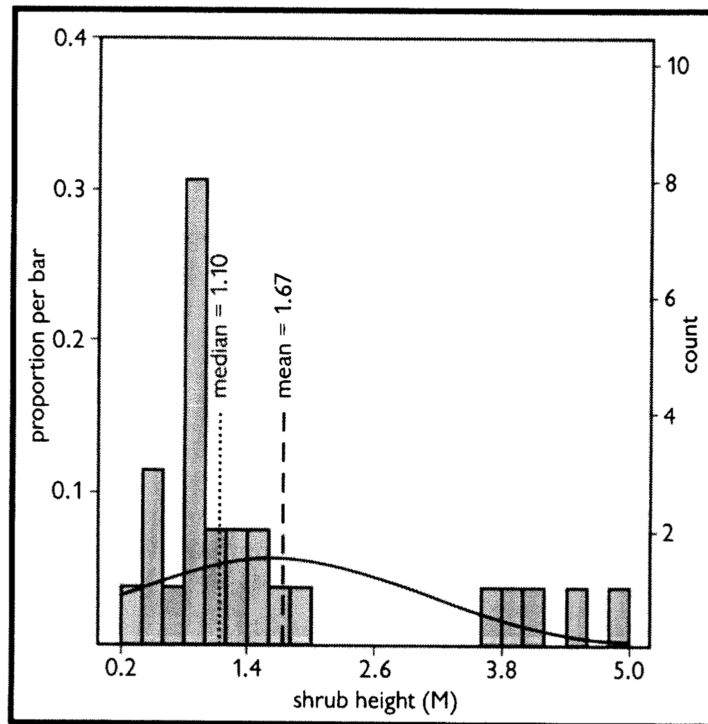


Figure 9.7. Histogram of shrub heights (n = 26) on which is superimposed a normal smoothing curve. Note that the distribution of values is positively skewed, with a long tail to the right. The median (*indicated by the dotted line in the figure*) for this sample is 1.10, while the mean (*indicated by the dashed line*) is 1.67. The mean is affected by the few large heights, while the median is not.

Problems with Significance Tests

All populations change over time, perhaps only slightly. A significant result from a statistical test is the result of three functions: the actual change of the population, the efficiency of the design, and the size of the sample (Johnson, DH 1999; Ellison 1996). If you measured too many sampling units, a statistical test may produce a *P* value that is considered significant when compared with your threshold, but the actual change in the population is less than the amount specified in your objective, and perhaps not biologically relevant.

Finding a change in a population statistically significant but not biologically significant is a rare problem in monitoring studies. The more common problem is the study has too few sampling units or a poor design that does not allow the detection of biologically important change (low power). In addition, if you have designed your study carefully and correctly determined the sample size needed to detect the amount of change you have specified as important in your management objective by conducting a pilot study, it is unlikely that you will oversample and detect a change that is not biologically relevant.

The solution we recommend is described in detail below. First, complete the statistical test as planned. If the test returns a significant *P* value, then evaluate whether the actual difference observed is biologically relevant. If the test returns a nonsignificant *P* value, conduct post hoc power analysis as described below.

Another option is to graph the confidence interval of difference between two population means. If this interval includes zero, you know that the difference between the two samples is not significant at the precision level you have selected. This is the recommended approach by some authors (e.g., Cohen 1994; Reichardt and Gollob 1997) and directions for doing this are given in Appendix III. The main benefit of this approach is its visual nature. You can visually assess the mean difference between the two samples and its biological relevance. If the confidence interval around the mean difference is very wide, you know you likely have a problem with the power of your study. This analysis is available on some but not all statistics packages and on some online and freeware analysis programs.[5] A drawback to this approach is that for most people this procedure is unfamiliar and not intuitively obvious.

---

[4] *The chi-square and McNemar's tests are nonparametric tests, but are only appropriate for frequency data. We cover these in this chapter.*

[5] *See our web site (address in Preface) for links to these and other online and freeware/shareware analysis packages.*

CONFIDENCE INTERVALS

- Target/threshold management objective
- Estimate a mean, proportion or total size for a single sample

If your management objective is a target or threshold objective, it is sufficient to estimate the parameter (mean, total, or proportion) and construct a confidence interval around the estimate. The analysis required is to calculate the sample statistic (mean, total, or proportion) and the confidence interval (the desired confidence level, such as 95% or 80%, should be specified in your sampling objective). Confidence intervals for estimates of population means and totals were introduced in Chapter 7. Appendixes III and IV give directions on calculating confidence intervals around estimates of means and totals, as well as around estimates of proportions. You can calculate sample statistics and confidence intervals in each year of data collection and graph these using bar or point graphs with the confidence intervals as error bars (graphing results, including the use of bar and point graphs, is discussed in further detail below). The sample statistic and confidence interval of each sample would be compared with the target or threshold to determine if action is necessary or if the objective has been reached.

Confidence Intervals for Means and Totals

For example, your management objective is to maintain a population of at least 2000 individuals of the wood frog (*Rana sylvatica*) at the Hemlock Creek Preserve over the next 5 years. Your sampling objective is to annually estimate the population size of wood frogs at the Hemlock Creek Preserve and be 95% confident that the estimate is within 250 frogs of the true population total. This is a threshold objective, because you are concerned with the population falling below the threshold. Therefore, data analysis consists of estimating the population size from the sample mean (by multiplying the total number of possible sampling units by the sample mean) and calculating the confidence interval for this estimate.

The estimated total and confidence interval are then compared with the threshold of 2000 frogs. If both the estimated total and lower bound of the confidence interval are above the threshold, you can be confident (relative to the level chosen) that you have met your objective. If both the estimated total and upper bound of the confidence interval are below 2000 frogs, you can be confident (again relative to the selected confidence level) that you have failed to meet your objective. Less clear are situations where the threshold value is included within the confidence interval, with the estimated total either above or below the threshold (see below). You should have prepared for this eventuality and prescribed the action you will take should this occur.

Confidence Intervals for Proportions

A target/threshold objective can also be framed using a proportion. For example, your management objective is to decrease the frequency (in $1 m^2$ quadrats) of the noxious weed yellow star thistle to 30% or less at Key Area 1 in the Cache Creek Management Area by 2006 (the current frequency is 70%). This is the same thing as saying that, out of all the quadrats you could place in the sampled area (with no overlap), you want the proportion of quadrats containing yellow star thistle to be 30% or less. Your sampling objective is to annually estimate the percent frequency with 95% confidence intervals no wider than 10% of the estimated true percent frequency. Because you are dealing with a proportion, as opposed to a mean or total, the confidence interval width (10%) is expressed as an absolute rather than a relative value. If your estimate of the true proportion is 40%, for example, your target confidence interval width is from 30% to 50%.

For the yellow star thistle example, data analysis entails estimating the percent frequency (by dividing the quadrats that contain yellow star thistle by the total number of quadrats sampled) and calculating a confidence interval around this estimate (see Appendix III for instructions

on doing this). The estimated frequency (proportion) and confidence interval are then compared with the target objective of 30%. If both the estimated proportion and upper bound of the confidence interval are below the target objective, you can be at least 95% confident that you have met your objective. If both the estimated proportion and lower bound of the confidence interval are above the target objective, you can be at least 95% confident that you have failed to meet your objective. If the target objective falls within the confidence interval, your interpretation is more difficult.

Interpreting the Results of Parameter Estimation

Following are two examples of management responses to threshold and target management objectives.

- Action X will occur if the mean density of rare species Y drops below value Z.
- We will judge our restoration efforts to be successful if we have raised the mean density of species A to value B by the year 2000.

Because you have taken a sample (as opposed to conducting a complete census), you will not know the true population parameter (e.g., the true mean value). You will have only your estimate of the parameter (e.g., the sample mean) surrounded by a measure of precision such as a confidence interval. Interpretation then requires you to compare the parameter estimate and confidence interval with the threshold value. There are four possibilities, illustrated in Figure 9.8, and discussed below:

1. Your threshold level has not been crossed by either the parameter estimate or the confidence interval (top arrow of Fig. 9.8). Here the interpretation is relatively simple. You can be confident, at least to the degree of the confidence level you have selected for your confidence interval, that the true parameter has not crossed the threshold. For example, if your confidence interval is 95%, then you can be at least 95% confident that the true parameter is still below the threshold (the actual confidence may be greater than this if the upper bound of the 95% confidence is some distance from the threshold).
2. Your threshold level has been crossed by both the parameter estimate and the confidence interval (bottom arrow of Fig. 9.8). Here again the interpretation is relatively simple. You can be confident, at least to the degree of the confidence level you have selected for your confidence interval, that the true parameter has crossed the threshold. If your confidence interval is 95%, then you can be at least 95% confident that the true parameter has crossed the threshold (the actual confidence may be greater than this if the lower bound of the 95% confidence interval is some distance from the threshold).
3. The parameter estimate does not exceed the threshold value, but the upper bound of the confidence interval does exceed the threshold value (as in the second arrow of Fig. 9.8). Now the interpretation is not nearly as clear. Because the true population parameter can be anywhere inside of the confidence

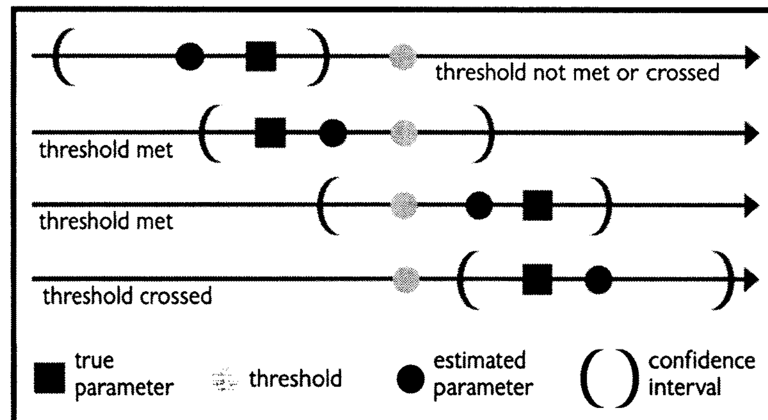


Figure 9.8. The four different possible outcomes when comparing a parameter estimate and confidence interval to a threshold level. The true parameter is shown only for illustrative purposes; we would never know it when conducting sampling. Adapted with permission from a figure prepared by Sylvia Mori, U.S. Forest Service, Pacific Southwest Research Station.

interval, it is quite possible that the true population parameter has, in fact, crossed the threshold.

4. The fourth possibility is that both the parameter estimate and the upper bound of the confidence interval have crossed the threshold, but the lower bound of the confidence interval has not. This is illustrated by the third arrow of Figure 9.8. Again, because the true population parameter can be anywhere within the confidence interval, it may have crossed the threshold. In this case, because the midpoint of the confidence interval (the parameter estimate) has also crossed the threshold, it is more likely than in situation (3) that the true parameter has crossed the threshold.

How you will interpret situations like (3) and (4) should be determined before calculating your parameter estimate and confidence interval. In fact, you should have decided on this before even initiating sampling. One approach is to decide that if any part of the confidence interval crosses the threshold, you will take action, based on the possibility that the true parameter has crossed the threshold. This minimizes the risk to the plant resource for which you are managing. Remember, however, that the size of the confidence interval depends on the confidence level you choose, the degree of variability in your sampling data (as expressed by the standard deviation), and your sample size. Thus, an inefficient sampling design and small sample size will result in much wider confidence intervals, which in turn will result in facing situations like (3) and (4) much more often. Good sampling design and reasonable sample sizes will therefore facilitate interpretation by making confidence intervals narrower and reducing the number of times you encounter the predicaments illustrated by (3) and (4).

INDEPENDENT-SAMPLE T-TEST

- Test for difference between two means
- Samples are independent (not paired)
- Measurement data
- Ranked data when the categories are similar in magnitude
- Frequency data grouped in transects or clusters[6]

The independent-sample *t*-test is employed to test for difference in the means of two samples. This test is applicable to the analysis of density data, height data, and biomass data. It can be used to analyze cover data estimated in independent quadrats or along line intercepts. It is also appropriate for the analysis of cover data collected with points if the sampling unit is a group of points such as points arranged along transects.[7] The independent-sample *t*-test can also be used to analyze frequency data when quadrats are arranged along transects and the transects (not the quadrats) are treated as the sampling units. When the frequency quadrats are the sampling units, then the chi-square test is the one to use. Similarly, responses or nonresponses to baits or broadcast calls in animal monitoring is frequency data, but if the baits or calls are arranged along a transect, the transect can be considered the sampling unit and analyzed with an independent *t*-test.

Many microcomputer software packages and some handheld calculators carry out the independent-sample *t*-test. The basic principle is that we examine the ratio (after Glantz 1997):

---

[6] *Frequency data can be collected in individual sampling units (quadrats or points) or as secondary sampling units, grouped as along a transect (the primary sampling unit). In the latter situation, the transect is considered the sampling unit, and analysis conducted on the transect values. These data are properly considered measurement data and analyzed as such.*

[7] *If the points are treated as the sampling units, the chi-square test is the appropriate test.*

$$t = \frac{\text{Difference of sample means}}{\text{Standard error of difference of sample means}}$$

When this ratio is small, we do not reject the null hypothesis that there has been no change in the true population mean. If the ratio is large, we reject the null hypothesis and conclude there has been a change in the true population mean. How "large" the *t* value must be to reject the null hypothesis depends on the *P* value we have previously chosen as our threshold of significance.

The fact that the *t* value is smaller than the value of *t* corresponding to our *P* value does not indicate that there has not been a change in the true population mean. It only means that we have not demonstrated this change at a given level of significance through our monitoring study. To see how likely we would have been to detect a real change of a given magnitude, we can (and should) conduct a post hoc power analysis as discussed below.

Two types of *t*-tests can be run on independent samples, a two-tailed test and a one-tailed test. The type of test selected depends on the type of null hypothesis being tested. If the null hypothesis is that there has been no change in the population mean, then a two-tailed *t*-test would be used, because you need to detect change in either possible direction (smaller or larger values of the mean). If, however, the null hypothesis is, for example, that the population mean has not increased, then a one-tailed test would be used because you only need to detect change in one direction (an increase). Note, however, that a nonsignificant *P* value after a one-tailed test could mean either that the population mean has decreased or stayed the same; there is no way of testing which.

In many cases, one-tailed tests are advantageous because they increase the statistical power for detecting a change in the direction of interest. If, for example, our management objective is to increase the density of a particular rare species, we may decide to frame our sampling objective in terms of detecting only whether an increase in density has occurred. If our monitoring study shows no increase between sampling periods, then we institute a management change. The appropriate test would be a one-tailed test. Similarly, we may be only concerned with a decrease in a species. Again, the one-tailed test is appropriate. In many cases, the increase in power is considerable. The one-sided test, however, only demonstrates significance in one direction.

Two-Tailed Example

Let us say that we have monitored the density of a rare plant species in each year over a 2-year period. We randomly place 50 quadrats, each 0.25m × 25m, in each of the years and calculate the mean and standard deviation for each of these two independent samples. In the first year our sample mean and standard deviation are 4.0 and 2.5, respectively (the units for both the mean and standard deviation are in plants/quadrat; the units are left out here for simplicity). In the second year our sample mean and standard deviation are 3.0 and 2.0, respectively. We now want to conduct a *t*-test to determine if this observed difference is significant. Before sampling we have decided to set our false-change error rate (α) at 0.10. Thus, our threshold *P* value is 0.10.

Before testing, we must formulate a null hypothesis. In this instance we are interested in detecting change in either direction (either an increase or decrease in density). Our null (H_o) and alternative (H_a) hypotheses are therefore as follows:

H_o: The population mean has not changed between Year 1 and Year 2.

H_a: The population mean has changed between Year 1 and Year 2.

To test these hypotheses we calculate the *t* statistic as follows:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

where
t = Test statistic
X = Mean (subscripts denote samples 1 and 2, respectively)
n_1 = Sample size of sample 1
n_2 = Sample size of sample 2
s^2 = Pooled estimate of variance, calculated as follows:

$$s^2 = \frac{(s_1^2 + s_2^2)}{2}$$

where
s_1 = Standard deviation of sample 1
s_2 = Standard deviation of sample 2

Entering our two sample standard-deviation values into the pooled estimate formula we obtain the following:

$$s^2 = \frac{(2.5)^2 + (2.0)^2}{2} = 5.13$$

We now enter our pooled variance estimate into the formula for t and obtain the following:

$$t = \frac{4 - 3}{\sqrt{\frac{5.13}{50} + \frac{5.13}{50}}} = \frac{1}{\sqrt{0.1026 + 0.1026}} = 2.208$$

To determine the likelihood of H_o being true, we compare this calculated t statistic of 2.208[8] with the critical value of t in a t table for an α of 0.10 (remember we decided prior to testing that an α of 0.10 [$P = 0.10$] would be our threshold for significance) and the appropriate degrees of freedom.[9] For an independent-sample t-test like the one we are conducting here, degrees of freedom are determined by applying the formula $2(n-1)$, where n is the size of each sample. In our example the sample size is 50 in each year. The degrees of freedom are therefore $2(50-1) = 98$.[10]

The critical value of t from a t table for $\alpha = 0.10$ (for a two-tailed test we use the α [2] row in the table, where the [2] stands for a two-tailed test) and 98 degrees of freedom (designated n in the t table) is 1.661. Since our calculated t value is greater than this critical value, we reject the null hypothesis of no change and conclude that there has been a downward change in the population mean (since the mean of the second year is less than the mean of the first year). We would also report our calculated P value, which we could interpolate from the t table, but could obtain more easily through a statistics program. For this example, the P value is 0.0296, well below the threshold P value of 0.10. We can say there is about a 3% chance that we have committed a false-change error (concluding that there has been a change in the population mean when no true change has occurred).

One-Tailed Example

Using the same example we used for our two-tailed test, we will evaluate whether the population has decreased. We have decided to take action if the population decreases, but to take no action if the population remains the same or increases. In this situation we have a different set of hypotheses as follows:

---

[8] *If we have sampled more than 5% of the population, we should apply the finite population correction factor to the* t-test. *This increases the* t *statistic and gives us greater power to detect change. See the section on the finite population correction factor of this chapter for instruction on how to do this.*

[9] *A* t *table can be found in most standard statistics tests. See our web page for links to online tables.*

[10] *Note that the size of the two samples does not have to identical. Degrees of freedom are then calculated* $(n_1-1)+(n_2-1)$. *It is best, however, if sample sizes are the same or nearly the same as discussed above.*

H_o: The population has not decreased.

H_a: The population has decreased.

The first thing we do with the one-tailed test is look at the sample means. If the Year 2 sample mean is greater than the Year 1 sample mean, we will not bother to conduct the *t*-test, since we already know we cannot reject the null hypothesis and say that the population has decreased (the population may have increased or it may have stayed the same—since we are conducting a one-tailed test, however, we will not be able to say which).

If the Year 2 sample mean is less than the Year 1 sample mean, we then conduct the *t*-test, using the same formula as for the two-tailed test. The only difference is that we compare our calculated *t* value with the critical value for the one-tailed test (the row labeled α [1] in a *t* table). The one-tailed critical *t* value for 98 degrees of freedom and $\alpha = 0.10$ is 1.290. Since this is less than our calculated *t* value of 2.208, we reject the null hypothesis in favor of the alternative hypothesis and conclude that the population has decreased. Using a statistical program we calculate the actual *P* value as 0.0148. Thus, we can state that there is about a 1.5% probability that we have committed a false-change error. Note that the *P* value for the one-tailed test is exactly one-half the *P* value for the two-tailed test. With the same data set this will always be the case. Thus, the one-tailed test is always more powerful than its two-tailed counterpart in detecting change in one direction.

ANALYSIS OF VARIANCE

- Test for difference between three or more means[11]
- Samples are independent
- Measurement data
- Ranked data when the ranks are of similar magnitude
- Frequency data grouped in transects or clusters[12]

The analysis of variance, often abbreviated as ANOVA, is used for testing for the difference between the means of three or more samples. All microcomputer statistics programs carry out this test.

Instead of *t*, ANOVA uses *F* as the test statistic. The test statistic *F* is calculated as follows (from Glantz 1997):

$$F = \frac{s^2_{bet}}{s^2_{wit}}$$

where

s^2_{wit} = Within-groups variance: population variance estimated from sample means

s^2_{bet} = Between-groups variance: population variance estimated as the average of sample variances

You will probably use some computer program to calculate *F*, so only the concept is presented here. The formulas for calculating the *F* test statistic can be found in most standard statistical text books such as Snedecor and Cochran (1989), Sokal and Rohlf (1994), Steel and Torrie (1980), and Zar (1999). The basic concept behind the ANOVA is that under a null hypothesis of no difference between true population means, the two variances are estimates of the same

---

[11]*Analysis of variance can also be used to test the difference between 2 years, but the* t-test *is less complicated.*

[12]*Frequency data can be collected in individual sampling units (quadrats or points) or as secondary sampling units, grouped as along a transect (the primary sampling unit). In this situation, the transect is considered the sampling unit, and analysis conducted on the transect values. These data are properly considered measurement data and analyzed as such.*

population variance. Therefore, the closer this ratio is to 1, the less likely there is a difference between population means. How large the *F* statistic must be before you reject the null hypothesis and conclude there has been a change in the true population mean depends on the threshold *P* value chosen.

ANOVA is a two-tailed test.[13] You should also realize that a significant *F* statistic leads to the conclusion that at least one of the sample means tested comes from a different population. It does not tell you which means are different, although you can usually get a reasonable idea from your estimates.

As an example, assume that we have collected 3 years of density data from the same macroplot in 1996, 1998, and 2000. Quadrats were randomly located in each year of measurement using different sets of random coordinates (their positions in any year are therefore independent of their positions in any previous year). Before sampling, we determined we would accept a false-change error rate of 0.05 and a missed-change error rate of 0.05. The summary statistics are as follows:

YEAR	SAMPLE SIZE (N)	MEAN	STANDARD DEVIATION	STANDARD ERROR
1996	30	21.467	10.136	1.851
1998	30	16.633	9.807	1.790
2000	30	14.800	9.539	1.742

The raw data are entered into a statistical computer program, and the analysis of variance option is chosen. The program creates an "ANOVA table," which gives the pertinent statistics for the analysis of variance test. This table may look slightly different from one computer program to another, but will have the same basic format as the one below. The following is an ANOVA table for our 3 years of data:

ONE-WAY ANOVA RESULTS

SOURCE	DF	SS	MS	F	P
Between groups	2	711.6667	355.8333	3.6822	.0292
Within groups	87	8407.2333	96.6349		
Total	89	9118.9000			
Alpha level = .05					
Critical *F* (0.05,2,87) = 3.1013					

The value of the test statistic, *F*, is 3.6822.[14] The *P* value, given in the last column, is .0292. Thus, there is about a 2.9% probability of obtaining an *F* value of 3.6822 or larger when in fact there is no difference between all 3 of the years. (The other values in the table are those used in the calculation of the *F* statistic. "DF," "SS," and "MS," stand for degrees of freedom, sum of

---

[13] *There are analysis of variance techniques that do not depend on the* F *statistic that can be used to test one-sided or directional hypotheses, but few, if any, statistical programs can perform these techniques. See Rice and Gaines (1994) for an introduction.*

[14] *If we have sampled more than 5% of the population, we should apply the finite population correction factor to the* F *statistic. This increases the* F *statistic and gives us a greater power to detect change. Instructions can be found in the section on the finite population correction factor.*

squares, and mean squares, respectively. The MS value between groups divided by the MS value within groups yields the *F* statistic. The alpha level is the one we entered into the program, and the critical *F* value is the one corresponding to an alpha level of 0.05, with 2 and 87 degrees of freedom for the between and within group sources of variance, respectively.)

Our sampling objective specified a false-change error rate of 0.05. Since the *P* value is less than this, we conclude that 1 or more of the years is significantly different from the others.

To test statistically which of these 3 years is different, we can compare each of the pairs of means using two-sided *t*-tests. However, we must modify the *P* value used for the ANOVA for each *t*-test performed, by dividing the *P* value used for the overall ANOVA by the number of *t*-tests to be performed. In this case, our overall *P* value is 0.05. If we want to compare all three mean values (mean 1 with mean 2, mean 2 with mean 3, and mean 1 with mean 3), we divide the overall *P* value by 3. Our new threshold *P* value for each of these tests is thus 0.05/3 = 0.0167.

When we do these pairwise *t*-tests, we come up with the following statistics:

YEARS COMPARED	DF	t VALUE	P VALUE
1996 vs. 1998	58	1.8771	.0655
1996 vs. 2000	58	2.6234	.0111
1998 vs. 2000	58	0.7340	.4659

Only the *P* value of 0.0111 for the years 1996 versus 2000 is less than our threshold of 0.0167. We therefore conclude that there has been a significant change between those 2 years (but not between any of the other pairs of years). Dividing our original threshold *P* value by the number of comparisons we wish to make is the Bonferroni correction. It works reasonably well when the number of comparisons are few (Glantz 1997). As the number of comparisons increases above 8 to 10, however, the value of *t* required to conclude a difference exists becomes much larger than it needs to be, and the method becomes overly conservative (Glantz 1997). Other multiple comparison tests are less conservative and preferable in these cases. Three such tests are the Student-Neuman-Keuls test, the Scheffe test, and the Tukey test, some or all of which are performed by many microcomputer statistical packages. There is debate over which of these is the preferable test; see Zar (1999:208–215) for a discussion of this. Another such test, the Duncan multiple-range test, is not conservative enough and should be avoided (Day and Quinn 1989).

THE CHI-SQUARE TEST

- Frequency data (proportions)
- Tests the difference between 2 or more years
- Samples independent

The chi-square test is used to analyze frequency or proportion data. Examples are individual quadrats used as presence/absence sampling units and point-intercept cover data when individual point intercepts are the sampling units.[15] If the frequency data are collected on more than one

[15]*Even though cover is expressed as a percentage, cover data are appropriately analyzed by calculating mean values except when individual points are the sampling units.*

species, each species is usually analyzed separately. Another alternative is to lump species into functional groups such as annual graminoids and analyze each of the groups.

A 2 × 2 Contingency Table to Compare 2 Years

To estimate the frequency of a beetle species in 2 separate years, we have taken two independent random samples of 400 quadrats each.[16] In each of these quadrats the species is either present or absent. For analysis we put these data into a 2 × 2 contingency table, as follows:

	2000	2005	TOTALS
Present	123 (0.31)	157 (0.39)	280 (0.35)
Absent	277 (0.69)	243 (0.61)	520 (0.65)
Total	400 (1.00)	400 (1.00)	800 (1.00)

In the contingency table above, 123 quadrats (31%) contained the beetle in 2000 and 157 of the 400 quadrats contained the beetle in 2005. The numbers in parentheses are frequencies of occurrence in 2000 and 2005, and, in the last column, for both years combined. The chi-square test is conducted on actual numbers of quadrats, not percentages. The chi-square test is not appropriately applied to percentage data.

Just as for the *t*-test and ANOVA, we must formulate a null hypothesis. Our null hypothesis states that the true proportion of the target beetle species (the proportion we would get if we placed all of the quadrats of our particular size that could be placed in the sampled area) is the same in both years. This is equivalent to saying there has been no change in the proportion of the species from 2000 to 2005.

Before we can calculate the chi-square statistic we must determine the values that would be expected in the event there was no difference between years. The total frequencies in the right hand column are used for this purpose. Thus, in both 2000 and 2005, 0.35 × 400 quadrats, or 140 quadrats, would be expected to contain the species, and in both 2000 and 2005, 0.65 × 400 quadrats, or 260 quadrats, would be expected to not contain the species. The following table shows these expected values:

	2000	2005	TOTALS
Present	140	140	280
Absent	260	260	520
Total	400	400	800

Now we can compute the chi-square statistic as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where

χ^2 = The chi-square statistic
Σ = Summation symbol
O = Number observed
E = Number expected

---

[16] *The number of sampling units does not have to be identical. Simply use the actual numbers in the tables and calculations.*

Applying this formula to our example, we obtain the following:

$$\chi^2 = \frac{(123\text{-}140)^2}{140} + \frac{(277\text{-}260)^2}{260} + \frac{(157\text{-}140)^2}{140} + \frac{(243\text{-}260)^2}{260}$$

$$= 2.06 + 1.11 + 2.06 + 1.11 = 6.34$$

We then compare the chi-square value of 6.34 with a table of critical values of the chi-square statistic[17] to see if our chi-square value is sufficiently large to be significant.[18] The *P* value we have selected for our threshold before sampling began is 0.10. Now we need to determine the number of degrees of freedom. For a contingency table, the number of degrees of freedom, v, is given by the following:

$$v = (r - 1)(c - 1)$$

where r = Number of rows in the contingency table
c = Number of columns in the contingency table

For a 2 × 2 table $v = (2-1)(2-1) = 1$. Therefore, we enter the table at degrees of freedom = 1, and the *P* threshold of 0.10. The critical chi-square value from the table is 2.706. Since our value of 6.34 is larger than the critical value, we reject the null hypothesis of no difference in frequency of the beetle species and conclude there has been an increase in its frequency. We would also report our calculated *P* value, which we could interpolate from the chi-square table, but could obtain more easily through a statistics program.[19] For this example, the *P* value is 0.012.

Statistics texts differ on whether to use the chi-square statistic as calculated above in the special case of a 2 × 2 contingency table. Some authors (e.g., Zar 1999) state this value overestimates the chi-square statistic and recommend that the Yates correction for continuity be applied to the formula as follows:

$$\chi^2 = \sum \frac{(|O - E| - \frac{1}{2})^2}{E}$$

Other authors (e.g., Steel and Torrie 1980; Sokal and Rohlf 1994) point out that the Yates correction is overly conservative and recommend against its use. Salzer (unpublished data) has shown through repeated sampling of simulated frequency data sets that the Yates correction is not needed. Munro and Page (1997) point out that the Yates correction is required only when the expected frequency of one of the cells in the table is less than 5. With the proper selection of frequency quadrat size (see Chapter 12) this should rarely occur in monitoring studies. Accordingly, we recommend calculating χ^2 without the Yates correction.

Statistical packages for personal computers calculate the chi-square statistic and give exact *P* values. For 2 × 2 tables, however, you should determine whether the program applies the Yates correction factor. Some programs such as SYSTAT give both the uncorrected and corrected chi-square values. Other programs such as STATMOST give only the corrected chi-square value. Because you want the uncorrected chi-square value, this presents a problem for 2 × 2 tables; no program applies the correction to larger tables.

---

[17] *See any basic statistics textbook for a table, or go to our web site (see Preface) to find links to online tables.*

[18] *If we have sampled more than 5% of the population, we should apply the finite correction factor to the chi-square test. This increases the chi-square statistic and gives us greater power to detect change. Instructions on how to do this are in the section on the finite population correction factor.*

[19] *See our web site for shareware and freeware programs that provide actual* P *values associated with a particular chi-square statistic.*

Larger Contingency Tables for More Than 2 Years

When you have more than 2 years of data to compare, you can increase the size of the contingency table accordingly. For 3 years of data, you would use a 2 × 3 table; for 4 years, a 2 × 4 table; and so on. The chi-square statistic is computed according to the directions given above for a 2 × 2 table. You also must calculate the degrees of freedom according to the directions given above when using a table of critical values of the chi-square statistic. Because there will never be more than two rows (present and absent), the number of degrees of freedom will always be 1 fewer than the number of years. Thus, for a 2 × 3 table, there are 2 degrees of freedom; for a 2 × 4 table, there are 3 degrees of freedom; and so on.

It is important to realize that, just as for an ANOVA, a significant result in a chi-square table larger than 2 × 2 indicates only that the frequency in at least 1 year is significantly different than expected. Which year(s) are different cannot be determined without further testing. This can be done by subdividing the larger contingency table into smaller 2 × 2 tables. Because this involves making multiple comparisons on the same set of data, however, the Bonferroni correction to the *P* value must be made before running these tests (directions on the use of the Bonferroni correction are given above).

PERMANENT QUADRATS, TRANSECTS, AND POINTS: THE USE OF PAIRED-SAMPLE SIGNIFICANCE TESTS

Independent Versus Paired Samples

Thus far we have discussed significance tests for independent samples. Independent samples are ones in which different sets of sampling units are selected randomly (or systematically with random starts) in each year of measurement. Now we will consider the case in which sampling units are randomly selected only in the first year of measurement. The sampling units are then permanently marked, and the same (or at least approximately the same) sampling units are measured in the subsequent monitoring year.

Because the two samples are not independent (the second sample depends on the first), the independent-sample significance tests discussed previously are not appropriate. Instead, a paired-sample significance test is used.

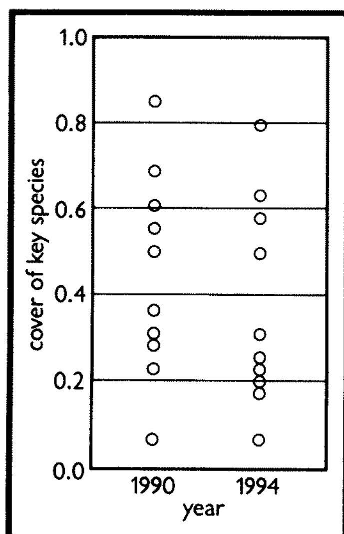


Figure 9.9. Cover estimates (in percent) for 1990 and 1994. Data from 10 permanent transects of 50 points each.

Paired *t*-Test: Use it When You Can

- Compare 2 years
- Sampling units permanent, samples are not independent
- Measurement data
- Ranked data when the ranks are of similar magnitude
- Frequency data grouped in transects or clusters

The appropriate significance test for two paired samples is the paired *t*-test (unless the samples are proportions, in which case McNemar's test, discussed below, is the test to use). There is often a great advantage to testing change using a paired *t*-test rather than an independent-sample *t*-test. This is because the paired *t*-test is often much more powerful in detecting change. To see why this is so, let us examine Figures 9.9 and 9.10 (adapted from Glantz 1997).

The data depicted in Figure 9.9 are cover estimates (in percent) for 10 transects in 1990 and 1994. The estimates were derived by placing 50 point intercepts at systematic intervals along a line (transect), recording whether

the target plant species was present or absent, and reporting a total cover for the species on the transect. For example, if 16 out of the 50 points on the transect were "hits" on the target species, the cover for that transect is 16 divided by 50, or 0.32.

The spread of cover estimates for both years is great, ranging from 0.06 to 0.86 in 1990 and from 0.06 to 0.80 in 1994. As might be expected from this variability, the estimates of the mean—0.44 for 1990 and 0.38 for 1994—are not very precise: the 95% confidence interval for 1990 is 0.27 to 0.61 and for 1994 is 0.21 to 0.55. Not surprisingly, an independent-sample *t*-test run on these samples results in a conclusion of no change. (The calculated *t* value is 0.617 and the actual *P* value is 0.55. This is not statistically significant at all.)

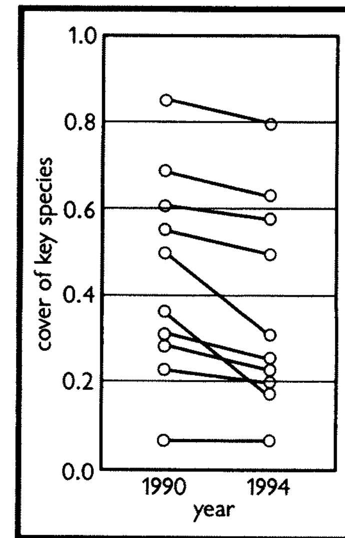


Figure 9.10. Cover estimates (in percent) for 1990 and 1994. Same data as in Figure 9.9 but by focusing on changes in each permanent transect, you can detect a change that was masked by the variability between transects obvious in Figure 9.9.

Consider now, however, Figure 9.10. These are the same data as shown in Figure 9.9, but now we can see that the same transects were measured in 1994 as in 1990. (The transect beginning and ending points were permanently marked in 1990, a measuring tape laid between the two points, and 50 cover points read at systematic intervals along the tape. In 1994 the same procedure was used, with the same transect locations and the same systematic interval of cover points read.) Thus, an independent-sample *t*-test is not appropriately applied to these data, because the 1994 sample is not independent of the 1990 sample. Each of the 10 transects read in 1990 is paired with one read in 1994.

Even if we could conduct an independent-sample *t*-test, we would not want to. To see why, notice that the cover values in 9 of the 10 paired transects have gone down between 1990 and 1994. A paired *t*-test ignores the between-transect variability in both years and looks at only the differences between the 1990 and 1994 values for each of the transects. Conducting a paired *t*-test on these same data results in a highly significant difference between years (the calculated *t* value is 3.34 and the actual *P* value associated with this is 0.009).

The message is clear: if you are interested only in documenting change, as is often the case in monitoring studies, paired *t*-tests are more powerful than independent-sample *t*-tests as long as the pairs of sampling units are correlated (i.e., a sampling unit with a large value the first year is likely to have a large value the second year, while a sampling unit with a small value the first year is likely to have a small value the second year; Zar 1999). The degree of correlation is measured by means of a correlation coefficient (see Zar 1999 for instructions on calculating a correlation coefficient; virtually all statistical programs and most spreadsheets will perform the necessary calculations). The closer the correlation coefficient is to 1.0, the higher the correlation between samples (a value of 1.0 represents perfect correlation; a value of 0 represents no correlation). The paired samples illustrated in Figure 9.10 have a correlation coefficient of 0.96. The test increases in power as the degree of correlation increases. Even if you do not know the degree of pairwise correlation, a paired *t*-test is still valid (Snedecor and Cochran 1989). The lower the correlation, however, the less the advantage of the paired *t*-test over the independent-sample *t*-test (the latter test is invalid for a design that measures the same sampling units in both years of measurement).

Just as for other significance tests, you should apply the finite population correction factor to the paired *t* statistic if you have sampled more than 5% of the population. The next section of this chapter provides instructions on how to do this.

Repeated-Measures Analysis of Variance

- Compare 3 or more years
- Sampling units permanent, samples are not independent
- Measurement data

- Ranked data when the ranks are of similar magnitude
- Frequency data grouped in transects or clusters

For 3 or more years of measurements on the same sampling units, there is a test analogous to the independent-sample analysis of variance discussed above. The test is the repeated-measures analysis of variance. An excellent introduction to the procedure, as it is used in medical experiments, can be found in Glantz (1997). Most statistical programs perform this test.

The repeated-measures ANOVA may not be the best choice in monitoring studies. One problem is the series of statistical decisions that must be made before tests of significance are calculated (Krebs 1998; Barcikowski and Robey 1984). One of the important assumptions of the repeated measures ANOVA is that the correlations between pairs of data for all the years analyzed are the same (Zar 1999). In other words, the correlation between the data of Year 1 and Year 2 is the same as that between Year 2 and Year 3, as well as that between Year 1 and Year 3, and so on. This condition of equal correlations is known as sphericity. Depending on the type of permanent sampling unit employed, this can be a problem. If the sampling unit is a quadrat, and the boundaries of the quadrat are permanently marked, this is less likely to be a problem than if the sampling unit is a line of point intercepts for cover estimation. Even if the endpoints of this line are permanently marked, and one takes care to place the point intercepts in the same place in each year of measurement, because the points themselves are not permanently marked there is more room for error. Therefore, correlations may not be the same between each year of measurement and the assumption of sphericity could be violated.

Just like the situation with ANOVA for independent samples, a significant result from the repeated-measures ANOVA indicates only that 1 or more years differ from each other, not which of these years is different. Most multiple comparison tests for an independent-sample ANOVA such as Tukey's test are not valid for the repeated-measures ANOVA.

We recommend using paired *t*-tests to compare pairs of years, instead of using the repeated measures analysis of variance. If, however, you compare more than 2 years you will need to apply the Bonferroni adjustment to your threshold *P* value. Let us assume that we have decided on a threshold *P* value of 0.20, meaning that if the paired *t*-test results in a calculated *P* value less than 0.20, we will conclude a change has taken place between the 2 years tested. We take measurements in permanent quadrats for 3 years. If we compare only Year 3 with Year 1 or only Year 3 with Year 2, then no correction to the *P* value of 0.20 is required; a calculated *P* value less than 0.20 would lead to a decision of significance. If, however, we compare Year 3 with Year 1 and Year 3 with Year 2, we need to adjust our threshold *P* value by dividing it by the number of comparisons we are making. In this case, we are making two comparisons so our threshold *P* value is $0.20/2 = 0.10$. If either of these comparisons results in a calculated *P* value less than 0.10, we can declare a significant difference.

McNemar's Test

- Compare 2 years
- Sampling units permanent, samples are not independent
- Frequency data

Frequency data, when individual quadrats are the sampling units, may also be analyzed as paired data using McNemar's test. The data are arrayed in a 2×2 table, similar to the contingency table discussed previously.

McNemar's test is used instead of chi-square to test for a difference in proportion between years when the same sampling units are measured each year. Unlike the chi-square test, McNemar's test is useful for comparing only 2 years; it cannot be used for more than 2 years.

Pairing of frequency quadrats can be accomplished by permanently marking quadrats the first year and resampling them the next year. This can be accomplished by positioning quadrats

systematically (with a random start) along randomly positioned permanent transect lines. Care must be taken, however, to permanently mark not only both ends of each transect, but intermediate points in between, and to stretch the tape to approximately the same tension at each time of measurement. You must then ensure that quadrats are placed at the same position along each transect in each year of measurement. It helps if in the first year at least two corners of each quadrat are marked with inexpensive markers such as long nails. See Chapter 5 for more information on monuments for permanent sampling units, and Chapter 12 for a discussion on permanent frequency quadrats for sampling vegetation.

Just as with the paired *t*-test, McNemar's test can be applied regardless of the level of correlation between the pairs of measurement, but the power of the test increases with the degree of correlation (J. Baldwin, 1996). When the degree of correlation between sampling units is high, the use of McNemar's test can be much more powerful in detecting change than the chi-square test used on independent samples. The following example illustrates this.

Let us first look at the situation with temporary frequency quadrats, where we decide to measure change in frequency by randomly locating 100 quadrats in a macroplot in each of 2 years. We decide that our *P* value for significance is 0.10. In the first year, 60 of the quadrats have one or more individuals of Species X in them. In the second year, 50 of the quadrats have Species X in them. The analysis in this case is a typical 2 × 2 contingency table using the chi-square statistic. The null and alternative hypotheses are as follows:

H_o: The proportion of quadrats containing Species X is the same in both years of measurement.

H_a: The proportion of quadrats containing Species X is not the same in both years of measurement.

Here is the contingency table:

	YEAR 1	YEAR 2	TOTALS
Present	60	50	110
Absent	40	50	90
Totals	100	100	200

A chi-square analysis of these data gives the following:

Chi-square statistic = 2.020
P value = 0.155

The observed change of 10 fewer quadrats is not significant at *P* = 0.10. We therefore do not reject the null hypothesis that the proportion of quadrats containing Species X is the same in both years of measurement.

If we decide to permanently mark 100 quadrats (we could either actually mark all 100 quadrats or mark the ends and intermediate locations of several transects and systematically place the quadrats at the same points along tapes in each year of measurement), our null and alternative hypotheses are set up exactly the same way they were in the case of temporary quadrats:

H_o: The proportion of quadrats containing Species X is the same in both years of measurement.

H_a: The proportion of quadrats containing Species X is not the same in both years of measurement.

Just as before, we decide on a *P* value of 0.10 as our threshold of significance. In this case, however, we are going to either accept or reject the null hypothesis based on what happens in permanently established quadrats.

In the first year we find that Species X is found in 60 of the quadrats. In the second year we measure the same 100 quadrats and find that 10 of the 60 quadrats that contained the species the first year no longer contain the species. We also find that the 40 quadrats that did not contain the species the first year still did not contain the species in the second year. A 2 × 2 table set up for a McNemar analysis is shown below. Note the difference between this table and the contingency table given above: 1) the cell values total only 100, instead of 200 as in the contingency table; and 2) the years are not independent of one another (consequently, the values in the cells represent quadrats that meet both row and column requirements: 50 quadrats had Species X present in both Year 1 and Year 2, 40 had Species X absent in both years, 10 had Species X present in the first year but absent in the second, and no quadrats with Species X absent in the first year had it present in the second).

		Year 1	
		Present	Absent
Year 2	Present	50	0
	Absent	10	40

McNemar's test ignores the quadrats that responded in the same way each year. Thus, the 50 quadrats with Species X present in both years and the 40 quadrats with Species X absent in both years are ignored.

$$\chi^2 = \frac{(|\,AP - PA\,| - 1)^2}{AP + PA}$$

where

AP = The number of quadrats in which the species was absent in the first year and present in the second

PA = The number of quadrats in which the species was present the first year and absent the second

Here are the results of McNemar's test on these data:

McNemar chi-square statistic = 8.1000
P value = 0.0044

The calculated *P* value is well below our threshold *P* value of 0.10. We therefore reject the null hypothesis of no change. Even though only 10 quadrats went from containing the plant to not containing it, we have determined this to be significant, something we would not have done if we measured temporary quadrats in each year.[20]

You can also calculate a confidence interval around the change in frequency (in this example from 60% to 50%). If this confidence interval contains zero, you know that at the selected confidence level the difference between the two frequencies is not significant. Appendix III shows you how to do these calculations.

---

[20]*If we have sampled more than 5% of the population, we should apply the finite population correction factor to the McNemar test. This increases the McNemar chi-square statistic and gives us greater power to detect change. Instructions can be found in the next section on the finite population correction factor. We present the continuity-corrected version of the McNemar's test. This is usually conservative, decreasing the probability of a false change error, but increasing the probability of a missed change error. You can eliminate the continuity correction by changing the numerator in the equation to* $|AP-PA|^2$. *Statistical analysis programs differ in use of the continuity correction.*

APPLYING THE FINITE POPULATION CORRECTION FACTOR TO THE RESULTS OF A SIGNIFICANCE TEST

If you have sampled more than 5% of an entire population, then you should apply the finite population correction factor (FPC) to the results of a significance test.[21] The formula for the FPC is 1 – (n/N). The procedure for applying the FPC depends on the nature of the test statistic. For tests that use the *t* statistic, the procedure involves dividing the *t* statistic from a significance test by the square root of the FPC. For tests involving the chi-square (χ^2) and *F* statistics, the procedure entails dividing the χ^2 or *F* statistic from a significance test by the FPC itself (not by its square root). The following examples illustrate the procedure for significance tests that use the *t*, χ^2, and *F* statistics.

Tests That Use the *t* Statistic

Independent-sample and paired *t*-tests calculate the *t* statistic, which is compared with the critical value of *t* from a *t* table (see most basic statistics books and our web page for online locations) for the appropriate degrees of freedom and the threshold *P* (α) value. If the calculated *t* value is larger than the critical *t* value, the null hypothesis of no change is rejected in favor of the alternative hypothesis that a change has taken place. The formulas for independent-sample and paired *t*-tests given earlier in this chapter do not include the FPC. Computer programs also do not apply the FPC to their calculated *t* values. If you have sampled more than 5% of the population, you should correct the calculated *t* statistic by applying the FPC as in the following example. This will increase the size of the *t* statistic, resulting in greater power to detect change.

Let us say that the *t* statistic from a *t*-test (either an independent-sample or paired *t*-test) is 1.645 and in each of 2 years you sampled n = 26 quadrats out of a total of N = 100 possible quadrats. The FPC is applied as follows:

$$t' = \frac{t}{\sqrt{1 - (n/N)}} \qquad t' = \frac{1.645}{\sqrt{1 - (26/100)}} = 1.912$$

where t = The *t* statistic from a *t*-test.
t' = The corrected *t* statistic using the FPC.
n = The sample size (the number of quadrats sampled in each year; note that you do not add the number of quadrats sampled the first year to the number of quadrats sampled in the second year).
N = The total number of possible quadrat locations in the population.

To calculate *N*, determine the total area of the population and divide by the area of each individual quadrat.

Following this calculation, you need to look up the *P* value of t' = 1.912 in a *t* table at the appropriate degrees of freedom.[22] If this is an independent-sample *t*-test, the appropriate number of degrees of freedom would be $(n_1 - 1) + (n_2 - 1) = (26 - 1) + (26 - 1) = 50$. If this is a paired *t*-test, the values analyzed are the observed changes in each permanent quadrat. Since there are 26 permanent quadrats, n = 26, and the appropriate number of degrees of freedom is $n - 1 = 26 - 1 = 25$.

---

[21] *Population as used here refers to statistical population. In the context of the types of monitoring addressed by this handbook, the FPC would be applied only to significance tests on data collected using quadrats. This is because there is a finite population of quadrats that can be placed in the area to be sampled (assuming quadrats are positioned as they should be to avoid any overlap). The FPC should never be applied to significance tests on line or point-intercept data because a population of lines and points is by definition infinite.*

[22] *A* t *table can be found in most basic statistics texts. See our web page (address in Preface) for links to online statistical tables.*

Looking up *P* values in a *t* table is difficult and inexact because it requires you to interpolate between values in the table. A more exact and convenient method is to use a computer program (shareware programs available from links on our web page).

Tests That Use the Chi-Square Statistic

The chi-square (χ^2) statistic is used to test the difference between years in a proportion when using temporary sampling units. McNemar's test, which tests the difference between 2 years in a proportion using permanent sampling units, also makes use of the chi-square statistic. In both cases, the chi-square statistic calculated using standard formulas and computer programs should be corrected using the FPC if you have sampled more than 5% of the population.

For example, the χ^2 statistic from a particular test is 2.706 and you sampled $n = 77$ quadrats out of a total of $N = 300$ possible quadrats. The FPC would be applied as follows:

$$\chi^{2\prime} = \frac{\chi^2}{1-(n/N)} \qquad \chi^{2\prime} = \frac{2.706}{1-(77/300)} = 3.640$$

where χ^2 = The statistic from a chi-square test or McNemar's test.
$\chi^{2\prime}$ = The corrected χ^2 statistic using the FPC.
n = The sample size (the number of quadrats sampled in each year; note that you do not add the number of quadrats sampled the first year to the number of quadrats sampled in subsequent years).
N = The total number of possible quadrat locations in the population.

To calculate *N*, determine the total area of the population and divide by the area of each individual quadrat.

Following this calculation, you need to look up the *P* value of $\chi^{2\prime} = 3.640$ in a χ^2 table at the appropriate degrees of freedom.[23] For McNemar's test, which can be used only to test for a difference between 2 years, there is always 1 degree of freedom. For a chi-square test applied to a contingency table, the number of degrees of freedom is always one less than the number of years being compared. Thus, for a 2 × 2 table comparing 2 years there is 1 degree of freedom, for a 2 × 3 table comparing 3 years there are 2 degrees of freedom, and so on.

Looking up *P* values in a χ^2 table is difficult and inexact because it requires you to interpolate between values in the table. A more exact and convenient method is to use a computer program (see our web site for links to shareware programs that do this).

Tests That Use the *F* Statistic

The analysis of variance and the repeated-measures analysis of variance use the *F* statistic to determine if 1 or more of the years sampled is different from the other years. The *F* statistic can also be corrected by the FPC. This is accomplished as illustrated in the following example.

An analysis of variance calculated by a computer program yields an *F* statistic of 3.077. In each of 3 years you sampled n = 50 quadrats out of a total of $N = 400$ possible quadrats. The calculated *F* is corrected as follows:

$$F' = \frac{F}{1-(n/N)} \qquad F' = \frac{3.077}{1-(50/400)} = 3.517$$

where F = The *F* statistic from an analysis of variance or a repeated-measures analysis of variance.
F' = The corrected *F* statistic using the FPC.
n = The sample size (the number of quadrats sampled in each year; note that you do not add the number of quadrats sampled the first year to the number of quadrats sampled in subsequent years).

---

[23] *A* t *table can be found in most basic statistics texts. See our web page (address in Preface) for links to online statistical tables.*

N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the area of each individual quadrat.

Following this calculation, you need to look up the P value of $F' = 3.517$ in an F table at the appropriate degrees of freedom. Looking up P values in an F table is difficult and inexact because it requires you to interpolate between values in the table. A more exact and convenient method is to use a computer program (see our web page for links to shareware programs that do this).

INTERPRETING THE RESULTS OF SIGNIFICANCE TESTS

A significance test is conducted when your management objective is to detect change from one period to another in some average value (such as a mean or proportion). Once that test has been performed, you must now interpret the results from the test. Figure 9.11 is a flow chart to help you in your interpretation. Interpretation entails answering the following questions:

Is there a statistically significant result? What is the likelihood that no true change occurred and that any observed difference is simply the result of random sampling errors?

The P value calculated from the significance test gives you the answers to these two questions. A threshold P value should be set before conducting the significance test so that the P value from the test can be assessed relative to the threshold. If the P value from the test is smaller than the threshold, it is considered "significant" and the null hypothesis of no-change is rejected in favor of the hypothesis that a change did actually take place. If the P value from the test is larger than the threshold, it is considered "nonsignificant" and the null hypothesis of no-change is not rejected. The P value calculated from the significance test is the likelihood that the observed difference is the result of chance sampling errors (false-change error).

Does the observed magnitude of change have any biological significance?

Given a large enough sample size, a statistical test can find even an extremely small difference between two populations to be significant. It is unlikely that any two populations or the same population over any two time periods will ever be exactly the same. Therefore, it is important that you determine whether a statistically significant change has any biological significance. People often get fixated on the idea of statistical significance. A helpful exercise is to pretend the

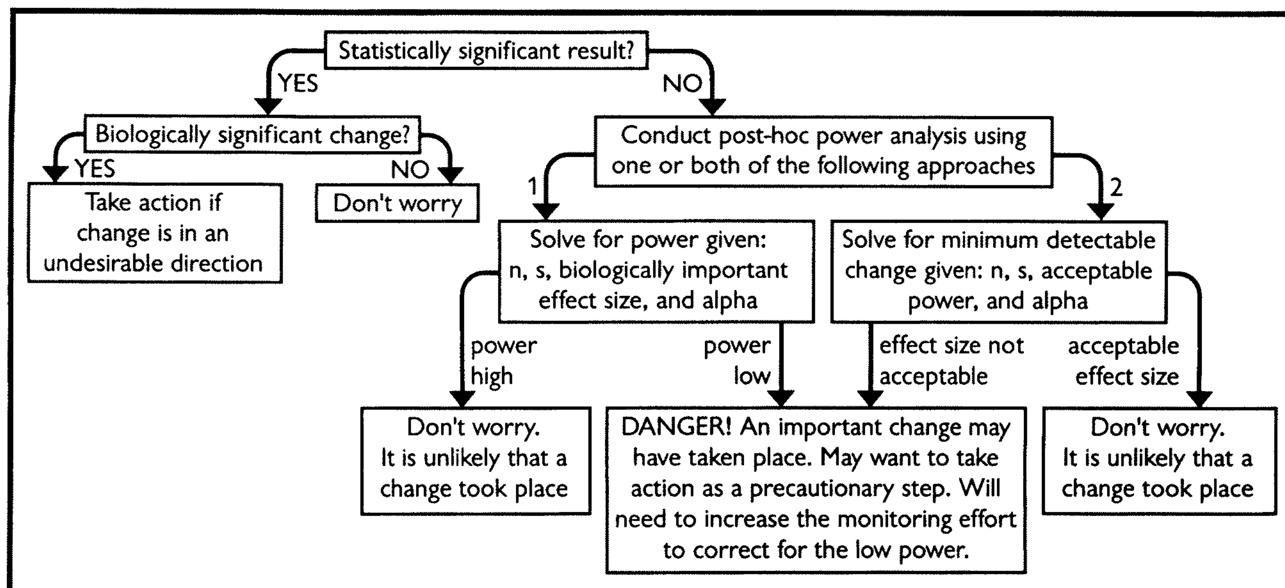


Figure 9.11. Interpreting the results from a statistical test comparing change over time.

difference observed through sampling is the true difference (i.e., pretend you conducted a complete census at each sampling period). Now ask yourself what action you will take if this observed difference is in fact the true difference. If your answer to this is that you would take no action, then the observed change, even though statistically significant, is not biologically significant.

If the test yields a nonsignificant result, what is the probability that a biologically important change actually occurred?

If your study results in a conclusion that an observed change is not significant, your interpretation is not complete until you have conducted a post hoc power analysis. The post hoc power analysis tells you the probability of your test failing to detect a true change (i.e., committing a missed-change error). Following are two approaches you can take in conducting this power analysis. Both of them are easy to do if you have a computer program developed for this purpose.[24]

1. Calculate a power value. This is option 1 on the right side of Figure 9.11. Using this approach, you plug in your sample size, the sample standard deviation, the threshold significance level (α) you have chosen for the significance test, and an effect size you consider biologically important. A power value is then calculated.[25] If the resulting power value is high, then it is unlikely that a change took place. If the resulting power value is low, then a biologically important change may have taken place. You need to improve your monitoring design immediately to ensure that you can detect the level of change you believe is biologically important. If, based on ancillary information, you have reason to believe a deleterious change may have taken place, you may need to take action as a precautionary step until the monitoring design can be improved to address the low power issue.
2. Calculate the minimum detectable change (MDC). This is option 2 on the right side of Figure 9.11. This approach requires you to plug in the values for sample size, sample standard deviation, threshold significance level (α) for the test, and an acceptable level of power. The program then solves for the minimum detectable change (MDC) that can be detected. If the MDC is smaller than the size of change deemed to be biologically important, then it is unlikely that the specified MDC actually occurred. If, however, the MDC is larger than this biologically important change, then an important change may have taken place. You need to improve your monitoring design immediately to ensure that you can detect the level of change you believe is biologically important. If, based on ancillary information, you have reason to believe a deleterious change may have taken place, you may need to take management action as a precautionary step until the monitoring design can be improved to address the inability to detect a change deemed to be biologically important.

Figure 9.12 illustrates a post hoc power analysis, comparing 2 years of density data for *Lomatium cookii* at the Agate Desert Preserve in Oregon. Note that even though the significance test yielded a nonsignificant result, we cannot be confident that no change has taken place. This is because of the extremely low power (0.13) of the test to detect the 30% change we have determined to be biologically significant. Note also that the minimum detectable change is 155%; we could lose our entire population and not detect the loss! Although we may want to take action as a precautionary step, we very definitely want to improve the study design to reduce the standard deviation (in this case our standard deviation is more than twice the size of the mean, a very undesirable trait indeed).

---

[24]*See our web site (address in Preface) for links to shareware or freeware programs that do this.*

[25]*A threshold power value should be set in advance so a decision can be made whether the power value calculated through post hoc power analysis is considered high or low.*

Results of a statistical analysis comparing 1989 and 1990 data on *Lomatium cookii* from the Agate Desert Preserve. False-change threshold value = 0.10. Desired magnitude of change is 30% from the 1989 value.

sample size	sample statistics				observed change (percent)	results of a statistical test (P)	calculated power ($1-\beta$) to detect a 30% change from the 1989 mean	minimum detectable change size with a power of 0.9, $\alpha = 0.10$, (% change from 1989)
	1989		1990					
	mean	sd	mean	sd				
50	3.12	11.16	1.30	2.92	1.82(58%)	0.85	0.13	4.82 (155%)

INTERPRETATION: cannot conclude that a change took place (cannot reject the null hypothesis). Low confidence in the results due to low power and high minimum detectable change size. May want to take action as a precautionary step and make changes in the monitoring design to increase power.

Figure 9.12. Example of a post hoc power analysis comparing two years of density data for *Lomatium cookii* at the Agate Desert Preserve in Oregon.

GRAPHING THE RESULTS OF DATA ANALYSIS

Graphs are important tools for displaying the results of data analysis and helping the investigator (as well as others) interpret the meaning of these data. When, as is usually the case, summary statistics such as a mean, total, or proportion are displayed, error bars must be used to display the precision of the estimate.[26] Commonly encountered error bars are the sample standard deviation, the sample standard error, and confidence intervals (such as a 90% or 95% confidence interval). Because it is the true parameter (mean, total, or proportion) that is of interest, we recommend that you use only confidence intervals as error bars. You must clearly state what error bar you are using, as well as the sample size on which the estimate and measure of error is based (Ellison 1993).

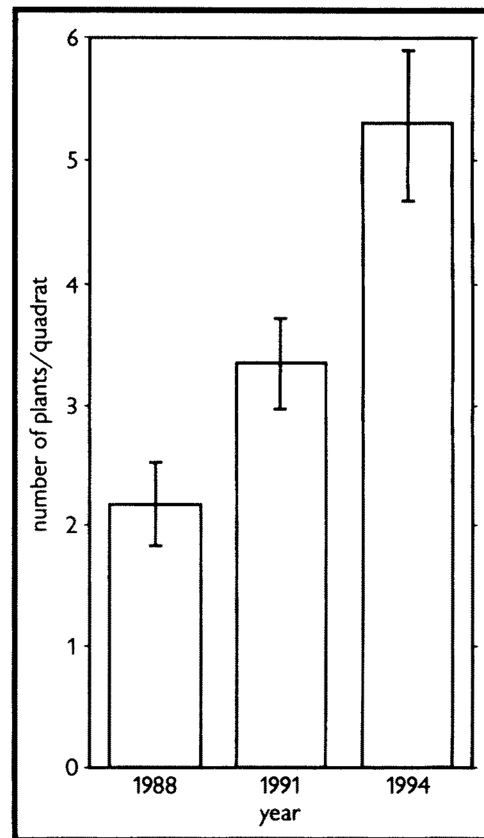


Figure 9.13. Bar chart of mean number of plants of the key species per 0.5m × 4.0m quadrat. Error bars are 90% confidence intervals. In each year n = 100.

Types of Graphs

Bar charts are commonly used to display the results of data analysis. They should not be confused with histograms. A histogram shows the density (or frequency) of the values occurring in the data set between the lower and upper bounds of each bar, whereas a bar chart is used to illustrate some summary measure (such as the mean, total, or percentage) of all the values within a given category such as the year of measurement (Ellison 1993).

Figure 9.13 is an example of a bar chart showing the results of 3 years of monitoring. Mean density per quadrat of a hypothetical key species is displayed, along with error bars corresponding to 90% confidence intervals. When displaying information about

[26] *This precludes use of pie charts and stacked bar charts, which we neither recommend nor describe in this handbook.*

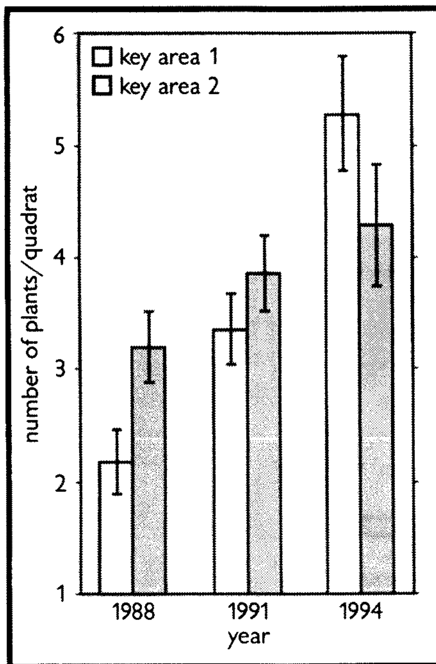


Figure 9.14. Side-by-side bar chart of mean number of plants of the key species per 0.5m × 4.0m quadrat, at key area 1 and key area 2. Error bars are 90% confidence intervals. All bars represent n = 100.

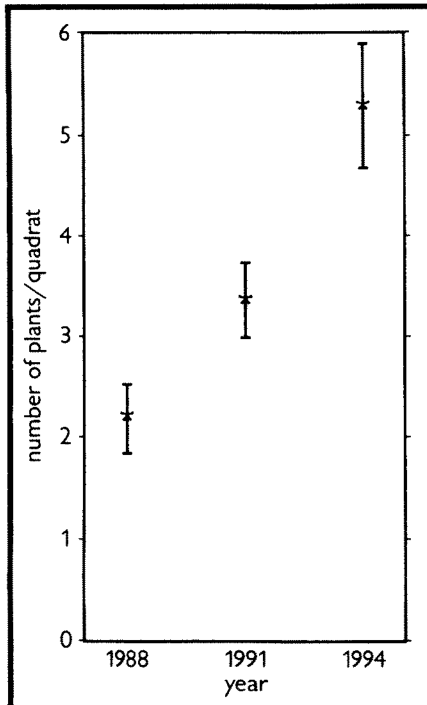


Figure 9.15. Point graph (also called category plot) of same data as shown in Figure 9.13. Error bars are 90% confidence intervals.

more than one summary statistic per year (as, for example, data on the same species measured at two or more key areas per year), side-by-side bar charts (Fig. 9.14) can be used (again, with error bars for confidence intervals). Stacked bar charts should not be used; they are unintelligible and provide no way to display error bars (Ellison 1993).

Point Graphs, with error bars corresponding to confidence intervals, can be used in lieu of bar graphs. An example of such a graph is given in Figure 9.15.

Sometimes lines are connected to each of the points as in Figure 9.16, although this is really unnecessary unless more than one summary statistic is presented in each year. Figure 9.17 illustrates means for two key areas in each year of measurement. Lines are appropriate here to clearly separate the two sets of means. Note, however, that the confidence intervals for the two key areas overlap in 1991, leading to some confusion. A side-by-side bar chart, as shown in Figure 9.14, would provide a clearer representation.

Box plots with "notches" for error bars are often used to display the median with its confidence interval. Box plots were discussed at the beginning of the chapter as a way of exploring your data before or during analysis. They can show the results of analysis, providing error bars for confidence intervals can be displayed. Some statistical packages offer the option to "notch" the box plots at a set confidence interval. Figure 9.18 shows such a notched box plot. These have the advantage of showing summary statistics (in this case the median and its 95% confidence interval), as well as other features relative to the distribution of data points. Note, however, that this option displays the median, not the mean and that the confidence interval is one that includes the true median with 95% probability, not the true mean.

Graphing Summary Statistics When Data Are Paired

Recall that significance tests are often much more powerful when data are collected in permanent plots or along permanent transects. The sampling units (plots or transects) in this case are said to be paired; that is, the data from the second year of measurement depend on the data from the first year of measurement.

Graphic presentation of these paired data is less straightforward than independent samples. Consider the data depicted in Figures 9.9 and 9.10. If one were to simply graph a summary statistic like the mean for each year of measurement, along with confidence intervals computed as if these data were independent, the graph would appear to illustrate no difference between the years 1990 and 1994. Figure 9.19 is a point graph that does just that.

Now consider the point graph shown in Figure 9.20. This graph is constructed with the same data used to produce Figure 9.19, but this time takes advantage of the fact that each of the 10 transects is paired. What is graphed in Figure 9.20 is the mean difference in cover between the paired transects. Because there was a decline in cover from 1990 to 1994 in all but one of the transects, the mean difference is negative. Also plotted is the 95% confidence interval around this mean difference. Because this interval does not include 0 (which would

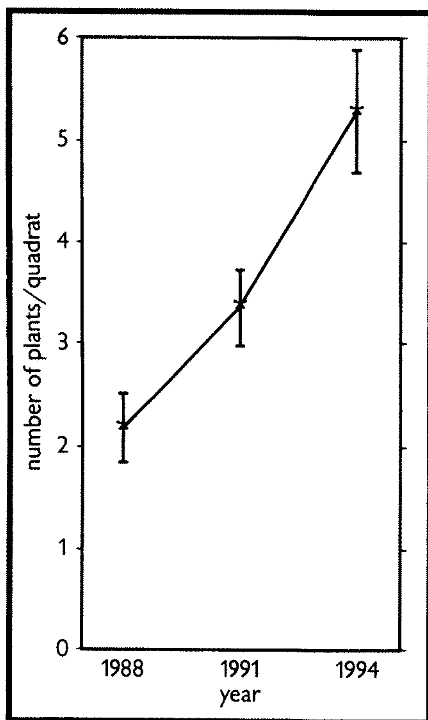


Figure 9.16. Point graph of same data as Figure 9.15, but with lines connecting points.

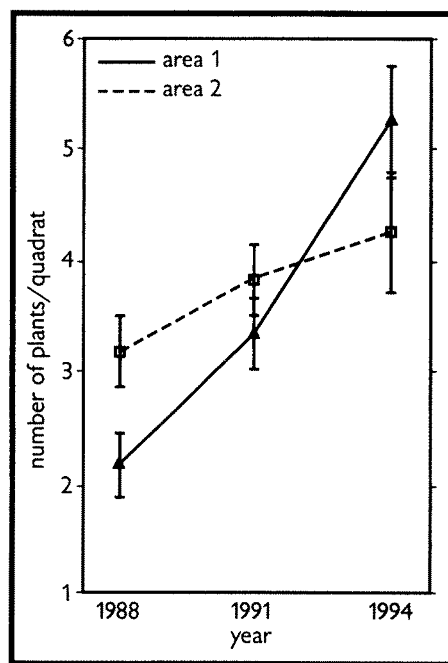


Figure 9.17. Point graph of same data as in Figure 9.14. Lines connect the means from each of the key areas. Error bars are 90% confidence intervals.

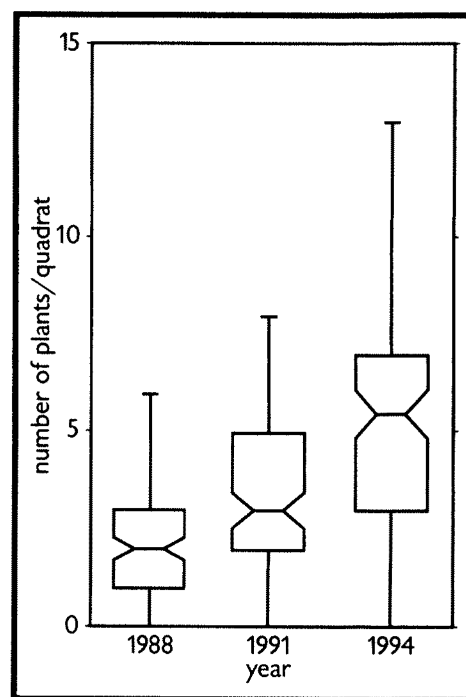


Figure 9.18. Notched box plots of the number of plants/quadrat in samples of one hundred 0.5m × 4.0m quadrats. The points at which the boxes reach full width on either side of the median represent the 95% confidence interval for the median.

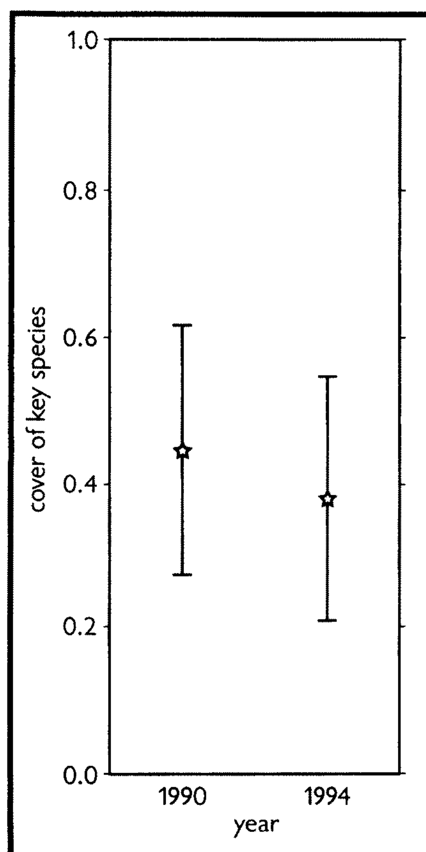


Figure 9.19. Point graph of cover data collected along permanent transects treated as if each year was independent. Error bars are 95% confidence intervals. See text for explanation.

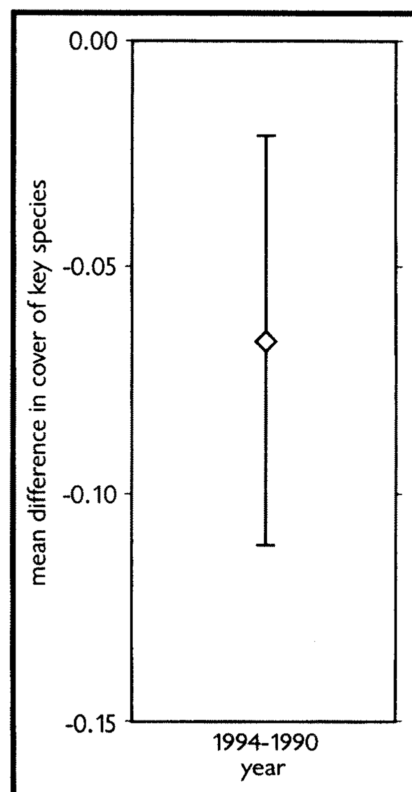


Figure 9.20. Point graph showing mean difference of cover in 10 paired transects of 50 points each. Error bar is 95% confidence interval.

indicate the possibility of no change), this difference is significant at the 95% confidence level (i.e., $P<0.05$).

Bar charts of the mean difference could be similarly constructed. If you are more interested in median differences, you could use a notched box plot of these differences. It is also valuable to plot several mean differences on a single graph. For example, the graph could show points and confidence intervals for each of the differences between 1990–1994, 1994–1998, 1998–2002, and so on.

MANAGEMENT IMPLICATIONS

Monitoring data must be analyzed if the information is to be effectively applied to management decision. Proper analysis is critical to credible reporting of results. The correct analysis depends on the type of management objective. Target/threshold management objectives are analyzed using confidence intervals. Change/trend objectives are analyzed using significance tests. If the test suggests that the change is not significant, a post hoc power analysis should be conducted to determine if an important change may have taken place but was missed by a study design with low power to detect change.

CHAPTER 10
Analysis of Trends

Sitta carolinensis
White-breasted nuthatch
Artist: D. Andrew Saunders

Most monitoring for adaptive management involves comparisons of a site before and after management, that is, an examination of changes that occur in a population or resource between two periods. In some situations, however, monitoring may be performed over longer periods, for example, for plants or animals with extended generation times that respond slowly to management. This is the sort of situation best addressed by looking at counts or measurements made over 5 or more years to ask whether the overall trend of the population is one that is increasing, decreasing, or stable over the long term. Such trends are evaluated with statistics different from those used for comparing just two time intervals. These trend tests and associated monitoring designs are the focus of this chapter.

SIMPLE LINEAR REGRESSION

A common approach to estimate trend is with a linear regression *t*-test, which is widely used and has been shown to be among the most statistically powerful of both parametric and nonparametric methods for detecting trends (Hatfield et al. 1996). To perform a linear regression, population values are plotted against the point in time (usually the year) when the measurements were made. Regression analysis of these data provides an estimate of rate of change or trend of counts with time, a quantity otherwise known as the slope of the regression. If the slope is not statistically different from zero, then a population is assumed to be stable (no trend is evident). Slope estimates significantly less than zero imply a negative trend, that is, that a population decline has occurred, whereas slopes greater than zero indicate a positive trend, that is, that a population increase has occurred. The power of the test is an important issue with this and other trend tests because of the common risk of wrongly accepting the null hypothesis of no trend because too few counts were made or because the counts were too variable to detect a trend that was indeed occurring. Thus, while a series of at least three counts is needed to perform a trend test, in practice it is difficult to reliably detect trends with less than five counts.

Table 10.1. Sample Data of Counts Over Time Typical of That Encountered in Trend Monitoring Studies

YEAR (X)	COUNT (Y)	Y^2	X × Y	X^2
1	2.8	7.8	2.8	1.0
2	3.0	9.0	6.0	4.0
3	4.4	19.4	13.2	9.0
4	4.8	23.0	19.2	16.0
5	6.2	38.4	31.0	25.0
6	6.4	41.0	38.4	36.0
7	6.4	41.0	44.8	49.0
8	7.8	60.8	62.4	64.0
9	8.2	67.2	73.8	81.0
10	9.4	88.4	94.0	100.0
11	9.0	81.0	99.0	121.0
12	10.4	108.2	124.8	144.0
13	10.0	100.0	130.0	169.0
Σ=91	Σ=88.8	Σ=685.2	Σ=739.4	Σ=819.0

Performing a linear regression by hand is not complicated but it is tedious. The requisite equations for doing so, along with a worked example, are provided (Tables 10.1 and 10.2; Fig. 10.1). In a nutshell, linear regression involves identifying the slope of the line that best fits the pattern of the counts versus time. The best fit is that which minimizes the distances between the line and all the data points. The slope is usually denoted as β and indicates the number of individuals being added or lost from the population at each time interval. This is the trend estimate, which is usually reported with a standard error that, in conjunction with an appropriate *t*-value for *n*-2 degrees of freedom (where *n* is the number of counts you have made), can be used to determine whether the trend estimate is different from zero. Conventional spreadsheets and statistical software packages will do the requisite calculations easily and quickly.

Table 10.2. Equations for estimating with least-squares regression the trend and significance of trend in counts over time, as a worked example using data from Table 10.1. The parameters calculated include the total sum of squares (totalSS), the regression sum of squares (regressionSS), the residual mean square (residualMS), the slope or trend estimate (b), and the standard error of the regression coefficient (S_b). We estimate that the counts in Table 10.1 exhibit an increasing trend of 0.65 individuals per year and conclude that the estimate is different from zero.

$$totalSS = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n} = 685.20 - \frac{(88.8)^2}{13} = 78.62769$$

$$regressionSS = \frac{\left(\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}\right)^2}{\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}} = \frac{(117.8)^2}{182} = 76.24637$$

$$residualMS = \frac{(totalSS - regressionSS)}{(n-2)} = (78.62769 - 76.24637)/11 = 0.216484$$

$$b = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}} = \frac{117.8}{182} = 0.647253$$

$$S_b = \sqrt{\frac{residualMS}{\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}}} = \sqrt{\frac{0.216484}{182}} = 0.034489$$

$$t = b / S_b = \frac{0.647253}{0.034489} = 18.7671$$

$$df = n - 2 = 13 - 2 = 11$$

$$t_{0.05(2),11} = 2.201$$

$$P(H_o : b = 0) << 0.001$$

When linear regression analysis of population counts is performed (a very common situation), counts are usually first log-transformed. A common transformation is ln (count + 0.5). The main reason that counts are transformed is that populations generally change over time at a particular rate, for example by 10% per year, rather than by a particular amount, for example, by 10 individuals per year. What this means is that at low numbers populations change slowly (for example, a population of 10 increasing annually by a factor of 10% would increase by one individual the next year), whereas at high numbers populations change more quickly (for example, a population of 100 increasing at the same rate would increase by 10 individuals the next year). This describes what is called an "exponential" change in populations—slow at low levels and fast at higher levels. As its name implies, however, linear regression assumes a constant (linear) relationship, not an exponential change, between counts and time. This problem is dealt with by "linearizing," with a simple log-transformation, the counts from a population changing in an exponential fashion. Unless you are working in a situation where you know the population is changing by a constant amount (for example, a contractor has been asked to remove a fixed number of individuals of an exotic species from a particular site each year, or harvest for a game species has been set at a fixed quota of animals), then a log-transformation of the counts is recommended. Whatever the case, at least make simple plots of your counts against time—if they

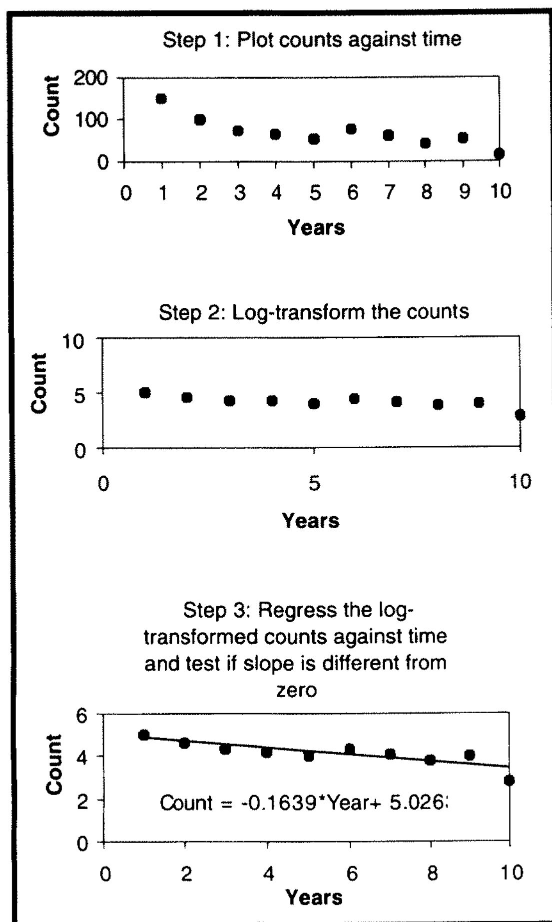


Figure 10.1. Three steps in performing a simple linear regression analysis of counts against time on a single plot or route.

follow an exponential pattern of change, then be sure to use a log-transform on the counts. A second reason that counts should be log-transformed is that it helps to stabilize the amount that they vary across years. An important assumption of linear regression is that the variances of counts are expected to be equal each year. For ecological data, however, count variation is almost invariably related to the size of the count, with smaller counts having smaller variances than larger counts. The log-transformation helps to bring ecological data into conformance with the assumption of constant variances over time.

More advanced approaches to trend analysis on single plots, which permit you to overlook the basic assumption made by simple linear regression that the trends within counts are constant and linear over time, include polynomial regression and additive models, which are elaborated on by Thomas (1996) and James et al. (1996).

ROUTE REGRESSION

Route regression is a variation on linear regression that is frequently used in monitoring for assessing aggregate trends across more than one plot. This approach is actually quite straightforward and addresses a common monitoring design that involves making counts repeatedly in different sampled areas, macroplots, or sites, each of which is termed a "route" (because the method was initially developed for bird surveys which are typically run along transects called routes). Route regression is used in an attempt to make a general inference about population trends across all routes, macroplots, or other sampled areas. In other words, each route may show its own unique trend, but route regression will permit you to obtain a general sense of what is occurring across all the routes you are monitoring. Such is the case, for example, when monitoring a metapopulation in which counts are made within a sample of its constituent subpopulations, or when monitoring with many macroplots dispersed widely in a population.

In route regression, a separate linear regression analysis is made for each route (macroplot, transect, or other sampled area) on which three or more counts were made to estimate the slope or trend on each route. You then combine the route-specific trend estimates to generate a mean slope across them all, as well as a variance in slopes across routes. These two statistics (the mean and variance of slopes across routes) permit you to build a confidence interval and determine whether trends are, in aggregate across your study area, different from zero. In the worked example presented in Table 10.3, 10-year trends in heron colony size (number of nests) were calculated using linear regression for each of nine different colonies. The average trend across these colonies was positive and statistically different from zero, suggesting that a general increase in the heron nesting population had occurred despite declines in some colonies.

Note that with route regression, slope estimates from individual routes can be weighted using the formula for the weighted average. This can be done to emphasize or de-emphasize their importance in the overall trend estimate. Weighting the mean slope by the average abundance on each route is one way to increase the contribution of routes with large populations (and decrease the influence of those with small populations) to the overall trend.

A related approach to analyzing trends across many routes is one that combines both parametric and nonparametric methods. First you determine the proportion of routes (macroplots, transects, or other sampled areas) with slope estimates from a regression analysis (this is the parametric part of the analysis) that are increasing (positive slopes) or decreasing (negative slopes). If populations were stable across a large area, then you would expect about half the routes to show increases and the other half to show decreases. You can test this expectation with a chi-square goodness of fit test. The chi-square procedure, which is the nonparametric part of the analysis, tests whether the sample comes from a population that conforms to the 1:1 distribution that would exist if there was no change in the population (i.e., 50% of the routes showing increases and 50% of the routes showing decreases). Box 10.1 gives an example of how to conduct the test.

Table 10.3. Route Regression Analysis Applied to Simple Data of Nest Counts Made Over 10 Years in Nine Great Blue Heron Colonies on Offshore Islands in Maine.

SAMPLE DATA	
SITE	**TREND IN NEST COUNTS (SLOPE OF REGRESSION OF NEST COUNTS VERSUS YEAR)**
Hardwood Island	+2.3 nests/year
Little Goose Island	+4.3 nests/year
Mark Island	−1.9 nests/year
Stone Island	+9.7 nests/year
Scraggy Island	−2.2 nests/year
Upper Birch Island	+8.6 nests/year
Graffam Island	+12.4 nests/year
Middle Douglas Island	+14.5 nests/year
Eaton Island	−3.3 nests/year
ROUTE REGRESSION ANALYSIS	
Mean trend across all islands	+4.9 nests/year
n	9
Standard error of the mean slope (SE)	2.2240
Degrees of freedom	$n - 1 = 8$
t(∀ = 0.05)	2.31
Mean slope ±tSE	10.071 > μ – 0.204 nests/year
CONCLUSION	
The 95% confidence interval about the mean trend in annual change in nests per colony does include zero; therefore, conclude that nesting populations of Great Blue Herons are not increasing on these islands.	

NONPARAMETRIC TREND TESTS

A nonparametric approach to assessing trends in populations involves rank models, which estimate the tendency for counts to increase or decrease, but do not calculate a rate of increase or decrease. These methods are extremely useful in that they make no assumptions about the distribution of counts. They are limited, however, by the very fact that they provide no inference about the rate a population is changing, indicating only that a population is changing. These nonparametric tests are not as powerful as parametric tests if necessary assumptions for parametric testing are approximately met. Nevertheless, nonparametric trends are the better choice when the data include outlying counts or other skewed patterns within them (Hatfield et al. 1996). Transformations of counts are unnecessary for rank methods because, applied equally to each count, the transformation will not change the rank order of the counts. Rank-based tests generally involve comparisons of all counts made, one pair at a time, to determine the number of positive and negative differences, which would be approximately equal in a stable population but otherwise skewed.

Perhaps the most common rank-based method for detecting trend in single count series is the Mann-Kendall test (Gilbert 1987). This test is based on an iterative approach that compares all possible pairs of counts or measurements in a series. The number of positive differences and the number of negative differences are then summed to produce the Mann-Kendall test statistic, S (ties, or

Box 10.1. *CALCULATION OF CHI-SQUARE GOODNESS OF FIT. DATA CONSIST OF INCREASES AND DECREASES SHOWN IN A SAMPLE OF 25 ROUTES.*

Threshold significance level: 0.05

H_O: *The sample data came from a population having a 1:1 ratio of increasing and decreasing routes*

H_A: *The sample data came from a population not having a 1:1 ratio of increasing and decreasing routes*

The data recorded below are the 25 observed frequencies in each of the two trend categories (increasing or decreasing). The frequencies expected under the null hypothesis are shown in parentheses below the observed values.

	TREND		
	INCREASING	DECREASING	*n*
Observed	18	7	25
Expected	(12.5)	(12.5)	

The computer program SYSTAT computes the following values:

Pearson Chi-Square statistic: 4.840

P *value: 0.0278*

Because the calculated P *value is below our threshold of 0.05, we reject the null hypothesis in favor of the alternative hypothesis that the sample data came from a population not having a 1:1 ratio of increasing and decreasing routes. Since the sample shows more routes increasing than decreasing, we conclude that the population in the sampled area has increased.*

zero differences, are ignored). If S is a negative number, counts tend to be smaller later in the series, and hence a decline is occurring, whereas if S is larger, then an increase is likely occurring.

As an example of the Mann-Kendall test, consider a series of five counts: 10, 12, 14, 13, and 15. The signs of differences between all unique pairs of these counts are: 12-10 is "+", 14-10 is "+", 14-12 is "+", 13-10 is "+", 13-12 is "+", 13-14 is "–", 15-10 is "+", 15-12 is "+", 15-14 is "+", and 15-13 is "+". These yield a value of $S = 8$, that is, nine positive differences minus one negative difference. To determine the significance of the trend, note your sample size (n, or number of counts) and value of S and refer to Figure 10.2 where all possible combinations of S and n are depicted for $n \leq 10$. In this example, $n = 5$ and $S = 8$, which falls slightly below a probability level of 0.05. Therefore, we conclude that a significant, positive trend ($P<0.05$) is occurring in these counts (assuming that our sampling objective specified a false-change error rate of 0.05).

Rank-trend analysis is a nonparametric analog to the route regression method and is useful for the same study designs (many plots or routes monitored for a long period for which one wants to obtain an estimate of overall change across plots). Rank-trend analysis provides a useful, and sometimes more powerful (Thomas and Martin 1996), method for detecting trends in a route regression

situation when the assumptions of linear regression cannot be met. An overview of the rank-trend method is provided by Titus (1990) and Titus et al. (1990).

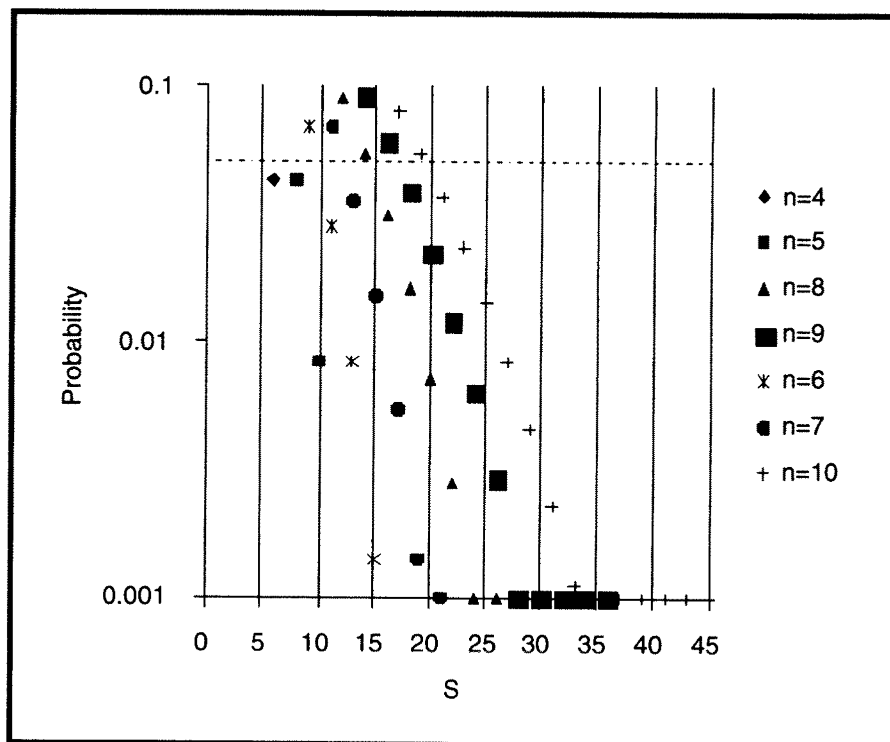


Figure 10.2. Probabilities for the Mann-Kendall trend test for all possible combinations of S and n (number of counts or measurements taken) for $n \leq 10$). To use this table compare your calculated S value (horizontal axis) with n (match with legend), and determine if their combination falls below the critical alpha-levels of 0.1, 0.05 (dotted line), 0.01, or 0.001 indicated on the vertical axis of the graph.

ANOVA-BASED APPROACHES FOR ESTIMATING TRENDS

In some circumstances a comparison of the beginning and ending periods of a series of monitoring data by an analysis of variance (ANOVA) (or *t*-test if only two periods) may provide a more sensitive test of long-term change in a population than will a trend test. This is particularly true for organisms that show abundant short-term variation that can obscure longer-term trends. For example, using an extension of a repeated-measures design, Lesica and Steele (1996) halved their monitoring data between an earlier and later segment and apportioned variation in their counts between plots, within sites, and within years to test whether mean counts differed between the first and second monitoring segments. Using this approach for arctic-alpine plants, Lesica and Steele (1996) estimated high levels of power to detect modest (>20%) variation in numbers of individuals based on a fairly typical sampling design (two sites with 60 microplots each). Although designed originally for plant population monitoring, the approach can be useful for any organisms in which short-term variation in counts can confound detection of longer-term trends.

PLANNING A TREND-ORIENTED MONITORING PROGRAM

For all monitoring involving trend analyses there are some important considerations to make, especially during the planning process. Anyone contemplating a monitoring program over a term longer than 2 years has to make some complicated sampling decisions relating to where to place plots, the number of plots to be monitored, how often to monitor each season, and whether counts should be made each year. Count duration in particular is an important issue. Populations often undergo fluctuations as a result of normal variation in the environment, population cycles, or changing age composition that, for short series of counts (<5 years), can obscure long-term trends. For these reasons detecting trends in counts series of less than 5 years can be very difficult.

Two software programs, MONITOR (Gibbs et al. 1998) and TRENDS (Gerrodette 1987), are helpful in the design of trend-monitoring studies by addressing interactions among the many components of a monitoring program and evaluating how each component influences the monitoring program's power to detect trends. The program MONITOR uses a Monte Carlo approach based on linear regression analysis. It permits a researcher to define the basic structure of a monitoring program. The researcher also provides a variance estimate for the population index used. The program then runs a simulation in which many sets of sample counts are generated based on

the structure of a proposed monitoring program (sample number, sample interval, monitoring duration) with trends of varying strength underlying them. The frequency with which trends are detected in the counts, despite the sampling error imposed by the population index and the structure of the monitoring program, reflects the power of the monitoring design to detect trends. The simulation procedure is useful for evaluating the trade-offs between sampling effort, logistic constraints, and power to detect trends. The simulation software (monitor.exe) has been adapted for general use on DOS-based microcomputers and is available through the Internet from our web site. The program TRENDS approaches the same problem from an analytic, rather than simulation, approach, with a somewhat more restricted set of parameters.

Table 10.4. Variability Estimates for Local Populations of Plants and Animals

GROUP	N	MEAN CV	SE (MEAN CV)	MINIMUM*	MAXIMUM*
Mammals, large	17	0.14	0.034	0.02	0.62
Grasses and sedges	16	0.21	0.055	0.01	0.61
Herbs, compositae	9	0.21	0.098	0.03	0.94
Herbs, Noncompositae	32	0.22	0.051	0.01	1.15
Turtles	7	0.33	0.095	0.07	0.77
Terrestrial salamanders	8	0.35	0.047	0.17	0.51
Large-bodied birds	25	0.36	0.038	0.01	0.70
Lizards	11	0.42	0.086	0.15	0.93
Fishes, salmonids	42	0.47	0.040	0.14	1.24
Caddisflies	15	0.50	0.072	0.24	1.23
Snakes	9	0.54	0.070	0.27	0.94
Dragonflies	8	0.57	0.083	0.33	1.09
Small-bodied birds	73	0.57	0.056	0.11	2.48
Beetles	20	0.58	0.093	0.03	1.48
Small mammals	14	0.60	0.064	0.31	1.09
Spiders	10	0.64	0.050	0.37	0.86
Mediium-sized mammals	22	0.65	0.075	0.21	1.38
Fishes, nonsalmonids	30	0.71	0.073	0.11	1.73
Pond-breeding salamanders	10	0.86	0.17	0.45	2.31
Moths	63	0.90	0.044	0.33	1.95
Frogs and toads	21	0.93	0.13	0.05	2.78
Bats	24	0.93	0.31	0.20	8.00
Butterflies	13	1.11	0.047	0.81	1.27
Flies, drosophilids	13	1.31	0.17	0.69	2.40

Minimum and maximum refer to the smallest and largest CVs reported among the studies.

Values are useful for performing a power analysis to design an effective population trend monitoring program. N = number of de-trended count series of at least 5 years duration obtained from the literature. Values are coefficients of variation (standard deviation/mean) for standardized, 5-year count series. See our website for more information.

Both programs TRENDS and MONITOR require as inputs estimates of how variable the population counts of different types of plants and animals tend to be. Approximations are provided by Gibbs et al. (1998) from a survey of some 500 published long-term counts of plant and animal populations (reproduced in Table 10.4). Lacking a pilot study of a particular population, these estimates can give you a sense of the amount of sampling likely needed to detect trends of different strengths for the organism with which you are concerned. These should be only used as starting points for planning your monitoring until actual counts from a pilot study are available. Specific sampling recommendations for various monitoring designs are available through our web site.

MANAGEMENT IMPLICATIONS

In some monitoring situations, monitoring may be performed over periods longer than two time periods to ask whether there is an overall trend in a population that is increasing, decreasing, or stable. Populations, especially of animals, often undergo fluctuations owing to normal variation in the environment and population cycles so that reliably detecting trends in counts series < 5 years can be very difficult.

Evaluating such trends usually involves some type of regression analysis, a particularly useful form of which is route-regression, suitable for the many monitoring situations that involve tracking permanent plots over long periods. Anyone contemplating a trend-focused monitoring program has to make complex sampling decisions relating to where to place plots, the number of plots to be monitored, how often to monitor each season, and whether counts should be made each year. Power analysis of pilot data or published data from similar situations can be very useful identifying an adequate level of sampling intensity that insures that trends important to management will be detected and not simply overlooked.

CHAPTER 11
Selecting Random Samples

Primula capillaries
Ruby Mountain Primrose
Northeast Nevada in moist seepy soils
Artist: Jeanne R. Janish

Sampling units must be selected without bias for the statistical methods described in Chapters 9 and 10 to be applied to the data. Most of the designs described in Chapter 8 incorporate random selection at some level. For example, we recommend that systematic samples initiate with a random start. What does selecting random sampling units mean, and how do we do it?

SELECTING A RANDOM SAMPLE

Let us imagine we have defined a 50m × 100m macroplot around a rare plant population. We have selected a 4m × 10m quadrat, the long edge of which will lie along the short edge of the macroplot. (We will call the short edge of the macroplot the x axis. You can see this design in Figure 8.5 of Chapter 8). A total of 125 quadrats of this size can be placed without overlap in the macroplot. You wish to draw a random sample of 10 quadrats.

One way to draw a random sample (n) of 10 quadrats from the population (N) of 125 possible quadrats is to number each one of the quadrats from 1 to 125, put numbers from 1 to 125 on small slips of paper into a box, shake thoroughly, and select 10 slips from the box. Although valid, this is a time-consuming method. A much more efficient method of quadrat selection would be to select random points along both the x and y axes to serve as beginning points for each quadrat. Here is how to accomplish this.

Along the x axis are five possible starting points for each 4m × 10m quadrat (at points 0, 10, 20, 30, and 40). Number each of these points 0 to 4 accordingly (in whole numbers). Along the y axis there are 25 possible starting points for each quadrat (at points 0, 4, 8 . . . 96). Number each of these points 0 to 24 (again in whole numbers) accordingly.

Now, using a random numbers table (see next section) or a random number generator on a computer or handheld calculator, choose at random 10 numbers from 0 to 4 for the x axis and 10 numbers from 0 to 24 for the y axis. At the end of this process, we will have 10 pairs of coordinates. If any pair of coordinates is repeated, we reject the second pair and pick another pair at random to replace it. We continue until we have 10 unique pairs of coordinates.

Assuming that the x axis is on the bottom and the y axis is at the left, each pair of coordinates would represent the lower left corner of each quadrat. Thus, if we came up with the coordinates 0, 0, the quadrat would be placed with its lower left corner at the origin.

Now let us use a sampling unit of 4m × 25m instead of 4m × 10m. Now only two possible starting points lie along the x axis (0 and 25). One way to select random positions along the x axis with a random numbers table would be to consider every even number as the 0 position and every odd number as the 25 position. Alternatively, numbers 0 to 4 could represent the 0 position and 5 to 9 the 25 position, or you could flip a coin, with heads representing the 0 position and tails representing the 25 position.

Now consider one last example. This time we decide to sample using 2m × 50m quadrats. The long edge of the quadrat stretches the entire length of the short edge of the macroplot. In this case drawing a random sample is simplified because we only have to choose random locations along one axis (the y axis).

USING RANDOM NUMBERS

The following are two methods for using random numbers to select random samples. Either can be accomplished using a random numbers table or a random number generator on a computer or handheld calculator. The first method is probably the most commonly used, but the second method is far more efficient, particularly with two-digit and three-digit numbers. The section concludes with a brief discussion of two additional ways to derive random numbers in the field when you have forgotten to bring along a random numbers table or a handheld calculator.

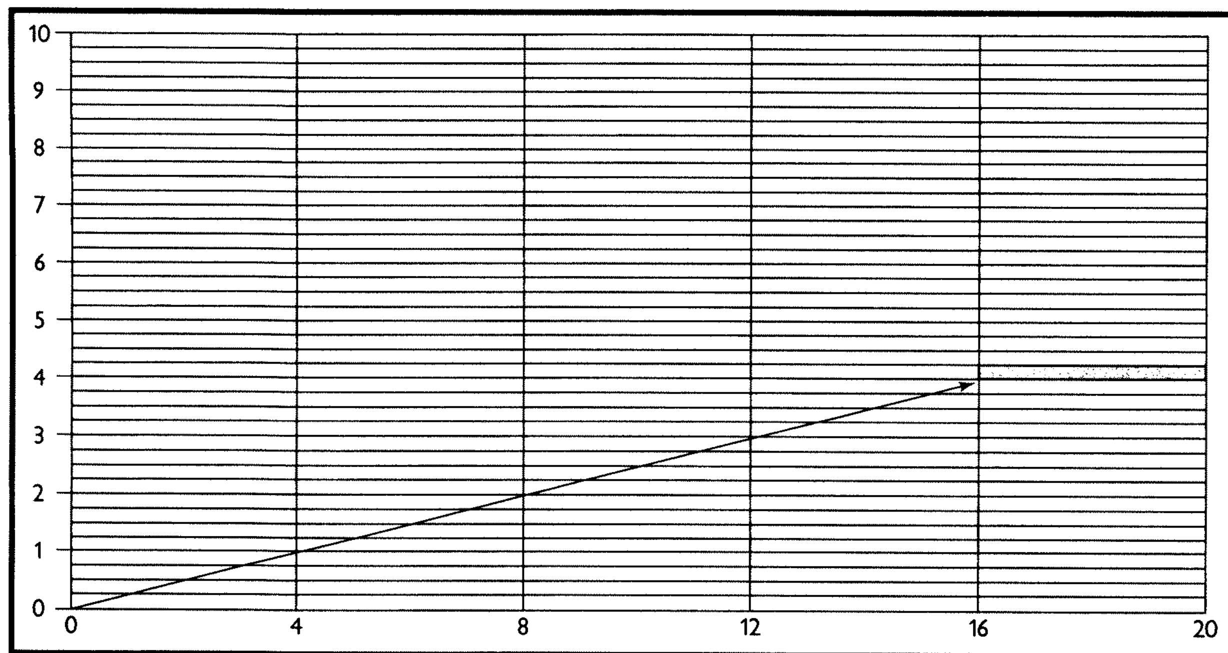


Figure 11.1. A 10m × 20m macroplot showing the 200 possible quadrats of size 0.25m × 4.0m that could be placed within it (assuming the long sides of the quadrats are oriented along the long side of the macroplot. A quadrat must be located at position 5 along the x-axis (16m) and 17 along the y-axis (4m). Quadrats can be located on the ground by pacing or taping from base lines at the edges of the macroplot or by converting the x-y coordinates into distance and azimuth from a single base point.

Method 1: Treating Random Numbers As Whole Digits

Example 1: Selecting Random Pairs of Coordinates

We have marked off a 10m × 20m macroplot within a key area, and we wish to randomly place forty 0.25m × 4.0m quadrats within that macroplot (in actual practice macroplots are usually much larger). We wish to place the quadrats so that the long side is parallel to the x axis and that the x axis is one of the 20m sides of the macroplot. The total number of quadrats (*N*) that could be placed in that 10m × 20m macroplot without overlap comprises the sampled population. In this instance, N is equal to 200 quadrats. The total population of quadrats is shown in Figure 11.1.

Using a Random Numbers Table

Along the x axis there are five possible starting points for each 0.25m × 4.0m quadrat (at points 0, 4, 8, 12, and 16). Number these points 1 to 5 (in whole numbers) accordingly. Along the y axis there are 40 possible starting points for each quadrat (at points 0, 0.25, 0.50, 0.75, 1.0, 1.25, and so on until point 9.75). Number these points 1 to 40 accordingly (again in whole numbers).

An abbreviated table of random numbers is shown in Table 11.1.[1] Using this random numbers table, you choose 40 numbers from 1 to 5 for the x axis and 40 numbers from 1 to 40 for the y axis. Because these numbers are in random order in the table, you can simply select these numbers in order, and they will be random. Arbitrarily begin at any five-digit number in the table. To pick numbers along the x axis, read either across or down (it makes no difference which) and select the first 40 one-digit numbers that correspond to the numbers 1 through 5. Reject any numbers that are not 1 through 5. For example, if we start at column 2, row 3, and read across the row, the first 10 numbers are 5, 0, 3, 1, 2, 9, 2, 6, 8, and 3. From this list, only the numbers 5, 3, 1, 2, 2, and 3 meet our criteria of being between 1 and 5. These become our

[1]*This table is presented only for instructional purposes; in practice you should use a much larger random numbers table. These can be found in most statistics texts. See our web page for links to on-line random numbers tables and generators.*

Table 11.1. A Table of 250 Random Digits

	1	2	3	4	5
1	55457	60189	95970	71641	75935
2	37232	58802	85478	23088	48214
3	29229	50312	92683	27179	98501
4	13135	53586	20722	77003	93064
5	20387	52649	66532	26770	88003
6	66611	22679	69735	40297	66715
7	71488	93726	54025	56130	36901
8	99078	11154	69689	62223	74431
9	57161	73561	33584	40186	22910
10	55220	37500	60530	36185	56969

first 6 randomly selected positions along the x axis (note that we do not reject the second occurrences of the numbers 2 and 3 because the y-axis numbers we select to go with these repeat numbers may be different). We write these numbers under a column heading "x axis" as follows:

X AXIS	Y AXIS
5	
3	
1	
2	
2	
3	

We continue until we have at least 40 numbers for the x axis (in anticipation of ending up with at least some duplicate pairs of coordinates, we may want to take 50 random numbers for the x axis). We then do the same for the y axis. We enter the random numbers table at a different location (or simply continue from where we left off after obtaining the x coordinates) and choose the first 40 (or more) two-digit numbers that correspond to the numbers 1 through 40. We reject any numbers that are not 1 through 40. For example, if one starts at column 1, row 9, and reads across the row, the first 10 numbers are 57, 16, 17, 35, 61, 33, 58, 44, 01, and 86.[2] From this list, only the numbers 16, 17, 35, 33, and 01 meet our criteria of being between 1 and 40. We write these down in the column marked y axis:

X AXIS	Y AXIS
5	16
3	17
1	35
2	33
2	01
3	

We continue until we have at least 40 (or more) numbers under the y axis. At the end of this process we will have 40 pairs of coordinates. If any pair of coordinates is repeated, we reject the second pair and pick another pair at random to replace it (because we are sampling without replacement). We continue until we have 40 unique pairs of coordinates.

[2] *Some people find it easier to read numbers down rather than across the table. It does not matter; you must simply be consistent, initiate your series randomly (you can drop a pencil decimal point), and make sure your system does not select the same number more than once.*

Using a Random Number Generator

Many handheld calculators have random number generators, making their use in the field very easy. Several computer programs also have the ability to generate random numbers. For example, Lotus 1-2-3 will generate random numbers using the @RAND function. With both handheld calculators and computer programs you need to consider whether you must reset the random number seed to generate different groups of random numbers. With some calculators and computer programs, failure to reset the seed will result in generation of the same set of random numbers (i.e., the numbers will not be "random" at all if you repeat the procedure more than once). Lotus 1-2-3 resets the random number seed automatically.

Random number generators yield numbers between 0 and 1 in decimals, usually to at least five places. In Method 1 the random numbers generated are used in the same way as numbers from the random numbers table, except that the decimals are ignored.

Example 2: Selecting Random Points Along a Baseline From Which to Run Transects

In our second example, we have laid out a 200m baseline oriented in a north-south direction, and we wish to randomly select points along the baseline (Fig. 11.2). The 0 point is at the south end of the baseline. At each point we will run a 50m transect perpendicular to the baseline. We can go in either of two directions, east or west, so we also need to randomly select the direction in which to run each transect. We intend to treat the transects as our sampling units and have determined from pilot data that 20 transects are required.

Along each transect we intend to lay 10 systematically spaced 1m × 1m quadrats for visually estimating cover of a rare plant species. These quadrats are larger than those usually used for estimating cover because the cover of the species is low, and we wish to encounter the species in at least a few of our quadrats along each transect line. The transect is the sampling unit.

These quadrats will always be placed on the south side of the transect. Since we are sampling without replacement, we want to avoid the possibility that any two quadrats could overlap. Thus, we want to select transect locations in whole-meter increments beginning at the 1-meter point along the baseline (if the 1-meter point were chosen through our random process, quadrats placed along that transect—since they will always be on the south side of the transect—would reach the outer boundary, the 0 point, of our sampled population).

Using a Random Numbers Table

Using the same process that we used for selecting random coordinates, we enter into the random numbers table at some arbitrary point and begin reading numbers from left to right. In this case, however, we must look at groups of three-digit numbers, since we are selecting points ranging from 1 to 200. Starting in column 2, row 2, and reading left to right, the first ten three-digit numbers would be: 588, 028, 547, 823, 088, 482, 142, 922, 950, and 312. We accept the numbers 028, 088, and 142, because they meet our criterion of being between 1 and 200; we reject the others. We then continue: 926, 832, 717, 998, 501, 131, 355, 358, 620, and 722. Of these numbers only 131 meets our criterion. We then continue until we have 20 numbers from 1 to 200 (do not worry; there is a far more efficient means of deriving our random set of points discussed under Method 2 below).

Once we have generated a list of 20 random points along the baseline, we now need to determine in which direction we will run the transect.[3] To determine direction we arbitrarily assign one-digit numbers to E and W; for example, E might be 0, and W might be 1. We then enter the random numbers table, read across and write down (next to the points we have already selected) the directions that correspond to every 0 and 1 we encounter (ignoring all numbers that are not

---

[3]*Just as for the random coordinates, we do not reject points that are the same; we only reject sets of points and directions that are the same. Thus, we can select point 75 and direction E, as well as point 75 and direction W. If we select another point 75 and direction E, we reject it and select another point and direction.*

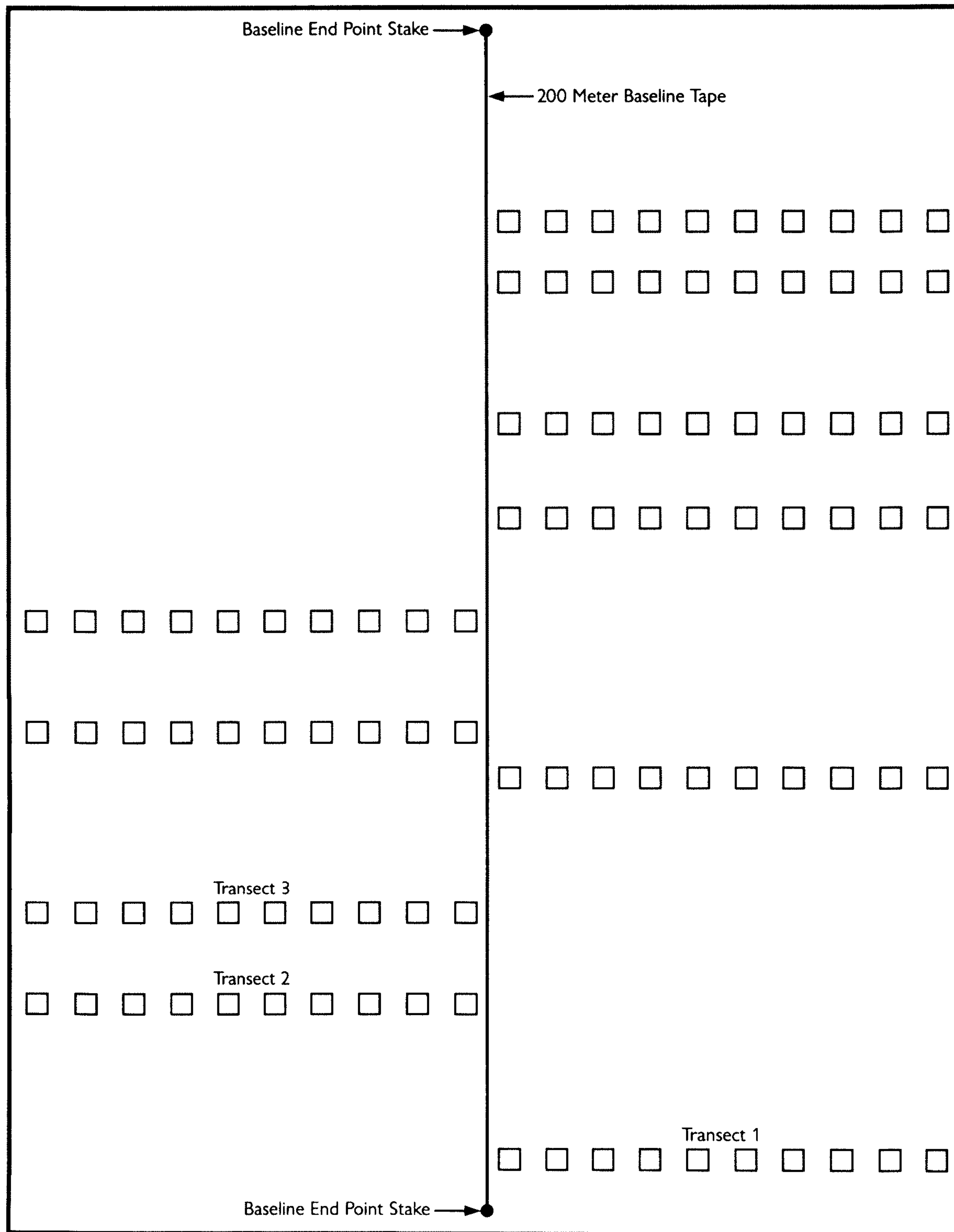


Figure 11.2. A 200m north-south baseline, showing 10 randomly positioned transects of 10 1m × 1m cover estimation quadrats. We have determined through pilot sampling that 20 transects will be required to detect the level of change we want to be able to detect at a particular significance level and power. We therefore will need to randomly select an additional 10 transects.

0 or 1). Alternatively, we could flip a coin, assigning heads to one direction and tails to the other. Or, we could consider every even number to correspond to E and every odd number to correspond to W.

Using a Random Number Generator

A random number generator would be used in the same way as the table of random numbers except that the decimal would be ignored.

Method 2: Treating Random Numbers as Decimals

This is by far the most efficient method of selecting random samples, particularly for numbers of two digits and higher. To use this method the random numbers must be treated as decimals. In our set of 250 random digits we would simply place a decimal point in front of every group of five digits and treat each group as one random number. Thus, if we entered the table at column 1, row 7, and read across, we would have the following six random numbers: 0.71488, 0.93726, 0.54025, 0.56130, 0.36901, and 0.99078. If we used a random number generator, it would be even easier since these provide random numbers as decimals falling between 0 and 1.

The formula for using these decimal random numbers for selecting a sampling unit or point is as follows:

$$[uN] + 1$$

where u = random number (expressed as decimal)
N = total population size
[ ] = indicates that only the integer part of the product is used in the calculation

To illustrate how this formula works, consider our baseline example. Here we need to select numbers between 1 and 200 as points along a baseline. Consider these points as a "population" of 200 possible points. Using the first of the six random numbers we came up with above, 0.71488, we calculate the following:

$$[0.71488 \times 200] + 1$$
$$= [142.976] + 1$$
$$= 142 + 1$$
$$= 143$$

Thus, 143 is our first point. Using the second random number we have the following:

$$[0.93726 \times 200] + 1$$
$$= [187.452] + 1$$
$$= 187 + 1$$
$$= 188$$

Now we have our second point, 188. We would continue in this manner until we had the 20 points we need. Although the formula may look difficult, a handheld calculator or computer program with a random number generator makes it easy. With a handheld calculator, for example, one could program the population size in memory, hit the button generating a random number, multiply it by the number in memory, and come up with the random point. Twenty such points could be produced in just a few minutes.

The reason for adding the 1 to the integer of the product of the random number and N may not be intuitively obvious. It is necessary because we are using only the integer of the product. Without adding 1, it would therefore not be possible to obtain the number 200. Consider the highest possible random number we could obtain, 0.99999. If we multiply this number by 200 we obtain 199.99800; taking the whole integer of this number yields the number 199. Adding 1 makes it 200. If, instead of choosing numbers from 1 to 200, you are choosing numbers between 0 and 199, there is no need to add the 1 to the integer of the product.

As a rule, you should make sure the random numbers have more digits on the right side of the decimal point than the number of digits in N. In the example above, N is 200, and we are using random numbers with five digits to the right of the decimal point, so we are okay.

This process is much more efficient than Method 1 because we do not need to reject any numbers. When selecting random points along our 200-meter baseline using Method 1, we had to look at 20 three-digit numbers just to come up with four numbers that met our criterion of being between 1 and 200. Given the fact that there is only a one-in-five chance of any three-digit number falling between 1 and 200, this means we would, on the average, examine 100 three-digit numbers to come up with 20 points. Using Method 2 we could use the first 20 random numbers to select the same 20 points. When we need to select one-digit numbers (as, for example, to determine direction), it may be just as efficient (or even faster) to use Method 1.

Generating Random Numbers Without a Calculator or Random Numbers Table

You have left for a 5-day trip to the field and realize, 4 hours away from the office, that you have forgotten to bring either a random numbers table or your calculator[4] with its random number generator. You will need to generate random numbers for the monitoring study you intend to design while in the field. How are you going to do it?

Using a Digital Stopwatch to Generate Random Numbers

If you have a digital watch with a stopwatch function, you can use the stopwatch to generate random numbers. Fulton (1996) describes how to do this and offers proof that the procedure is actually random. The procedure is rather simple. Just start the stopwatch, and let it run. Whenever a random digit is needed, simply stop the watch and read either or both of the two numbers in the tenths and hundredths of seconds places. For example, you stop the watch and it reads 27.45 seconds (ignore minutes and hours). You can use either the 4 in the .45 if you need only one random digit between 0 and 9, or you can use the 45 if you need two random digits between 0 and 99. If you need more than two random digits you can repeat the procedure as many times as necessary until you have the required number of digits. For example, we need three random digits between 0 and 999. After we stop the watch the first time and get .45, we restart the watch and stop it again. This time, say, it stops at .82. Then we use the .8 as our random digit and end up with a random number of 458. We can repeat this procedure until we have the required quantity of random numbers.

Two things are important in this procedure: 1) to avoid bias, do not look at the watch before you stop it, and 2) wait long enough between starts and stops to allow a few seconds to elapse, making each reading independent and ensuring your selections are truly random. It is also important to note that you should use only the tenths and hundredths of seconds as your random digits, or the digits you choose will likely not be truly random, unless you wait a long time (several tens of seconds) between starts and stops. Even then, however, you could only use numbers in the ones of seconds place, because the numbers in the tens of seconds place range only from 0 to 6. The safest bet is to use only the tenths and hundredths of seconds places.

Using a Telephone Directory as a Source of Random Numbers

Another source of random numbers is a telephone directory. You can use the last four digits of the telephone numbers listed in the white pages as random numbers. Do not use the first three digits (prefixes), however, because these do not represent the full range of digits available between 0 and 999, they are not in random order, and they are not independent of one another.

---

[4]*Although you can certainly generate random numbers without a calculator, you are going to need a calculator to calculate means and standard deviations during pilot sampling. This means you are probably going to have to stop somewhere and buy an inexpensive calculator, which may not have the capability to generate random numbers.*

LOCATING SAMPLING UNITS IN THE FIELD

Once these random coordinates have been identified, how do you locate the points in the field, and how accurately do these points need to be located?

Tapes along all four sides of a macroplot will increase your efficiency because you can measure to a plot from any of the four sides. You will likely need to set some type of visible boundary around your sampled area anyway, so you may as well use measuring tapes. You may even want to place pin flags regularly along your boundaries. You could place them every 10m if you are still in the pilot stage and might be trying plots of different configurations, or at the increments determined by your selected quadrat size. These flags form the grid in which you will locate your sampling units. You can, however, locate sampling units from a single measuring tape along the baseline.

Let us assume we are establishing quadrats for measuring density and wish to locate our first quadrat at the 40m point along a baseline and the 9m point along the y axis. You have several alternative methods for locating this quadrat:

1. Pacing. You can simply pace from the 40m point along the baseline up approximately 9m using a compass. If you did this, you would only need the baseline tape from which to pace all of your quadrat corners. In practice, the additional tapes may save you steps by allowing you to pace from any of the four sides. Once you have located and measured the first quadrat, you can pace to the next quadrat from the first, again by using a compass. For example, if the next quadrat was at 60m along the baseline and 10m along the y axis, from the corner of the quadrat that you just completed, you could pace 20m along the x axis and 1m up along the y axis. Pacing is an acceptable way to find sampling unit corners; sampling units do not have to be located exactly. Pacing will not work, however, if the "slop" inherent in pacing allows you to place a sampling unit with bias. If you are working in an area with scattered pockets of prickly brush, for example, you might (inadvertently, of course!) shorten your pace somewhat to avoid a brush pocket. Another example of bias is to adjust your pace to try to intersect the target species. Placement bias is less of a problem with very long or large sampling units because you probably cannot judge whether features will be included in the sampling unit once it is established. Bias can also be lessened by consciously avoiding looking at the ground as you near the destination. If, in spite of your best efforts, you think you have biased placement with pacing, consider taping your distances.
2. Pin Flags. If you have placed pin flags along your four boundaries and can see all four boundaries, you can approximate your coordinate location using those pin flags. If, for example, you placed pin flags every 10m along all four boundaries (alternating colors is helpful), you could then move until you were between the two flags at the 40m mark along the top and bottom x axis, and about a meter short of the flags that marked the 10m points along the left and right y axis. Again, if you think you might bias sampling-unit locations using this method, opt for taping.
3. Taping. You can measure sampling unit locations accurately (within 10cm to 30cm) with a tape measure or a pocket electronic distance measurer (see Chapter 5). Measuring is more time-consuming than pacing or pin flags, but will help eliminate bias.
4. GPS unit. Global Positioning Units with moderate accuracy (e.g., to within a meter or so) can be used to locate sampling units. These systems can save time in brushy areas where travel with a tape is problematic. In meadows or savannahs, where taping or pacing can be done quickly, the time required by some units to lock onto a location may be greater than that needed for "low-tech" approaches.

Taping or using a GPS system with high accuracy can also function as a backup monumenting system for permanent plots and transects. If markers of the individual quadrats or transects are lost, but the baseline markers remain, you may be able to quickly relocate the buried markers that accompanied the permanent marker by measuring from the baseline.

An alternative approach to the x-axis, y-axis grid described above is translating x and y coordinates into distance and azimuth from a single base point (see Fig. 11.2). The distance from the baseline is paced (or taped), and the direction from the base point is measured by compass. Awbrey (1977) provides a complete discussion of this approach.

The cost of converting x-y coordinates to distance and azimuths is offset by the time saved by having to measure only one distance, rather than the two measured distances required in the grid method. The time savings can be significant in brushy or wooded areas where establishing a baseline is difficult. In very dense brush, or where increased accuracy is desired, an observer at the base point with a Sonin (for distance measures—see Chapter 5) and a compass (for azimuth) directs another observer to the correct location. Two observers could also use a survey instrument with an electronic distance measurer set up at the base point, providing both direction and distance to the sampling-unit point. Sampling-unit locations found with these instruments, especially survey instruments, are very accurate and much faster than taping through brush. Surveyed plots that are permanent sampling units can be relocated accurately if individual plot monumentation is lost. These instruments can, however, require significant setup time.

A similar concept is to dispense with x-y coordinates and simply use a random distance and azimuth as the plot location. One problem with this approach is that the distances along azimuths radiating from a central point are clustered near the center (like the spokes of a wheel near the hub) and farther apart toward the outside edge. Laferriere (1987) provides a complete discussion of this approach and some solutions to the clustering problem. Another problem with this approach is that sampling-unit corners located in this way may result in projecting transects and long quadrats beyond the boundaries of the sampling area (see Chapter 8).

MANAGEMENT IMPLICATIONS

Statistical analysis techniques assume sampling units are selected without bias. Identifying random samples is not difficult, nor does it require much more time than locating sampling units selectively. In the field, sampling units do not have to be located exactly, but they do have to be located without bias. Once random locations are determined using random numbers tables or other means, the sampling units may be located on the ground using pacing and compass. In some areas where pacing may result in bias (e.g., dense brush), sampling units may be located using tape measures or electronic distance measurers.

CHAPTER 12
Field Techniques for Measuring Vegetation

Pseudacris crucifer
Spring peeper
Artist: D. Andrew Saunders

The techniques described in this chapter are primarily used for herbs, graminoids, and shrubs. Specific methods for tree species were deliberately excluded because many references describing methods for measuring trees of commercial importance are available. We recommend several texts that describe density and basal-area estimation for trees. These include Dilworth and Bell (1973), Husch et al. (1982), Dilworth (1989), Schreuder et al. (1993), Avery and Burkhart (1994), and Shivers and Borders (1996). While this chapter is directed at techniques for vegetation, many of the comments are applicable to measuring any stationary object such as sessile animals and habitat features.

Several classes of measures are available to use in monitoring vegetation. Density (and population size, easily calculated from density) is the number of individuals per unit area. Frequency is the proportion of sampling units containing the species. Cover is the amount of ground surface covered by the plant, as observed from above the plant (usually expressed in percentages). Biomass is the amount of plant material produced over a given period (e.g., annual production), usually expressed as dried weight/area. Finally, measures of plant vigor such as height, biomass, or number of flowers may be made on individual plants.

COMPLETE POPULATION COUNTS OF PLANTS

Some populations are small enough that they can be completely counted or censused. No statistics are required to analyze the results or the precision of the estimate. The change or number observed is real (provided the count is accurate and plants are not missed); sampling error is not a concern. The only question is whether the change is biologically significant.

To use a census approach, a counting unit must be consistently recognizable (ramet, genet, or some consistent arbitrary unit). For a nonclonal species, individual plants (genets) may be relatively easy to delineate and recognize, but for a clonal species such as a rhizomatous grass or sedge, it is much more difficult to define a counting unit. For a clonal species like aspen (*Populus tremuloides*), the count may focus on obvious units like trunks (which, since aspen is clonal, are actually ramets), but you must still decide whether the count excludes any size classes such as small ramets or seedlings. If consistency of the counting unit is a problem, an alternative sampling approach (such as cover or frequency) is a better option.

In theory, any population that can be counted can be censused. In practice, however, accuracy of counts may be very poor because of missed individuals. This can occur even when the plant is large and obvious. For example, a census of a large (up to 60cm tall) herbaceous species done with four individuals walking a grid pattern resulted in a count of 163 plants. When approximately 20% of the area was sampled, the sampled area alone contained 93 plants (Elzinga, unpublished data). The discrepancy was likely the result of misses of nonreproductive and small individuals.

Factors that make accurate counts unlikely include a large population area, a large population, dense associated vegetation, the presence of similar species, small stature of the target species, and many of the target species being in cryptic stage classes (such as seedlings). Before using a census approach, ensure that counts are accurate by using two or more observers and comparing the results.

You can improve census counts by using some type of systematic search of the population area (e.g., 0.1-hectare macroplots or parallel lines marked by pin flags) and by setting standards (parallel swathes of a certain width; macroplots searched for a given amount of time each year). Boundaries of the population or macroplot should be permanently marked so future counts cover the same area.

SAMPLING VEGETATION—SAMPLING UNITS REVISITED

The concept of sampling units was introduced in Chapter 7, and a variety of sampling unit types were described in Chapter 8. These sampling units may be individual plants or, more commonly, quadrats, transects, or points. Many vegetation studies use quadrats; measures of density, measures of frequency, measures of cover use small cover estimation plots, and measures of plant production or biomass use small clipping plots. Individual plant characteristics are also usually estimated in quadrats, using the quadrat as the primary sampling unit and individual plants as secondary sampling units within the quadrats (a two-stage or cluster design—see Chapter 8).

Two design issues are applicable to several types of vegetation measures and sampling units. Boundary protocols describe the consistent manner in which observers determine whether plants are in or out of a sampling unit. For two-stage sampling designs using small sampling units distributed along transects, a decision must be made about spacing of these secondary sampling units.

Boundary Decisions

Boundary decisions are few for thin-stemmed plants, but plants with large basal diameters (e.g., bunchgrasses and trees) may often straddle a boundary, presenting a more formidable problem (Fig. 12.1). Most observers will consistently include boundary plants within the quadrat,

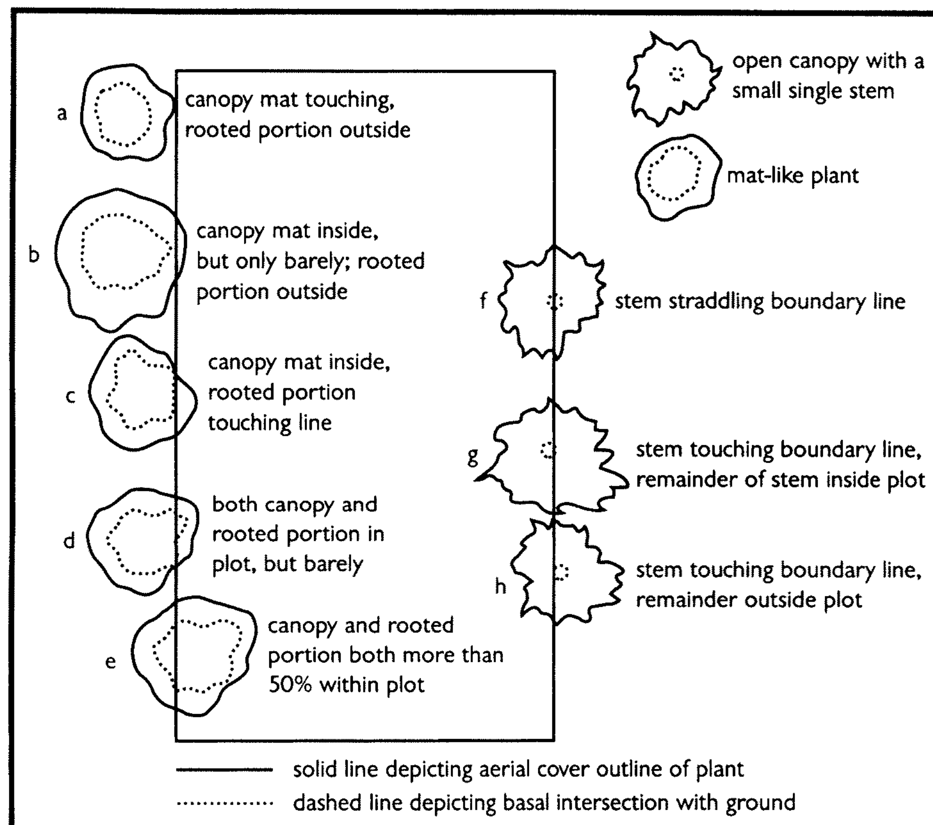


Figure 12.1. Boundary decisions. Which plants should be considered within the plot? Most investigators use one of two rules: (1) all boundary plants are counted in on two contiguous sides and out on the other two sides; or (2) every other boundary plant is counted. Plants c-h would be considered by most observers to be boundary plants. Plants a and b would generally not be considered boundary plants (because only the canopy intersects the plot), although occasionally a specific situation may require that the canopy boundary be used rather than the basal boundary (see text for additional discussion).

overestimating density, frequency, cover, or biomass. Establishing rules for boundary plants minimizes differences between observers and between years of observation. You must establish boundary rules and apply them consistently each time the monitoring is done. How will these be addressed?

Some viable alternatives are as follows:

1. Plants are considered "in" if any part of the plant boundary is touching the plot boundary along two adjacent sides of a rectangular plot, and they are considered "out" if any portion of the plant boundary is touching the other two sides of the plot. This provides an accurate estimate of density and is the recommended approach for reducing boundary bias. For monitoring in permanent plots, you must specify which sides are interpreted in which way (compass direction works well) and measure along those sides consistently from year to year. The sides must be split so an equal portion of the perimeter is treated as the "in" sides compared with those considered the "out" sides. In other words, if the plot is rectangular, you would consider boundary plants "in" along one long side and one short side of the rectangle and "out" along one long side and one short side (adjacent sides).
2. Boundary plants are counted as "in" or "out" alternately along the boundary. This provides an unbiased estimate of density, but in a very large or long quadrat, you may have trouble keeping track of whether you last counted an "in" or an "out" plant.
3. Plants are considered in if more than 50% of the plant boundary (canopy or basal) is within the plot. This is illustrated by plants e and f in Figure 12.1. While this method will give an accurate measure of density, we do not recommend it because additional subjective observer decisions are required. Observers may have consistent bias in their estimates of 50% (overinclusion is the most common), introducing an unknown observer error. Plants with irregular basal outlines are especially difficult to consistently determine if they are to be counted "in" or "out" by this method.

Some nonviable alternatives are as follows:

1. Count all plants that touch the line, even if most of the plant boundary is outside the plot. This is illustrated by plants c and h in Figure 12.1. If you use this approach, you will overestimate density (number of individuals per unit area) because the length and width of the plot are essentially increased by the average diameter of the boundary of the plant. This is easiest to visualize with the matted plant in Figure 12.1.
2. Include only plants that are completely within the plot, including those that just touch the line. This is illustrated by plant g in Figure 12.1. This approach underestimates true density.

Both approaches have been used in monitoring studies, and it is not a fatal error if you have a current study using one of these designs. If the purpose of the study is to measure change over time, and if the boundary rule resulting in overestimation or underestimation has been consistently applied, you may still be able to interpret changes in terms of trend in the population. Note that as plant boundaries change because of changes in vigor, the impact of that change on density estimated by using either of these boundary rules will be much larger than if boundary decisions were made using an unbiased approach. Thus, interpreting changes in density measured in a monitoring project using one of these boundary rules will be partially obscured by changes in vigor. Another problem is that both methods create difficulties in comparing density estimates at different sites since the estimate of density is partially a function of plant diameter, which can vary from site to site.

Independence in Two-stage Designs

A common practice in vegetation sampling is to arrange smaller plots[1] (such as those used for frequency, production, or visual estimates of cover) along transect lines. Which are the sampling units—the small quadrats or the transects? In Chapter 8 the answer given was either or both. If the quadrats are far enough apart to be considered independent, they can be treated as the sampling units, dramatically increasing the sample size and the precision of the estimate while retaining the field efficiencies of locating sampling units along transects. Briefly, independence means that the sampling units are not correlated, that the response of the species in Quadrat A is not related to the response of the species in Quadrat B because of their proximity to one another. If sampling units are separated by short distances, it is unlikely they are independent. Correlated sampling units result in an underestimation of the standard error and questionable results.

How far apart must the small quadrats be for them to be independent? It is easier to define what is not far enough apart. Clearly, quadrats that are positioned contiguous to one another along a transect are not far enough apart to be considered independent. The same can be said of quadrats or points that are spaced so closely that they may fall on the same individual (when sampling plants or stationary animals). What should the minimum spacing be? Some factors to consider are the average size and pattern of gaps or microsites in the habitat (especially in forests), the average size of individuals, and the size of clones. In general, sampling units should be far enough apart that they do not fall into the same microsite, gap, or clone. This, however, is scale dependent. If you are sampling an area that only covers a typical gap, your plots by necessity will all fall within that gap.

SAMPLING TO ESTIMATE DENSITY

Density is the number of counting units per unit area, usually estimated by counting plants in quadrats. A counting unit must be consistently recognized by all observers for density to be used as a monitoring method.

Advantages and Disadvantages

Density is a suitable measure for any plant species that can be counted with a consistently recognizable counting unit. It cannot be used for plants that are difficult or impossible to separate into individuals or counting units such as mat-forming perennials and bunchgrasses. It may also not be the most useful measure for monitoring annual plants if populations vary dramatically from year to year. Density is impractical for species with extremely high densities in tight patches because a sampling unit that intersects a patch of plants may contain too many to count.[2] Density is also sometimes impractical for species that are distributed very sparsely across a large area, because it is difficult to make a quadrat large enough to include any individuals (e.g., some rare tree species).

Estimated density (in terms of mean number per unit area) is theoretically the same for all quadrat shapes and sizes, although the precision of the estimate will vary (sometimes dramatically) among sampling units of different configurations (see below and Chapter 8). The fact that density is reported as a per-area measure allows comparison between sites even if the quadrat

---

[1] *While this discussion focuses on small quadrats, the same issues apply to any sampling unit arranged along a transect such as point intercepts for measuring plant cover (see below) or points for conducting singing bird counts (see Chapter 13).*

[2] *This can often be addressed by sampling-unit size and shape (see Chapter 8), but for some plants any quadrat configuration will still include many plants. An example of this is the ubiquitous Kentucky bluegrass (*Poa pratensis*), which can form small patches with shoots numbering in the thousands.*

size and shape used for sampling differs. This is in contrast to another measure described below, frequency, which is dependent on quadrat size and shape and is not comparable across studies using different quadrat sizes.

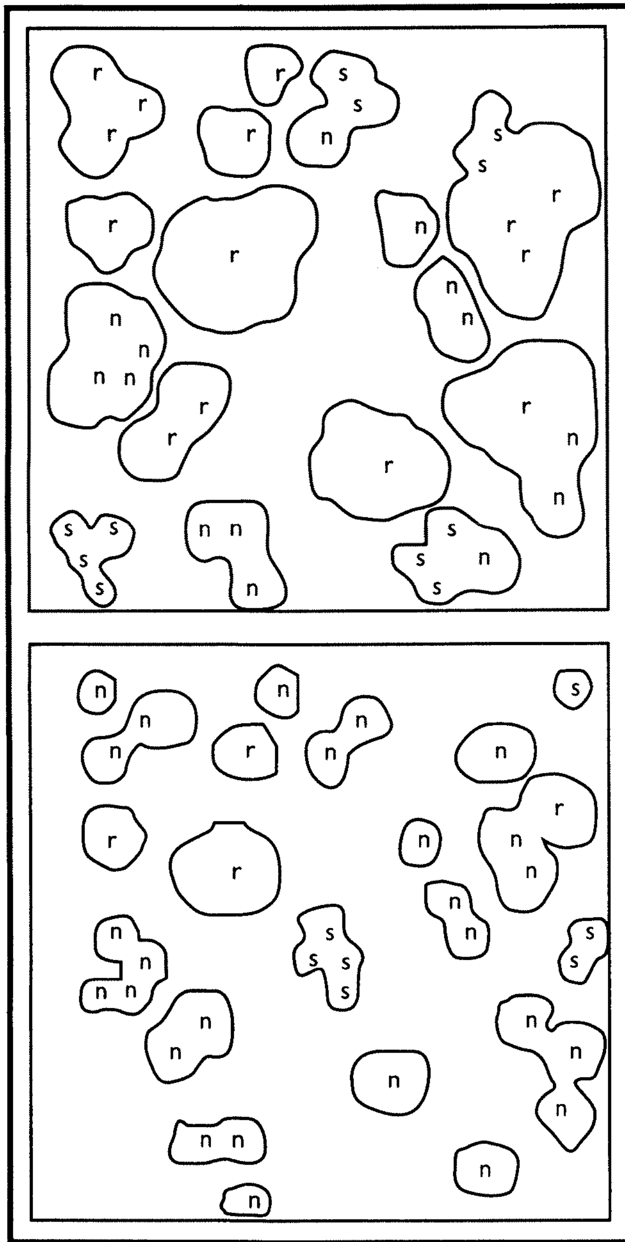


Figure 12.2. Two views of the same quadrat, the top measured in 1995 and the bottom in 1996. Outlined polygons denote canopy cover; letters represent individuals. Note that density declined from 39 individuals to 37 individuals. In 1995, there were 14 reproducing individuals (r), 14 non-reproducing individuals (n) and 11 seedlings (s) in the plot. In 1996, there were 4 reproductive individuals, 26 non-reproducing individuals and 7 seedlings. Note also the dramatic decline in cover from 1995 to 1996. The changes illustrated in this plot are not well captured by density measures of total individuals. Even a count of seedlings versus adults would not have captured the dramatic change in reproductive fraction.

Density is most sensitive to changes caused by mortality or recruitment. It is less sensitive to changes that are vigor-related, especially those that are sublethal (e.g., a reduction in production that is not accompanied by an increase in mortality or a decrease in recruitment). Figure 12.2 illustrates how a population can change dramatically without a large change in density. In this example, cover and the ratio of reproductive to nonreproductive individuals have declined dramatically, but simple counts would have detected a decline of only two individuals. Density may be an especially poor monitoring measure when individuals are long-lived and respond to stress with reduced biomass or cover, rather than mortality.

Design

The density of herbaceous plants is usually counted within the boundaries of a quadrat, each of which is a sampling unit. Quadrat design is discussed at length in Chapter 8. A few of the points are reiterated here:

1. The size of the quadrat should not be impractical; that is, the quadrat should not be too large in terms of either number of individuals to be counted or search time required.
2. Size and shape of the quadrat must be tailored to the specific plant distribution observed in the field. The most efficient quadrat shape will usually be an elongated rectangle.
3. Consider quadrat widths of 1m or less to allow searching the entire quadrat from one side. This reduces double-counting errors that result from counting from both sides of the quadrat and increases the speed at which a quadrat can be measured. Narrow quadrats are also quickly established in the field, requiring a single tape along one of the long edges of the quadrat and a meter stick to determine if plants are in or out of the quadrat along the opposite boundary.
4. You should attempt to include at least some "clumps" of the target species in your initial trial of quadrat sizes and shapes. The most efficient plot shape and size in terms of number of quadrats

needed will be one in which the density in each quadrat is very similar (little variability between quadrats). Reasonable guesses on size and shape can be made by first observing the distribution of the plants in the field. Pin flags can be placed throughout the population in areas of concentration to get a better picture of the distribution of the species at the site. You want to design a plot size and shape that intersects those areas of concentration. Chapter 8 (Box 8.3) describes a procedure to compare the efficiency of different quadrat sizes and shapes based on pilot sampling.

While different quadrat configurations should produce the same estimate of density (although efficiency and precision will vary), in practice the density estimate may vary with plot size. This variation is because of the effects of boundary decisions, which are most pronounced in small or long, narrow quadrats. Because most observers will consistently include boundary plants, estimates of density in quadrats with high perimeter-to-area ratios are usually higher than estimates from larger or square quadrats. A key monitoring design decision when using density is to select a quadrat size and shape that will efficiently estimate density with acceptable precision (see Chapter 8), while controlling these boundary errors (see above). Protocols for boundary decisions should be described and followed consistently.

Density is usually based on a count of plants rooted within a quadrat, but this may be problematic for some species. For example, if the counting unit is a shoot of grass, individual tillers are sometimes not clearly rooted because they remain partially attached to the main plant, but these tillers are clearly individual shoots. For many matted plants, trying to determine the rooted zone requires lifting and pulling at the top mat, possibly causing injury to the plant; thus, for these matted plants using the canopy outline for boundary decisions may be better than the rooted area.[3] For most species, however, avoid using the outline of the canopy as the boundary to determine whether a plant is in or out of the quadrat, because changes in canopy (vigor) will affect the density measure and increase the complexity of the interpretation. For most species the best counting unit is a rooted individual, but for some species other rules may have to be developed (and documented).

In addition to inconsistent boundary decisions, differences between observers and nonsampling errors often arise when quadrats contain cryptic individuals or numerous plants. The most common nonsampling errors originate in rapid counts that overlook small individuals. Establishing a minimum search time per quadrat can reduce the temptation to hurry the measurements, although the actual time required per quadrat will vary depending on the number of counting units occurring within it.

Consider the value of using stage classes such as seedling, nonreproductive, and reproductive in density counts. Doing counts by stage class requires more time, but in many situations the additional information warrants the extra effort. Figure 12.2 clearly shows that measuring density in stage classes can rectify the insensitivity of density measures to some kinds of change. In this example, the number of plants in the quadrat only declines by two—from 39 plants the first year to 37 the second—but demographic structure displays a dramatic change, declining from 14 reproducing plants the first year to four the second year. Dividing the population into seedlings and nonseedlings can provide additional information for interpreting changes in density, although in this example adults increased by two individuals and seedlings declined by four.

Quadrats for measuring density are usually distributed through a sampled population using either a simple random, stratified random, systematic, or restricted random sampling design. Quadrats can be permanent or temporary. For most perennial plant species, permanent density quadrats are often more efficient than temporary ones. Using permanent quadrats may also enhance interpretation of the monitoring data because the additional spatial information often allows relating a change in the population to a change in the local environment (e.g., tree fall and

---

[3]*Cover is usually the most appropriate measurement technique for matted perennial plants.*

canopy opening or animal impacts). The expense of permanent plots, however, may exceed their value when monitoring the density of annual plants if the distribution of plants is not correlated from year to year (see Chapter 8).

Density data are generally suitable for parametric statistical analysis if collected in a well-designed sampling effort. Chapter 9 describes all of the analysis methods in detail.

DISTANCE MEASURES FOR ESTIMATING DENSITY

An alternative to estimating density in quadrats is a set of techniques called distance measures. Several variations on the theme have been developed, but they all involve the measure of the distance of an individual from a point or from another individual and estimating density from the average distance measure. Figure 12.3 shows the four most commonly used distance measures in vegetation sampling. These measures are most often used for large or scattered taxa such as trees, for which the use of quadrats is not practical. They have, however, occasionally been used in grasslands on common herbaceous plants (Becker and Crockett 1973). Distance measures are based on the concept of a mean area per plant. Once this is known, the value can be used to calculate a density per unit area.

These techniques, however, are only suitable for use on plants with random distributions. Most plants do not grow randomly in space, but occur in clumps, the result of short-distance dispersal of propagules or microvariation in habitat. One technique, the wandering quarter method

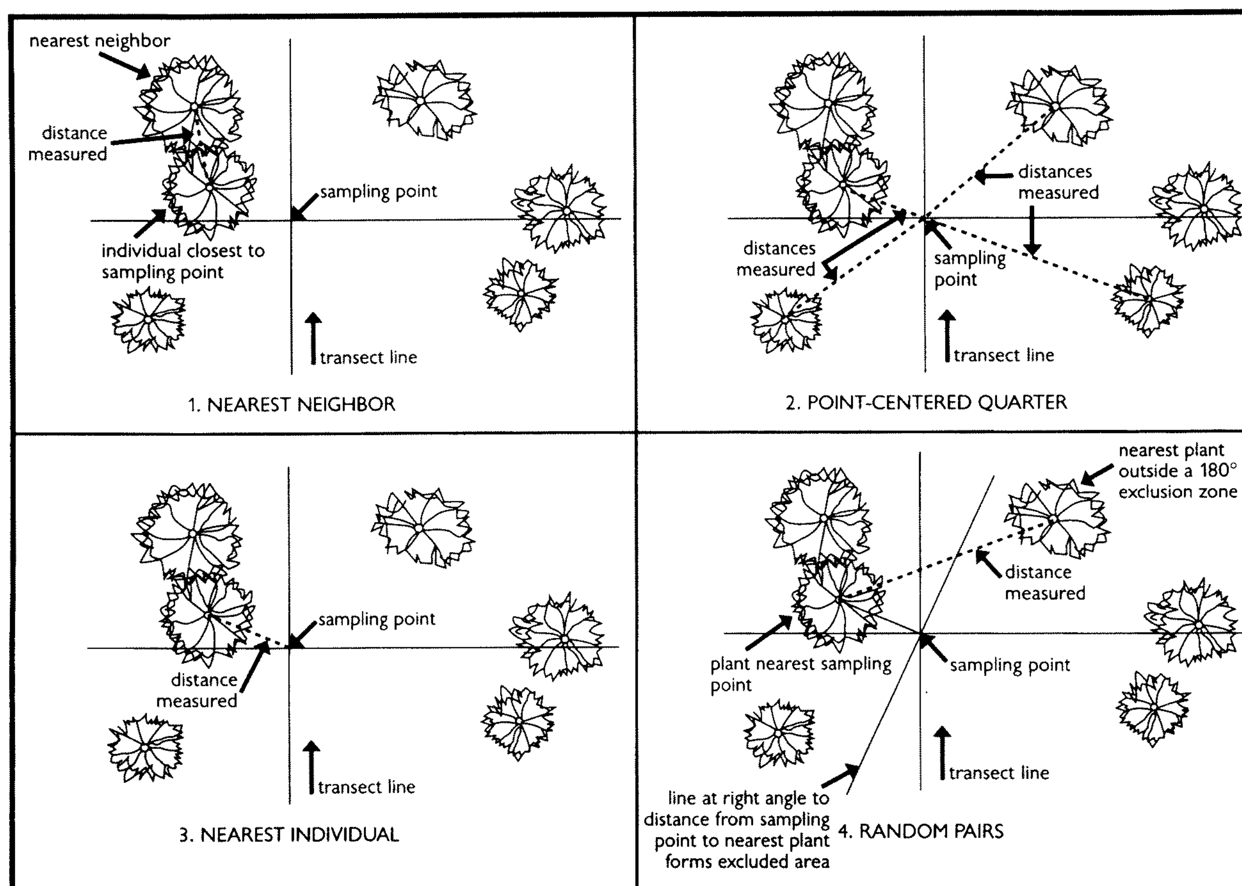


Figure 12.3. Four distance methods used for measuring density in plant populations with randomly distributed individuals: (1) nearest neighbor; (2) point center quarter; (3) nearest individual; and (4) random pairs. None of these methods are appropriate for species that have contagious (clumped) distributions.

(Fig. 12.4), was designed for plants with non-random aggregated distributions (Catana 1963). A similar approach, the T-square method, was proposed by Diggle (1975) and by Blyth (1982).

Field tests of the latter two methods give mixed results. Lyon (1968) found that the wandering quarter method gave an accurate estimate of density in a shrub community in which all individuals had been enumerated. To achieve a reasonably precise estimate of density, however, actual counts of the entire population were quicker than sampling with points and distance measures. McNeill et al. (1977) sampled an area in which all individuals had been marked and mapped. They found that quadrats were superior in terms of the accuracy of the estimate and field efficiency compared with several distance measures. Becker and Crockett (1973) concluded that the wandering quarter method underestimated a clumped species and overestimated a single-stalked, well-dispersed species.

In a simulation study of 24 distance-based density estimators, Engeman et al. (1994) determined that the approach proposed by Diggle (1975) did not provide unbiased estimates of the mean when sampling clumped distributions. They also argued that the method is relatively inefficient in the field because of the difficulty in defining the area of exclusion (see Fig. 12.4). They concluded that the best estimators were those that measured three distances per point (point to nearest individual, nearest individual to nearest neighbor, and nearest neighbor to its nearest neighbor) and those that measured from the point to the third nearest individual. The field efficiencies of these complex distance methods are questionable.

The value and performance of distance measures depends on the field situation. The best estimators for clumped distributions are complex, either requiring three measures per sampling point or determining which individual is the third farthest from the sampling point. This complexity dramatically reduces the field efficiency of these methods. Distance measures may be appropriate when the individuals are so widely spaced that using quadrats is not practical (as for some trees), but for most monitoring situations involving rare plants, quadrat-based density estimates are more efficient and free from the potential biases of distance methods.

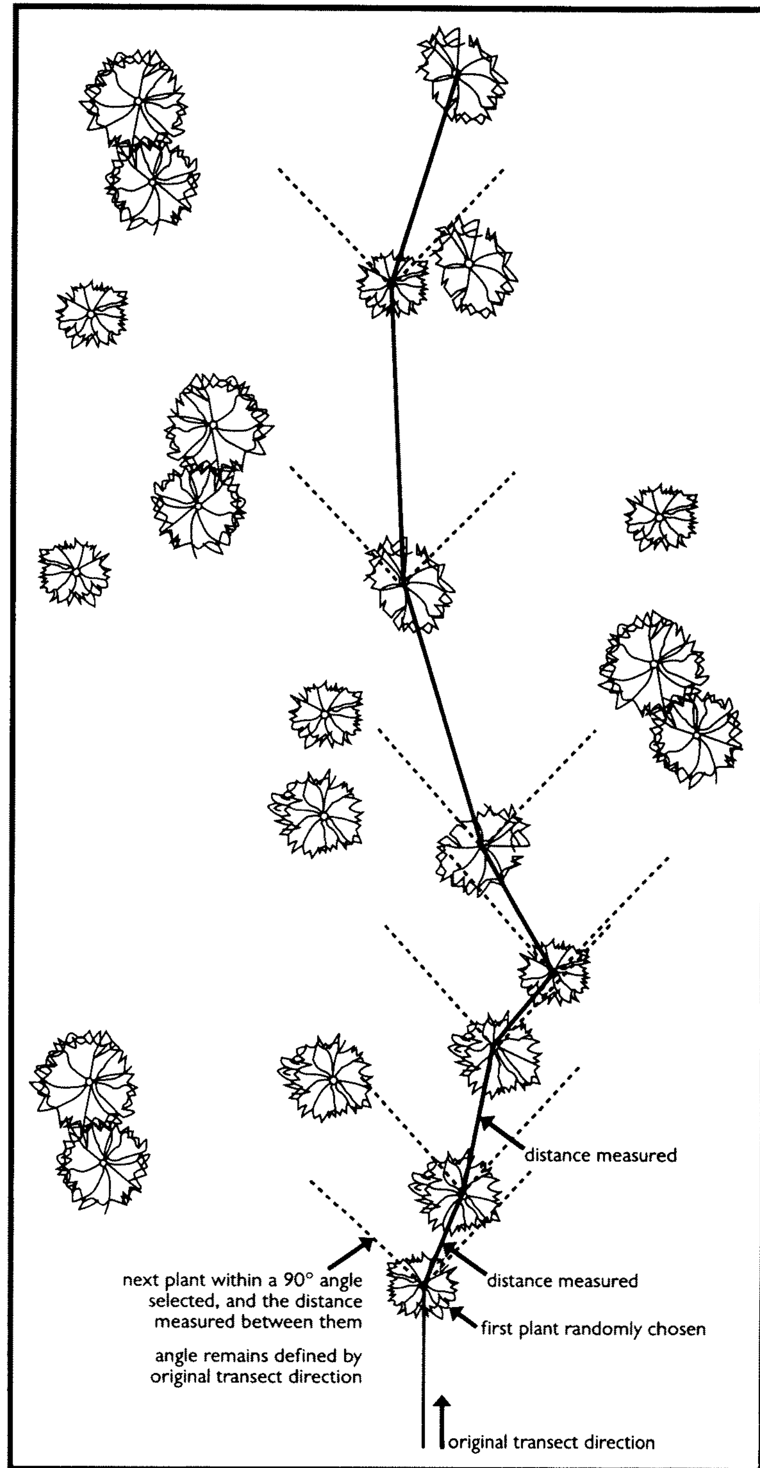


Figure 12.4. Wandering quarter distance measure, which can be used in plant populations with individuals contagiously distributed. This is the only distance measure recommended for general use since few plant populations are randomly distributed.

CLUMPED-GRADIENT POPULATION WITH 2m X 2m QUADRAT DESIGN

Figure 12.5. The clumped-gradient population with a grid of 2m × 2m quadrats overlaid on it. There are 1,250 possible quadrat locations for this size and shape of quadrat. Note that the quadrats do not overlap, yet cover the entire sampled population (the macroplot).

FREQUENCY

Frequency is usually measured in plots and can be defined as the percentage of possible plots within a sampled area occupied by the target species. With frequency sampling you are estimating the proportion of all possible quadrats in the population containing the species (or other attribute of interest) in them. You can visualize frequency by imagining the sampling area overlaid with a grid of cells the same size as the frequency plot. Figure 12.5 shows the clumped-gradient population introduced in Chapter 8 overlaid by a 2m × 2m grid. Of the total number of 1250 grid cells in this population, 540 cells have one or more plants in them. Thus, the true frequency is 43.2%. When you sample this population, you will randomly select some subset of these 1250 quadrats. Let us say you sample 100 quadrats. If 45 of the 100 quadrats contain the plant, then your estimate of the true percent frequency would be 45%. The percentage of cells occupied by the species is the frequency. Occupation is defined by occurrence; the abundance of the species within the plot does not matter, only whether it is present. Because the target species will more likely occur in very large plots compared with small ones, frequency is a measure that depends on plot size and shape. Frequency values from different studies are not comparable unless the plots used were identical.

Uses, Advantages, and Disadvantages

Frequency is a versatile measure, unconstrained by species growth form, and suitable for monitoring a variety of plant types. It is often used in community studies. Frequency is especially sensitive to changes in spatial arrangement. It

may be appropriate for monitoring some annuals, whose density may vary dramatically from year to year, but whose spatial arrangement of germination remains fairly stable unless the population is changing. It is also a good measure for monitoring invasions of exotic plants, where changes in spatial extent are often a key management concern. Rhizomatous species, especially graminoid species growing among similar vegetation, are often measured by frequency because no counting unit is required as with density.

An advantage of frequency methods over methods for measuring cover (see below) is the longer time window for sampling. Once plants have germinated, frequency measures are fairly stable throughout the growing season, compared with cover measures, which can change dramatically from week to week as the plants grow.

Another advantage of frequency methods is the high degree of repeatability among different observers because the only decision required of the observer is whether the species occurs within the plot. Other less-objective methods, such as ocular estimates of cover, require more extensive training to reduce variability among observers. Frequency can usually be measured consistently with minimal training on methodology. If the species is distinctive, frequency plots can be evaluated very quickly.

A disadvantage is that frequency is a measure affected by both the spatial distribution and the density of a population (Grieg-Smith 1983).Therefore, changes can be difficult to interpret biologically since we will not know if a change is caused by changes in density, distribution, or both (Fig. 12.6). Unlike other vegetation measures such as density or cover, frequency is difficult to visually estimate for a whole site. Thus, the biological significance of changes may also be difficult to communicate to managers and user groups.

Design

Quadrat size has a strong influence on the resulting percent frequency values. The larger the plot, the greater the likelihood that an individual will occur within the plot, resulting in a larger overall frequency value. If you make the quadrat large enough, you will have some individuals in every quadrat, giving you a frequency of 100%. This precludes you from detecting any upward changes in frequency. On the other hand, if your quadrat is very small, you will end up with very low frequency values that will not be sensitive to declines in frequency. Good sensitivity to change is obtained for frequency values between 30% and 70%, but if you are concerned about change in only one direction, or that the change may be dramatic, you may wish to change these target percentages. For example, if you are only concerned about declines, you may want to target your initial measure to between 50% and 80% to provide a wide margin of sensitivity to declines.

The advantages of long narrow plots for estimating density, cover, and biomass are clearly shown in Chapter 8. For these types of estimates, square plots are inefficient for many plant populations because their clumped spatial distributions lead to few sampling units with large values of the species, and many sampling units with few or none of the target species (a situation that increases the variability of values and decreases the precision of the estimate). With frequency data, however, only two values are possible—present or absent—and you want at least 30% of your plots to contain no plants. For this reason, use square plots when sampling frequency and adjust the size of the plot to reach the desired frequency range. An exception to this rule may be sparsely distributed rare taxa for which square frequency plots would have to be very large to contain plants 30% of the time. For these plants, a more rectangular frequency plot may be advantageous. If you use a rectangular quadrat, you must be sure to orient the long sides of the quadrat in the same direction in each year of measurement.

Because frequency values are measured separately for each species, the optimum size quadrat for one species may be less than optimum or even inappropriate for another. If you are measuring the frequency of more than one species, this problem is partially resolved by the use of a quadrat frame that includes nested quadrats of different sizes. Nested plots are often used in

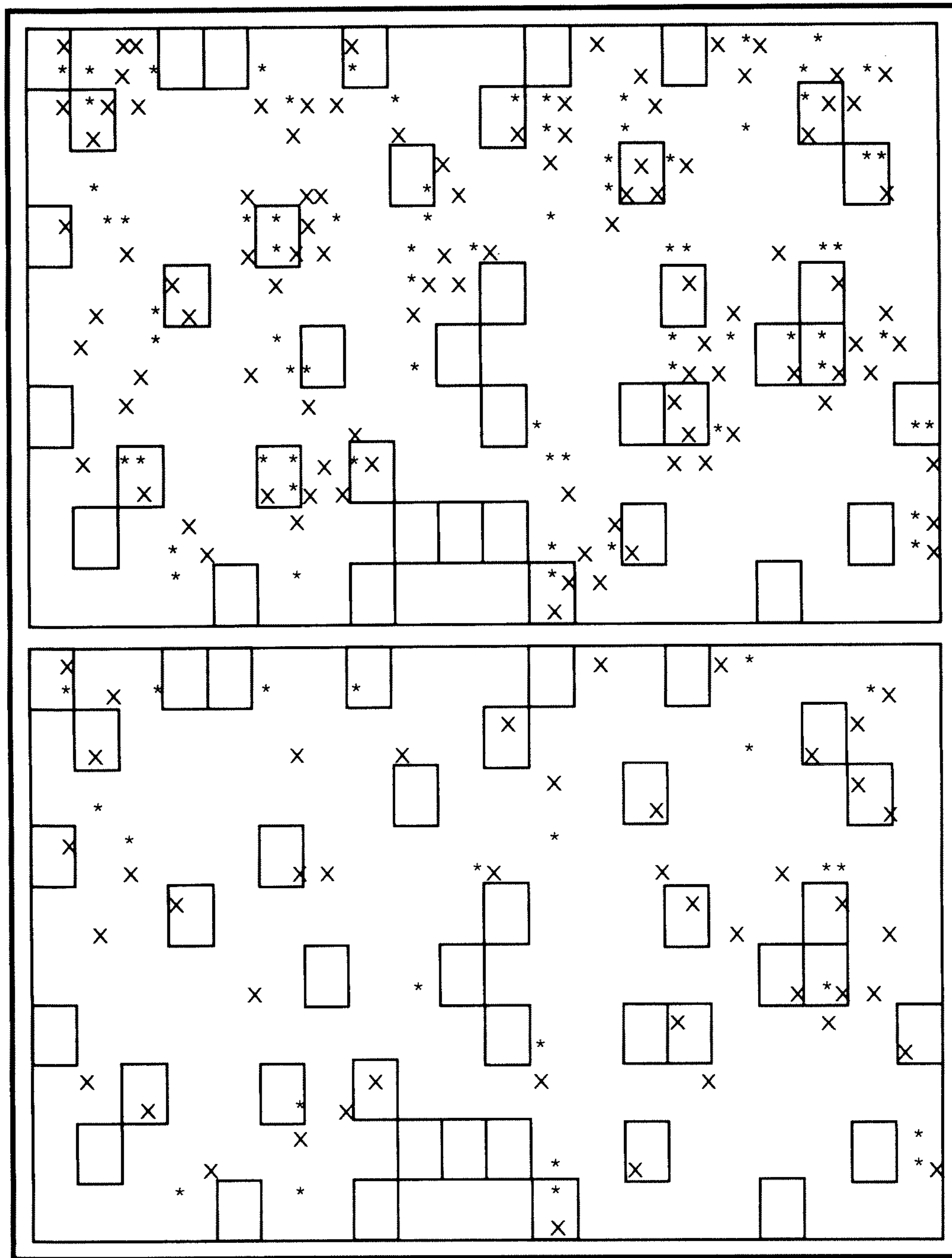


Figure 12.6. This macroplot was sampled with 40 permanent frequency plots. The first year, density in the macroplot was 198 individuals—72 seedlings (*) and 126 adults (X). The second year, density declined to 71 individuals—23 seedlings, 48 adults. Frequency between the two years declined from 57.5% to 52.5%.

sampling range vegetation; a common frame size is 50cm × 50cm, with four smaller plot sizes nested within the 50cm × 50cm frame (5cm × 5cm, 25cm × 25cm, and 25cm × 50cm). A nested frequency frame with square frames measuring $0.01m^2$, $0.1m^2$, and $1.0m^2$ for plant community monitoring studies is another common configuration. When reading nested frequency quadrats, always start with the smallest quadrat size. Any species that occurs in the smallest quadrat does not need to be searched for in the larger quadrats because its presence in the smallest quadrat indicates presence in all larger quadrat sizes because the quadrats are nested.

Nested quadrats may also be useful for monitoring a single species if frequencies are expected to change substantially over time. For example, if the frequency of a species increases from 50% to 90% in a particular quadrat size, that quadrat size will no longer by useful for detecting further increases. In this case, you would want to switch to a smaller quadrat size for future assessments. If frequency data had also been gathered in a smaller quadrat as part of a nested design, you would already have data for the smaller quadrat size, and you would not lose continuity by changing to a smaller quadrat size. A nested design that gives about 20% for one plot size and 80% for another provides a greater range for measuring large upward or downward changes the following year compared with a single plot size that gives approximately 50% frequency the first year. Nested plots may also be advantageous for measuring populations by stage classes. If, for example, seedlings are more abundant than adult plants, a smaller plot for seedlings nested inside a larger plot for adult plants may be very efficient. Finally, using nested plots in a pilot study is the best approach to determine the optimum plot sizes to use.

A special case is a plot size reduced to a point. These data can be considered a frequency measure but are most often interpreted as a measure of percent cover. Cover estimated by point intercept is described in the section Comparison of Plots, Points, and Lines.

The key decision in frequency measures is whether the species occurs in the plot. While this is relatively straightforward for small, single-stemmed plants, it is more difficult for larger plants and matted ones. You must establish boundary rules and apply them consistently. Some researchers have used the rule that if a perennating bud occurs within the plot, the plant is included (Bonham 1989). Under this rule, shrubs and trees with live buds that fall within the volume of the plot (as projected upward in space) would be considered in the plot. Most researchers, however, use rooted occurrence. Developing boundary rules similar to those described for density is important for all frequency studies, but is especially so for plants with wide bases such as bunchgrasses.

Frequency plots may be located randomly throughout the sampled area or may be placed along transect lines. Plots arranged systematically along randomly located transects are much more efficiently located in the field than plots that must be located individually at random coordinates. If frequency plots are located along transects, you can consider either the transect or the individual frequency quadrats as the sampling unit (see Chapter 8). Because sample size of the individual frequency quadrats will be much larger than the sample size of the transects, your estimates will be more precise and significance tests more powerful if you consider the individual frequency plots the sampling unit.[4] You can only do this, however, if the frequency plots are spaced far enough apart along the transect to be considered independent observations (see above). We recommend locating frequency plots along transects using systematic placement with a random start (see Chapter 8). The distance between frequency plots should be about the same as the average distance between transects, ensuring adequate interspersion and independence.

Frequency plots can be either temporary or permanent. Permanent plots should be placed along transects. Monumenting individual permanent frequency plots can be very time-consuming, but it is the most reliable means of ensuring that the same quadrat position is accurately relocated in the future. Alternatively, by monumenting the ends of a transect and providing periodic monuments along the transect to ensure later accurate relocation, quadrats placed along a transect can be relocated fairly accurately and can be considered permanent (see Chapter 5). This permanent design is usually (but not always) much more powerful for detecting change (see Chapter 8). Permanent plots may also provide greater biological understanding if you can relate changes to spatial data. Because you know the location of the changed quadrats, you may be able to hypothesize causes of change (opening of the canopy, wet microsite, invasion of weeds, etc.).

---

[4]*Frequency data are analyzed by the chi-square test (if plots are temporary) or McNemar's test (if plots are permanent). If you consider the transect the sampling unit, however, you would average the frequency values of the plots along a transect to calculate the value for the transect and then use parametric tests (e.g.,* t-*test and paired* t-*test) to analyze the sample of transects. Chapter 9 discusses these issues in detail.*

As with density, you must decide whether to evaluate occurrences by size or stage class (such as seedling, nonreproductive, reproductive). Using stage classes increases the amount of time required to evaluate each plot, but it can dramatically increase the understanding of frequency change in many cases. At a minimum, consider separating into seedling and nonseedling classes.

COVER

Cover is the vertical projection of vegetation from the ground as viewed from above. Two types are recognized. Basal cover is the area of the intersection of the plant with the ground; aerial cover (also called canopy cover) is the vegetation covering the ground above the ground surface. You can visualize aerial cover by considering a bird's-eye view of the vegetation.

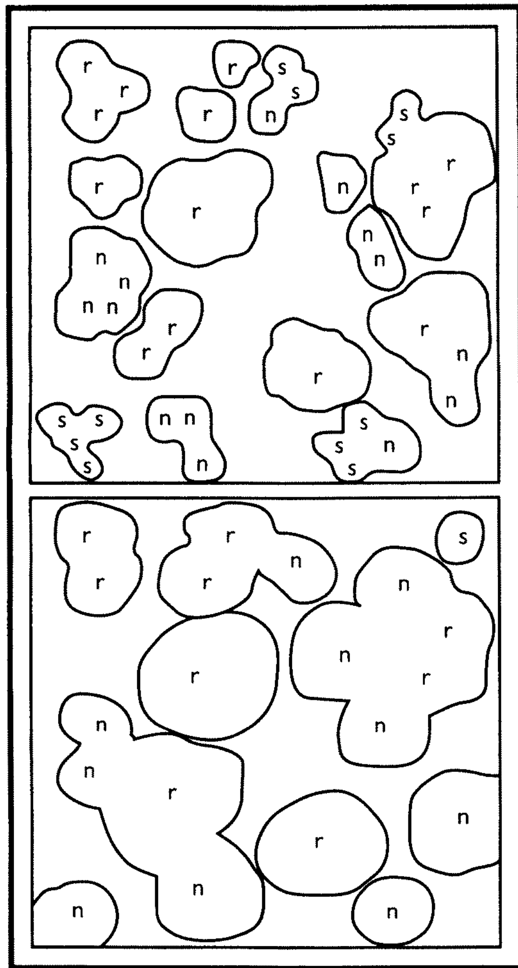


Figure 12.7. Two views of the same quadrat, the top measured in 1995 and the bottom in 1996. Note that the density declined (from 39 individuals to 21), while cover actually increased. Also note the scarcity of seedlings in 1996, which would not have been detected by cover methods unless cover was measured separately for adults and seedlings.

Uses, Advantages, and Disadvantages

Cover measurements can be made on plants of any morphology, although not all methods of measuring cover are equally applicable to all types of plants. Plants with extremely low cover (e.g., 5% or less) are probably better sampled by another method such as density or frequency. At such low cover values, estimates must be quite precise to have any sensitivity to change.

Cover is one of the most common measures of community composition because it equalizes the contribution of species that are very small but abundant and species that are very large but few. Of the three measures—density, frequency, and cover—cover is the most closely related to biomass or annual production. Measuring cover does not require the identification of the individual (as density does), yet it is an easily visualized and intuitive measure (unlike frequency).

A disadvantage of cover measures (especially canopy cover) is the potential for dramatic change over the course of a growing season, while both frequency and density measures are fairly stable after germination is complete. Growing-season changes in cover may make it hard to compare results from different portions of large areas where sampling takes several weeks or a few months. Sampling must be done at the same stage of the growing season during each measurement event. Because comparable stages will probably not occur on similar calendar dates given variation in annual weather, planning field work may be difficult.

Another disadvantage is that cover measures are sensitive to both changes in number (mortality and recruitment) and in vigor (annual biomass production). Because you may be unable to determine whether measured cover changes are the result of density or production changes, cover trends can be difficult to interpret. Real trends in density may be obscured in species with highly variable annual production. For example, increases in cover can obscure significant mortality (Fig. 12.7). For plants with less annual variability such as shrubs and matted perennials, cover changes will be caused primarily by mortality or recruitment. Because basal cover is generally less responsive to annual weather events than canopy cover, annual variability for all species will be highest with measures of canopy cover.

A final disadvantage is that cover estimates are susceptible to variability caused by wind. Fine-leaved vegetation, such as grasses, is especially prone to "lay over," exposing greater surface area and increasing canopy cover.

Design

Cover data are usually collected in one of three sampling units. Small quadrats are used for visual estimates of canopy cover (often in classes). Transects are used as line intercepts, measuring the percentage of the transect intercepting the canopy cover of the species. Finally, points are measured to determine if they intersect the species. All three approaches have been used in plant ecology for over 50 years, and many studies have compared their relative strengths (Bonham 1989). Growth form and the objectives of the study are the key determinants of the best sampling unit for a particular situation.

Table 12.1. Cover estimation classes recommended by Braun-Blanquet, 1965 (B-B), Daubenmire, 1959 (DAUB), and Jensen et al., 1994 (EcoData).

CLASS	B-B	DAUB	ECODATA
			<1% (+)
	Very small		1–5% (c)
1	1–5%	1–5%	6–15%
2	6–25%	6–25%	16–25%
3	26–50%	26–50%	26–35%
4	51–75%	51–75%	36–45%
5	>75%	76–95%	46–55%
6		96–100%	56–65%
7			66–75%
8			76–85%
9			86–95%
10			>95%

MEASURING COVER BY VISUAL ESTIMATES IN QUADRATS

Cover data collected in plots are usually based on a visual estimate of cover class. Many cover-class systems have been developed (Table 12.1); all are fairly similar, but the Daubenmire (1959) and the Braun-Blanquet (1965) systems are probably the most commonly used. Many later systems (e.g., Bailey and Poulton 1968; Jensen et al. 1994) split the lowest classes into even finer units. This is because in community studies, which are the most common application of plot cover methods, many species fall into these low cover classes.[5] For rare-plant monitoring studies, a cover-class system that is specific to the species may be more appropriate than any presented here.

Uses, Advantages, and Disadvantages

Cover can be estimated in small quadrats for most small-to-moderate-sized herbaceous plant species, although it is most easily and consistently estimated for plants with solid cover such as broad-leaved perennials. The method is difficult to apply to shrub species and to large herbaceous plants.

The key problem with visual estimation of cover in plots is the introduction of an unknown level of observer bias. Kennedy and Addison (1987) determined that changes in cover must be greater than 20% before the change can be attributed to factors other than observer bias and annual variation. Greig-Smith (1983) states that observer bias can be as high as 25% of the mean. Hope-Simpson (1940) concluded that a cover change of up to 23% could be attributed to observer disagreements. In a comparison of estimates by two trained observers measuring 5-m × 5-m plots, it was found that for 39.5% of the species there was a difference of one class assigned

---

[5]*For example, in a prairie ecosystem, Stohlgren et al. (1998) found that almost half of the plant species had less than 1% cover.*

by each observer, and for 3% of the species the observers differed by two classes (Leps and Hadicova 1992). Clymo (1980) found that estimates of cover of wetland vegetation in 25cm × 25cm plots could vary tenfold among observers. Fine-leaved and lacy-leaved species are more difficult to estimate consistently compared with broad-leaved species (Goebel et al. 1958; Clymo 1980; Sykes et al. 1983). Accurate estimates are especially difficult when the target species is intermingled with similar species such as a rare sedge that occurs in a meadow with dense cover of several similar grasses and sedges. Estimates are most variable among observers at moderate levels of cover (40% to 60%), but are least accurate at the lowest cover values (Hatton et al. 1986).

In spite of these limitations, using cover estimation in quadrats remains popular because of the ease and speed at which data are collected and because of its familiarity to many researchers.

Design

Several techniques have been used to improve the reliability and repeatability of visual estimates. The plots should be sized to allow evaluation of the entire plot at once; small quadrats (50cm × 50cm or less) produce more consistent visual estimates than larger ones (Sykes et al. 1983). Use of frames that include a known number of grid squares can also increase the consistency of estimates among observers. In a study of sessile marine species, Dethier et al. (1993) used a 50cm × 50cm frame divided into twenty-five 10cm × 10cm squares, each of which was considered 4% cover. Incompletely filled squares were grouped. This method resulted in visual estimates that were more similar among observers than estimates made with 50-point intercepts in each frame and required only half the field time. Another approach that has been successful in reducing the variability between observers is training with pieces of cardboard of known cover values. You should assess observer variability during a pilot study by conducting trials using several observers. If variability is extremely high, either take steps to reduce variability or use another method of estimating cover.

As with most methods of measuring vegetation, boundary decisions are important. Species occurrences within a plot are usually based on canopy occurrence—if a portion of the plant overhangs the plot, it is considered in the plot. Boundary decisions in this case are made in space above the plot at the point the overhang intersects the plot volume (if you imagine the plot projected upward). You can use a meter stick with a level to carefully project the boundary of the plot vertically to determine how much of the plot is covered by the overhanging canopy of the species, or you can use a frame with vertical guides placed at each corner of the plot.

While visual estimates are most consistent within small square quadrats, from the perspective of statistical precision, the same types of considerations as those given for density apply: long, narrow quadrats will likely provide better estimates than circular, square, or shorter, wider rectangular quadrats. The best approach is usually to randomly position transects in the population to be sampled and to systematically (with a random start) place square or small rectangular quadrats of a size that facilitates accurate cover estimation along each transect. The transects, not the quadrats, are treated as the sampling units. The transects will intersect several clumps of the population, thus ensuring much of the variation will be incorporated within each sampling unit. This design is really a two-stage sampling design, with the transects serving as the primary sampling units and the quadrats serving as the secondary sampling units, but you can treat the data as a simple random sample of transects (see Chapter 8). Cover percentages from the individual plots are simply averaged over the transect, and the average value is treated as the value for each transect.

Because each transect is a single sampling unit, the precision of cover estimates will depend on the variation among transects. Transects should be long enough to cross most of the variability in the vegetation being sampled (for the same reasons discussed relative to quadrat size and shape for density sampling). Just as for density quadrats, the optimum transect length should be determined from pilot sampling (see Chapter 8).

In rare cases, a small plot may be a good plot design (size and shape) to use and may be appropriate as the primary sampling unit. If so, you may still want to arrange plots along transects

for ease in locating them in the field. If they are far enough apart to be considered independent (see above), and if distributing them along transects results in adequate interspersion, you can consider each quadrat the sampling unit (see Chapter 8).

MEASURING COVER WITH LINE INTERCEPTS

Canopy cover is measured along a line intercept transect by noting the point along the tape where the canopy begins and the point at which it ends (Fig. 12.8). When these intercepts are added, and then divided by the total line length, the result is a percent cover for that species along the transect.

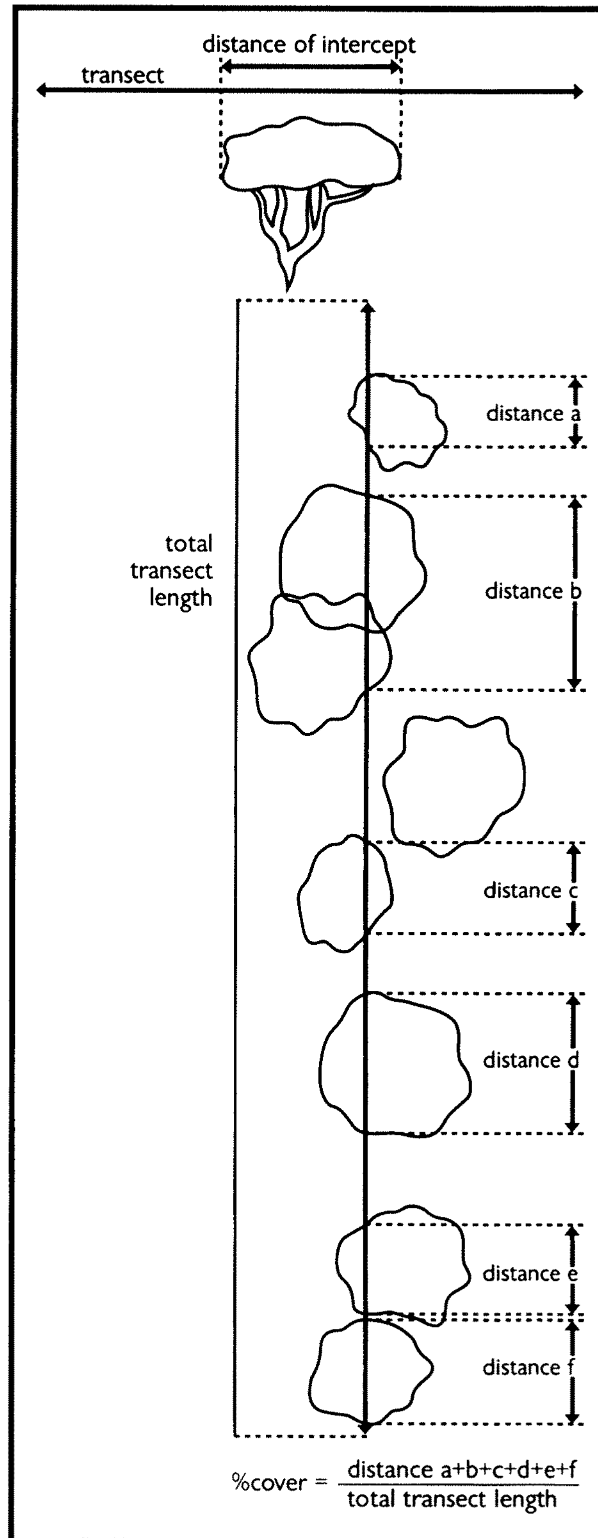


Figure 12.8. Line intercept method of measuring cover for a single shrub species.

Uses, Advantages, and Disadvantages

Line intercept techniques are most effective for species with dense unbroken canopies such as some shrubs and matted plants. Line intercept is less effective for plants with lacy or narrow canopies, such as grasses and some forbs and shrubs, because of the large number of gaps and small interceptions requiring evaluation. Line intercept is commonly used for shrubs that are less than 1.5 m tall because a tape can be suspended above the shrub canopy and the interception measured. Visually estimating canopy cover of shrubs of this height in small quadrats would be difficult. Line intercept can also be used on shrubs with taller canopies if the observer has a means to project the intercept upward such as an optical sighting device, and if the transect can be physically placed through the shrubs. In dense shrub vegetation this may be quite difficult.

Line intercepts are quickly measured for plants with low but densely clustered cover. Plants that are sparse, small, and well-distributed along a line will require more meticulous evaluation of the transect.

Repeatable measures are difficult to achieve with line intercepts if the wind is blowing. Not only must you locate the intersection of the tape with a moving target, but the tape also bows in the wind. While blowing tapes can be secured somewhat by clipping to stakes every 5m to 10m, a more difficult problem is that in hard wind the vegetation lays over at an angle (especially in grassland systems) and presents a larger surface area than would be available under still conditions.

Cover measured by line intercept is less prone to observer variability than visual estimates in quadrats, but more susceptible than using point intercepts, which are discussed in the next section. Good design of line intercept studies requires developing protocols to deal with canopy gaps and ensuring that observers measure along the intercept without bias.

Design

Few plants form complete canopies that lack any gaps. Typical gaps are formed by dead centers in bunchgrasses, fractured canopies in matted plants, gaps between blades of grass, and gaps between branches of shrubs. Lacking protocols, observers treat gaps variably. To improve consistency among observers, a maximum gap width should be established during design. Observers assume a closed canopy until the gap exceeds this width. Bonham (1989) suggests 2 cm for a maximum gap width, but gap protocols should be designed for particular species and situations and should be clearly documented in the description of the sampling methodology to ensure consistency among observers over time.

The theoretical basis of line interception depends on reducing the width of the lines to zero (Lucas and Seber 1977; DeVries 1979; Floyd and Anderson 1987). Line intercepts should be read only along one edge of a measuring tape. Ensure that the tape is not inadvertently moved to include or exclude certain plants. A related source of observer bias stems from a sighting line that is not perpendicular to the tape. If the tape is suspended at some distance above the canopy, movement of your head can change the location of the intercept because you are unable to maintain an exact perpendicular orientation without tools to help. One option is to use two tapes, placing one above the other and sighting along the two edges. Another option is to suspend the tape over the vegetation and use a plumb bob to locate canopy starts and stops. For overhead vegetation, a pole with a level can be used. None of these are extremely accurate, but they help reduce observer bias (which for most people is to include as much of the canopy as possible within the intercept). The most accurate method for locating canopy boundaries of both low and overhead vegetation is to use some type of optical sighting device (described under points, below).

The sampling unit for line intercept is always the transect, and an important sampling design issue for line-intercept sampling is the length of the transect. Longer transects will cross more small-scale variability, reducing the number of transects needed for a given precision of the cover estimate. Longer transects, however, require more time to measure and are difficult to establish in dense vegetation. Another design issue is deciding whether to make line intercepts permanent or temporary. Because most species measured with line intercepts are long-lived, with high correlation between years (taxa for which permanent designs are most efficient), line intercepts are often permanent. See section on Production and Other Vigor Indicators for more information.

POINT INTERCEPTS FOR MEASURING COVER

Cover is measured by point intercept based on the number of "hits" on the target species out of the total number of points measured (Fig. 12.9). Because at each point the only decision is whether the point intersects the species, measuring cover by points is considered the least biased and most objective of the three basic cover measures (Bonham 1989). Point intercepts are not subject to observer variability from canopy gaps or visual cover estimations.

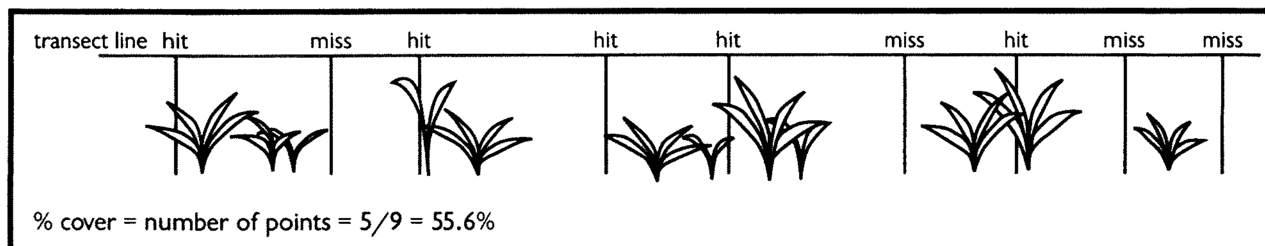


Figure 12.9. Point intercept method of measuring cover.

Uses, Advantages, and Disadvantages

Point intercept can be used for species of any morphology. It may be the best way to measure cover for fine-leaved plants such as grasses or species with open, lacy canopies, which are difficult to visually estimate in quadrats or measure along line intercepts.

The cover most often measured by points is canopy cover. Cover can also be measured within defined layers (i.e., the cover of individuals over 50cm tall and those less than 50cm tall) or by different species. For both approaches you will likely record more than one interception at each point, depending on how many layers have been defined or how many species occur at the point. This can become quite time-consuming. You can also use points to measure multiple layers of a single species by recording all the interceptions with the species as a pin is lowered to the ground (Goodall 1952). Note that this measure is no longer a measure of canopy cover since the pin may intercept the same individual or same species more than once at each point. Most researchers have interpreted multiple interception measures as an index of biomass, volume, or composition (Goodall 1952; Poissonet et al. 1973).

Design

Most cover measures are perpendicular to the ground, but species with narrow, upright leaves are rarely encountered with this angle. Other angles have been used to increase the number of "hits" on these types of plants (Bonham 1989), although angled pins eliminate the value of visualization of canopy cover as a bird's-eye view. The monitoring methodology should always specify the angle used.

Points are measured either with pins that touch the vegetation or with a crosshair optical sighting method. Most optical sighting devices employ a mirror system to allow the observer to remain standing while looking at the ground. Crosshairs in the field of view identify a nearly dimensionless point. These devices are usually mounted on a tripod, and can be set for a specific angle of intercept. They are quick, accurate, and fairly easy to use, but are costly ($500 to $1500). Their only other disadvantage is that a second observer is needed to move canopy vegetation from the line of sight if the target species is an understory plant. This movement may change the probability of intersecting the target species.

Inexpensive sighting tubes with crosshairs made of fine wire, fishing line, or dental floss can be constructed, but often require the observer to bend awkwardly to look downward and are also difficult to maintain at a constant angle. Buell and Cantlon (1950) and Winkworth and Goodall (1962) give complete directions for their versions. For braver do-it-yourselfers, Morrison and Yarranton (1970) describe the construction of a high-quality optical device from a rifle telescope, a right-angle prism, and a homemade frame.

Pins are inexpensive and easy to use. Their key disadvantage is the error associated with the diameter of the pin, resulting in overestimation of actual cover, especially for narrow or small-leaved species (Warren-Wilson 1963). This is generally not a problem in most monitoring situations where change is of interest rather than the actual cover value. It is important, however, to use the same pin diameter for successive measurements. Pins should also be used in some type of frame that eliminates the bias that results from attempting to manually place a pin vertically through vegetation (a tripod frame can be constructed that holds only a single pin).

The sampling unit depends on the arrangement of points. Points can be sampled in frames (which then form the sampling unit), as single, randomly located points (each point a sampling unit), or as points located along a transect (either points or the transect forming the sampling unit). The original method of point interception used a linear point frame of 10 pins as the sampling unit (Levy and Madden 1933), with suspended pins that could be gradually lowered until contacting the vegetation. Point frames can also be rectangular, with a grid of points (Floyd and Anderson 1982). The literature is replete with variations on themes of point frame size and shape, but frames are rarely the most efficient sampling unit. As early as the 1950s, Goodall

(1952) demonstrated that sampling point intercepts using frames was much less efficient than sampling random points, and later studies have supported his conclusion (Evans and Love 1957). Depending on the vegetation, the time required for measuring single points as the sampling unit can be one-third to one-eighth that required for point frames, since many more points must be measured in the latter (because the frame is the sampling unit) to achieve the same precision as independent random points. Unfortunately, point frames continue to be regularly used in vegetation sampling.

Rarely is each point-intercept sampling unit located randomly. Point-intercept sampling often involves sample sizes greater than 100 points, a prohibitive number to locate randomly throughout an area. The better approach is to arrange points along transects. If points are far enough apart, they can be considered independent sampling units (see Chapter 8). If fairly close together, the transects can be considered the sampling unit. It is almost always better to space points far enough apart to consider each point an independent sampling unit.[6]

If the transect is the sampling unit, how many points should be placed along each transect? Fisser and VanDyne (1966) found that it was best to sample with fewer points and more lines when using the transect as the sampling unit. This design maximizes the number and interspersion of sampling units throughout the sampled area. The number of points you place along the transect, however, controls the resolution of the cover value if your sampling unit is the transect. For 10 points, for example, only cover values of 0%, 10%, 20%, 30%, etcetera are possible. With 50 points, cover values can be measured in increments of two: 2%, 4%, 6%, etcetera. At a minimum, you want enough points so that you will intersect at least some individuals of the species of interest along each transect line. This may require many points (50 to 100 or more) for some species with very low cover.

If the point is the sampling unit, using more points and fewer transects may be advantageous as long as points remain well dispersed throughout the sampled area. When points are located along transects, the largest field expense is incurred by the location and establishment of each transect. Maximizing the number of points per transect would minimize setup time. Transects may, however, have to be quite long in this design to have points located far enough apart to be considered independent sampling units (see Chapter 7). If the structure of the vegetation presents challenges to establishing long transects, more short transects with fewer points per transect may actually require less time.

Point intercepts are almost always temporary rather than permanent sampling units. The section Production and Other Vigor Indicators discusses this in more detail.

Many points may need to be sampled. For example, to estimate cover of about 50% within ± 5% absolute cover at a 95% confidence level, 384 points would be needed. To estimate within ± 1% absolute cover, however, over 9600 points would be needed (see Appendix II for equations to calculate sample size). Estimating cover within ± 5% would probably be acceptable at higher cover (e.g., 50% cover), but estimating cover within ± 5% cover may not be acceptable at lower cover (e.g., 7% cover).

Such large sample sizes are mostly a problem in community sampling, where each point requires recording the species intercepted. In a study that monitors the cover of a single species such as a rare species with low cover, many points would require no evaluation beyond the fact that the species is not there. You can imagine that a 100-m transect of 20 point intercepts could be evaluated very quickly if it only crosses a clump of the target species once. Most points could simply be checked visually; only those points that are close to intercepting the target species would require a setup of the point frame or optical sighting device.

---

[6] *If the individual point intercepts are the sampling unit, the data are analyzed by the chi-square test (if points are temporary) or McNemar's test (in the unlikely event that points are permanent). If you consider the transect of points the sampling unit, however, you would use the percentage of points intersecting the samples as the value for the transect and then analyze the sample of transects using parametric tests (e.g.,* t*-tests and paired* t*-test). Chapter 9 discusses these issues in detail.*

PERMANENT SAMPLING UNITS FOR MEASURING COVER

All three methods of monitoring changes in cover—plots, line intercept, and point intercept—can be used with either permanent or temporary sampling designs. Remember that permanent sampling designs will be far more efficient than temporary designs if there is a high between-year correlation between the sampling units. Two exceptions are: 1) when points are part of a permanent sampling design, in which case the transects should always be treated as the sampling units, and 2) when the plants in the sampled area respond dramatically to an environmental gradient and transects can be arrayed in a manner that incorporates the resulting variability, providing transect values that are similar. In this situation, treating the transects as the sampling units is likely to be more efficient than treating the points as the sampling units.

A key consideration with measuring cover with permanent sampling units is whether you can actually make them permanent. Measuring the exact same line intercept each time is much more difficult than measuring within a permanent density plot in which all four corners of the plot are permanently monumented. Changes in tape tension (sag), bowing in the wind, and slightly different placement because of brush are examples of factors that may reduce the correlation between each measurement. If you intend to use a permanent design, consider the following factors:

Plant morphology. Failure to intersect a particular plant on the second measurement—a plant that was recorded during the first measurement—results from 1) tape and point movement and a miss of the exact location, 2) decline in cover of the plant so that your sampling unit no longer intersects it, or 3) death (or dormancy) of the plant. Only the first is a problem; the second two scenarios are true changes. You can perhaps guess which scenario is most likely based on the size and morphology of the target species. If the plant you are sampling has a fairly large area, you will likely intersect the same individual at the second measure because you have the area of the plant as room for "error." Thin-leaved species are more problematic than matted species because only minor movements of a point or a line will result in missing a plant that has not changed.

Field conditions. The exact relocation of a permanent transect in places that are difficult to traverse such as dense brush is unlikely. You can be more confident of accurate tape relocation in a short-grass prairie.

The sampling unit. If points are the sampling unit, your individual points must be correlated from year to year. If transects or plots are the sampling unit (either as a line intercept, a line of points, or a collection of plots), the transects or plots must be correlated. Points are much more difficult to relocate than transects or plots. For this reason we do not recommend the use of permanent designs that treat points as the sampling units.

Several field techniques can reduce placement error (see Chapter 5). You should monument transects with permanent markers at each end and at intermediate positions along the transect. The number of intermediate markers depends on the field circumstances. In dense brush a transect may require a marker every few meters to ensure accurate relocation, while at a meadow site every 10 or 20 meters may be sufficient. Shorter transects are less affected than long transects by tape stretch, bowing, and sagging or by using alternative pathways around large vegetation. For cover estimation in plots, marking one or two individual plot corners, as well as the transect ends, will ensure that plots are relocated accurately. This monumentation adds to the time required to establish a study, and these costs must be weighed against the benefits gained from a permanent design compared with a temporary design (see Chapter 8).

In general, permanent sampling units for cover, especially using points as the sampling unit, may be very difficult to achieve in field settings, although they usually do increase the efficiency of the design for measuring change (Goodall 1952). If you intend to use a design with permanent

sampling units, test the degree of physical correlation by conducting a measure, picking up the tape, and then having a second observer reestablish the transect and complete the measurements. If the correlation between the two measures is not good, you should use a sampling design with temporary sampling units.

COMPARISON OF PLOTS, POINTS, AND LINES

You must choose transects, plots, or points as the sampling unit for measuring cover. The best sampling unit depends on the total cover of your species, its distribution in the field, and its morphology.

Transects Versus Plots

Daubenmire (1959) found that the cover estimates from 40 to 50 quadrats was nearly identical to that measured by 350m of line intercept. Standard error of the quadrat samples, however, was high (likely because many did not contain the target species). Bonham (1989) states that line intercept is more accurate than quadrats when working with different-sized plants. Hanley (1978) found that at low cover (8%), line intercepts required about half the time to achieve the same precision as randomly placed quadrats, but at 26% cover, the two methods became more comparable (34 minutes for quadrats compared with 29 minutes for lines). Note that these studies evaluated efficiencies of estimating cover of all species in community studies. When sampling a single species, the most efficient method will be determined by the morphology of that particular species.

Points Versus Plots

Dethier et al. (1993) created simulated plots containing a known cover of 13 species and compared cover measured by point intercept with cover visually estimated to the nearest percent in the plot. Cover estimations done with the aid of subdividing the plots into 4cm x 5cm rectangles were close between observers, and they were closer to the true value of cover than measured points. In the field, point intercept failed to detect 19% of the species that were detected by cover estimation, all of which had cover values of 2% or less. Differences among observers were less for cover estimation than for point measurements. Meese and Tomich (1992) similarly noted a problem with point intercept for detecting rare species.

Transects Versus Points

Floyd and Anderson (1987) found that point interception achieved the same precision as line interception in one-third the time. Line transect and points gave similar results in a study by Heady et al. (1959), but points required only about half the field time and less office time compared with line intercept. At low cover (3% or less), line intercept gave better results. Brun and Box (1963) found that point intercepts required less than two-thirds the time of line intercepts to achieve the same precision. In contrast, Whitman and Siggeirsson (1954) found that points and line were similar in time requirements.

PRODUCTION AND OTHER VIGOR INDICATORS

Production is the annual output of vegetative biomass. It is most commonly measured as a harvest of aboveground standing crops, usually at peak (before plants start senescing and losing leaves). This approach underestimates total annual production, missing biomass consumed by herbivores, loss that occurs throughout the growing season, below-ground production, and regrowth after harvest.

Vigor indicators are many and include height, basal diameter, number of flowers, number of inflorescences, number of leaves, number of stems, number of leaf whorls, diameter of rosette, and volume of plant (height × cover).

Uses, Advantages, and Disadvantages

Production varies each year depending on the favorability of growing conditions and therefore may not be sensitive to the type of trend that is of interest in plant monitoring projects. Production is also usually sampled destructively by harvesting, drying, and weighing, and for most rare plants this type of monitoring is not appropriate. Because of these constraints, production is not discussed at length here. If you are monitoring a more common species and wish to use a production measure, information on the subject is abundant (Malone 1968; Sandland et al. 1982; Ahmed et al. 1983; Bonham 1989; Ruyle 1991; Catchpole and Wheeler 1992).

Vigor indicators are also strongly influenced by annual weather patterns, but they may be appropriate for some monitoring questions. As nondestructive measures of vigor, they are appropriate for use on rare plants. Most are easy to measure, with little observer bias.

Design and Field Considerations

The key consideration often ignored in production and vigor studies is explicit definition of the sampling unit. Sampling units can be plots (e.g., grams produced/m^2, number of flowers/m^2), an individual (e.g., grams/individual, number of flowers/individual), or a part of the individual (e.g., seeds/fruit). The different sampling units have very different design considerations. See Chapter 8 for an extensive discussion of the difficulty of selecting a random sample of individuals and the use of cluster sampling or two-stage sampling to address that difficulty.

Biomass on a unit area basis is usually estimated by clipping or visually estimating in small square or rectangular plots. We have already noted that from the standpoint of statistical precision we are better off with long, narrow quadrats when estimating density, biomass, or cover. We have also noted the impracticality of using long, narrow quadrats for anything but density. A good compromise for biomass is to place square (or small rectangular) quadrats systematically (with a random start) along transects and to treat the transects as the sampling units. Thus, we are able to clip plots or estimate biomass efficiently in the quadrats, while at the same time crossing the variability in the population, making for more precise estimates of means. For analysis, we would take the mean of the quadrat values for each transect and use this set of transect means as our sample.

Edge effect is important in clipped plots. Measurement bias (usually overestimations) can be significant if the quadrats are too small (Wiegert 1962). Since aboveground vegetation must be clipped in some quadrats, circular quadrats should be avoided because of the difficulty in cutting around the perimeter of the circle with hand-shears and the nonsampling errors that will likely result.

CHOOSING AN ATTRIBUTE AND TECHNIQUE

The best measurement technique for monitoring plants depends on the morphology of the plant and the type of change you expect. Some species are difficult to monitor with any of the standard techniques. In this section, we give you some ideas for monitoring different kinds of plant species.

Fine-Leaved Rhizomatous Species (Grasses and Grass-likes)

Density is usually not appropriate for these species because tillers or ramets may be quite dense, difficult to distinguish, and difficult to count. Exceptions are coarse-stemmed grasses (such as the cosmopolitan common reed [*Phragmites communis*]) and sedges.

For fine-leaved rhizomatous species, frequency may be a good measure. You must consider the type of change expected, however, since many of these grasses respond to stress by thinning, which may not be detected in frequency plots. Cover is also a useful measure, but only point intercept is practically applied. Fine-leaved species may be difficult to visually estimate in plots, especially at low density, and line intercepts require very careful evaluation to accurately measure length of intercepts.

Matted Plants

Most matted plants are difficult to separate into individuals and thus are not easily counted for a density estimate. If frequency is used, it may be difficult to determine the rooting location for boundary decisions. You can use the canopy boundary of the plant instead. Frequency may be quite sensitive to mortality of this type of plant since it is likely that a single plant may be the only occupant of a frequency plot (although this relationship depends on relative plant and plot size).

Cover may be the best measure of matted plants. Because the canopy of these plants is usually fairly solid, estimates in quadrats, line intercepts, and points are all appropriate methods. Long-lived matted plants often change in canopy cover very slowly. For these types of plants, large changes in cover are probably caused by mortality or recruitment.

Annuals With a Long-Lived Seed Bank

One of the most difficult situations for monitoring is an annual species that only appears above ground once every few years or even once every few decades. Measuring aboveground expression such as density may provide some insight on the weather patterns that create a "good" year but little information on long-term trends of the population. Most of the population is out of sight, below ground, expressing itself only occasionally.

The study of seed banks is a fairly new discipline (Leck et al. 1989). The biggest problem with studying seed banks is that their distribution under ground is usually quite clustered. Because of this spatial distribution, when the small soil cores are used to sample the seed bank, a large number of cores will contain none of the target species, and a few cores will contain many. This creates a serious problem for determining seed bank density with any reasonable precision (Benoit et al. 1989), but can be addressed by grouping cores collected along transects in a two-stage sampling design (the transect is the primary sampling unit and the soil core is the secondary sampling unit; see Chapter 8). A second problem with studying seed banks is the labor expense. Extracting cores is time-consuming, but estimating the number of seeds in each core is even more so. Two methods are generally used: 1) growing out the cores in a greenhouse and counting the number of germinants, and 2) extracting seeds from the soil core by flotation and physically counting the seeds extracted (Gross 1990). Both are obviously labor-intensive. Both are also fraught with problems. The grow-out method is sometimes unsuccessful because dormancy-breaking and germination requirements are not met. The flotation method may extract dead seeds as well as live ones (Gross 1990).

An alternative that may be more successful than monitoring the species itself is to focus on the habitat. Habitat features such as level of human activity, invasion of exotics, and changes in community composition caused by succession may identify problems for an annual species. Note that for many annuals, some level of disturbance is necessary for exposure of the seed bank and germination; thus, change in disturbance level may be a sensitive attribute to monitor.

Extremely Long Lived Plants

These are species such as cacti, trees, shrubs, and some perennial herbs. Changes in density or population size may occur very slowly. For some, habitat conditions can change significantly before mortality occurs. Cover of these species is also often slow to change.

For some species, a vigor measure may reflect stress or decline. These include reproductive effort (number of flowers or fruit produced), canopy condition, number of new leaders, and extent of stress-related signs (such as disease). Monitoring may also effectively focus on the seedling or reproductive class, which is often much more dynamic and responsive to management than established adults. Monitoring changes in habitat condition or threats may also be useful in providing information for management modification.

Dense Shrubs

Most shrub species can be monitored using either density or cover measures, but it is extremely difficult to work with those that grow in dense colonies (especially if they are spiny). It is nearly impossible to stretch a tape reliably. Travel between sampling units is extremely slow.

The best method to sample this type of vegetation is remotely. Canopy cover could be monitored on low-level aerial photography. If, however, you are interested in a specific species growing among dense shrub vegetation, you might not be able to discern the species of interest on the photograph. In addition, if the species of interest is in the understory, this method would clearly not be useful.

Physically sampling the stand on the ground must employ a design that minimizes travel, both between sampling units and around sampling units. Density estimates, for example, would require placing a transect (as one side of a narrow quadrat) and then covering the transect distance again while measuring the quadrat. Point intercepts could be placed along paced transects (to avoid needing to stretch a tape), but it is unlikely that paces could be done without bias. Many points may be needed, requiring extensive travel through the stand.

If these shrubs are dense but short, a line intercept may be the easiest sampling unit to establish. With two observers, the tape could be suspended over the vegetation to tighten and straighten and then secured to fenceposts.

None of these sampling methods will be easily implemented, however. A better approach may be to focus on threats or habitat features that are more easily monitored.

Plants That Act As Metapopulations

Species that exhibit metapopulation behavior occur on the landscape with both temporal and spatial variability. These plants may be viable as a metapopulation over the entire landscape, but individual populations may be short-lived. Dispersal of seeds or propagules and available colonization sites are the two most important factors in the success of a metapopulation. A good example of a plant metapopulation is the Furbish's lousewort (*Pedicularis furbishiae*), found along a major river system in Maine. Populations of the species are eliminated by ice scouring and spring flooding, but new populations appear on suitable sites left bare by receding floods (Menges 1986, 1990). Because the plant has no seed bank, colonization depends on the dispersal of the fall seed crop to new sites. Metapopulation dynamics depend on a dispersal mechanism so that available habitat can be colonized as existing populations become extinct.

Many in the conservation community contend that consideration of metapopulation dynamics is crucial to any conservation strategy (Hanski 1989), while others argue that the importance of metapopulations has been overstated (Doak and Mills 1994). While a few empirical studies have shown the importance of metapopulation dynamics for some invertebrate and animal species, plant studies are much rarer. A review of the literature found only nine plant studies in which a parameter important to the theory of metapopulation dynamics—migration, extinction, or colonization—was actually measured (Husband and Barrett 1996).

While there are exceptions, most plant species disperse propagules locally (Harper 1977; Silvertown and Lovett-Doust 1993). In the absence of an obvious long-distance dispersal mechanism (such as the river in the Furbish's lousewort example), it is difficult to hypothesize how a plant species could function as a metapopulation and how to design management to allow that function to occur. It is also questionable whether the dispersal mechanisms important to

metapopulation dynamics that may have operated in the past can still operate in today's fractured and fragmented landscape.

In the absence of obvious potential for metapopulation dynamics, the most conservative strategy is to maintain both existing populations and some potential habitat areas. The latter can then provide opportunities for both natural colonization and deliberate reintroductions.

MANAGEMENT IMPLICATIONS

Several vegetation measures are available for monitoring plants, including density, cover, frequency, and vigor. Design features to optimize efficiency vary among measures. The selection of a measure depends on the morphology and life history of the plant and the type of change expected from management activities.

CHAPTER 13
Specialized Sampling Methods and Field Techniques for Animals

Penstemon moriahensis
Mt. Moriah Penstemon
Western Nevada in open woodlands
Artist: Jeanne R. Janish

Table 13.1. Monitoring efforts for different animal groups, 1980–2000. Data are number of reports cited in Wildlife Worldwide, an index to literature on wild mammals, birds, reptiles, and amphibians, and Fish and Fisheries Worldwide, an index to literature on ichthyology and fisheries that include the key words "population" and "monitoring."

GROUP	NUMBER OF REPORTS	PERCENT OF TOTAL REPORTS
Waterfowl, shorebirds, and wading birds (e.g., ducks, herons, plovers)	44	18
Songbirds (e.g., warblers, sparrows)	42	18
Large herbivorous mammals (e.g., deer, antelope, elephants, kangaroo)	26	11
Seabirds (e.g., puffins, gannets, boobies)	17	7
Amphibians (frogs, toads, salamanders)	14	6
Medium-sized mammals (e.g., mustelids, opossum, beaver, bandicoot)	14	6
Reptiles (e.g., sea turtles, tortoises, crocodilians, lizards)	13	5
Raptoral birds (e.g., hawks)	13	5
Large carnivorous mammals (e.g., wolves, bears, lions)	11	5
Small mammals (e.g., mice, voles, shrews)	10	4
Upland gamebirds (e.g., pheasant, woodcock)	7	3
Aquatic mammals (e.g., seals, manatees)	6	3
Warmwater fishes (e.g., bass, perch)	6	3
Insects (e.g., butterflies, crickets, beetles)	5	2
Coldwater fishes (e.g., salmon, trout)	4	2
Marine fishes (e.g., cod)	4	2
Flying mammals (bats)	2	1
Total	238	

Which animals do biologists most frequently monitor? A search of the current scientific literature (Table 13.1) indicates that the bulk of the monitoring effort is focused on birds (particularly songbirds, shorebirds, waterfowl, and seabirds) and large mammals. Other animals, including fishes, reptiles, amphibians, small- and medium-sized mammals, and insects, receive less monitoring attention. In some circumstances these groups of animals are observed and counted directly. Seabirds nesting on a cliff, bison grazing on a prairie, or large tortoises in a desert are examples. More often, however, animals are highly mobile and quite secretive. A variety of specialized techniques have been developed to sample such animals in ways that overcome the chronic problem that plagues animal monitoring studies, that is, incomplete detectability of individual animals. The specialized sampling and field techniques used for this purpose are the focus of this chapter. Related and useful references for designing animal population monitoring studies include Thompson et al. (1998), Sutherland (1996), and Krebs (1998).

COMPLETE AND SAMPLED COUNTS

For some animal species that are highly visible and not particularly mobile during a counting period, populations can be reliably estimated through complete or sampled counts. Generally speaking, complete counts (censuses) are effective only for small areas and only if animals are conspicuous and easily distinguished from one another. Deer and moose on an open range surveyed from an airplane are an example. In practice, there are few situations where animals are completely visible to observers, and hence such counts should always be considered minimum counts. Direct counts of animals on sample plots (sample counts) are usually preferable to attempted censuses. These generally require less effort, can reduce problems with double counting, and cause less disturbance, although they are susceptible to sampling errors and incorrect statistical assumptions about random distributions of individuals. For sampled counts of animals, all of the issues described in Chapter 8 concerning replication, precision, and plot size and shape apply.

MARK-RECAPTURE METHODS

Mark-recapture methods are frequently used in animal monitoring to adjust for incomplete detectability and thereby obtain estimates of true population size. These methods are time-consuming and expensive and involve some variation on capturing and marking a sample of animals, releasing them, and recapturing some fraction of them at a later point. They also rarely deliver precise estimates of population size unless a large fraction of a population is handled and marked. For example, to achieve a relative precision level of ±10% for an estimate of a population of 200 animals, 90% (nearly a complete census) must be captured (Table 2.4 in Greenwood 1996). The general imprecision of mark-recapture approaches applies to all methods no matter what their sophistication. A fine example of the pitfalls of basing monitoring on mark-capture methods is provided by Wilson et al. (1999). A further drawback to mark-recapture is that it may also be stressful to a population.

Often there are no other means available, however, to estimate populations of secretive animals other than through mark-recapture studies. It is also important to note that valuable ancillary information can be obtained once the investment has been made to perform a mark-recapture study. That is, many of these methods provide, in addition to an estimate of population size, inferences about movement patterns (immigration and emigration) and survival rates, thus permitting an assessment of not only the trends in populations but perhaps also the causes underlying those trends.

The basic idea behind all mark-recapture is that a sample of animals is captured, marks are attached to them, and the animals are released. Sometime later, once the marked animals have had the opportunity to mix with the rest of the population, another sample of animals is taken and some of the marked animals recaptured. The proportion of animals marked in the second sample should be equal to the proportion marked in the entire population, and an estimate of the total population is thereby obtained. All methods use some form of this basic relationship, although many are based on multiple capture occasions and fairly complicated calculations. Each method makes certain assumptions about the number of capture-recapture occasions, whether a population is open or closed (whether migration is occurring), and whether animals are equally likely to be caught. A biologist chooses the appropriate method (described below) accordingly.

Whatever the method used, before beginning a mark-recapture study the following questions (Bibby et al. 1992) should be asked:

> Will enough individuals be caught to obtain useful results? Substantial numbers of animals must be captured and recaptured to obtain precise and useful estimates. If overall

captures will be small (<50 captures per session), then a mark-recapture study is generally not worth undertaking because the precision of estimates will be so poor (Krebs 1998).

Will marking an animal harm it or affect its behavior? If an animal become less healthy or changes its social status as a result of marking, population estimates will be biased.

Will marking an animal affect its chances of being caught again? This is an important issue because previously trapped animals that become more wary of traps (or perhaps more attracted to them) will skew population estimates.

Will the marks used be distinguishable despite wear and tear and different observers? Marks must not be lost and must remain interpretable throughout the study by all observers involved in the monitoring.

Except for the Petersen method, calculations for all capture-recapture methods are quite onerous and tractable only with computers. Detailed descriptions of the mathematics behind these methods are provided by Seber (1982), Pollock et al. (1990), Lancia et al. (1994), Greenwood (1996), Krebs (1998), and Thompson et al. (1998). Fortunately, specialized software, all freely available, has been developed for each method (see our web site to access these programs). Thus, our goals are to highlight the limitations and context within which each method is applicable and to provide guidance for securing the requisite software to undertake the appropriate analysis.

Petersen—Closed Populations

The Petersen method, also known as the Lincoln method, is the simplest of the population estimation methods based on mark-recapture. The Petersen method 1) assumes that a population is closed (experiences no movement into or out of it during the study period and hence is of constant size), 2) requires just one session of capture and marking and a second session of recapture to check for marks, 3) assumes that capture probabilities of all individuals in the population are uniform, and 4) requires a single mark not unique to individuals. Different proportions of the population can be captured at each occasion with no effect on population estimates. In other words, capture effort does not need to be constant at each occasion. Keep in mind, however, that at least 50 marked individuals should be captured during the second occasion to obtain reasonably accurate estimates. In terms of precision of the estimate, the higher the proportion of recaptures in the second sample, the more precise will be the final population estimate. The basic formula (from Chapman 1951) for estimating population size (N) is as follows:

$$N = \frac{(n_1+1)(n_2+1)}{(m_2+1)} - 1$$

The variance of the estimate of population size is as follows:

$$\text{var}(n) = \frac{(n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)}{(m_2+1)^2(m_2+2)}$$

where

n_1 = the number captured and marked on the first occasion

n_2 = the number captured on the second occasion

m_2 = the number captured on the second occasion that are marked

From the variance estimate, standard confidence intervals (see Chapter 9 and Appendix III) can be generated using *t*-values. See Krebs (1998) for further details, particularly for situations where numbers of recaptures are small.

Program CAPTURE—Closed Populations

Program CAPTURE (Otis et al. 1978, White et al. 1982, Rexstad and Burnham 1992) 1) assumes that a population is closed (experiences no movement into or out of it during the study period and hence is of constant size), 2) requires several capture/recapture sessions, 3) can accommodate heterogeneous capture probabilities among individuals within the population, and 4) requires use of multiple or individual-specific marks. The assumption of uniform catchability made by other mark-recapture methods is frequently violated; program CAPTURE's recognition that individuals are not equally likely to be captured underpins its usefulness and popularity among biologists. The method depends on relatively constant trapping effort over at least four (preferably more) capture occasions to make multiple estimates of population size and to choose the most precise one. The number of capture-recapture sessions may have to be increased if capture probabilities are low or populations small. For example, Rosenberg et al. (1995) estimated that more than 12 capture sessions would be needed to reliably estimate abundance for a low-density population with low capture probabilities (about 0.10), a situation that is fairly typical of many mark-recapture studies of animal populations.

To meet the assumption of population closure, trapping over a brief period on grids large enough to minimize edge areas is recommended (White et al. 1982). Particular attention must be paid to the design of trapping grids, which should be large relative to the size of the home range of the animal under study. Trapping grids used in practice are often too small (multiple traps are placed within single home ranges) and thereby undercut sample sizes. A good discussion of sampling issues is provided by White et al. (1982). Information on Program CAPTURE can be obtained over the Internet via our web site.

Jolly-Seber—Open Populations

The Jolly-Seber method 1) assumes that a population is open (experiences some movement into or out of it during the study period and hence is not of constant size), 2) requires at least three sessions of capture/recapture, 3) assumes that capture probabilities of all individuals in the population are uniform, and 4) requires use of batch-specific or individual-specific marks. Such marks permit the trapper to determine the occasion when an animal was most recently trapped—critical information for this method. This is the best method for estimating population size for the common situation in which some immigration and emigration is likely occurring. The Jolly-Seber method produces estimates of population size, as well as the number of animals entering the population (through births and immigration), and estimates of survival rates in the population (which amalgamates but cannot distinguish between rates of mortality and emigration). Unless the number of individuals marked in each sample is greater than 10, the population estimates can become imprecise to the point of not being useful. Study designs that include an intensive series of mark/recapture sessions over a short time frame (e.g., within a breeding season), to ensure that population immigration and emigration is minimized within years, can greatly improve the precision of estimates of population size and turnover. Variations on standard Jolly-Seber analyses (program JOLLY) include assuming that mortality or capture rates are constant, thereby producing more precise estimates of population size and simultaneous analyses by age classes (program JOLLYAGE). The software can be obtained via our web site.

Mark-Resight Methods

Methods have been recently developed that permit analysis of data from animals marked during an initial period and then subsequently resighted over later periods. Such an approach is preferred when recapturing animals is problematic or undesirable (e.g., large mammals or raptors) yet marked animals remain visible in the field to observers. Mark-resight methods often require some knowledge of the number of marked animals alive in the population. Thus mark-resight studies are often performed in conjunction with having a sample of animals radiocollared. Pro-

gram NOREMARK performs mark-resight analyses and is useful both for analyzing mark-resight data and for planning mark-resight studies (White 1993, 1996a, 1996b). The program and information on its use can be obtained via our web site.

Removal

Repeatedly trapping and removing animals from a population results in fewer and fewer new animals captured each session and a marginally increasing cumulative catch. A plot of catch per unit effort for each trapping session in relation to accumulated catch (total animals already caught in previous sessions) generates a declining pattern that can be analyzed with linear regression analysis to extrapolate the point at which capture efforts would yield no new animals. At this point, the entire population would be removed, and hence the accumulated catch would equal the size of the original population. Though not widely used, the removal method may provide a useful means of monitoring population change in the context of management where an attempt is being made to control a pest through removal. In other words, the management action and the estimation method can be combined.

The removal method 1) assumes that a population is closed (experiences no movement into or out of it during the study period and hence is of constant size), 2) requires several sessions of capture, and 3) assumes that capture probabilities of all individuals in the population are uniform. Population estimates from the removal method are particularly vulnerable to bias associated with variation among capture sessions because of different conditions (like weather or changes in trapping methods) that affect the likelihood of animals getting trapped. Traps that lose their efficiency, perhaps because remaining animals learn to avoid them or earlier caught animals were the most vulnerable in the population, violate the assumption that catch per unit effort declines linearly with accumulated catch. Capture effort can vary among days, given that the dependent variable is capture per unit effort. However, one must be cautious to not oversaturate a study area with too many traps (e.g., more than one per home range), such that a trap's capture influences the capture in adjacent traps, or too few traps, such that captures are limited by trap numbers. Population estimates generated from the removal method tend to be quite imprecise. Consult Seber (1982), Krebs (1998), or Greenwood (1996) for further guidance.

DISTANCE SAMPLING

Distance sampling is analogous to plot (quadrat) and line (transect) sampling (discussed in Chapter 8), except that incomplete counts are made and the visibility bias of the undetected animals is adjusted for by using the distances between the observer and the detected animals to make the correction. Thus, distance sampling uses similar methods to traditional sample counts. With a modest increase in effort required to record observer-animal distances, however, it allows far more statistically rigorous estimates of population size to be obtained. Note, however, that in some situations where many animals are counted (e.g., numerous waterbirds in wetlands), distance estimation can become quite time-consuming.

The sightability functions that underpin distance sampling assume that all animals on the transect line or counting point are visible, that animals are sighted near their original locations (flushing and other movements are minimal), and that distances are measured without substantial error. Groups, as well as individual animals, may be sampled with this technique, although certain biases arise with counting groups such as higher visibility of groups compared with solitary individuals. Individuals at any distance may be more likely to be detected if in a group than alone. Points, as well as transects, can be used. Constraints on distance sampling include a requirement for at least 60 animals to be sighted for the use of distance models. If actual distances to animals are not recorded, a minimum of five to seven distance categories is needed to adequately establish the sightability functions. A comprehensive treatment of distance sampling is given by Buckland et al. (1993), who provide examples, as well as useful formulas for estimating

the minimum length of transects or number of points to achieve desired precision levels. The software DISTANCE, a manual, and other guidance can be obtained via our web site.

INDICES

Estimates of absolute abundance through mark-recapture or distance methods are time-consuming and difficult to make. Wildlife biologists therefore often resort to relative indices of animal abundance as a surrogate. An index to population size is simply any "measurable correlative of density" (Caughley 1977). Examples include track densities of mammals, tadpole captures in sweep nets, or counts of singing birds. Although we implicitly assume that the index and actual abundance have a positive, linear relationship with the slope constant across habitats and over time, these relationships are sometimes but not always true and are usually quite "noisy" (Gibbs 2000).

One common problem is an index that becomes "saturated" at high population densities. Frog populations monitored using an index of calling intensity are an example. The index is sensitive to changes at low relative densities of calling male frogs in breeding choruses because calls of individuals can be discriminated by the frog counters. At higher relative densities, however, calls of individual frogs overlap to an extent that size variation of choruses cannot be discriminated by observers.

Another example of a nonlinear, index-abundance relationship concerns the use of presence/absence data, such that the proportion of plots occupied by a given species is the index of abundance. At low relative population densities, changes in population size can be reflected in changes in degree of plot occupancy. Once all plots are occupied, however, further population increases are not reflected by the index because the index becomes saturated at 100% occupancy. Presence/absence information is very sensitive to the size of the sample plots (the proportion occupied increases with the plot size). The problem of saturation can be minimized by using a nested plot design where presence/absence information is gathered within subunits of the larger plots (see Chapter 12 for more information on nested plots for presence/absence sampling). Presence/absence indices usually level off strongly after 70% and lose much of their ability to reflect population changes. Note, however, that while presence/absence data may have a weak ability to reflect changes in abundance, it may have a powerful ability to reflect changes in distribution, habitat use, and other spatial aspects of population change.

Wildlife biologists should also be aware that developing indices with a 1:1 relationship with abundance will most reliably reflect changes in abundance. If the slope describing the index-abundance relationship is low, then large changes in abundance will be reflected in relatively small changes in the index. Such small changes in the index are more likely to be obscured by variation in the index-abundance relationship than if the slope of the index-abundance relationship were steeper.

Index variability may be reduced and the precision of the index-abundance relationship increased by adjusting the index by accounting for auxiliary variables such as weather and observers (Greenwood 1996). In an ideal situation each index would be validated, adjusted for sampling error by accounting for external variables, and corrected to linearize the index and to make it comparable across habitats and over years. This will rarely be an option, however, for regional-scale surveys conducted across multiple habitats over many years by many persons and involving multiple species, although it may be for local monitoring programs focused on single species. In practice, these factors may be overlooked if many years of data are gathered, because the short-term bias they introduce (e.g., weather-caused variation) typically is converted simply to "noise" in long-term datasets.

In summary, index surveys can be improved through the following steps. First, the basic relationship between the index and abundance should be ascertained through a validation study to determine whether the index might yield misleading results and therefore should not be imple-

mented. Second, any results from analysis of index data changes should be considered in light of potential limitations imposed by the index-abundance relationship. Particular attention should be paid to "saturated" indices that can result in a failure to detect population changes. Most importantly, wildlife biologists must be cautious about concluding that a lack of trend in a time series of index data indicates population stability. Often an index may be unable to "capture" population change as a result of a flawed index-abundance relationship or simply excessive "noise" caused by sampling error in the index.

FIELD METHODS FOR BUTTERFLIES

Butterflies are colorful, readily identified, and fly by day. They also are sensitive indicators of changing environmental conditions (Kremen 1994). For these reasons they are perhaps the most popular of insects and hence one of the most frequent targets of insect monitoring programs. Because most species in the temperate zone can also be identified in flight, they are readily surveyed in a nondestructive fashion. Butterflies have traditionally been surveyed by catching them in nets and killing them prior to identifying and enumerating them. This is unnecessary. With a pair of binoculars (with at least 7× magnification), a butterfly guidebook, and some practice you can identify most butterflies as they flit about an area or rest or feed on plants. An excellent guide to using binoculars to identify butterflies is Glassberg (1993).

In designing surveys to monitor butterfly populations, one must be aware of butterfly flight periods. Each species will generally be most active during a particular part of the season. Flight periods depend on the timing of the life cycles in each species in relation to weather conditions but mostly in relation to the status of the plants that host the eggs and larvae. Some species are early spring fliers and others late summer fliers. Additionally, one must be aware of daily cycles of activity. Some species are active in early morning, others in late afternoon. All species are most active on warm, sunny days. Surveys on chilly, cloudy, or rainy days will be unproductive.

Quantitative surveys of butterfly communities typically involve counts of individuals made from fixed survey routes. Routes should lie within a single habitat type, should generally be quite long (>1 km in length is typical), should be about 10m in width, and should be run during favorable weather conditions. Each route represents an individual sampling unit, and counts made on each route represent an index of butterfly abundance. Counts cannot be considered to be an estimate of absolute numbers because butterflies are continually entering and leaving the sampling area while the count is being made. The total number of individuals seen per species along the route is recorded as the index of abundance for that visit. Over a single flight season, the counts will typically start out low in the early part of the season, gradually increase to some peak number, and then taper off. A standard monitoring approach involves making weekly counts and then summing the weekly counts for the flight season to obtain an overall index of abundance for a particular year. Data from this method can be used to compare the relative occurrence of different species among habitats and years at a site. The key reference for butterfly monitoring is Pollard and Yates (1993). More information on field methods and sampling approaches for butterflies can be found at our web site.

FIELD METHODS FOR TERRESTRIAL BEETLES

Terrestrial beetles are a common focus of insect monitoring program, in part because they can be captured readily and passively in pitfall traps. Even in nontropical areas, however, beetle diversity can be daunting. Therefore, distinctive, low-diversity groups of beetles are frequently focused on as indicators for monitoring changes in terrestrial invertebrate communities, for example, the carabids (Rykken et al. 1997) or the silphids, also known as the carrion beetles (Gibbs and Stanton 2000).

Both carabids and silphids can be captured in pit traps, which are plastic containers or cups that can be purchased inexpensively. For carabids, groups of a few to several pitfalls are typically established in clusters, of which several to many such clusters comprise the sample for a particular site (Rykken et al. 1997). For silphids, a single, baited trap usually serves as the sampling unit, which may sample an area of radius 500m (traps must be separated accordingly).

For each pitfall, a small hole is dug in the ground so that the lip of the container is level with the ground surface. The bottom of the container is filled with about 1 inch of water to which a few drops of formalin are added. The formalin is added to discourage omnivorous mammals from eating the contents of the traps and to kill the captured beetles quickly and prevent them from damaging other specimens. A highly saline solution can be used as an alternative. Traps need not be lethal if checked daily. Traps are overlain with a piece of bark or other cover to protect them from flooding by rainfall and should not be located in depressions where ponding during rain events is likely. Unbaited traps work well for carabids, but for carrion beetles a bait, dung or rotting meat, is suspended in a small cup over the trap, supported by a wire or string. Baited traps should be covered with a wire mesh, or even hung from a tree branch well above ground, to discourage predators. Beetles can be pinned in the field or preserved in 70% ethanol solution for later identification. Long forceps greatly facilitate setup and collection of these beetles.

Traps should be checked weekly and monitored for some standard period to generate comparable indices of abundance among sites and over time. Sampling should be undertaken over an extended period (often the entire active season) because the phenology of species are highly variable and related particularly to climatic conditions, which may also interact with geographic location. For example, emergence may be delayed for a particular taxon by cold weather and even more so in valley bottoms. Sampling throughout the period of seasonal activity for all species of interest is the only way to account for this variability.

Further information on sampling beetles (and other insects) can be found in Southwood (1978) and Ausden (1996). Our web site also contains more information on field methods and sampling approaches for invertebrates.

FIELD METHODS FOR AQUATIC INVERTEBRATES

Many methods are available for generating indices of aquatic invertebrate populations useful for monitoring. For aquatic stages, a simple method in moving waters involves the kick screen, a 1-meter-square piece of window screen stretched between two dowels and held in the stream current. The bottom is disturbed with one's foot or with an implement for a fixed period (1 minute is typical), and invertebrates uncovered, along with debris, are captured by the screen. A similar index for still waters involves forcing the bottom edge of a long-handled dip net along the bottom for a fixed period. Each kick or dip generally represents a single sample that must be replicated to characterize populations or communities at a particular site.

Conceptually similar but more-standardized, area-based indices involve the use of Surber and portable invertebrate box samplers (PIBS) to sample fixed areas and thereby generate density estimates. Alternatively, if several weeks are available for sampling, artificial substrates can be used, which include bricks, wooden discs, and stones enclosed in wire mesh. Sample substrates, placed in different areas, are examined at regular intervals to monitor species and numbers of individuals that have colonized them. Generally speaking, each PIBS sample or artificial substrate serves as a sampling unit, placement of which can be stratified within a lake or stream to increase the precision of the abundance estimate.

For invertebrates of lakes and rivers, core samplers and dredges are commonly used. The dredges (e.g., the Eckman dredge) can be lowered out of a boat, with closure of their jaws triggered on contact with the bottom. Core samplers generally consist of a plastic tube mounted beneath a metal mount, such that the weight of the mount forces the corer into the substrate.

Invertebrates and debris secured must be later sorted. Each core or dredge represents a single sampling unit.

For aerial stages, emergence traps, sometimes lighted, are useful. Some adults such as dragonflies and mayflies are quite conspicuous in flight, and collections may be feasible over water with butterfly nets. The sampling issues discussed for butterflies pertain also to aerial stages of aquatic insects. Specific guidelines for monitoring dragonfly populations are provided by Moore and Corbet (1990).

Merritt and Cummins (1996) provide a useful overview of sampling methodologies for aquatic invertebrates, whereas Resh (1979), Kerans et al. (1992), and Klemm et al. (1990) provide useful overviews of sampling issues for aquatic invertebrates. More information on field methods and sampling approaches for aquatic invertebrates can be found at our web site.

FIELD METHODS FOR FISHES

Many techniques are available to monitor fish populations (e.g., Perrow et al. 1996). Occasionally, you can remain dry and make direct counts of fish from overhead in a boat or from the side of a bank. Direct counts can also be made from within the water using a mask and snorkel or SCUBA gear, in conjunction with specialized paper for recording observations under water. Unfortunately, direct counts of fishes are plagued by double-counting individual fishes, which dart in and out of the observer's field of view.

Monitoring fish populations more commonly involves some type of capture method. The very simplest is the conventional fishing rod, which if deployed for a fixed period with standardized gear can actually provide a useful index of fish abundance, particularly for large fishes, which may otherwise avoid nets and other gear. For smaller fish, minnow traps are useful, inexpensive, and widely available. These attract fish unbaited and are most effective in the littoral zone of lakes at depths of 1m or less and in streams if placed near large rocks or logs in pools. These are checked usually daily over a standard interval to generate monitoring data.

Mesh nets are very commonly used to track fish populations. The most common is the seine, a weighted net of variable length (5m to 30m), with mesh appropriate to the size of fish sought, attached to poles and stretched taut and pulled through the water by two people. Seine sampling may be standardized by limiting seine hauls to a fixed duration. Trap nets are useful mainly in standing water and involve a leader perpendicular to shore that guides the fish into an anchored trap where they pass through a set of mesh cones and are retained within a mesh box. Another common netting method is the gill net, a weighted net supported by floats on the water surface, that intercepts fish and entangles them in the net's mesh. Frequent checking of gill nets (every 2 to 3 hours) greatly reduces fish mortality. All netting is problematic in rapid waters or areas choked with vegetation or filled with submerged trees.

Electroshocking provides an alternative approach to monitoring fish populations. Battery powered backpack units are most common, particularly those with direct current, which may cause less injury to fish. Fish temporarily stunned by the electric current are collected by an insulated dip net and placed in buckets for identification and counting. Standardizing the voltage used and area searched can generate reliable indices of fish abundance.

Most of these capture methods, if practiced nonlethally, can provide the opportunity for marking and recapturing individual fishes and thereby rigorously estimating population sizes using mark-recapture methods (see above), a common practice in fish population monitoring.

From a sampling perspective, the primary difficulty in estimating fish populations is that fish tend to be clumped in their distribution across sampling units, more so than other vertebrates (Hilborn and Walters 1992). In other words, some locations will generally contain many fish, whereas others will contain none. Furthermore, fishes are generally quite mobile, horizontally among sites, as well as vertically within the water column, adding further to sample variation.

There are two approaches to cope with the high sampling variation that plagues efforts to monitor fish populations. The first is to always tend to allocate sampling effort extensively rather than intensively. In other words, monitoring fishes will nearly always be more effective if it is based on less intensive sampling at many sites than more intensive sampling at fewer sites. This is somewhat opposite of the traditional practices of managers and biologists who prefer to intensively study a few "index" sites. The second approach to coping with high sampling variation in fish populations is to stratify sampling (see Chapter 8), thereby drastically reducing unexplained variation and increasing the power to detect trends. Stratifying a sample may involve two levels by designating strata at the regional level and again within each stream. For example, at a regional level streams may be stratified based on stream type (e.g., steepness, size, and order), while at the stream level habitats may be further stratified (e.g., riffles, runs, and pools). A similar approach can be used in rivers and lakes, except that the second stage of stratification (habitats) is usually based on depth (e.g., near-shore, pelagic, profundal zones for lakes, and near-shore versus midchannel for rivers).

As with all monitoring projects, but for fishes in particular, pilot studies are critical to estimate the amount of sampling variability and costs of sampling to achieve a desired level of precision. Because of the uniformly high level of sampling variability with fishes, fish-monitoring programs are always at risk of failing to deliver useful information within existing time and budgetary constraints. Only through a pilot study and sample-size estimation will it become clear if a given fish-monitoring program is worth pursuing or if the scope of monitoring should be revised. A useful overview of fish population sampling issues is provided by Thompson et al. (1998: 192–224) and Hillborn and Walters (1992). See also our web site for further information on methods and sampling issues for fishes.

FIELD METHODS FOR AMPHIBIANS AND REPTILES

Amphibians as a group are increasingly the focus of monitoring studies. They add considerable biomass to animal communities. In undisturbed eastern United States forests, for example, amphibian biomass can equal or exceed the biomass of other vertebrate groups (Burton and Likens 1975). Amphibians may further be particularly responsive to environmental changes because of the permeability of their skins that underpins their sensitivity to water, soil, and air quality. Reptiles, particularly lizards, though less sensitive to environmental changes, can also represent the bulk of animal biomass in some ecosystems, particularly in drier habitats (for example, up to 50 kg/ha in some dry tropical forests) (Bullock and Evans 1990).

In terrestrial areas, both groups (although not all species) can usually be surveyed using a combination of two simple techniques: visual encounter surveys along transect lines and total counts in quadrats. Visual encounter surveys in particular can be used for efficiently generating an index of abundance of the herpetofauna. Line transects are a refinement of visual encounter surveys and are particularly useful. One observer walks a fixed distance (generally 0.1km to 1 km) along a transect line, recording observations of amphibians and reptiles, especially active frogs and toads, seen from the transect line. Be cautious of differences in sightability caused by changes in vegetation density. The observer can also turn over (and carefully replace) any "cover objects" (logs, rocks, and other debris) intercepted by the transect line. These cover objects often host amphibians and occasionally reptiles during the day. It is best to premeasure the line, perhaps using a hip chain (see Chapter 5), and lay down a string. This permits the observer to concentrate on looking for amphibians and reptiles rather than on staying on the transect, and it also unambiguously identifies the cover objects intercepted by the line. Because the data are obtained by sampling a straight line of measured distance, they allow a comparison of relative abundance and species composition among different parts of the site. Using this approach, the transect line becomes the unit of sampling.

Transects are excellent for surveying surface-active species. However, litter quadrats are preferred for finding those species, especially salamanders and small lizards, that tend to lurk away from direct sunlight but that nevertheless often represent the most numerous members of many amphibian and reptile communities. Quadrats, usually $1m^2$ to $25m^2$ in area, can be placed in a systematic or random array (Heyer et al. 1994), with each quadrat representing the sampling unit. You can delimit each plot using stakes and a premeasured line. Once you have identified a plot, then begin to search slowly through leaf litter and ground detritus, sifting through the layers to locate amphibians. Work from the outside to the inside of the plot, and record the animals as they are found. Litter should then be redistributed across the disturbed site at the close of the search. These data can be used to describe the species composition and to estimate the absolute density of the more common amphibians at a site. A drawback to litter searches is that disturbance may be too great to use them in conjunction with permanent plots.

Permanent plots are perhaps better served by artificial cover objects of standard dimensions such as bricks and planks with moist undersides for salamanders and boards and sheet metal with dry undersides for reptiles (e.g., geckos). These can be distributed across the ground surface (usually buried slightly in shaded areas) in a random or systematic fashion (often on grids to facilitate checking) and monitored at regular intervals (usually every 2 weeks). Generally speaking, the individual cover object is the sampling unit, although in practice groups of up to five such objects clustered together (an array within 10m of one another), across which individuals detected are summed for each visit, also can serve as an operational sampling unit that generates a more stable index. Untreated pine or fir boards 30cm × 30cm × 5cm have been used successfully by Fellers and Drost (1994) for monitoring salamanders, although larger boards or buried cedar shingles may be better for retaining moisture.

For amphibians of aquatic areas, other, more specialized techniques are useful. These include auditory surveys (Zimmerman 1994), important for surveys for calling frogs and conducted in much the same manner as for birds, although at night rather than during the morning. Numbers of calling males or aggregations of males are counted at survey points (stops), which involve waiting 1 minute and then recording observations during the subsequent 3 minutes. Records are made of the number of individuals or calls per species heard at each station. The sampling unit can involve single stops at a given wetland (the stop being the sampling unit) or multiple stops (typically up to 10) along a route extending within a large wetland or among several small wetlands, with the route becoming the sampling unit. Keep in mind that stops aggregated into routes should be adequately spaced to avoid overlap between stops—some large frogs can be heard up to 500m away and stops should be spaced accordingly. Another consideration is that calling periods for some species can be very brief and special effort must be made to target surveys around their periods of activity. Special efforts must also be made to detect those species with softer calls (e.g., those calling from in the water) whose signal is often overwhelmed by the louder species calling from perches. Also, counts can be difficult for large choruses as the calls of individuals blend together and index of abundance essentially becomes saturated.

Drift fences in conjunction with multiple pitfall traps (Gibbons and Semlitsch 1981) are often suggested for monitoring studies. The entire fence or perhaps an array of three to five of them is considered the sampling unit. It is important to note that drift fences are expensive and time-consuming to construct and tedious to check (typically twice daily, morning and evening), thereby severely limiting the numbers of such sampling units that can be deployed for the purposes of monitoring. Captures in drift fences also are highly dependent on weather conditions, which limit their stability as an index of abundance. Generally speaking, captures on any given day at a given drift fence are very sparse, and captures are pooled across several days or even weeks of sampling to generate data for a sample. That said, for intensive monitoring of selected sites such as vernal pools that draw large numbers of breeding frogs and salamanders from across large areas, drift fences can be extremely useful and efficient for monitoring local populations.

For snakes, time- and area-constrained searches on foot or transect surveys are generally the most efficient ways to generate an index of abundance at a site (Fitch 1987, 1992), although

quadrat searches are sometimes used where snake densities are high. Trapping with specially adapted drift fences (with pit traps altered to prevent snake escape) are sometimes useful for monitoring populations at sites of aggregation (e.g., denning sites). Freshwater turtles can be most easily monitored with hoop-nets secured to the shore (to leave breathing space for captured animals) and baited with a partially opened can of sardines. Seine nets can be useful for estuarine species such as terrapins, with nets blocking tidal creeks and other popular turtle travel routes. Basking traps also should be considered in freshwater and estuarine situations, which are essentially floating pitfall traps (see Jones et al. 1996). These traps consist of a floating wooden frame often rimmed on the inside with aluminum flashing. Inside the frame is a submerged trap or net bag. Turtles climb onto the edge of the trap to bask, and those that reenter the water toward the inside of the trap frame are captured in the net and prevented from climbing out by the flashing. Such traps are usually opened (and checked twice daily) for many days (trap-days) and serve as independent sampling units.

For a highly informative document on amphibian survey techniques, see Heyer et al. (1994). For a brief description of reptile counting techniques see Blomberg and Shine (1996). Thompson et al. (1998:234–255) also provide a valuable discussion of field methods and sampling issues for amphibian and reptile monitoring. Consult our web site for further guidance on monitoring amphibian and reptile populations.

FIELD METHODS FOR BIRDS

Birds show close affinities to particular habitats and even to subtle variations in what we might consider the same habitat, for example, the vertical diversity component of vegetation structure in forests. Thus, they are good indicators of local ecological conditions. Birds also are generally colorful, charismatic, easily identified, and popular with the public. For these reasons, birds, especially songbirds, are a major focus of monitoring efforts (see Table 13.1).

There are many methods described for monitoring landbirds, but point counts are the most efficient way to make counts and gather data for forest and grassland birds. The key reference to this technique is Ralph et al. (1993). Remaining at a point for a fixed period and counting birds heard and seen provides the basis for point counts. Point counts are made during the breeding season and take advantage of the singing behavior of birds as they vocalize in an attempt to repel intruders to their territories and to attract mates. Obviously, you need to be familiar with the songs of birds in your region before undertaking a monitoring program. Many resources exist for learning bird songs (e.g., Peterson 1990).

Landbird populations are often monitored using fixed-radius point counts. You should first identify points, or counting stations, scattered throughout your site at which you will make counts of singing birds. Points should be systematically located with a random starting point, and separated by a distance of at least 100m (200m is often recommended) to avoid counting the same individuals at different points. All points should be located at least 50m away from a sampling site's border. At each sampling point, an experienced observer identifies all birds seen and heard within as well as beyond a radius of 50m. Points are often arrayed along "routes" or transects consisting of a few to many points, which in aggregate become the sampling unit because low counts for most species at any given point usually limit the usefulness of point-level data for detecting population changes. The number of points adequate to sample a site or characterize populations and communities at the scale of a watershed should be determined during the pilot study.

For establishing plot boundaries, considerable practice is needed to become proficient at estimating distances under field conditions. In particular, estimating distances to calling birds is notoriously difficult. Distance sampling is sometimes used for bird counts but primarily only with visually detected birds in conjunction with a range finder to periodically check on distance estimations. At a minimum you must become comfortable with estimating the distance representing

the fixed radius of your plot. Once you are adept at establishing a plot's boundaries, count data can most easily be recorded on a point location mapping data sheet. This is a single sheet of paper used at each point with a large circle on it. Within the circle you note the relative location of each bird detected within your fixed radius plot, using standard abbreviations for each species.

At each point, you will spend a fixed period (typically 5 minutes but usually between 3 to 10 minutes), within 5 hours of dawn, generally 05:30 to 10:30 A.M., when birds are most actively singing. Avoid making counts when it is windy or raining. Also, make repeated visits to your site during the primary breeding season (minimally three visits), or throughout the year if in a relatively nonseasonal environment. Many repeated visits (10 to 15) over a short period to a site can also be used to analyze clusters of observations within the mapping area, each of which likely will represent the locations of an individual territory holder. This is the basis of spot-mapping or territory-mapping, which is sometimes used for intensive studies of bird populations at particular sites.

Point counts will not provide reliable data for secretive birds, large soaring birds such as hawks, or waterbirds. Some members of these groups are nevertheless important to include in a monitoring program, and you may have to devise specialized methods to find and count them. For those species that are not detected by point counts (waterbirds and raptors), consider broadcasting tapes of their calls to provoke a response and then tally these responses as an index of abundance. For species that aggregate, flock counts, roost counts, lek counts or colony or group counts are valuable, although observer disturbance to such aggregations is always a problem. Some species such as large game birds produce distinctive droppings that can be counted. Mist-netting can be useful for tracking songbird populations, especially of those species that sing infrequently, but, like drift fencing for amphibians, is subject to the following limitations: 1) the procedure yields low returns for substantial effort expended, 2) counts are highly weather dependent, 3) often only a small fraction of available habitat is sampled (usually only the ground stratum unless elevated nets are used), and 4) net locations are often detected and avoided by birds. For these reasons, mist-netting has not been widely used to monitor bird populations (but see DeSante 1992), although it is an important tool for bird research. Both Gibbons et al. (1996) and Thompson et al. (1998:262–293) provide useful guidelines for measuring and sampling bird populations. Our web site provides further guidance on monitoring bird populations.

FIELD METHODS FOR MAMMALS

Mammals as a group are quite charismatic, although this applies more to larger than smaller mammals. Many mammals also are important carnivores (e.g., coyotes and shrews), and others, both large (e.g., deer) and small (e.g., voles), are herbivorous and may exert large effects on the vegetation. Still others are much sought after by hunters. These attributes lend a high profile to the group and underpin the frequency with which they are monitored (see Table 13.1). Mammals are quite discreet in their habits, however, and robust counts of mammals can be difficult to make. The general skittishness of mammals is why many monitoring studies rely on indices based on "sign" or on mark-capture methods to track populations. One exception is small mammals, which can be readily captured in traps, identified, and released. Another is bats, which often represent the majority of mammalian species diversity at a site, for which specialized mist-netting techniques are generally required. An exhaustive guide to survey methods for mammals, including bats, is provided by Wilson et al. (1996).

For large mammals, line transect sampling is the most efficient way to record their sign (browse, dung, tracks) and make direct observations of individuals. The transect is usually the unit of sampling and its length can range from 1km to 10 km or whatever standard length is needed to avoid frequent zero counts. Such transects can be walked, driven, or flown at a constant, average speed. Ground transects are most frequently used at the site level, and often follow secondary roads or trails of predetermined direction. Where feasible, these should be systematically

placed throughout a survey area, with transects orientated perpendicular to contours following a compass line. Morning ground surveys should be completed within 2 hours after sunrise, and evening surveys within the last 2 hours of daylight. Data recorded for each observation should include: time of day, distance along transect, and perpendicular distance from transect to animal or sign observed. The data can then be used to calculate relative abundance of animals or sign along the transect, which represents the sampling unit. Useful guidelines for line transect sampling are given by Rudran et al. (1996). Aerial surveys are useful in some circumstances, but are beyond the scope of most local monitoring programs. Consult Bartman et al. (1987) and Bowden and Kufeld (1995) for guidance on aerial surveys for large mammals.

For small mammals, population size is easily determined by using traps to capture rodents and insectivores. Trap lines should be established, along which trap stations are located. Two live traps are generally placed at each trap station, with stations spaced more than 15m apart. Traps should be placed on the ground along natural features such as fallen logs, but avoid any site with flooding potential. Traps can be baited with a mixture of peanut butter and rolled oats, and can remain open for several consecutive nights (if checked daily). Wadding placed inside the traps provides insulation for captured individuals. All traps should be checked two times daily (early morning and early evening). Captured animals should be transferred to plastic bags to facilitate identification, and subsequently released near where they were captured. Seek some combination of trap numbers and trap effort to attain 500 trap-nights per site if you want to adequately characterize small mammal abundance for a variety of species at a site (Jones et al. 1996).

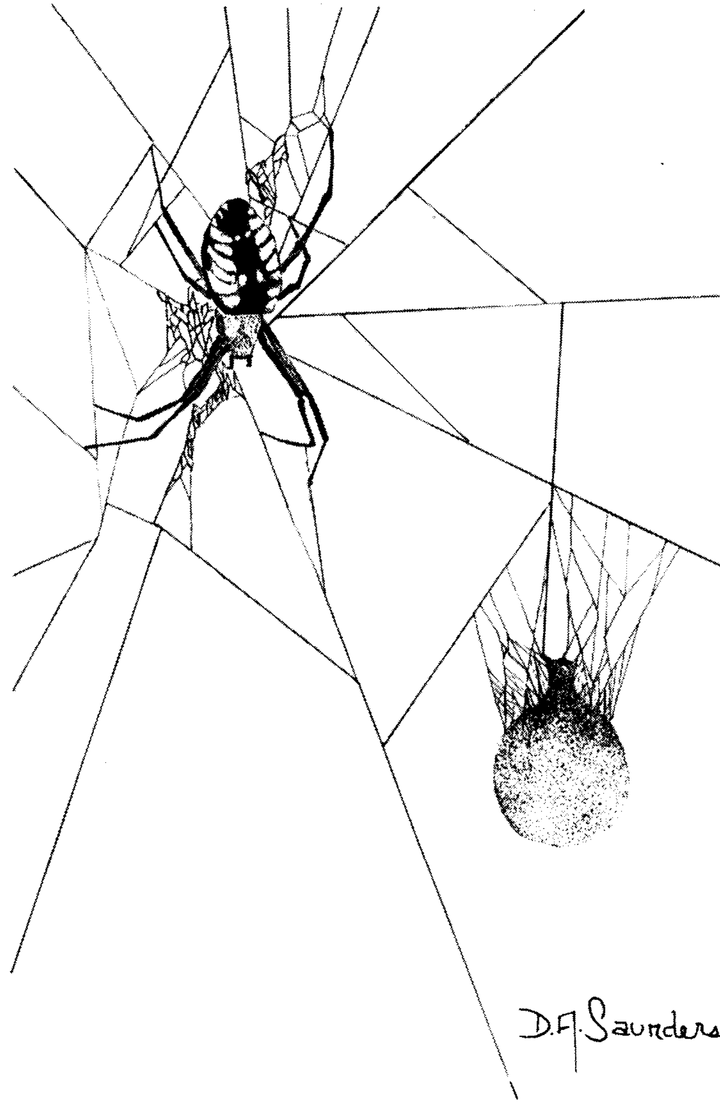
Specialized techniques for monitoring mammal populations include hair tubes and hair catchers, counting footprints (e.g., pug marks) on soft soil and sometimes in tracking stations often baited to increase their attractiveness, counting numbers of burrows, and counting dung (e.g., deer pellets) and feeding marks (e.g., beaver activity) for species that produce distinctive signs. Pellet counts have traditionally been used to estimate and monitor large mammal populations but have fallen on disfavor in the recent decade owing to difficulties in establishing a firm linkage between pellet densities and absolute abundance (e.g., Fuller 1992). Occasionally, calling animals can be counted and mapped indirectly (e.g., wolves, primates). Direct counts of marine mammal and bat colonies, roosts, or nurseries are particularly valuable for species that occasionally form aggregations. Field techniques are further elaborated on by Wilson et al. (1996) and Sutherland (1996) and sampling issues by Thompson et al. (1998:301–317). More information on field methods and sampling approaches for mammals can be found at our web site.

MANAGEMENT IMPLICATIONS

Animal populations are the frequent targets of monitoring and management. They are, however, problematic to track with precision because the elusive nature of most animals precludes directly counting them. Many specialized sampling procedures based on mark-recapture methodology have been developed to estimate animal populations, but these must be used with caution because their implementation often entails great effort but yields imprecise estimates.

Animal monitoring instead relies more often on tracking indices of populations. Such indices are useful for monitoring purposes as long as there is a known and preferably linear relationship between the index and actual abundance that is constant through time and across habitats. Caution must be exercised because an index may be unable to "capture" population change owing to a flawed index-abundance relationship or simply to excessive noise caused by sampling error in the index. Validating the index is therefore always a recommended component of the pilot phase of any animal monitoring program. Because of the mobile and highly dynamic nature of animal populations, extensive rather than intensive sampling of population indices (sampling less frequently in more places rather than sampling more frequently in fewer places) is generally most useful for inferring population response to management activities.

CHAPTER 14
Objectives



Argiope aurantia
Black and yellow argiope
Artist: D. Andrew Saunders

In this handbook we are promoting objective-based monitoring within an adaptive-management framework (see Chapter 1). We believe successful monitoring depends on developing specific management objectives. Objectives are clearly articulated descriptions of a measurable standard, desired state, threshold value, amount of change, or trend that you are striving to achieve for a particular population or indicator. Objectives may also set a limit on the extent of an undesirable change.

In this chapter, we describe a process for developing clear management objectives. We also describe a process for writing sampling objectives, which are companion objectives to be included whenever monitoring involves sampling procedures. The sampling objectives include information on desired levels of precision, minimum detectable change, and acceptable false-change and missed-change error rates. The information contained within the sampling objectives is essential for completing the sampling design.

As part of the adaptive-management cycle, management objectives accomplish the following:

- Focus and sharpen thinking about the desired state or condition of the resource.
- Describe to others the desired condition of the resource.
- Determine the management that will be implemented, and set the stage for alternative management if the objectives are not met.
- Provide direction for the appropriate type of monitoring.
- Provide a measure of management success.

As the foundation for all of the management and monitoring activity that follows, developing good management objectives is probably the most critical stage in the monitoring process (MacDonald et al. 1991). Objectives must be realistic, specific, and measurable. Objectives should be written clearly, without any ambiguity.

COMPONENTS OF AN OBJECTIVE

Six components are required for a complete management objective:

- Species or Indicator: identifies what will be monitored
- Location: geographic area
- Attribute: aspect of the species or indicator (e.g., size, density, cover)
- Action: the verb of your objective (e.g., increase, decrease, maintain)
- Quantity/Status: measurable state or degree of change for the attribute
- Time frame: the time needed for the management strategy to prove effective

Management objectives lacking one or more of these components are unclear. Box 14.1 gives examples of typical incomplete objectives and identifies their missing components.

Species or Indicator

Monitoring may involve measuring the change or condition of some aspect of the species itself. If you are monitoring the species, the objective should include its scientific name. If the objective will address a subset of the species (e.g., only flowering individuals, only females), this should be specified.

Monitoring may also measure indicators that function as surrogate measures of species success. We described four general classes of indicators in the first chapter: 1) indicator species that correlate with the success of the target species and are easier to measure; 2) habitat characteristics; 3) threats; 4) indices of abundance such as tracks and sign.

Box 14.1. *EXAMPLES OF OBJECTIVES MISSING ONE OF THE SIX COMPONENTS OF A MANAGEMENT OBJECTIVE: SPECIES OR INDICATOR, LOCATION, ATTRIBUTE, ACTION, QUANTITY/STATUS, AND TIME*

Most of these objectives are missing a component. What's missing?

1. *Decrease cowbirds (*Molothrus ater*) at the Three Creeks Wildlife Management Area by 2005.*
2. *Exclude livestock from the Summit Creek* Primula alcalina *population.*
3. *Exclude livestock from the Summit Creek if cattle are impacting* Primula alcalina.
4. *Increase percent cover of living hard corals on the Scott's Head Bay reef by 50%.*
5. *Decrease the percent of* Astragalus aquilonius *individuals trampled by livestock at the Grandview site by the 2003 grazing season.*
6. *Maintain a population of at least 500 breeding giant tortoises on Isla Pinzon between 2000 and 2025.*
7. *Allow no more than 30% herbivory of inflorescences in any two years in a row between 2000 and 2008.*
8. *Increase the white-tailed deer population at the Huntington Wildlife Forest by 25% between 2000 and 2005.*
9. *Increase the habitat occupied by* Gymnosteris nudicaulis *by 300 hectares.*
10. *Increase the viability of the onion.*
11. *Maintain, at a minimum, 300* Happlopappus radiatus.
12. *Increase the number of waterfalls with viable populations of Chittenango ovate amber snail (*Succinea chittenangoensis*) under protective management by 3 by 2010.*

Missing component:

1. *Decrease what attribute of cowbirds? Decrease from what level or from which time?*
2. *This is a management response, not an objective.*
3. *Not an objective, more similar to a management response. The term "impacting" is ambiguous. Need to identify some measurable parameters.*
4. *Increase by 50% over current value? By when?*
5. *How large a decrease in percent? From when?*
6. *Looks OK.*

(Continued)

Box 14.1. *EXAMPLES OF OBJECTIVES MISSING ONE OF THE SIX COMPONENTS OF A MANAGEMENT OBJECTIVE: SPECIES OR INDICATOR, LOCATION, ATTRIBUTE, ACTION, QUANTITY/STATUS, AND TIME (Continued)*

7. *What plant? What site or population? Is an inflorescence included in the 30% if it is only partially eaten, or does it have to be completely consumed?*
8. *What attribute of deer populations? Total numbers? Numbers of reproductive females? Number of antlered males? Something else?*
9. *Where? In a certain population or watershed or throughout the resource area? By when should this increase occur?*
10. *What is viability? How much increase? What onion? What population? By when?*
11. *Where? Time frame? Maintain 300 of what attribute (individuals, stems, flowering plants)?*
12. *What is protective management? What is a viable population? Where should these sites occur?*

Monitoring indicators may be less expensive, provide more immediate monitoring feedback to management, and focus on the aspect of the species that you actually have management control over (habitat quality or intensity of threat). Monitoring indicators may also be problematic because the relationship between an indicator and a particular species is usually hypothetical, or at best only partially understood. Monitoring an indicator may thus result in false conclusions about the condition of a species population. The benefits and potential problems with using indicators is discussed at length in Chapter 1.

Location

Clear delineation of the specific entity or geographic area of management concern allows all interested parties to know the limits to which management and monitoring results will be applied. The spatial bounds of interest defined in a management objective will vary depending on land management responsibilities (e.g., you may only have access to a portion of a particular population because of land ownership patterns) and particular management activities (e.g., you may only be interested in individuals located within recently logged forests). The location is related to the selected scale of monitoring (see Chapters 3 and 8), which is affected by conservation goals and responsibilities, the biology of the species, and the realities of limited monitoring resources.

Attribute

The best attribute to use in monitoring depends on the management situation, the species, and the monitoring resources available. Population size is a common attribute when monitoring rare species. Population size of plants and some animal species may be counted directly in a census (if you can count them all) or estimated by making counts in plots within an area of known size.

For some species, monitoring population size may be difficult. Animals that are secretive and hard to count may be estimated by the techniques described in Chapter 13, but another attribute may be easier to monitor and just as effective for assessing management (e.g., indices

of population size, habitat characteristics, threats—see Chapter 1). Box 14.2 describes some examples of these. For plants, population size is sometimes difficult to measure when individuals (or some other counting unit) are difficult to distinguish (e.g., rhizomatous plants). For these plants another measure such as cover may be a better attribute to monitor (Box 14.3). Chapter 12 describes several vegetation measures suitable for use as monitored attributes. Qualitative estimates of abundance, presence/absence and aerial extent are all useful attributes to consider (described in Chapter 4). Attributes of habitat indicators or threats may be similar to quantitative measures for species (e.g., density of tire tracks or cover of woody species) or may be particular to the indicator chosen (e.g., level of a trace contaminant expressed in parts per million).

When selecting an attribute, first narrow the list of potential attributes given constraints of species morphology and site characteristics (e.g., density is not an option if your species lacks a recognizable counting unit). Then narrow the list further by considering the following criteria:

- The measure should be sensitive to change (preferably the measure should differentiate between human-caused change and "natural" fluctuation).
- Biologically meaningful interpretations of the changes exist that will lead to a logical management response.
- The cost of measurement is reasonable.
- The technical capabilities for measuring the attribute are available.
- The potential for error among observers is acceptable.

Box 14.2. *ATTRIBUTES THAT CAN BE MONITORED FOR ANIMAL POPULATIONS*

- *Age and sex composition (based on differences in plumage, pelage, or other external characteristics often visible to the eye during counts)*
- *Size of individuals (antler beam in deer, snout-vent length in amphibians and reptiles, tarsus or wing length in birds)*
- *Mass*
- *Condition (often some combination of mass divided by length, or, in birds, overall electrical conductance or visual inspection of subcutaneous fat deposits as a measure of lipid reserves, or presence of deformities in amphibians)*
- *Reproductive status (often based on evidence of lactation in mammals or swellings of the reproductive organs in birds and amphibians, x-rays of reptiles to detect eggs)*
- *Parasite loads (e.g., tick-loads in mammals, mite-loads in birds)*
- *Reproductive rates (for example, number of offspring accompanying their mother—ungulates, whales, young per nest in birds, or egg masses deposited by frogs and some salamanders)*
- *Survival rates and immigration/emigration (require intensive studies)*
- *Area occupied*

Box 14.3. *PLANT ATTRIBUTES SUITABLE FOR MONITORING*

- *Density*
- *Cover (canopy cover or basal cover for trees, matted plants, and some bunchgrasses)*
- *Frequency*
- *Biomass (on a plot basis rather than an individual plant basis)*
- *Condition (vigor, color, percentage of damaged or dead parts)*
- *Size of individuals (height, basal diameter, biomass)*
- *Reproductive output (number of flowers, percentage flowering success, number of fruits, seed production)*
- *Seedlings (survival, density)*
- *Flowering plants (density or percentage of the plants that are reproductive)*
- *Density or percentage of plants exhibiting herbivory or injury*
- *Mortality (density or percentage of dead plants)*
- *Population area*

Action

There are three basic actions: increase, decrease, and maintain. There is a tendency when managing rare things to want to have them increase. Some populations, however, may already be at the maximum potential for their habitat or suffer from no apparent threats. For these, a more realistic objective would be to maintain current condition. For other populations you may wish to set a threshold that will trigger a management action if the population falls below the threshold. The following are some questions to consider:

- Are current populations viable or have recovery needs such as increased population size, improved vigor, or change in demographic distribution been identified? Species with potential for rapid declines or existing significant degradation of habitat may deserve a more aggressive approach than simply maintaining the current condition.
- Are management options available that you believe will increase the abundance or improve the condition of the species?
- Will increases occur with removal of threats, or will more active management efforts be necessary (e.g., prescribed fire, augmentation by transplants, control of competing exotics).

The following is a list of common action verbs used in management objectives and guidelines describing when each is appropriate:

- **Maintain.** Use when you believe the current condition is acceptable or when you want to set a threshold desired condition (e.g., maintain a population of 200 individuals).

- **Limit.** Use when you wish to set a threshold on an undesirable condition or state of the species or habitat (e.g., limit Noxious Weed A cover to 10%; limit mortality to 10% per year).
- **Increase.** Use when you want to improve some aspect of the species or indicator (e.g., increase the average density by 20%; increase the number of populations to 16).
- **Decrease.** Use when you want to reduce some negative aspect of the species or indicator (e.g., decrease livestock utilization of inflorescences to 40% or less; decrease cover of Noxious Weed A by 20%).

Quantity/State

The condition or change must be described with a measurable value. This can be a quantity (e.g., 500 individuals, 20% cover, 30% change), or a qualitative state (e.g., all life stages present at the site, cover class 4).

Determining these quantities or states requires consideration of a number of factors:

- How much can the species respond? Populations of long-lived species (such as tortoises or trees) may be very slow to respond to management changes. Changes may be small and difficult to detect, or take many years to express. (Consider using an indicator as an alternative).
- What is necessary to ensure species or population viability (e.g., how much change, what population size, what qualitative state)?
- How much change is biologically meaningful? Some species (such as annual plants) can have tremendous annual variability, and an objective that specifies, for example, a 10% increase in density is meaningless.
- What is the intensity of management? Will you continue existing management, remove current threats, or implement a radical alternative?
- What is the implementation schedule of management? If the monitoring project is scheduled to last 5 years, but new management will not be implemented until the second year of the study, the change results from only 3 years of management.
- What are the costs and problems associated with measuring the amount of change specified? Small changes are often difficult and expensive to detect (see Chapters 7 and 8).

The task of specifying a measurable quantity or state is usually a challenging one. The ecology of many species, especially rare ones, is poorly understood. Predicting the response of a population to particular management activities is often difficult. Many populations undergo natural fluctuations as they respond to varying climatic conditions or to the fluctuating populations of pollinators, herbivores, predators, or prey. Most populations have been subject to impacts from human activities; thus, historic conditions or natural population levels are unknown. Few species have been studied in enough detail to reliably determine minimum viable population levels, and theoretical problems with the concept of minimum viable populations remain even in species that have been intensively studied. These challenges should not serve as obstacles to articulating measurable objectives. Use the tools described below and do the best that can be done. If you do not articulate a measurable management objective, you have no means to assess if current management is beneficial or deleterious to the species of interest.

Time Frame

The time required to meet a management objective is affected by the biology of the species, the intensity of management, and the amount of change desired. Populations of short-lived species that reproduce annually may respond quickly, but long-lived species and those with episodic

reproduction may require more time. Intense management will result in more rapid changes than low intensity or no special management. Large changes will require more time than smaller ones, unless a management action will have immediate, large impacts (e.g., timber harvest).

Objectives with time frames as short as a few months to a year may be appropriate in some situations. We recommend that time frames be as short as possible for several reasons:

- Changes in management budgets and personnel often doom long-term monitoring projects.
- Short-term objectives promote regular reassessment of management and implementation of management changes.
- Monitoring often uncovers unexpected information; short-term objectives encourage modification of objectives and monitoring based on this information.
- Short-term objectives circumvent the trap of monitoring ad infinitum while avoiding difficult decisions.
- The adaptive-management cycle must occur within a short enough period that opportunities for species recovery or alternative management are not lost.

TYPES OF MANAGEMENT OBJECTIVES

Objectives can be described in one of two ways:

- A **condition** (e.g., increase the population size of Species A to 5000 individuals; maintain a population of Species B with at least 2500 individuals; maintain Site B free of noxious weeds X and Y). We will call these target/threshold management objectives.
- A **change** relative to the existing situation (e.g., increase mean density of Species A by 20%; decrease the frequency of noxious weed Z by 30%). We will call these change/trend management objectives.

For target/threshold objectives, you assess your success in meeting your objective by comparing the current state of the measurement attribute to the desired state or to an undesirable state that operates as a red flag or threshold. With a change/trend objective you measure the trend over time. The two types of objectives are appropriate for different situations. You may choose a change/trend objective when you have insufficient information to describe a realistic future condition but you can describe a realistic rate of change. You would also use a change/trend objective when you believe the current state is less important than the trend over time. For example, whether a population has 8000 individuals or 6000 individuals may not matter; a decline from 8000 individuals to 6000 individuals (a 25% decline) may be very important to detect. Usually change objectives are more appropriate than target/threshold types of objectives when management has changed and you want to monitor the response (trend) of the selected attribute.

The two types of objectives also require different considerations in designing the monitoring methodology and analyzing the results, especially when the monitoring of the objective requires sampling. Chapters 8 and 9 describe these issues in detail.

Management objectives can be written to describe either desirable or undesirable conditions and trends. You would frame your objective in desirable terms if you believe improvement of the population or indicator is necessary and if you have implemented management that you believe will result in improvement. These objectives are sometimes referred to as "desired condition objectives" because they describe the target condition or trend of the resources (e.g., increase to 2000 individuals, decrease cover of a noxious weed by 40%).

If you believe the current condition is acceptable, and that a continuation of current management will likely maintain that condition, you could frame your objective using undesirable thresholds of condition or trend. These are sometimes termed "red flag objectives" because they state the level of an undesirable condition or change that will be tolerated (e.g., no fewer than 200 individuals; no more than 20% cover of the noxious weed; no more than a 20% decrease in density). These objectives act as a warning signal that management must change when the threshold is exceeded. Red flag objectives can be written to identify an unacceptable decline in a rare species or a surrogate habitat variable, or an unacceptable increase in a negative factor (e.g., an exotic species, encroaching shrub cover, the percentage of habitat disturbed by recreational vehicle traffic, etc.).

Different types of management objectives require varying intensities of monitoring (see Chapter 3). Qualitative objectives can be monitored using techniques that assess condition or state without using quantitative estimators. Simply finding if the species still occurs at a site is a type of monitoring that can be very effective for some situations. Another approach is to use estimates of abundance such as "rare," "occasional," "common," and "abundant," or to map the aerial extent of the population. Objectives may also be written so they can be monitored by complete counts. Other populations may require monitoring by sampling. If so, the management objective is paired with a sampling objective (see below). We give you examples of plant and animal management objectives (paired where needed with sampling objectives, described later in this chapter), arranged in order approximating increasing intensity and including desired condition and red flag types (Box 14.4 for plants and 14.5 for animals).

Box 14.4. *THESE EXAMPLES OF OBJECTIVES FOR PLANTS ARE DIVIDED INTO TWO MAIN CATEGORIES: TARGET/THRESHOLD AND CHANGE/TREND. WITHIN EACH CATEGORY, OBJECTIVES ARE ARRANGED IN ORDER THAT ROUGHLY CORRESPONDS TO INCREASING MONITORING INTENSITY. EXAMPLES OF DESIRED CONDITION AND RED FLAG TYPES OF OBJECTIVES ARE INCLUDED.*

Many of the following management objectives illustrate examples where sampling is not occurring and therefore no sampling objective needs to be articulated. For management objectives where sampling is likely to occur, an example sampling objective is included.

TARGET/THRESHOLD OBJECTIVES

Management Objective — *Increase the estimated total cover (plotless visual estimate) of* Astragalus leptaleus *in Macroplot A at Birch Creek from Class 1 (1–10%) to Class 3 (21% to 30%) by 2010.*

Management Response — *Grazing will be changed to fall use only if an increase is not observed.*

Management Objective — *Eliminate OHV (off-highway vehicle) tracks in* Xanthoparmelia idahoensis *habitat (illustrated on habitat areas Map 1) beginning in 2002.*

Management Response — *If OHV evidence is found, implement educational efforts to reduce OHV traffic in habitat areas. If these are unsuccessful, area closures will be identified and fences constructed.*

(Continued)

Box 14.4. *THESE EXAMPLES OF OBJECTIVES FOR PLANTS ARE DIVIDED INTO TWO MAIN CATEGORIES: TARGET/THRESHOLD AND CHANGE/TREND. WITHIN EACH CATEGORY, OBJECTIVES ARE ARRANGED IN ORDER THAT ROUGHLY CORRESPONDS TO INCREASING MONITORING INTENSITY. EXAMPLES OF DESIRED CONDITION AND RED FLAG TYPES OF OBJECTIVES ARE INCLUDED. (Continued)*

Management Objective	*Increase the number of population areas of* Penstemon lemhiensis *within the Iron Creek Drainage from 8 to 15 by 2010.*
Management Response	*If new populations fail to establish under current management, a transplant reintroduction program will be considered and, if approved, implemented by the year 2011.*
Management Objective	*Maintain a minimum cover of 30%* Xanthoparmelia idahoensis *(plotless visual estimate) in at least 7 of the 10 macroplots established in the Bent Hills population area between 2003 and 2009.*
Management Response	*If, in any year, cover decreases below this threshold, reduce OHV use by area closures and erecting fences.*
Management Objective	*Maintain an estimated cover of at least 20% (plotless visual estimate) of* Xanthoparmelia idahoensis *in Macroplot A in the Warm Springs drainage between now and 2003.*
Management Response	*If cover declines below an estimated 20%, institute a more extensive, quantitative monitoring project that assesses the trend of the entire population in the Warm Springs drainage.*
Management Objective	*Increase the number of individuals of* Penstemon lemhiensis *in the Iron Creek population to 160 individuals by the year 2005.*
Management Response	*Failure to detect an increase to 160 individuals will result in more intensive monitoring to determine if the current population of 122 is stable and viable (demographic analysis), and the implementation of alternative management by 2010 if it is not.*
Management Objective	*Maintain at least 50 reproductive individuals of* Thelypodium repandum *at the Lime Creek population during mining operations.*
Management Response	*Collect seed the first year the population falls below 50 individuals, and for 3 years following.*
Management Objective	*Maintain a population of at least 200 individuals of* Thelypodium repandum *at the Malm Gulch site between 2003 and 2010.*
Management Response	*Failure to maintain a population of the minimum size will trigger additional monitoring and study to determine the reason for failure. Alternative management will be implemented by 2012.*

Management Objective	*Increase the mean density of* Viola adunca *in Macroplot A at the Clatsop Plains Preserves to 1.0 plants/m*^{2} *by 2005.*
Sampling Objective	*Be 95% confident that estimates of density are within ±30% of the estimated mean density.*
Management Response	*If the desired increase does not occur, additional monitoring of the population will be implemented, and alternative management implemented by 2008. If the mean density is equal to or greater than the target density, current management will continue and the population will be monitored again in 2008.*
Management Objective	*Allow herbivory of inflorescences on no more than 20% of the individuals of* Primula alcalina *at the Birch Creek population in any year.*
Sampling Objective	*Be 90% confident that estimates of inflorescence herbivory are within ± 8% of the estimated percent grazed.*
Management Response	*Exclude cattle use from the Birch Creek* Primula alcalina *site by constructing a buck and pole fence within 6 months of the time the threshold is exceeded.*
Management Objective	*Maintain a frequency of 20% (0.10-m*^{2} *square quadrats) or less of the noxious weed* Taeniatherum caput-medusae *in Macroplot A at the Agate Desert Preserve in any year between 2002 and 2007.*
Sampling Objective	*Be 95% confident that frequency estimates are within ±5% of the estimated frequency values.*
Management Response	*Initiate chemical weed control the following field season if the frequency of* Taeniatherum caput-medusae *exceeds 20% in Macroplot A.*
Management Objective	*Allow no more than 30% of the population of* Silene scaposa *var.* lobata *to be killed by logging operations at the Wood Creek site.*
Sampling Objective	*Obtain estimates of percent mortality with 95% confidence intervals that are no wider than ± 10% of the estimated percent mortality.*
Management Response	*Logging will not be allowed in other population areas of* Silene scaposa *var.* lobata *if the mortality at this site exceeds the threshold.*
Management Objective	*Maintain a minimum population of 1000 clumps of* Sarracenia oreophila *at the Eller Seep Preserve between 2003 and 2010.*
Sampling Objective	*Estimate the number of* Sarracenia orephila *clumps with 95% confidence intervals no wider than ±10% of the estimated number of total clumps.*
Management Response	*Additional monitoring will be initiated if the population falls below the threshold of 1000 clumps.*

(Continued)

Box 14.4. *THESE EXAMPLES OF OBJECTIVES FOR PLANTS ARE DIVIDED INTO TWO MAIN CATEGORIES: TARGET/THRESHOLD AND CHANGE/TREND. WITHIN EACH CATEGORY, OBJECTIVES ARE ARRANGED IN ORDER THAT ROUGHLY CORRESPONDS TO INCREASING MONITORING INTENSITY. EXAMPLES OF DESIRED CONDITION AND RED FLAG TYPES OF OBJECTIVES ARE INCLUDED. (Continued)*

CHANGE/TREND OBJECTIVES

Management Objective	*Increase the density of* Lomatium cookii *at the Agate Desert Preserve by 20% between 2003 and 2008.*
Sampling Objective	*Be 90% sure of detecting a 20% change in density with a false-change error rate of 0.20.*
Management Response	*If the density fails to increase, additional research of potential management options will be initiated and alternate management implemented by 2012.*
Management Objective	*Allow a decline in the mean density of* Primula alcalina *at the Summit Creek site of no more than 20% between 2002 and 2010.*
Sampling Objective	*Be 95% sure of detecting a 20% change in density with a false-change error rate of 0.10.*
Management Response	*A decline of 20% will trigger a more intensive study of the interaction of livestock grazing and* Primula alcalina, *with the implementation of alternative management within 4 years after the first year an unacceptable level of decline is measured.*
Management Objective	*Allow a decrease of no more than 20% of the 2001 cover of* Astragalus diversifolius *at the Texas Creek population between 2001 and 2008.*
Sampling Objective	*Be 90% certain of detecting a 20% decrease in the percent cover with a false-change error rate of 0.10.*
Management Response	*Exceeding the decrease will trigger a change in grazing management to a fall-use only system, implemented the season after a 20% decrease is exceeded.*
Management Objective	*Allow a decrease of no more than 30% in the number of individuals of* Conradia glabra *at Apalachicola Bluffs and Ravines Preserve over a 2-year period after implementing prescribed fire.*
Sampling Objective	*Be 80% certain of detecting a 30% decrease in the total number of individuals with a false-change error rate of 0.20.*
Management Response	*If the reduction exceeds 30%, populations will be protected from subsequent burns at the Apalachicola Bluffs and Ravines Preserve. Prescribed fire in population areas of* Conradia glabra *at other preserves will be designed to affect 20% or less of the population and implemented only if resources are available to monitor the response of the species to fire.*

Box 14.5. *EXAMPLES OF OBJECTIVES FOR ANIMALS.*

TARGET/THRESHOLD OBJECTIVES

Management Objective	*Allow harvest of no more than 100 capybara (*Hydorchaeris hydrochaeris*) adults per year between 2000 and 2005 at Hato Cedral.*
Management Response	*Assuming all harvest is accurately and completely reported, maintain current harvest program as long as reported kills each year are < 100.*
Management Objective	*Maintain the current low population level of the yellow-footed rock-wallaby (*Petrogale xanthopus*) population at the Melville Hall property from 2000 to 2005. Currently 95% of permanent plots show no evidence of rock-wallaby sign (pellets or browse).*
Management Response	*No rock-wallaby control measures will be undertaken unless a sign occurs in more than 10% of the permanent plots.*
Management Objective	*Increase local recruitment of young into the aging population of Galapagos giant tortoises (*Geochelone elephantopus*) on Isla Pinzon so that >25% of the population is in the immature age class by 2050.*
Sampling Objective	*Obtain estimates of the proportion of the population in the immature age class with a 90% confidence interval no wider than ± 5%.*
Management Response	*Continue efforts to control introduced rat populations on the island until local recruitment target has been achieved.*
Management Objective	*Increase the Sandy Point Marsh's population of territorial least bitterns (*Ixobrychus exilis*) to 25 males by 2010.*
Sampling Objective	*Be 90% confident that numbers of calling bitterns are estimated to within ±5 individuals.*
Management Response	*Continue efforts to increase cattail growth in the marsh to create favorable habitat for least bitterns until 25 males are detected, by 2010 at the latest.*

CHANGE/TREND OBJECTIVES

Management Objective	*Double the number of beaches occupied by populations of Puritan tiger beetles (*Cicindela puritana*) on the lower Connecticut River by 2010.*
Management Response	*Cease population translocation efforts after beetle populations are thriving on twice the current number of occupied beaches*
Management Objective	*Allow a decrease of no more than 10 in the population of eastern spadefoot toads (*Scaphiopus holbrookii*) attempting to breed in the Sand Plains Pond between 2005 and 2006.*
Management Response	*If numbers of toads captured (and censused) in drift fences surrounding the pond drop by more than 10 individuals between years, the stricter protective measures outlined in the management plan will be implemented.*

RESOURCES AND TOOLS FOR SETTING OBJECTIVES

Existing Plans

General goals for a particular species may be described in other planning documents such as conservation plans, watershed plans, regional or local land use plans, forest plans, or activity plans. Linking a monitoring project to these higher-level planning documents may increase management support and funding for the project. The goals in these plans may also serve as a useful starting point for developing more complete and specific objectives.

Ecological Models

Ecological models are simply conceptual visual or narrative summaries that describe important ecological components and their relationships. Constructing a model stimulates thinking about the ecology and biology of the target species. You do not have to be mathematically inclined to develop and use a model; the type of model described here rarely involves complicated formulas or difficult mathematics.

Ecological models have three important benefits. First, they provide a summary of your knowledge of the species, helping you to see the complete picture of the ecology of the species. For example, because livestock grazing affects a plant species negatively by direct herbivory, you may consider that relationship first. Grazing may, however, also affect the species positively through indirect effects on community composition by reducing competition. Trampling by livestock may positively affect the population by exhuming seeds from the seed bank and increasing germination. During the development of an ecological model, you will have to think about these indirect and sometimes hidden relationships. The model will often identify several factors that can cause the change you hope to detect by monitoring, and perhaps help isolate the most important mechanism.

Second, ecological models identify the gaps in your knowledge and understanding of the species. Your model may suggest that these gaps are not important, in which case you may choose to ignore these unknowns. Conversely, the model may suggest that an unknown relationship is extremely important for understanding the total ecological and management scenario. You may need additional studies before effective monitoring can begin.

Third, ecological models help identify mechanisms and potential management options. If the ecological model suggests, for example, that seedling establishment appears rare, that successional processes of canopy closure may be occurring, and that litter buildup on the ground provides few germination sites, you may be inclined to think about prescribed fire, or some other management strategy that induces germination or reverses succession. Lacking an ecological model, you may have focused on only a single attribute such as the lack of seedling establishment, which can result from a multitude of causes.

An ecological model can be as simple or complex as you wish. You can focus on a single management activity, as shown in Figure 14.1, or you can attempt to summarize all the interactions, as shown in Figure 14.2.

Reference Sites

Reference sites can serve as comparison areas to help set quantitative targets in objectives. These are areas with minimal human impact such as designated natural areas or reserves, parks, or wilderness areas. Undesignated areas with populations that appear thriving and healthy may also function as reference sites.

Reference sites can be valuable, but use them with caution. Simply because a population is located in a protected area does not ensure that it is viable or healthy. Lack of management activities within protected areas may be allowing natural processes to occur that are detrimental to a species. In addition, populations that appear "healthy and thriving" to casual observation may actually be declining.

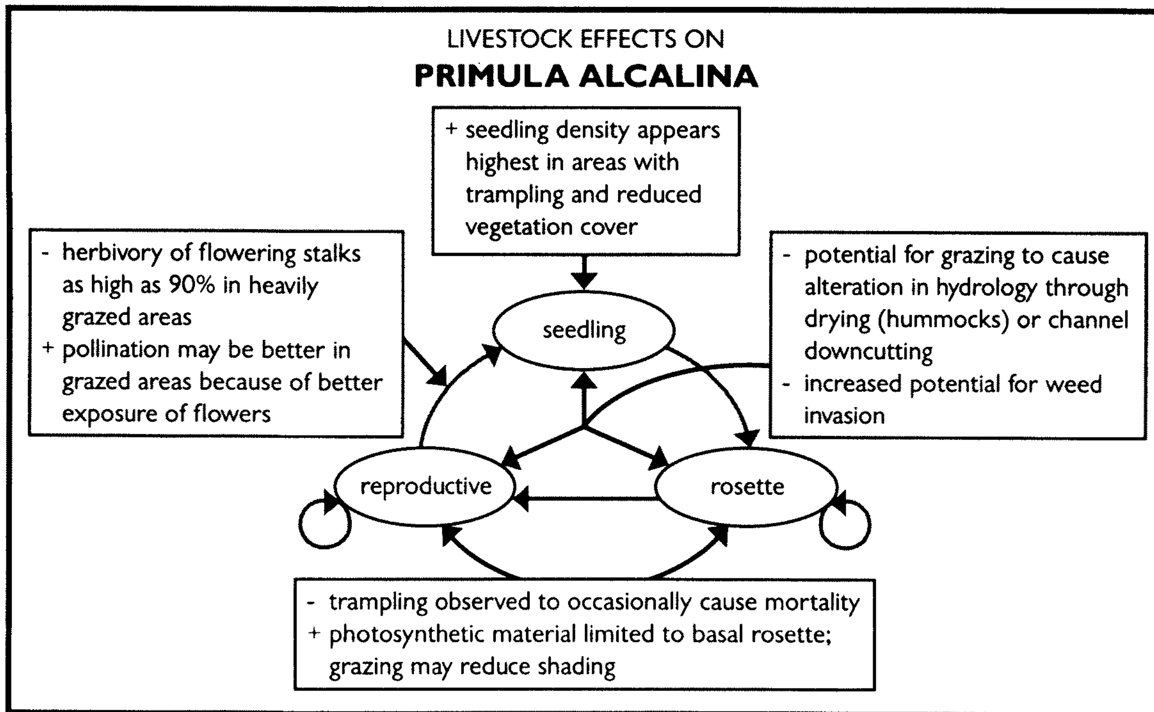


Figure 14.1. An ecological model showing positive and negative effects of grazing on an Idaho (United States) endemic plant species, *Primula alcalina.*

The protected status of these sites may actually limit their usefulness for setting objectives for some species. While the behavior of a species in the absence of human activity may provide useful information, in many cases we are managing populations in areas where human activity is occurring. Populations may respond differently than in pristine conditions, but still be "healthy." Look at the *Penstemon* modeled in Figure 14.2. Populations in disturbed open habitats (whether caused by fire or human disturbance) contain a higher percentage of reproductive plants and exhibit increased germination compared with populations in protected areas. One could argue in this example that the population dynamics exhibited in areas disturbed by human activities may actually function as the target for introducing disturbance (e.g., prescribed fire) in natural areas.

Related or Similar Species

Comparisons with more "successful" related species or with species that appear ecologically similar may help set objective quantities that are biologically reasonable (Pavlik 1993). For example, Pavlik (1988) compared nutlet production in an endangered borage, *Amsinckia grandiflora*, with a weedy *Amsinckia*. In another series of studies, the demography of the rare *Plantago cordata*, which grows in freshwater tidal wetlands along the East Coast and along nontidal streams in Indiana and Illinois, was compared with the widespread *P. major* (Meagher et al. 1978). A comparable approach has been used to examine causes of endangerment in animals, for example, primates (Jernvall and Wright 1998) and neotropical migratory songbirds (Whitcomb et al. 1981). This approach has obvious limitations. Rare species are often rare because they do not have the reproductive capacity, dispersal potential, or growth potential of more common species.

Experts

Experts can provide additional information and opinions on the assumptions within the ecological model. Within the agency or organization, experts include regional and national ecologists, biologists and botanists, as well as specialists in other disciplines such as forestry, range management, and riparian management. External specialists include academic, professional, and amateur

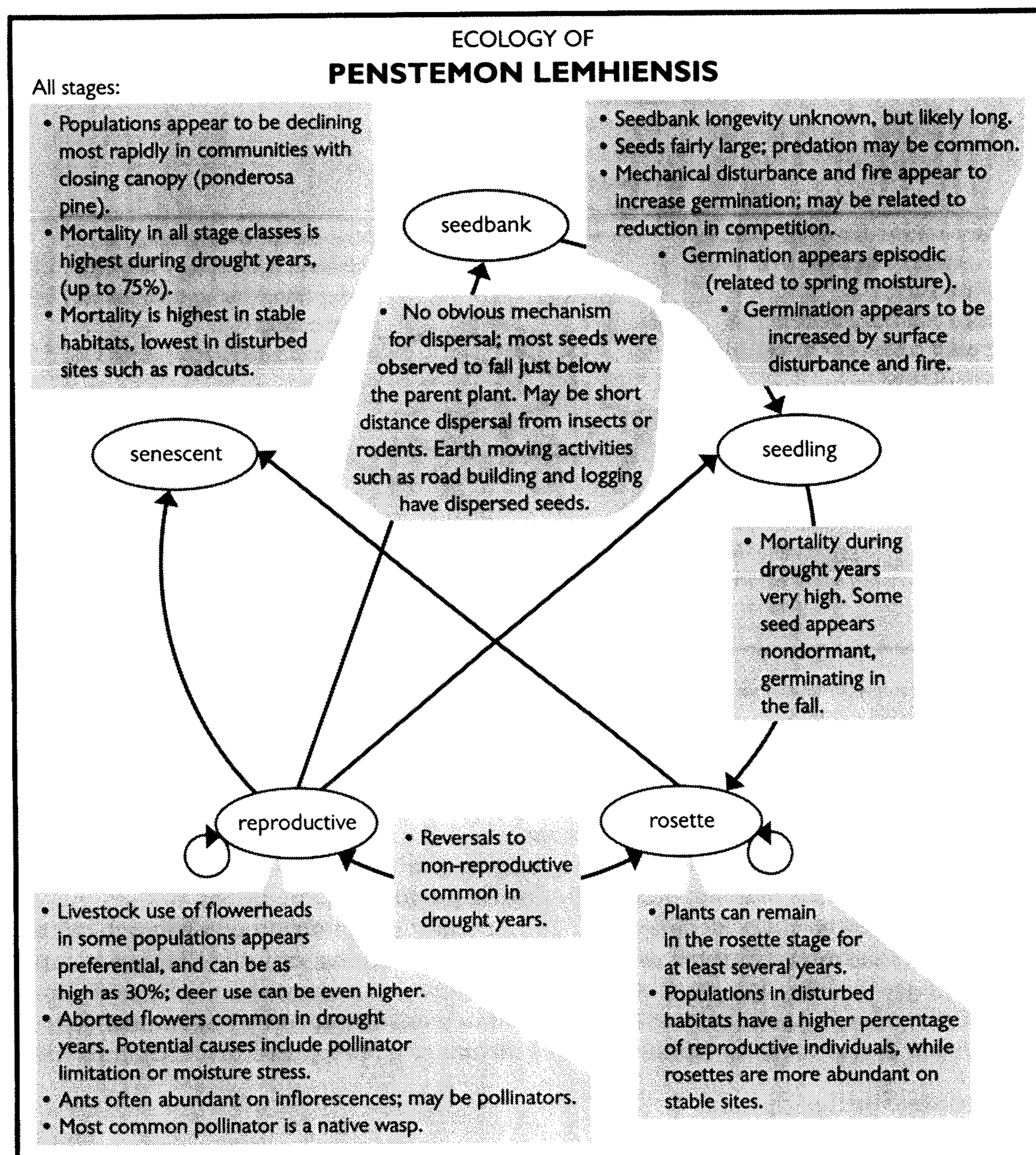


Figure 14.2. An ecological model of all known or suspected interactions for a rare *Penstemon* species.

ecologists, biologists and botanists who may know about the species of interest, or a closely related one, or may be knowledgeable about the ecological system in which the species resides. These people can help set realistic, achievable objectives.

Historical Records and Photos

Historical conditions at a site may have been captured in old aerial photos or in historic photos or other historical records housed in museums or maintained by local historical societies. Human disturbances such as roads, trails, and buildings may be visible. Woody species density and/or cover may also be visible. Early survey records often contained descriptions of general vegetation and habitat characteristics. Long-term elderly residents can be a fascinating source of information on local historical conditions.

DEVELOPING MANAGEMENT OBJECTIVES—AN EXAMPLE

The following provides an example of developing a management objective for a rare plant. This is intended as a generic demonstration of setting objectives. Please note that all the steps and concepts elaborated on are equally applicable to other plants or animals.

Our position is botanist with the United States Forest Service. *Collomia debilis* var. *camporum* is a long-lived, mat-forming perennial that occurs in 12 discrete locations (occurrences) along a 7-mile stretch of the North Fork of the Salmon River. Occurrences occupy stable slopes of blocky talus. Plants grow in soil pockets among the talus. Size of each occurrence ranges from 0.5 to 3 acres, each with 50 to 500+ pockets of plants. The number of plants cannot be determined because mats grow into each other and are difficult to separate into individuals. A two-lane highway runs along the base of the slope for the entire 7 miles. Any expansion of the highway (wider shoulders or more lanes) would severely impact all *Collomia* occurrences. Expansion is unlikely, however, given the status of the North Fork as a Wild and Scenic River and the controversial nature of any major road reconstruction. Two noxious weed species, cheatgrass (*Bromus tectorum*) and knapweed (*Centaurea repens*), occur along the highway right-of-way and are controlled annually. Some *Collomia* occurrences have sparse knapweed and cheatgrass. The effects of these weeds on the rare species are not known.

Review Upper-Level Direction

We first evaluate goals and objectives pertinent to *Collomia* in upper-level plans. The existing Forest Plan does not even recognize the occurrence of *Collomia* on USFS lands because the populations were discovered after the Forest Plan was finalized. The only direction provided by the Forest Plan is a standard operating procedure that states the effects of all projects on sensitive plant species will be evaluated through a field examination. An Allotment Management Plan (AMP) that describes cattle grazing management is in place for the area containing *Collomia*. It contains no references to sensitive plants nor is cattle grazing an issue on *Colomia* sites because the steep and rocky nature of its habitat precludes livestock use. The AMP is scheduled for evaluation and revision in 2010, and is the appropriate vehicle for describing management for all resources on that management unit (not just cattle management).

Identify the Species or Habitat Factor

An objective could focus on some aspect of *Collomia* or on the most immediate threat, weed infestation. You select the species itself for the following reasons:

- Although weeds are a concern, they currently are quite sparse in population areas, and current weed control efforts in the highway right of way appear fairly effective. You also have no information of the effects of weeds on *Collomia*, so monitoring weed density would not serve as a reliable indicator for population health.
- You have no data on trends or current condition of the *Collomia* occurrences except estimates of aerial extent and number of clumps of plants for each of the 12 occurrences. Although plants appear to be long-lived (many mat-forming species are), you noted in your field surveys that there seemed to be many dead individuals and no seedlings. You are concerned that some unknown factor may be causing these undesirable demographic dynamics. Because of the lack of information on trend or health of the occurrences you prefer to monitor the species directly.

In this situation, monitoring only the *Collomia* population and ignoring the potentially serious threat of weed infestation places the population at risk. If resources are available to monitor both the species and the weeds, you should develop a separate objective addressing the weed problem, rather than trying to combine the species and weeds into a single complex objective.

Draft objective: *Collomia debilis* var. *camporum*

Specify the Location

You decide to address all 12 occurrences because of the following reasons:

- All of the occurrences are administered by the Forest Service.
- You believe all 12 occurrences are important to the viability of the *Collomia* because this variety is so rare, and limited to such a small total area.
- This species is your top priority for monitoring, and will receive about half of your monitoring resources.

Draft objective: all 12 occurrences of *Collomia debilis* var. *camporum* along the North Fork

Describe the Attribute

Because of the high conservation priority of *Collomia*, you plan to quantitatively monitor this species at each occurrence. You select cover as an appropriate attribute for mat-forming perennials that cannot be separated into individuals.

Draft objective: Cover of all 12 occurrences *Collomia debilis* var. *camporum* along the North Fork.

Specify Action

Because you know so little about the species, you are unable to design management actions that would increase any aspect of this species. The current habitat exhibits no obvious impacts from humans (except for sparse weeds); thus, you assume that current levels are "natural." You decide that maintaining the current population would be acceptable.

Draft objective: Maintain cover of all 12 occurrences *Collomia debilis* var. *camporum* along the North Fork.

Specify Quantity

You want to maintain the current cover of *Collomia*, but you expect some natural fluctuation around a mean cover value even if *Collomia* populations are healthy and stable. You must specify the level of change that you will allow before you implement alternative management. You have no data suggesting an acceptable level of fluctuation. Because the species is so rare, you do not want to specify an allowable level of fluctuation so large that real and worrisome changes are not detected, but you also do not want your allowable limits of fluctuation so narrow that you are implementing new management unnecessarily. You decide to allow a decrease of 15% from current cover before you will implement alternative management. You base this value on your knowledge of natural fluctuations in unrelated perennial mat-forming species measured in a nearby range monitoring study.

Draft objective: At each of the 12 occurrences along the North Fork, limit any decrease in cover of *Collomia debilis* var. *camporum* to no more than 15%.

Specify Time Frame

Your objective is still unclear. As currently written, it suggests that an annual decrease of 10% from the previous year would be acceptable. You must identify the starting point from which you will measure the threshold decline of 15%. You also need to specify the period for which your objective is effective. Most objectives should include a final date that triggers a complete evaluation and final report.

You decide you want to measure the population for several years before writing a final report. You select the year 2005 because the AMP is scheduled for reevaluation in 2010 and because you are concerned about the percentage of dead plants in the population and the lack of seedlings. If you see a worrisome decline by 2005, you will have a few years for further study or

implementing trial management (such as weed control if weeds appear to be increasing) before the AMP is rewritten in 2010. You also decide that the baseline cover will be the cover measured in 2001, and that a decrease of more than 15% from that level would be unacceptable.

Final objective: At each of the 12 occurrences along the North Fork, limit any decrease from current (2001) cover of *Collomia debilis* var. *camporum* to no more than 15% between 2001 and 2005.

MANAGEMENT RESPONSE

The response of management to the outcome of monitoring must be identified before monitoring begins. If there are no management alternatives or options, monitoring resources are better spent on another species or population. Usually, however, there are options, but some of them may be expensive, or politically difficult to implement. There is a tendency in resource management agencies to continue monitoring, even when objectives are not met, rather than make the difficult decisions associated with changes in management. Because of this inertia, we recommend that management responses be an integral part of premonitoring planning. Management alternatives are more likely to be applied if they are identified before the monitoring begins, and if all parties agree to the objectives, monitoring methods, and response to monitoring data (see more on this in Chapter 15).

Identifying alternative management is difficult because in many situations the needed management changes are unknown. At a minimum, a management commitment can be made before monitoring begins that additional, more intensive investigation into the management needs of the species will begin if objectives are not achieved. For examples of management objectives paired with management responses, see Box 14.4 for plants and 14.5 for animals.

SAMPLING OBJECTIVES

Sampling objectives should be written as companion objectives to management objectives whenever monitoring includes sampling procedures. As described in Chapter 7, sampling involves assessing a portion of a population with the intent of making inferences to the sampled population as a whole. If you are weak on basic principles of sampling and have not yet read Chapter 7, please do so before reading this section on sampling objectives.

Sampling objectives specify information such as target levels of precision, power, acceptable false-change error rate, and the magnitude of change you are hoping to detect. Unlike a management objective, which sets a specific goal for attaining some ecological condition or change value, a sampling objective sets a specific goal for the measurement of that value. For example, consider the following management objectives, with corresponding sampling objectives:

> **Management objective:** We want to maintain a population of *Lomatium bradshawii* at the Willow Creek Preserve with at least 2000 individuals from 2002 to 2010 (target/threshold objective).
>
> **Sampling objective:** We want to be 95% confident that estimates are within ± 25% of the estimated true value.
>
> **Management objective:** We want to see a 20% increase in the average density of *Lomatium bradshawii* at the Willow Creek Preserve between 2002 and 2005 (change/trend objective).
>
> **Sampling objective:** We want to be 90% sure of detecting a 20% change in the density and we are willing to accept a 1 in 10 chance that we will say a change took place when it really did not.

The principal reason to add sampling objectives to management objectives is to ensure that you end up with useful monitoring information. If this additional information is not specified, you risk ending up with an inadequate sampling design that makes it difficult or almost impossible to assess whether you have achieved your management objective. For example, without setting sampling targets, you may end up with an estimate of population size with confidence intervals nearly as wide as the estimate itself (e.g., 1000 plants ± 950 plants) or you may find that you have low power to detect some biologically meaningful change (e.g., only a 15% chance of detecting the change you were hoping to achieve). The information specified in a sampling objective is also necessary to determine adequate sample sizes using the procedures described in Chapter 8 and Appendix II.

For monitoring that does not involve sampling, your ability to assess success at meeting your management objective should be obvious from the management objective itself without the need to specify additional information. Consider the following management objectives that involve monitoring without sampling:

- Maintain the current knapweed-free condition of the *Penstemon lemhiensis* population in the Iron Creek drainage for the next 10 years.
- Maintain at least 100 individuals of *Penstemon lemhiensis* in the Iron Creek drainage over the life of the Iron Creek Allotment Management Plan.

To determine success at meeting the first objective, you simply need to visit the site at some specified interval and search for the presence of knapweed. To assess success for the second objective, you will likely be able to count all the plants in the population (or at least the first 100 that you find). Thus, the management objectives for these nonsampling types of monitoring do not require the additional components that are discussed in this chapter.

Sampling objectives are classified into two types that correspond to the two major categories of management objectives: 1) target/threshold management objectives and 2) change/trend management objectives.

Target/Threshold Management Objectives

The sampling objective in this case is to estimate some parameter in the population (e.g., mean density per unit area, mean percent cover, or mean height or weight), to estimate a proportion (e.g., the frequency of a particular species within a set of quadrats placed within a sampled area), or to estimate total population size (total number of individuals within a sampled area). These estimates are then compared with the target/threshold value to determine if the management objective is met. Sampling objectives for this type of management objective need to include two components related to the precision of the estimate:

- **The confidence level.** How confident do you want to be that your confidence interval will include the true value? Is 80% confidence high enough or do you want 90%, 95%, or even 99% confidence?
- **The confidence interval width.** How wide a range are you willing to accept around your estimated value? For example, is ± 20% of the estimated mean or total value adequate or do you want to be within ± 10%?

The following is an example of a target/threshold management objective with a corresponding sampling objective:

Management objective: Increase the number of individuals of *Penstemon lemhiensis* in the Iron Creek Population to 1000 individuals by the year 2010.

Sampling objective: We want to be 95% confident that population estimates are within 20% of the estimated true value.

This sampling objective specifies a relative confidence interval width (± 20% of the estimated true value) so the targeted confidence interval width in absolute units will depend on the estimated population size. For example, if the first year of monitoring yields a population estimate of 500 plants, the targeted confidence interval half-width is 500 plants × 20% = ± 100 plants. Information from pilot sampling can be used to determine how many sampling units need to be sampled to achieve a confidence interval width of ± 100 plants.

Why should you set sampling objectives for target/threshold management objectives? Most importantly, it helps you avoid designing monitoring studies that provide unreliable estimates that are of little value for making management decisions (e.g., a population estimate of 1200 ± 950). The values set in your sampling objective will be used after the pilot study to determine the sample size needed to meet the sampling objective. Sampling objectives set a quantitative measure of the quality of your monitoring design.

See Box 14.4 (plants) and 14.5 (animals) for additional examples of sampling objectives paired with target/threshold management objectives.

Change/Trend Management Objectives

The sampling objective in this case is to determine whether there has been a change in some population parameter such as a mean value (e.g., mean density per unit area of a particular species, mean percent cover, mean weight), a proportion (e.g., the frequency of a particular species within a set of quadrats placed within the sampled area), or the total population (total number of individuals within a sampled area) between two or more periods. This category of sampling objective must include the following three components:

- **The acceptable level of power (or the acceptable level of the missed-change error [Type II error] rate).** How certain do you want to be that, if a particular change does occur, you will be able to detect it? If you want to be 90% certain of detecting a particular magnitude of change, then you are specifying a desired power of "90%" (power and missed-change error rates are complementary, so in this example, the missed-change error rate is 0.10).
- **The acceptable false-change error (Type I error) rate.** What is the acceptable threshold value for determining whether an observed difference actually occurred or if the observed difference resulted from a chance event? This represents the chance of concluding that a change took place when it really did not. While the $\alpha = 0.05$ level is frequently used, you should carefully consider the impact of this decision on the probability of making missed-change errors before selecting a false-change error rate. In many monitoring studies, a higher false-change error rate (e.g., $\alpha = 0.10$ or $\alpha = 0.20$) is appropriate.
- **The desired MDC (minimum detectable change).** The MDC specifies the smallest change that you are hoping to detect with your sampling effort. The MDC should represent a biologically meaningful quantity given the likely degree of natural variation in the attribute being measured.

The following is an example of a change/trend type of management objective with a corresponding sampling objective:

Management objective: I want to see a 20% increase in the density of *Lomatium cookii* at the Agate Desert Preserve between 2002 and 2010.

Sampling objective: I want to be 90% certain of detecting a 20% increase in density between 2002 and 2010 and I am willing to accept a 10% chance that I will make a false-change error.

This sampling objective specifies a power of 90%, a false-change error rate of 10%, and an MDC of 20%. The MDC is specified in relative terms, so the targeted MDC in absolute units

will depend on the estimated density in 2002. For example, if the mean density in 2002 is 10 plants/quadrat, the desired MDC is an increase of two plants/quadrat.

Why bother specifying false-change error rates, power, and some desired MDC when you are writing a sampling objective designed to detect change over time? The main advantage is that it helps you avoid designing monitoring studies with low power. The sample size determination procedures discussed in Chapter 8 require the specification of false-change error rate, power, and the size of the change you are interested in detecting before you can determine how many sampling units to sample. If your pilot data indicate that you have low power to detect a biologically important change (high probability of a missed-change error), you can then correct your sampling design before you have gathered many years of monitoring data.

See Box 14.4 (plants) and 14.5 (animals) for additional examples of sampling objectives paired with change/trend management objectives.

Setting Realistic Sampling Objectives

Sampling objectives should be written during the planning phase of a monitoring study. Targeted levels of precision, power, false-change error, and MDC should be based on the following:

- The biology of the species. How much can it change? How fast? How much does it fluctuate from year to year?
- The risk of being wrong. You are not required to use the 5% level, so common in research studies. Evaluate the relative risks of false-change and missed-changed errors. Remember false-change and missed-change error rates are inversely related to each other, although not proportionately (see Chapter 7). Remember too that a smaller MDC is more difficult to detect. Consult with decision-makers and stakeholders interested in the monitoring results to ensure that they are comfortable with the targeted levels of precision, power, etcetera, specified in the sampling objectives.
- The resources available for monitoring. Higher levels of precision and lower acceptable miss-change and false-change errors require more resources for monitoring (usually more sampling units because you have already done your best to develop an efficient design—see Chapter 8).

Writing sampling objectives for target/threshold management objectives is fairly straightforward. You must decide the width of the confidence interval, and the risk you are willing to take that your estimate is not actually within that interval (that risk equals one minus the confidence level).

Setting error rates in sampling objectives for change/trend management objectives is more complicated. Both false-change and missed-change error rates can be reduced by sampling design changes that increase sample size or decrease sample standard deviations, but missed-change and false-change error rates are inversely related, which means that reducing one will increase the other (but not proportionately) if no other changes are made. The decision of which type of error is more important should be based on the nature of the changes you are trying to determine, and the consequences of making either kind of mistake. Because these errors have different consequences to different interest groups, there are different opinions as to what the "acceptable" error rates should be: The following examples demonstrate the conflict between false-change and missed-change errors.

- Testing for a lethal disease. When screening a patient for some disease that is lethal without treatment, a physician is less concerned about making a false diagnosis error (analogous to a false-change error) of concluding that the person has the disease when he does not than failing to detect the disease (analogous to a missed-change error) and concluding that the person does not have the disease when in fact he does.

- Testing for guilt in our judicial system. In the United States, the null hypothesis is that the accused person is innocent. Different standards for making judgment errors are used depending on whether the case is a criminal or a civil case. In criminal cases, proof must be "beyond a reasonable doubt." In these situations it is less likely that an innocent person will be convicted (analogous to a false-change error), but it is more likely that a guilty person will go free (analogous to a missed-change error). In civil cases, proof only needs to be "on the balance of probabilities." In these situations, there is a greater likelihood of making a false conviction (analogous to a false-change error), but a lower likelihood of making a missed conviction (analogous to a missed-change) error when compared to criminal cases.
- Testing for pollution problems. In pollution monitoring situations, the industry has an interest in minimizing false-change errors and may desire a very low false-change error rate (e.g., $\alpha = 0.01$ or 0.001). Companies do not want to be shut down or implement expensive pollution control procedures if a real impact has not occurred. In contrast, an organization concerned solely with the environmental impacts of some pollution activity will likely want to have high power (low missed-change error rate) so that they do not miss any real changes that take place. They may not be as concerned about occasional false-change errors (which would result in additional pollution control efforts even though real changes did not take place).

Missed-change errors may be as costly or more costly than false-change errors in environmental monitoring studies (Toft and Shea 1983; Peterman 1990; Fairweather 1991). A false-change error may lead to the commitment of more time, energy, and people, but probably only for a short time until the mistake is discovered (Simberloff 1990). In contrast, a missed-change error, as a result of a poor study design, may lead to a false sense of security until the extent of the damages are so extreme that they show up in spite of a poor study design (Fairweather 1991). In this case, rectifying the situation and returning the system to its preimpact condition could be costly. For this reason, you may want to set equal false-change and missed-change error rates or even consider setting the missed-change error rate lower than the false-change error rate (Peterman 1990; Fairweather 1991).

There are many historic examples of costly missed-change errors in environmental monitoring. For example, many fish population monitoring studies have had low power to detect biologically meaningful declines so that declines were not detected until it was too late and entire populations crashed (Peterman 1990). Some authors advocate the use of something they call the "precautionary principle" (Peterman and M'Gognigle 1992). They argue that, in situations where there is low power to detect biologically meaningful declines in some environmental parameter, management actions should be prescribed as if the parameter had actually declined. Similarly, some authors recommend shifting the burden of proof in situations where there might be an environmental impact from environmental protection interests to industry/development interests (Peterman 1990; Fairweather 1991). They argue that a conservative management strategy of "assume the worst until proven otherwise" should be adopted. Under this strategy, developments that may negatively impact the environment should not proceed until the proponents can demonstrate, with high power, a lack of impact on the environment.

The sampling objectives serve as a critical aid during the preliminary or pilot field sampling phase. Once pilot sampling data are available, information on the variability of the data can be plugged into sample size equations (see Chapter 8 and Appendix II) along with the information specified in the sampling objectives to determine how many sampling units should be sampled. If you are faced with a monitoring situation with high variability between sampling units (despite all of your sampling design efforts to lower this variability) and the components of your sampling objective lead to a recommended sample size of more sampling units than you can afford to sample, then you need to reassess the monitoring study. Is it reasonable to make changes to some components of the sampling objective? For target/threshold types of management objectives, this

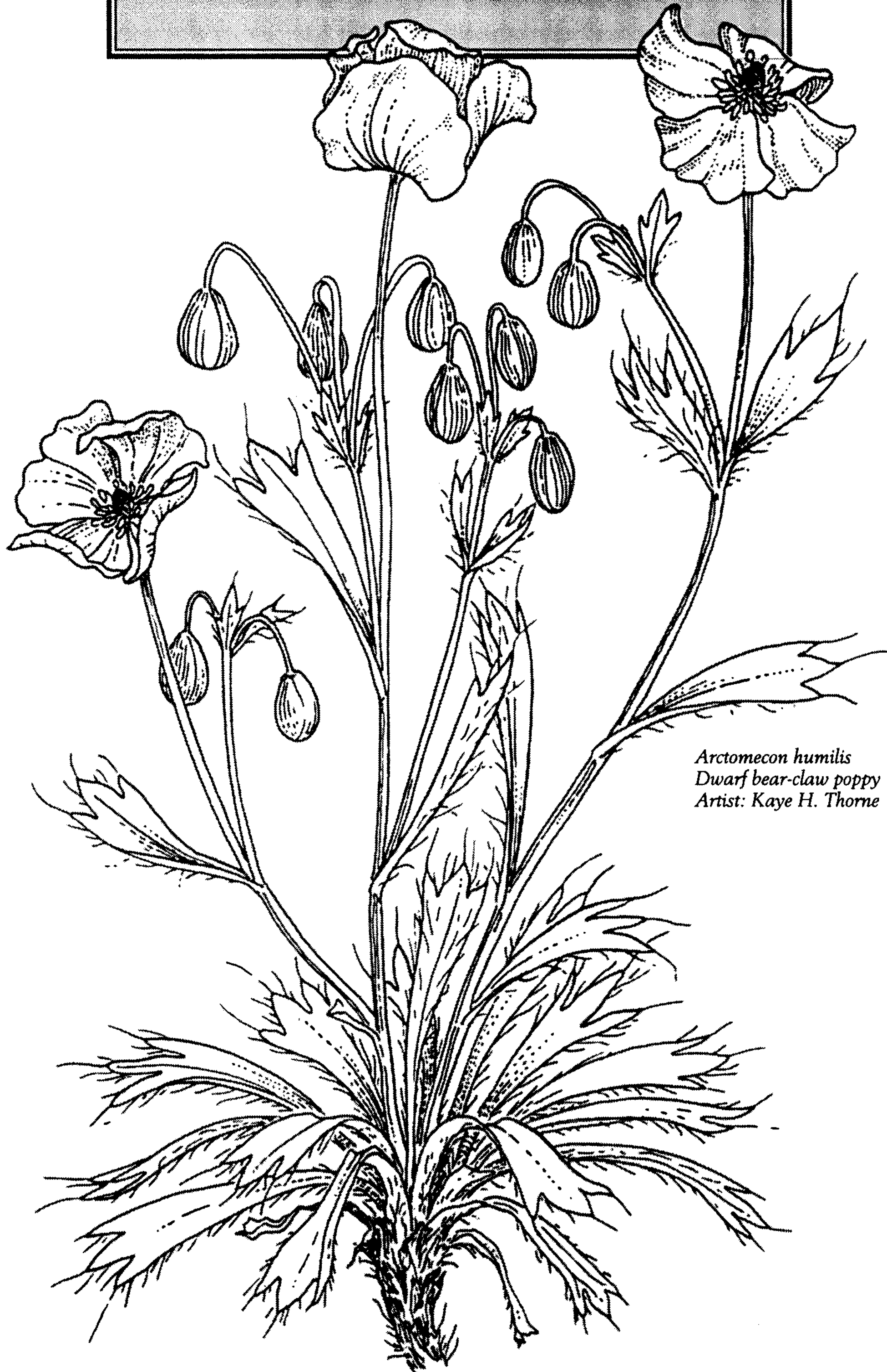
may mean lowering the level of confidence or decreasing the precision of the estimate (i.e., increasing the confidence interval width) or both. For objectives directed towards tracking change over time, this may mean increasing the acceptable false–change error rate, decreasing the targeted power level, or settling on a larger specified MDC. Will these changes be acceptable to managers and other stakeholders? If you feel that making these modifications to the sampling objective is unreasonable, then you should take an alternative monitoring approach rather than proceed knowing that your monitoring project is unlikely to meet the stated objectives.

MANAGEMENT IMPLICATIONS

Management objectives are the foundation of a monitoring study. As measurable descriptions of desired state or condition of the resource, objectives promote communication, give direction for management actions and monitoring approaches, and provide a means to measure management success. Objectives should include the following components: species or indicator, location of management, attribute of species or indicator expected to respond to management, amount of change or desired condition or the species or indicator, and the time frame during which management will be applied and results expected. Objectives can be generally classified into two types. Target/threshold management objectives state the desired condition or state of the species or indicator. Change/trend management objectives describe the amount and direction of desired change.

Management objectives should be paired with a clearly defined management response that will be implemented if the objective is not met. Management objectives must also be paired with sampling objectives when the monitoring study involves sampling. Sampling objectives ensure that the monitoring design will provide useful and meaningful data when using sampling. These describe the target level of precision (confidence level and confidence interval width) for target/threshold management objectives, and the desired power, acceptable false-change error rate, and minimum change that should be detected by a study that monitors a change/trend management objective.

CHAPTER 15
Communication and Monitoring Plans



Arctomecon humilis
Dwarf bear-claw poppy
Artist: Kaye H. Thorne

A successful monitoring project is characterized by three traits. First, it is well designed and technically defensible. Second, it is implemented as planned in spite of personnel changes, changes in funding, and changes in priorities. Third, the information from a successful monitoring program is applied, resulting in management changes or validation of existing management (Gray and Jensen 1993).

Monitoring projects that are implemented to completion and applied to decision making will complete the adaptive-management cycle described in Chapter 1. All three of these traits depend on good communication and documentation over the life of the project. Good design is usually the result of collaboration with stakeholders and other specialists and help from experts. Consistent implementation requires the support and knowledge of managers and documentation of methods to survive personnel changes. Finally, application to management decisions requires communication of results. A monitoring project that simply provides additional insights into the natural history of a species, or that languishes in a file and is read only by the specialist, does not meet the intent of monitoring.

COMMUNICATION

Communication does not start when the monitoring results have been analyzed. Beginning with the planning stage, those who will be making decisions based on the monitoring and those who may be affected by those decisions must be included in the design of the monitoring project. You will increase the likelihood of seeing needed management actions implemented by involving all interested parties in developing the management objective and in designing the monitoring, as well as by reaching agreement that all parties will abide by the results (Hirst 1983; Johnson 1993). Objectives, written as Management Objectives–Management Response pairs (see Chapter 14), should clearly identify the management changes that will be implemented based on monitoring results (Gray and Jensen 1993). This point cannot be stressed enough, especially when potential decisions may adversely affect some parties or interests. If you fail to include all who should be involved in the initial stages of objective setting and monitoring design, adversaries against implementing new management may appear once monitoring is completed.

Several classes of participants that may be involved in the development of a monitoring project are described in Box 15.1. The number of people and groups to involve in a monitoring

Box 15.1 *PARTICIPANTS IN A MONITORING PROJECT*

Decision-makers *(managers or management teams) are the most important audience. They will decide the amount of resources to devote to the monitoring project and, once monitoring is completed, decide whether management should change or continue. Each manager's "comfort level" varies for making decisions based on monitoring data. Some managers feel confident making decisions based on photographs and their specialist's judgment. Others require much more information.*

Agency specialists *(in-house). Other resource specialists may have information critical to the design of the monitoring (e.g., the area containing the population is likely to be rested from grazing for the next 3 years; the timber stand is set aside from cutting because it is in a protected watershed). These other specialists also tend to be advocates for the resource they manage and may potentially disagree with the management changes resulting from monitoring. Including these specialists in the design creates ownership in the monitoring and reduces the potential for in-house disagreements later.*

Regulatory decision-makers *(e.g., U.S. Fish and Wildlife Service, state agencies). Participation by these agencies is required for species listed under the Endangered Species Act or state laws and may be helpful for other species of concern.*

Nonregulatory agencies. *State agencies that maintain statewide conservation databases such as the Heritage Program or conservation programs often have information about the same species on private lands, on other federal lands, or on lands in other states. Many of these database agencies also maintain a monitoring database; participation in it can reduce redundancy in monitoring efforts. In the United States, for example, local Natural Resource Conservation Service personnel and County Extension agents may function as advocates for agricultural interests. Their participation and support of the monitoring project increases the credibility of the monitoring data with traditional federal land users such as grazing permittees.*

Traditional land users. *These are primarily commodity producers such as miners, loggers and timber companies, and livestock operators. If the monitoring potentially will affect these interests, you should include them throughout the process. Not only does their involvement from the beginning diffuse much of their disagreement when assessing results, but their involvement will also improve the quality of the monitoring. Because their economic interests are potentially at stake, they will be interested more in false-change errors (e.g., concluding that a decline took place when it really did not), whereas you may be more concerned with missed-change errors (e.g., failing to detect undesirable changes that in fact did occur). The explicit balancing of the two errors is important. In addition, individuals involved in commodity production on federal lands often know facts about a population area or an activity that you do not. A rancher, for example, may know that cows have not used an area for the last 10 fall seasons because of a nonfunctioning water source. A logger may know that his grandfather cut a patch of timber using horses in the 1930s. These bits of information may improve your ecological model.*

Nontraditional land users. *Newer users of public lands, for example, off-road recreationists, hikers, hot-spring visitors, and others whose use of the federal land may be affected by changes in management resulting from monitoring, should be included.*

Environmental and conservation groups. *Include groups that have an interest in native flora and biodiversity, especially if local representatives are available. Native plant societies not only have a special interest in the preservation of the diversity of native vegetation within a state, but may also have specialized skills or volunteer labor that will improve the quality of monitoring. Hunting and fishing enthusiasts are often willing to help with conservation projects. Environmental groups may be the source of information and support.*

Professional and academic botanists and biologists. *These people may have much to contribute to the development of ecological models, objectives, and monitoring designs. Their contribution to and review of the monitoring strategy will improve the quality and increase the credibility of the monitoring effort.*

project depends on the potential impacts of the management changes that may occur based on monitoring results. Developing objectives for populations in areas that are not affected by commodity extraction or recreational use may require little interaction with interest groups or other agency specialists. Large populations, populations with controversial land use activities, or populations in high-use/high-visibility areas, may require extensive communication efforts before management and monitoring is initiated.

Establishing communication and considering alternative points of view can be time-consuming and difficult. An apparently easier route often followed by some specialists is collecting "really good data" to prove a point and attempting to use the data to influence management changes. In practice, monitoring that is specialist-driven rarely results in a management change for three reasons. The most common is that the specialist who spearheaded the monitoring leaves and the managers suspend the monitoring project because it never had institutional support and now lacks an advocate. A second reason is that, without commitment by management, other priorities take precedence over the monitoring project and divert the limited resources and time available to the specialist. Third, a lack of consensus on objectives and methodology almost ensures that a decision-maker will not use the monitoring data because of continued controversy. You need to involve people early in the process to ensure a cooperative effort and the application of monitoring results to the decision-making process (Hirst 1983). At a minimum, potential internal antagonists should be involved and supportive by the time field data collection begins.

Effective facilitation of public participation is not easy, but successful examples do exist (Yaffee and Wondolleck 1997). Shindler and Aldred Cheek (1999) identified six characteristics of a program of effective communication and solicitation of public involvement:

1. Allows all who wish to participate open access
2. Includes people with leadership and interpersonal skills who are able to develop ongoing relationships with participants
3. Demonstrates institutional and personal flexibility by agency and technical personnel to explore innovative approaches
4. Carefully designs a process that describes how public participation will be incorporated into the decision-making process and applies that process at the initial stages
5. Delivers tangible products demonstrating that participation leads to meaningful progress (e.g., the construction of a project, a change in a service)
6. Creates trust by demonstrating consistent, open and honest behavior; by delivering the promised products; and by clearly incorporating the participant's ideas

You may not have the particular responsibility of ensuring effective public participation, but you must recognize the importance of the process to successful application of the adaptive-management cycle of which your monitoring is a part, and you must support the process in any way you can. If you do find yourself responsible for the public participation process (perhaps because in your position as, for example, a preserve manager, you bear responsibility for the whole program), recognize the process as integral, rather than ancillary, to the success of an adaptive-management and monitoring project.

Communication about monitoring projects associated with noncontroversial management actions can safely be limited to decision-makers and internal resource specialists. For example, often you will know too little about populations and their interactions with management activities to develop Management Objective–Management Response pairs that identify a specific management response. Many management responses in the examples in Chapter 14 specify a second stage of more intensive monitoring and perhaps research if the population is declining or failing to increase. Such two-tier monitoring requires only the involvement of the decision-maker and resource specialists within the administrative unit in the first stage because implementing increased monitoring or research may be expensive but is rarely controversial among stakeholders outside the agency.

Even in noncontroversial situations, however, you may want to enlist involvement and/or review by a broader spectrum of participants. Review by user groups during the development of objectives will inject fresh perspectives and often provide useful, local knowledge. Review during the design phase by academic specialists, statisticians, experienced professional biologists, and peers may help you avoid potential technical problems.

MONITORING PLANS AS A COMMUNICATION TOOL

Communication with these participants is facilitated by a monitoring plan that explains the rationale for the monitoring project, documents objectives and the management response, and describes the monitoring methodology in enough detail to direct continued implementation.

Monitoring plans serve five important functions:

1. A draft plan provides a full description of the ecological model, the objectives, and the proposed methodology.
2. Draft monitoring plans provide a means to solicit input from many participants.
3. A final monitoring plan consolidates all information into a single document that can be easily accessed and referenced.
4. A final monitoring plan documents the location and techniques of the monitoring in sufficient detail that a successor can continue the monitoring.
5. A final monitoring plan documents the agency's commitment to implementing a monitoring project and the management that will occur based on monitoring results. A monitoring plan can also be signed by all participants to demonstrate their support for the project and acceptance of the proposed management changes that may result.

Monitoring plans must be complete, providing all the information needed to judge the quality of your proposed monitoring and to continue it in your absence. Box 15.2 summarizes the elements to include in an extensive monitoring plan for a complex project. Less complex projects may require less extensive explanations and fewer elements. A short (one- to two-page) summary at the beginning of the plan will be useful to decision-makers, other specialists, and user groups.

Do all monitoring projects require a monitoring plan? Does a qualitative monitoring project that simply involves taking a picture of the population each year require a full-scale document such as the one summarized in Box 15.2? Some form of documentation of the management objective, sampling objective (if sampling), management response, location, and methodology is necessary for all monitoring projects, no matter how small or simple. (The field-monitoring cover sheet described in Chapter 6 lists many of these elements and may be adequate for some situations if an introduction that describes the objectives is included.)

The flow chart in Chapter 2 suggests writing the monitoring plan before the pilot study. There is a valid concern, however, that if the pilot study demonstrates that the monitoring approach needs significant revisions, the monitoring plan will need to be rewritten. We suggest drafting the plan early in the process, perhaps just including the analysis of the problem, the ecological model, and some ideas on management, objectives, and monitoring methods. Use the draft as a communication tool and a means of soliciting comments and suggestions. Finalize it after the monitoring methodology proves effective.

Clearly, a significant investment of resources is required to complete all the elements of a monitoring plan, and most biologists prefer fieldwork to writing plans. The temptation is great to skip this stage and get on with "more important" work such as counting plants in plots or frogs in ponds. Resist the temptation. A monitoring plan is critical to successful, long-term implementation of monitoring.

COMMUNICATING RESULTS

Evaluating Results at the End of the Pilot Period

In this handbook, we have advocated the use of pilot studies to avoid the expense and waste of a monitoring project that yields inconclusive results. After the pilot period you should consider several issues before continuing the monitoring project:

Box 15.2. *ELEMENTS OF A MONITORING PLAN*

I. *Introduction (general)*

Species, need for study, management conflicts.

II. *Description of ecological model*

Life history, reproductive biology, causes of distribution, habitat characteristics, known or suspected threats (e.g., herbivory of flower heads by cattle, competition from invasive species, off-highway vehicle impacts). The model should describe known biology (based on natural history observations) and conjectural relationships and functions. Relationships that are hypothesized and sources of information should be identified. The purpose of this section is to identify the sensitive attribute to be measured and to describe the relationships between species biology and management activities. This section is the biological basis for the development of objectives.

III. *Management objective*

Includes rationale for the choice of the attribute to be measured and the amount of change desired or target population size.

IV. *Monitoring design*
- A. *Sampling objective (if sampling)*
  Include rationale for choice of precision and error rates.
- B. *Sampling design*
  Describe methods clearly. What size are the sampling units? How are sampling units placed in the field? How many sampling units?
- C. *Field measurements*
  What is the unit counted (for density)? How are irregular outlines and small gaps treated (for line-intercepts)? How are plots monumented (if permanent)? Include all the information needed for someone else to implement or continue the monitoring in your absence.
- D. *Timing of monitoring*
  What time of year, both calendar and phenologically? How often?
- E. *Monitoring location*
  Include clear directions, maps and aerial photographs describing the study location, and the location of individual sampling units (if permanent).
- F. *Intended data-analysis approach*

V. *Data sheet example*

VI. *Responsible party*

VII. *Funding*

VIII. *Management responses to potential results*

Can the Monitoring Design Be Implemented as Planned?

The pilot period should answer several questions about field design and implementation: If sampling units are permanent, can they be relocated? Are sampling units reasonably sized, or do they contain hundreds of individuals? Is it difficult to accurately position a tape because of dense growth? Are the investigator impacts from monitoring acceptable? Is the skill level of field personnel adequate for the fieldwork, or is additional training needed? Projects rarely work as smoothly in the field as anticipated in the office. Nearly all monitoring projects require some modification for effective field implementation. Occasionally, you may find that the planned method does not work at all, and a major overhaul of the monitoring project is required.

Are the Costs of Monitoring Within Estimates?

The pilot period is important as a reality check on required resources: Does the monitoring take much longer than planned? Will the data entry, analysis, and reporting work take more time than allocated? If the monitoring project as designed requires more resources than originally planned, either more resources must be allocated to the project, or you will need to redesign the monitoring to be within budget.

Do the Assumptions of the Ecological Model Still Seem Valid?

Your understanding of the biology and ecology of a species may improve as you spend time on the site collecting data. Does new information suggest that another attribute would be more sensitive or easier to measure (cover instead of density, for example)? Is the change that you have targeted to monitor biologically significant, or is the natural annual variability that results from weather conditions so extreme that it masks the target change? Does the frequency of monitoring still seem appropriate?

Were Precision and Power Objectives Met (for Sampling Situations)?

After analyzing the pilot data, you may discover that you need many more sampling units than you planned to achieve the standards for precision, confidence, and power that you set in your sampling objective (see Chapter 14). You have six alternatives:

1. Reconsider the design. The pilot study should improve your understanding of the population's spatial distribution. Will a different sampling-unit shape or size improve the efficiency and allow you to meet the sampling objective within the resources available for monitoring?
2. Reassess the scale. Consider sampling only a single management unit or perhaps only one or a few macroplots.
3. Lobby for additional resources to be devoted to this monitoring project. Power curves such as those shown in Chapter 7 may help to graphically illustrate the tradeoffs of precision, power, and sampling costs for managers (Brady et al. 1995).
4. Accept lower precision (in target/threshold objectives) or larger levels of minimum detectable change (in change/trend objectives). It may be prohibitively expensive, for example, to be 90% confident of being within 10% of the estimated true mean, but it may be possible to be 90% confident of being within 20% of the estimated true mean using available monitoring resources. You may not be able to detect a 5% change with a power of 90%, but you may be able to detect a 10% change.
5. Accept higher error rates. You may not, with the current design and expenditure of monitoring resources, be 90% certain of detecting a specified change, but you may be 80% certain. You may have to accept a 20% chance that you will make a false-change error, rather than the 10% level you set in your sampling objective. You may not be within 10% of the estimated true mean with a 95% confidence level, but your current design may allow you to be 90% confident of being within 10% of the

estimated true mean. Look at the results from your pilot study, and consider whether the significance levels that can be achieved with the current design are acceptable, even though the levels may be less stringent than you originally set in your sampling objective.

6. Start over. Acknowledge that you cannot meet the sampling objective with reasonable precision or power within the budgetary constraints of the project.

The results from the pilot period should be reported even if your design and project require significant revision. Your audiences for this report would include all those who reviewed your initial project proposal or monitoring plan. A report to managers is especially important to describe the recommended changes in design. Your report is also important to your successor and possibly other ecologists or botanists who work with similar situations or species. Reporting failures of techniques will help others avoid similar mistakes.

Assessing and Reporting Results After the Pilot Period

Three possible conclusions result from a monitoring study: 1) objectives are (being) met; (2) objectives are not (being) met; or 3) the data are inconclusive (see Chapter 9 for interpretation of statistical analyses). The pilot period should eliminate the problem of inconclusive results caused by a poor design, but such results may occur even with an excellent design.

Objectives Are Met

Two management responses should result for objectives that have been met. First, the objective should be reevaluated and changed based on any new knowledge about a species and population. Second, both management and monitoring should be continued, although the latter perhaps less frequently or less intensely.

It is important that monitoring does not automatically stop when objectives are first met. Measured success over short time frame may not be related to management, but simply a lucky correlation of an increasing population size or a condition within the management period that is caused by unknown factors. Fluctuations in population size caused by weather can give the appearance of success, especially with annuals and short-lived perennials. If monitoring data shows stable or increasing trends, you may scale back the frequency and intensity of monitoring, but do not consider the job done and ignore the population or species permanently. Current management may in fact be detrimental, but its negative effects masked by fluctuations related to weather. In addition, conditions change— weeds invade, native ungulate populations increase, livestock-use patterns change with the construction of a fence or water trough, and recreational pressure increases. All these things and more may pose new threats.

Objectives Are Not Met

As described in Chapter 1, according to the adaptive-management approach, failure to meet an objective should result in the change in management that was identified as the management response during the objective development phase (see Chapter 14). Rarely, however, is resource management that simple. We need to remember that the inertia that resists changing management is very difficult to overcome. Managers will generally continue implementing existing management, the path of least resistance, unless monitoring or some other overriding reason clearly indicates a change.

Unfortunately, the data from most monitoring will not conclusively identify causes of failure to meet objectives or the corresponding corrective action (see the discussion on monitoring versus research, Chapter 1). The biologist who is monitoring the population may feel confident of the cause, but decision-makers may be uncomfortable making changes in management, especially unpopular ones, that have a basis only in the biologist's professional opinion.

Thus, the most common response in land-management agencies is to first reevaluate the objective. Was the amount of change too optimistic and biologically unlikely? Was the rate of

change too optimistic? While such assessment is necessary, it can too often result in changing the objective rather than implementing necessary management changes.

This scenario is extremely common, but often may be avoided by two techniques. The first is to articulate the management response along with the management objective (as suggested in Chapter 14). This clearly states the response to monitoring results before monitoring is even started. It represents a commitment by the agency to stand by its monitoring results and to use them to adapt management. The second technique is to reach consensus among all interested parties concerning the monitoring and the management response before monitoring data are collected (Johnson 1993).

You should analyze results of monitoring each year (or each year data are collected) and report them in a short summary. Analyzing data as soon as they are collected has several benefits. The most important is that analysis is completed while the field work is still fresh in your mind. Questions always arise during analysis, and the sooner analysis takes place after the field work, the more likely you can answer those questions. You may also find after analysis that you would like supplementary information, but it may not be possible to collect this in the middle of the winter or 5 years after the monitoring data were collected. You will have lost a valuable opportunity. Analysis after each data collection episode also means that you will assess the monitoring approach periodically. Although many problems will surface during the pilot period, some may not until after a few years of data collection. Periodic assessment ensures a long-term monitoring project against problems of inadequate precision and power and problems of interpretation.

Final Monitoring Reports

At the end of the specified monitoring period, or when objectives are reached, you should summarize the results in a formal monitoring report (Box 15.3). Much of the information needed for the report can be lifted directly from the monitoring plan (Box 15.2), although deviations from the proposed approach and the reasons for them will need to be described. The final report should be a complete document, so you should include all pertinent elements from the monitoring plan. You can either cut and paste electronically from the monitoring plan or simply append the report to existing copies of the monitoring plan. The preparation of the report should not be a major task. If you have been completing annual data analysis and internal reporting (as you should), summarizing the entire monitoring project should be straightforward.

Completing the monitoring project with a final formal report is important. This report provides a complete document that describes the monitoring and its results for distribution to interested parties. It provides a complete summary of the monitoring activity for successors, avoiding needless repetition or misunderstanding of the work of the predecessor. Finally, a professional summary lends credibility to the recommended management changes by presenting all of the evidence in a single document.

If the results would be interesting to others, consider sharing those results through a technical paper or symposium proceedings. Much of the preparation work for a presentation has already been done with the completion of the monitoring plan and monitoring report documents. Sharing the results has three important benefits: 1) it increases the audience, possibly helping more people and improving other monitoring projects (similar problems, similar species, etc.); 2) it increases the professional credibility of the agency or organization that conducted the monitoring; and 3) it contributes to your professional growth.

MANAGEMENT IMPLICATIONS

Successful monitoring projects are part of an adaptive-management cycle. To function within that cycle, monitoring must be technically defensible, consistently implemented to completion, and applied to decision-making. Communication with all parties facilitates the development of

Box 15.3. *MONITORING REPORTS*

I. Introduction

II. Description of ecological model

III. Management objective(s)

IV. Monitoring design
- *A. Sampling objective*
- *B. Sampling design*
- *C. Field measurements*
- *D. Timing of monitoring*
- *E. Monitoring location*
- *F. Intended data-analysis approach*

V. Data sheet example

VI. Responsible party

VII. Funding

VIII. Management response to potential results

IX. Summary of results

Include tables and figures communicating the results, as well as general natural history observations.

X. Interpretation of results

Describe potential causes for the results observed, sources of uncertainty in the data, and implications of the results for the resource.

XI. Assessment of the monitoring project

Describe time and resource requirements, efficiency of the methods, and suggestions for improvement.

XII. Management recommendations
- *A. Change in management*
  Recommend changes based on results and the management implications identified in Section VIII.
- *B. Change in monitoring*
  Analyze costs versus information gain, effectiveness of current monitoring system, and recommended changes in monitoring.

XIII. References

Include grey literature and personal communications.

XIV. Reviewers

List those who have reviewed drafts of the report.

management objectives, the design of the monitoring study, and the interpretation and application of the monitoring data. Failure to include stakeholders in development of a monitoring project generally results in failure of the adaptive-management cycle. Monitoring plans are effective tools for communication, both among stakeholders and among internal specialists. Monitoring plans represent a commitment to completing a monitoring study and using the data. They also ensure that monitoring design specifications are not lost as a result of personnel changes. Monitoring reports are important tools for summarizing and disseminating results.

APPENDIX I
Monitoring Communities

INTRODUCTION

We define communities simply as assemblages of plants and animals living in the same place. Communities are thus described by the presence, abundance, and pattern of all the species found within a specified area. Although each assemblage is unique and species presence and abundance are individualistic responses to the abiotic and biotic environment, the patterns of response often create recognizable and repeated similar assemblages of plants and animals across the landscape. Thus, we can name, consider, and alter these repeated assemblages; we can manage and study communities.

Management actions are often directed at maintaining or changing these assemblages. For example, we may wish to restore the native plant communities typically found in an undisturbed riparian corridor or to maintain in a forest managed for wood products an understory community of salamanders similar to that found in minimally managed forests. Sometimes, we wish to preserve a unique or unusual assemblage, a rare community. If communities are the focus of our management, how will we monitor the effectiveness of our management on the community?

Monitoring communities poses different challenges than monitoring individual species. The first challenge is identifying communities. Classifying communities and mapping their locations and boundaries is an attempt to organize and represent the complexity of assemblages found in an area, but boundaries between communities are rarely sharp, and classification is rarely clear-cut. Recognizably similar communities are nevertheless different, and the point at which the differences are great enough to classify as different communities is a subjective decision. Community classifications are also scale dependent, and the selection of scale is subjective. While many methods are available to provide some mathematic basis for classifying and delineating communities (see Box A-1), the final product depends on the best judgment of the ecologist.

A second challenge is that our lack of understanding of community processes makes it difficult to construct ecological models necessary for developing objectives (see Chapter 14). We often are unable to predict how a community as a whole will respond to our management. While predicting response of individual species is, of course, not trivial, communities are much more complex, with individual species often responding to management and interacting with each other in unexpected ways.

The typical response to the first challenge has been placement of studies in areas that are clearly homogeneous communities. This reduces many of the problems encountered when identifying community boundaries and is a reasonable simplification for many, if not most, monitoring situations. A typical response to the second challenge has been the application of standardized methods that monitor all species (or as many as possible) to capture within the monitoring data unexpected and surprising results. This approach is problematic, however, for three reasons:

1. A monitoring design cannot be optimal for all species (Kenkel et al. 1989; Kenkel and Podani 1991). As discussed in Chapters 7 and 8, good sampling design is critical to efficient, cost-effective monitoring. If the intent is to monitor the density of all

Box A-1. *A BRIEF SIMPLE INTRODUCTION TO MULTIVARIATE METHODS IN ECOLOGY*

The methods for multivariate analysis are many, and the literature is voluminous. Here, we introduce the basic concepts in a simplified manner, ignoring the nuances and complexity.

Consider the dataset shown in Figure A-1, where cover data are listed for 20 species from 10 sampling units. Each sampling unit is independently placed within the sampled area. (Note that this is an abbreviated matrix. Typical numbers of species, even in species-poor plant communities, would be more than 20, and in species-rich communities may number in the hundreds. Most studies would also require more than 10 sampling units.) A univariate approach would usually involve the selection of a single species for monitoring (here Penlem). Only cover of that species is recorded. Alternatively, cover of all species might be recorded during data collection but each species would be analyzed separately in a univariate manner. A multivariate approach would evaluate the cover of all species simultaneously.

In preparation for multivariate analysis, data are often displayed in a matrix where the rows are the sampling units and the columns are species (it can also be displayed with the sampling units the columns and the species the rows). Often this primary matrix is accompanied by a secondary matrix of environmental variables. For example, in Figure A-1, two environmental variables were recorded in categories: shade and a soil moisture index (SMI). These environmental variables are used to examine potential causes of the patterns observed in the species compositional data.

It can be helpful to visualize the information in a matrix such as that shown in Figure A-1 as sampling units plotted in species space. For example, consider the first two columns of the species matrix. If Penlem is the x-axis and Brotec is the y-axis, you could plot the sampling units based on the values for each species. Now add Psespi as the z-axis, and you could plot the sampling units in this three-dimensional space. Conceptually, sampling units can be

UNIVARIATE DATA		MULTIVARIATE DATA									
			SPECIES MATRIX							ENVIRONMENTAL MATRIX	
SU	Cover Penlem	SU	Cover Penlem	Cover Brotec	Cover Psespi	Cover Fesida	Cover Crerun	Cover Artwyo	Cover Artvas	Shade	SMI
1	0.5	1	0.5	35	15	5	0.5	15	0	0	0
2	10	2	10	25	10	5	0	5	15	0	0
3	3	3	3	10	20	3	0	45	15	2	2
4	15	4	15	15	15	0.5	0	5	0	0	2
5	5	5	5	15	5	3	0	0	0	1	5
6	20	6	20	20	5	3	0	0	15	1	1
7	15	7	15	5	0.5	5	0.5	35	0	0	1
8	0.5	8	0.5	0	10	10	0	15	10	0	0
9	3	9	3	15	10	15	0	15	35	0	0
10	3	10	3	5	20	5	0	5	55	1	1

Figure A-1. Portion of two datasets derived from the same quadrats. In the first, only the species Penlem is recorded. In the second multivariate dataset, all species encountered are recorded, as well as two environmental variables.

plotted in m-dimensional space, with m as the number of species.[a] *The goal of multivariate methods is to reduce this complexity to a few dimensions, so that it can be interpreted and analyzed.*

Many multivariate methods in ecology can be simplified as follows:

1. *Calculate a similarity or dissimilarity (distance) measure between each sampling unit and each of the other sampling units by comparing each of the species pairs within the sampling unit. The number of similarity and distance measures are many, each with strengths and weaknesses (Faith et al. 1987; Beals 1984; Gower 1985; Gower and Legendre 1986; Legendre and Anderson 1999). See Chambers (1983) for a fairly non-technical discussion in a monitoring context.*

2. *Portray the sampling units graphically with the most similar close together and the least similar far apart (ordination). OR*

3. *Group the sampling units into classes based on their similarity (classification).*

Other multivariate ordination methods use various mathematic approaches to reduce the multidimensional space into fewer dimensions (say two or three) that capture as much of the variability of the many dimensions as possible.

A number of methods exist for both ordination and classification (most of which are much more complex than this simplified explanation). Ecologists can become quite loyal to a particular method, and much of the literature examines the relative strengths and weaknesses of each method. For our examination of multivariate methods in community monitoring, this simple explanation is sufficient. For more information about multivariate techniques, consult Causton (1988), Digby and Kempton (1987), Gauch (1982), Greig-Smith (1983), Kent and Coker (1992), Krebs (1998), and Ludwig and Reynolds (1988) as introductions, and Legendre and Legendre (1983 and 1998), Jongman et al. (1995), and Pielou (1984) as more advanced texts.

---

[a] *Several types of multivariate space can be defined such as sample space (e.g., plot the species in the sample space), environmental space (only two-dimensions in this example), and others (see Gauch 1982 for a complete introduction to the concept).*

species occurring within a community, for example, quadrats large enough to "capture" some individuals of rare or highly patchy species may necessitate counting thousands of individuals of a common ubiquitous species within that community. For many measures of abundance (e.g., density, cover, frequency), a given sampling design will usually oversample some species and undersample other species with respect to the sampling objective (e.g., be 90% confident that the estimation of density is within ±20% of the true mean). Oversampling is a waste of time and personnel resources; undersampling is similarly wasteful because the data provide poor estimates that may not even be usable. Nested sampling designs that include several sampling units of varying size and shape can improve this situation, but require careful evaluation during the pilot stage.

2. Observer variability in a study that attempts to monitor all species is usually very high, unpredictable, and rarely quantified. This variability is a function of variable taxo-

nomic skills and experience of observers and differential detectability of species. Community studies of plants, for example, that require identification of all species are highly susceptible to missed species because observers with broad taxonomic skills are very rare and because plants often lack diagnostic parts such as flowers. Some types of species are also difficult to detect, perhaps because of small size or nondescript morphology, and observer detection levels will be variable. In one study, for example, in samples repeated by the same observer within a few days, the percent of the total plant species found on both occasions was 83% (Hope-Simpson 1940).

Similar comparisons in plant community studies have found this observer error to range up to 25% (Clymo 1980; Hall and Okali 1978; Nilsson and Nilsson 1983; 1985). Similar problems exist in animal community studies, especially for groups that are poorly known taxonomically (such as insects) or that are difficult to detect. In a study of birds, for example, it was found that observer errors resulted in an average increase in the width of the 95% confidence interval of about 40% (Cunningham et al. 1999). Such variability has important implications for a monitoring study, in which changes in an avian community over a typically short duration monitoring study may be comparably small. While observer differences were greatest for small birds foraging low in shrubs, surprisingly, it was also quite high for frequently calling, active birds and distinctive birds, perhaps a function of double counts on the part of inexperienced observers (Cunningham et al. 1999).

3. Community studies in which data are recorded for all species are very expensive in terms of data collection, data management, and data analysis. For example, you may be estimating cover of plant species in small plots distributed along a transect. If you are estimating a single species, each plot can be evaluated quite quickly. If, however, you are estimating cover of all plant species, each plot becomes quite tedious to evaluate. These tediously collected data must then be entered into a data-management system and analyzed.

Because of these problems (and others that we will address below), community monitoring is often less expensive and easier to design, implement, and interpret when using a single variable, a univariate approach. Box A-1 illustrates the differences between univariate and multivariate data. If community monitoring uses a univariate approach, then all of the concepts and most of the methods described in this handbook are applicable.

Sometimes, however, questions concerning community change are multivariate in nature and require the collection of multivariate data. In other situations, collecting data for all species does not exact much of a penalty, creating a cost-effective multivariate dataset. For example, much of the effort expended in collecting birds in a mist net is setting up the collection system; it would be wasteful to not record all species collected.

Unfortunately, an extensive treatment of monitoring communities is beyond the scope of this handbook. In this appendix we will summarize ways univariate and qualitative methods described in this handbook can be applied to monitoring communities. We also briefly describe diversity and multivariate approaches, outlining problems and providing references to the pertinent literature.

UNIVARIATE METHODS OF MONITORING

Indicators

An indicator is simply a surrogate whose characteristics are used as an index of an attribute of interest that is expensive or difficult to measure. Indicators can be species, environmental or habitat factors, or threats. All three types of indicators were discussed in Chapter 1 as they pertain to species monitoring. Issues are similar for community monitoring.

Indicators can be very effective for monitoring community change. They can also fail. The primary reason for failure stems from a weak relationship between the indicator and the attribute of interest. Correspondence between the two may be lacking, with change in the indicator unrelated to change in the attribute, or unexpected, with change in the indicator corresponding well at one level, but poorly at a different level. Predicting the correspondence between an indicator and a community attribute is much more complicated than, for example, between an indicator species and a species of interest. There are uncertainties associated with both situations, but because we often have a very poor understanding of community dynamics, relating those dynamics to an indicator can be very difficult. Selecting an indicator for a specific and local situation in a particular community, however, is less difficult and more likely to be successful than selecting an indicator for a widespread or complex management issue.

When selecting an indicator, first develop general goals for the community, then evaluate potential indicators by considering the following:

1. Is it sensitive to management? Do you know enough about the ecology or temporal behavior of the indicator to relate change to management efforts?
2. Does it correspond to the management goals for the community, those changes in the community that you are trying to influence with your management?
3. Is the change expected in the indicator fairly large, so that the change is easily detected by monitoring (see Chapter 7)?
4. Is the indicator convenient and inexpensive to monitor? Garton (1984) provides a long list of indicators suitable for monitoring natural areas and gives a difficulty and cost rating. Strive to identify the simplest indicator that will be effective in answering the question of whether management was effective.
5. Can different observers with different skill levels consistently evaluate the indicator?

Once an indicator is selected, develop an objective for the indicator following the concepts in Chapter 14. Some examples are as follows:

- An objective for improving stream condition in a grazing allotment: Increase willow (*Salix* spp.) density on point bars along Pratt Creek from the current condition of 5 stems/m^2 to 30 stems/m^2 between 2001 and 2010.
- An objective for an amphibian community: Increase the number of amphibian species in the Cove Creek watershed by 50% between 2002 and 2008.
- An objective for a wet meadow community containing several rare plants and threatened by an exotic plant: Reduce cover of the noxious shrub Russian Olive (*Elaeagnus angustifolia*) in the Bishop Creek meadow by 70% between 2002 and 2005.

Structural Characteristics, Functional Groups, Guilds

Another option for monitoring communities is to select a group of species or habitat characteristics that perform a particular function within the community. These either can be the focus of management such as trying to reduce the density of noxious weeds or can function as an indicator of a community attribute.

Structural characteristics are the building blocks of a habitat, without respect to species (Spies 1998). Changes in structure can often be a sensitive measure of community change, monitored without requiring careful identification of species. Examples of structural characteristics that can be monitored include overall tree cover, percentage of standing dead trees, number of vernal pools, animal grazing or browsing levels, extent of coral reefs, total vegetation cover, total woody or shrub cover, vegetation height, and total biomass. Structural characteristics are usually univariate measures. For example, you can measure the cover of downed woody material or

shrub cover using a line intercept technique (see Chapter 12). You could estimate the density of standing dead trees in quadrats.

Functional groups or guilds are comprised of species that perform similar functions within a community or that have similar requirements of a habitat. Groups of species are nearly limitless, depending on the management situation. Examples include all the birds that eat insects, leaf-eating insects, mammals that depend on tree cavities, noxious weeds, disturbance-related herbaceous plants, shrubs, and grass species highly desirable to grazing animals. For plants, the *Journal of Vegetation Science* (Volume 10, Issue 5) highlighted delineation and use of functional groups in a special feature in 1999. A general discussion of use of guilds for wildlife assessment is provided by Block et al. (1986), for birds by Szaro (1986), and for fisheries by Austen et al. (1994).

One approach to using functional groups in monitoring is to record data by species group (ignoring species designations). For example, you may classify grassland plants into native perennial grasses, annual grasses, noxious weeds, annual herbaceous plants, and perennial herbaceous plants. Monitoring functional groups of birds might focus on long-distance migrants, short-distance migrants, residents, and exotics. When examining sampling units, you would record the occurrence of a member of a group instead of recording by species. If you were using a frequency measure, you would simply record whether a member of that group occurs in the sampling unit. If you were measuring density, you would count all of the individuals of each group.[1]

This approach results in time-savings only when species can be quickly classed into a functional group without identification to species. If you must identify each species in order to place it in the correct functional group, you should record data by species rather than by functional group. Doing so requires little additional time if you must identify the species for correct group classification and the individual species information provides greater interpretive power of changes in the community. You also give yourself the flexibility of grouping species in different ways during analysis.

Even when data are gathered for many (or all) species in a community, it can be useful to aggregate the species data into functional groups for analysis purposes. For example, species presence in plant communities can be assessed with either point-intercept or nested frequency sampling techniques (see Chapter 12). As long as data are recorded separately for each species at each point or quadrat location, the individual species data can be classified into functional guilds such as non-native annual grasses, native annual grasses, non-native perennial grasses, native perennial grasses, etc. This analysis by guild can provide a valuable community-level assessment. When important changes are detected (e.g., an increase in a non-native species guild) the individual species data can be assessed to see which species are contributing to the guild-level change. This two-stage assessment procedure allows a manager to tailor a specific management response (e.g., implement control measures for a particular non-native species).

Another approach to using functional groups is to monitor a single group of species, recording the changes in each species occurring within that group. For example, you may monitor changes in perennial grasses, recording the abundance or occurrence of all species of perennial grasses within each sampling unit, but ignoring other plants such as annual grasses or perennial herbs. As another example, wildlife biologists concerned with forest or wetland fragmentation may chose to focus monitoring on area-sensitive and edge-sensitive species of songbirds.

Monitoring structural characteristics or functional groups can save significant time over monitoring all species. Structure or groups of species may also respond more predictably to management as a group than individual species. For example, if cattle are removed from an area, we can confidently predict an increase in those species that are usually heavily grazed because of their desirability. We may not, however, be able to predict the particular individual response of

---

[1]*This may be problematic because the morphology of some species within a group may not lend themselves to density counts (e.g., rhizomatous and matted plants). Cover and frequency are usually better measures of plant abundances in communities because of morphologic differences.*

each species within that group with as much confidence. It is likely that at least some of those species will respond unexpectedly, perhaps because of other factors such as insects or weather. If we selected one of those species to monitor, interpretation of management effects would be difficult; monitoring the group of species would capture a better representation of the real response of the community.

QUALITATIVE METHODS OF MONITORING

Many of the qualitative methods described in Chapter 4 are applicable to monitoring communities:

1. The site condition assessment described in Chapter 4 is essentially a community monitoring approach. Important aspects of the community that can be monitored by site visits and observations include the presence of nonnative species, human threats, large structural changes such as shrub encroachment into a meadow, and the condition of protective structures such as fences.
2. Boundary mapping can be used to monitor communities at the landscape scale. Communities may be mapped on aerial or satellite imagery and the boundaries ground-truthed. Boundary mapping of communities for monitoring can be a frustrating exercise, however, if boundaries between communities are gradual and indistinct.
3. Photoplots can be used to monitor changes in plant composition and structure within communities. Refer to the discussion of uses and problems with photoplots described in Chapter 4.
4. Photopoints can provide a visual record of gross community changes.
5. Aerial photographs and satellite imagery can be used to monitor some structural features within a community (e.g., woody species cover) and some threats (e.g., off-highway vehicle use). Chapter 4 discusses this in more detail.
6. Checklists of species may provide a gross measure of community change (Droege et al. 1998) that may be adequate for management actions.

INDICES FOR MONITORING COMMUNITIES

The simplest index is simply species richness: the number of species that occur within a community. Clearly, however, a community containing one dominant species and 99 rare species is quite different from a community containing 100 species of similar abundance. To differentiate, the concept of "evenness" or the relative abundances of the species is often incorporated with species richness into a diversity index. The number of diversity indices used in ecology are many, and each has proponents and detractors (Hurlbert 1971; Peet 1974; Solow et al. 1993; Washington 1984). For an extensive introduction to diversity indices, see Magurran (1988), Pielou (1975), and Grassle et al. (1979); for a shorter introduction see chapters in Krebs (1998) and Ludwig and Reynolds (1988).

Using diversity indices in monitoring community change has several problems. A fundamental problem is that use of diversity indices requires recording all encountered species, resulting in the three problems described earlier in the introduction. Problems specific to diversity indices and species richness estimates include the following:

1. Change in diversity indices or species richness is difficult to predict and relate to management. After decades of exploring the relationships between diversity and

community functions, many questions remain (Tilman 1999; Schlâpfer and Schmid 1999).

2. Indices assume that all species within a community are known (or can be accurately estimated), but this is extremely unlikely, especially in species-rich communities.
3. Diversity indices use a measure of abundance such as density, cover, or biomass. Different measures will reflect different kinds of changes in the community (Chiarucci et al. 1999; Cousins 1991; Qinfeng and Rundel 1997).
4. Different diversity indices give different results when applied to the same datasets. This is clearly a problem in monitoring where you wish to detect change, and the existence of change depends on the choice of the diversity index.
5. Diversity indices and species counts (richness) are often treated as normal data and are analyzed using standard statistical methods such as ANOVA, *t*-tests, and confidence intervals (see Chapter 9).[2] In many cases, however, the assumption of normality is often violated; this problem is especially problematic with small sample sizes. The statistical behavior of richness and diversity indices is unpredictable, meaning that standard statistical distributions (such as the normal or binomial distribution) are often not followed. Heltshe and Forrester (1983) present a method based on resampling (see Chapter 9) to estimate sampling error for species richness measured in a sample of quadrats, and Heltshe and Diconzio (1985) illustrate the method on simulated plant communities. Gove et al. (1996) use a resampling method to compare diversity indices of two treatment areas. Hellmann and Fowler (1999) compare four resampling methods for estimating the precision and accuracy of species-richness measures. Additional sources of information on methods for analyzing richness and diversity measures are Bunge and Fitzpatrick (1993); Dixon (1993); Grassle et al. (1979); Palmer (1990, 1991); Smith and van Belle (1984); and Solow (1993, 1994).

As an alternative to standard diversity and richness indices you can develop an index specifically for the monitoring situation. The most common approach is the multimetric Index of Biotic Integrity (IBI). Metrics are attributes known by research within a particular system or hypothesized based on ecological principles to be reliably correlated with disturbance. The original IBI, developed to assess warm-water streams in the central United States (Karr 1991), rated 12 metrics (e.g., species richness, number of sensitive species, number of species representing different trophic levels, presence of disease or deformity). Five points were assigned for each metric when conditions were similar to a system with little human influence, one point for a system strongly affected by human activity, and three points for intermediate conditions. The total IBI index score was the sum of the 12 metrics, ranging from 12 (worst) to 60 (best). All metrics in the original version of the IBI were based on the characteristics of the fish assemblage and thus could be rated from a single sample of fish (i.e., none of the metrics assessed, for example, macroinvertebrates or water quality).

Since the original metric was proposed, a number of metrics have been developed and tested in field assessments of benthic macroinvertebrates and fish communities (Karr 1991; Simon and Lyons 1995). Proponents argue that multimetric indices of biologic integrity have been found to better indicate human influence on these communities than principle components analysis of species composition[3] (Fore et al. 1996) or using single species indicators (Karr and

---

[2]*Krebs (1998) presents a modified* t-test *for the difference in two samples for which the Shannon index of diversity has been calculated based on Hutcheson (1970).*

[3]*Principal components analysis is a common ordination technique—see below.*

Chu 1998). While most of the use of the IBI has been in assessment of aquatic systems, the method has recently begun to be applied to terrestrial animals (Moyle and Randall 1998). Considerations in developing a monitoring project using a multimetric approach include the following:

- Select metrics that are known to correlate with the management change of interest.
- Select metrics that are inexpensive and easy to measure consistently.
- Select metrics that represent a range of biologic organizational levels such as conditions of individual species, occurrence and abundance of indicator species, composition and occurrence of functional groups, and community characteristics such as richness and trophic levels.
- Exercise caution in analysis; statistical behavior of multimetric indices may be unpredictable, similar to the problems described for diversity indices. While IBI indices may sometimes be appropriately analyzed by standard analysis techniques such as ANOVA or *t*-tests (Fore et al. 1994), the assumptions required by these tests may be violated by index data. In such cases, resampling techniques such as those used for diversity indices would be appropriate.

A conservation index for plants was proposed by Wilhelm and Ladd (1988), but has not been widely applied (see Wilhelm and Master 1994; Swink and Wilhelm 1994). Each species is assigned a "coefficient of conservatism" (CC) ranging from 0 (typical of species from highly disturbed habitats) to 10 (for narrow endemics of pristine habitats). A Natural Quality Index (also called Floristic Quality Index) is calculated:

$$\left(\frac{\sum CC}{N}\right)\sqrt{N}$$

where:
CC = the total coefficient of conservatism values for all species within the community
N = the number species within the community

Note that the method is based on simple presence and absence of species; no abundance information is required. It does, however, require a good knowledge of the ecological characteristics of all species within a local flora in order to assign meaningful coefficients of conservatism. When used to compare sites, the index is scale-dependent, sensitive to inclusions of small unique sites with high species diversity as well as the larger numbers of species typical of larger sites. For monitoring, however, where the sampled area does not change, the index may be sensitive to changes in species occurrence. Initial tests of the index suggests that it may follow a normal distribution and be suitable for standard analysis techniques described in Chapter 9 (Salzer, unpublished data).

MULTIVARIATE MONITORING OF COMMUNITIES

Multivariate data may result from the sampling method used for monitoring. For example, many animal studies involve capture of individuals to estimate population sizes (e.g., traps for small rodents or mist nets for birds). These capture techniques usually do not discriminate among species. Monitoring all species in some studies may not cost much more than monitoring one species such as point counts for calling birds in which there is little penalty for recording all birds heard compared with listening for a single species. For other studies, however, collecting data on all species dramatically increases the cost of data collection and management. For example, recording all plant species in a quadrat requires much more time than recording the abundance of only one species. This cost must be weighed against the information gained. Be careful not to

create a multivariate dataset when a simpler, univariate dataset will be adequate for making management decisions. If, however, data can be collected on many species with little additional cost, how should they be analyzed?

The simplest method for analyzing these data is to simply treat them as samples of each species and to analyze each species separately using standard techniques (see Chapter 9; Chapter 13 for capture studies of animals). For example, if you encountered 10 species in your monitoring study, you would analyze each of the 10 species separately.[4] Because it is difficult to design a sampling unit that functions well for all species occurring within a community, this approach will probably only be useful for the more common species. Rare species may only occur in a few sampling units, and you will be unable to get a reasonably precise estimate of their abundance (see Chapter 7). Nested sampling designs can be used to address this problem.

Another question that can be answered with data on all species is the change in the proportion of the total represented by a particular species. For example, if you captured 100 insects in a light trap and 5 were Species A, and the following year you captured 45 insects and 6 were Species A, you could test the change in the proportion of Species A in the total number of insects captured using a chi-square analysis (see Chapter 9).[5] You could also examine more than one species by constructing a contingency table (see Chapter 9) of size $2 \times c$, where the 2 rows represent the 2 years being compared and where c is the number of columns corresponding to the number of species you wish to compare. The number of species that can be considered is constrained by the small quantities of the rare species. These may need to be combined into a single class (see Snedocor and Cochran, Chapter 9).

Both of these approaches are applications of univariate analysis methods to species community data. Some monitoring questions, however, are truly multivariate in nature. For example, you may wish to compare current community composition to an idealized potential natural community, or you may want to know whether the community as a whole has changed. These types of questions are multivariate in nature and must be analyzed by multivariate methods. An extensive description of application of multivariate methods to monitoring is beyond the scope of this handbook, but the introduction in Box A-1 provides enough background for this discussion.

The first step is to define and measure the composition of species within a community. For animals, composition is usually based on relative numbers or density, but occasionally may be based on biomass. For plants, composition is usually measured by cover or biomass. While biomass is sometimes considered to be the ideal measure of plant community composition (Wilson 1991), it can be quite difficult to accurately measure even for a single species, much less for all the species within a community (see Chapter 12). In most cases, cover is a reasonable approximation of composition and is much easier to estimate. Cover also has the advantage that it can be applied to any species morphology. Of the various methods of estimating cover discussed in Chapter 12 (line intercept, point intercept, visual estimation), point intercept is the most easily and efficiently applied to all plant types.[6] Presence/absence data can also be analyzed by some multivariate methods.

---

[4] *Two concepts are of note in such a design. First, recognize that some species will show significant change simply because of chance. If, for example, you set a false-change error rate of 10%, approximately 10% of the species comparisons that you conduct may show significant change. Second, note that data should not be relativized or standardized (two common approaches in community multivariate analysis). For example, if you have measured species density, do not express that density in terms of a percentage of the total density of all species within a plot (relativized by plot total).*

[5] *This analysis assumes that the insects move independently, so each insect captured is an independent observation. This would not work for animals that travel in groups.*

[6] *While shrub species are often measured by line-intercept techniques, point intercepts (which can be projected upward to intercept trees and tall shrubs, or down to intercept shorter plants) are equally applicable for all types of plants. You will need to decide whether you will measure only the canopy intercept (the first species intercepted by a point), all species intercepted by a point, or all intercepts on vegetation (which can include multiple layers of the same plant). Although the last probably corresponds most closely to biomass, it is time-consuming and difficult to measure in the field and difficult to visualize. Measuring*

The next step is to compare the composition measured at $Time_1$ with that measured at $Time_2$.[7] The statistical question is whether the assemblages are more different than could be expected as a result of random chance sampling errors, at a given probability. Many of the methods used to answer that question are based on the use of similarity or dissimilarity (distance) measures between sampling units (Box A-1). Graphically, they can be visualized by imagining a cloud of points on the ordination diagram with the distances between points representing the distance calculated in the dissimilarity measure. Are the centers of the clouds of points from $Time_1$ and $Time_2$ separated by an acceptable distance ("acceptable" is a function of the probability level), or are the centers of the clouds close together with points in the clouds overlapping? Several methods have been proposed to evaluate this temporal difference, all of which require rather advanced statistical procedures (Anderson and Gribble 1998; Ault and Johnson 1998; Belbin 1992; Clarke 1993; Dietz 1983; Legendre and Anderson 1999; Smith et al. 1990; and Smith 1998).[8]

The need for familiarity and expertise in multivariate analysis is only one of the problems with using a multivariate approach in monitoring. As with other methods that require complete species assessment, expense and observer variability limit the use of these methods for monitoring. Other factors that should be considered when using multivariate approaches to monitoring communities include the following:

1. Different measures of abundance (e.g., density, cover, frequency, biomass) produce different results in analysis (Chiarucci et al. 1999). This is not surprising. We discussed in Chapter 12 how different vegetation measures provide a different picture of the change in vegetation.
2. Developing an ecological model to predict multivariate change in community composition is extremely difficult. How will you describe the direction and amount of change to include in your management objective?
3. Determining what is a significant level of change is also difficult. How will you determine how much you want the community to change and level of significance in your sampling objective?
4. Sampling design for multivariate monitoring is nearly completely unexplored terrain (see Kenkel et al. 1989; Kenkel and Podani 1991). You can imagine that if your sample each year is variable because of spatial variability (i.e., sampling units are compositionally different from each other), then the clouds of sampling units described above will be large and diffuse and detecting change between years difficult. Effective sampling design minimizes the compositional variability between sampling units. Many of the concepts in Chapter 8 are applicable (e.g., using large rectangular sampling units that incorporate as much of the variability within the

---

only the canopy will result in an underestimation of cover of all understory species (imagine a forest community). For these reasons, we recommend recording each species when first intercepted by a particular point.

These point intercepts must be grouped by transects or plots or some other method to create the sampling units used in the multivariate matrix. Individual points should not generally be used as the sampling unit.

[7] *You can also compare sites.*

[8] *A simple, nonstatistical approach for several years of data from permanent plots is to conduct an ordination of the plots and track the movement of each plot graphically on an ordination diagram. You could create a dummy plot (or several showing some variability) based on the theoretic composition of the potential natural community and observe whether the movement of the sample plots was toward the composition of the dummy plots.*

Parametric multivariate analysis of variance (MANOVA) has been used to analyze community change (Ratliff and Mori 1993; Stroup and Stubbendieck 1983). This approach is conceptually simpler, is available on many standard statistical packages, and produces an output similar to the familiar ANOVA (see Chapter 9). While some multivariate datasets may be appropriately analyzed by a MANOVA, datasets with many rare species (many zeros in the data matrix) will violate the assumption of normality required by this test (Legendre and Legendre 1983; Legendre and Anderson 1999). Datasets with more species than sampling units are also problematic (Legendre and Anderson 1999).

units as possible and minimizing that between sampling units), but design is much more difficult than for the single-species case.

Relating changes in community composition to management is difficult. What does it mean when a community has been found to change significantly? Usually, interpretation of change will be based on an examination of the change in individual species or groups of species. These species will often be the ones about which we know the most ecologically and whose change can be related to management actions. The question then is—why not simply monitor those species as indicators in the first place, rather than measuring the response of all species in the community? It must be recognized that all monitoring is a simplification of the system, a reduction of the complexity to a manageable question or objective. For example, we can measure the density of a particular plant species, but doing so misses changes in biomass, height, cover, and spatial pattern—any or all of which may be "important." We select the best attribute of the species that is measurable and that represents changes we believe we can interpret. Similarly, in community monitoring, we must accept that we cannot measure everything and focus on selecting a measurable attribute for monitoring that we can interpret as a basis for management decisions.

SUMMARY

Monitoring communities can be challenging because they can be difficult to identify and classify and because we often lack a clear understanding of community-level processes. It is difficult to develop efficient community sampling designs because no design works optimally for many different species or indicator variables. Efforts to track many separate community attributes often require high levels of observer expertise or extensive observer training. This type of community monitoring is often expensive to implement. The monitoring of communities can be simplified by using qualitative monitoring techniques, by focusing on indicator species or abiotic indicator variables, by directly assessing threats, by tracking structural characteristics, or by tracking functional groups of species. Species diversity indices and multivariate techniques are available to evaluate multi-species community monitoring datasets, but there are problems in analyzing and interpreting the results from these studies. In many cases, the most effective community monitoring approach will involve tracking a small number of measurable attributes that have a clear linkage to identified management concerns.

REFERENCES

Anderson, M.J., Gribble, N.A. 1998. Partitioning out the variation among spatial, temporal and environmental components in a multivariate dataset. Australian Journal of Ecology 23:158–167.

Ault, T.R., Johnson, C.R. 1998. Spatially and temporally predictable fish communities on coral reefs. Ecological Monographs 68:25–50.

Austen, D.J., Bayley, P.B., Menzel, B.W. 1994. Importance of the guild concept to fisheries research and management. Fisheries 19:12–18.

Beals, E.W. 1984. Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. Advances in Ecological Research 14:1–55.

Belbin, L. 1992. Comparing two sets of community data: a method for testing reserve adequacy. Australian Journal of Ecology 17: 255–262.

Block, W.M., Brennan, L.A., Gutierrez, R.J. 1986. The use of guilds and guild-indicator species for assessing habitat suitability. In: J. Verner, M.L. Morrison, C.J. Ralph (eds.), Wildlife 2000: modeling habitat relationships of terrestrial vertebrates (pp. 109–113). Madison, WI: University of Wisconsin Press.

Bunge, J., Fitzpatrick, M. 1993. Estimating the number of species: a review. Journal of the American Statistical Association 88:364–373.

Causton, D.R. 1988. An introduction to vegetation analysis. London: Unwin Hyman.

Chambers, J. C. 1983. Measuring species diversity of revegetated surface mines: an evaluation of techniques (p. 15). Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. (Research Paper, v. INT-322).

Chiarucci, A., Bastow, W.J., Anderson, B.J., De Dominicis, V. 1999. Cover versus biomass as an estimate of species abundance: does it make a difference to the conclusions? Journal of Vegetation Science 10:35–42.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18:117–143.

Clymo, R.S. 1980. Preliminary survey of the peat-bog Knowe Moss using various numerical methods. Vegetatio 42:129–148.

Cousins, S.H. 1991. Species diversity measurements: choosing the right index. Trends in Ecology and Evolution 6:190–192.

Cunningham, R.B., Lindenmayer, D.B., Nix, H.A., Lindenmayer, B.D. 1999. Quantifying observer heterogeneity in bird counts. Australian Journal of Ecology 24:270–277.

Dietz, E.J. 1983. Permutation tests for association between two distance matrices. Systematic Zoology 32:21–26.

Digby, P.G.N., Kempton, R.A. 1987. Population and community biology series: multivariate analysis of ecological communities. London: Chapman & Hall.

Dixon, P.M. 1993. The bootstrap and the jackknife: describing the precision of ecological indices. In: S.M. Schneiner, J. Gurevitch (eds.), Design and analysis of ecological experiments. New York: Chapman & Hall.

Droege, S., Cyr, A., Larivée, J. 1998. Checklists: an under-used tool for the inventory and monitoring of plants and animals. Conservation Biology 12:1134–1138.

Faith, D.P., Minchin, P.R., Belbin, L. 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69:57–68.

Fore, L.S., Karr, J.R., Conquest, L.L. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. Canadian Journal of Fisheries and Aquatic Sciences 51:1017–1087.

Fore, L.S., Karr, J.R., Wisseman, R.W. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212–231.

Garton, E.O. 1984. Cost-effective baseline inventories of research natural areas. In: J.L. Johnson, J.R. Franklin, R.G. Krebill (coordinators). Research natural areas: baseline monitoring and management: proceedings of a symposium (pp. 40–45), 1984 March 21, Missoula, MT. General Technical Report INT-173. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Gauch, H.G., Jr. 1982. Multivariate analysis in community structure. Cambridge, England: Cambridge University Press.

Gove, J.H., Banapati, P.P., Taillie, C. 1996. Diversity measurement and comparison with examples. In: R.C. Szaro, D.W. Johnson (eds.), Biodiversity in managed landscapes (pp. 157–175). New York: Oxford University Press.

Gower, J.C. 1985. Measures of similarity, dissimilarity and distance. In: N.L. Johnson, S. Kotz, C.B. Read (eds.), Encyclopedia of statistics, Vol. 5 (pp. 397–405). New York: John Wiley & Sons.

Gower, J.C., Legendre, P. 1986. Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification 3:5–48.

Grassle, J.F., Patil, G.P., Smith, W.K., Taillie, C. (eds.). 1979. Ecological diversity in theory and practice. In: G.P. Patil (ed.), Statistical ecology series, Vol. 6. Burtonsville, MD: International Cooperative Publishing House.

Greig-Smith, P. 1983. Quantitative plant ecology (3rd ed.). Berkeley, CA: University of California Press.

Hall, J.B., Okali, D.U.U. 1978. Observer bias in a floristic survey of complex tropical vegetation. Journal of Ecology 66:241–249.

Hellman, J.J., Fowler, G.W. 1999. Bias, precision, and accuracy of four measures of species richness. Ecological Applications 9:824–834.

Heltshe, J.F., DiConzio, J. 1985. Power study of jack-knifed diversity indices to detect change. Journal of Environmental Management 21:331–341.

Heltshe, J.F., Forrester, N.E. 1983. Estimating species richness using the jackknife procedure. Biometrics 39:1–11.

Hope-Simpson, J.F. 1940. On the errors in the ordinary use of subjective frequency estimations in grasslands. Journal of Ecology 28:193–209.

Hurlbert, S.H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586.

Hutcheson, K. 1970. A test for comparing diversities based on the Shannon formula. Journal of Theoretical Biology 29:151–154.

Jongman, R.H.G., ter Braak, C.J.F., van Tongeren, O.F.R. (eds.). 1995. Data analysis in community and landscape ecology. Cambridge, England: Cambridge University Press.

Karr, J.R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications 1:66–84.

Karr, J.R., Chu, E.W. 1998. Restoring life in running waters. Washington, DC: Island Press.

Karr, J.R., Dudley, D.R. 1981. Ecological perspectives on water quality goals. Environmental Management 5:55–68.

Kenkel, N.C., Juhasz-Nagy, P., Podani, J. 1989. On sampling procedures in population and community ecology. Vegetatio 83:195–207.

Kenkel, N.C., Podani, J. 1991. Plot size and estimation efficiency in plant community studies. Journal of Vegetation Science 2:539–544.

Kent, M., Coker, P. 1992. Vegetation description and analysis: a practical approach. London: Belhaven Press.

Krebs, C.J. 1998. Ecological methodology, 2nd ed. New York: Harper & Row.

Legendre, P., Anderson, M.J. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecological Monographs 69:1–24.

Legendre, P., Legendre, L. 1983. Numerical ecology. 1st English edition. Amsterdam: Elsevier.

Legendre, P., Legendre, L. 1998. Numerical ecology. 2nd English edition. Amsterdam: Elsevier.

Ludwig, J.A., Reynolds, J.F. 1988. Statistical ecology: a primer on methods and computing. New York: John Wiley & Sons.

Magurran, A.E. 1988. Ecological diversity and its measurement. Princeton, NJ: Princeton University Press.

Moyle, P.B., Randall, P.J. 1998. Evaluating the biological integrity of watersheds in the Sierra Nevada, California. Conservation Biology 12:1318–1326.

Nilsson, I.N., Nilsson, S.G. 1985. Experimental estimates of census efficiency and pseudoturnover on islands: error trend and between-observer variation when recording vascular plants. Journal of Ecology 73:65–70.

Nilsson, S.G., Nilsson, I.N. 1983. Are estimated species turnover rates on islands largely sampling errors? American Naturalist 121:595–597.

Palmer, M.W. 1990. The estimation of species richness by extrapolation. Ecology 71:1195-1198.

Palmer, M.W. 1991. Estimating species richness: the second order jackknife reconsidered. Ecology 72: 1512-1513.

Peet, R.K. 1974. The measurement of species diversity. Annual Review of Ecology and Systematics 5:285–307.

Pielou, E.C. 1975. Ecological diversity. New York: John Wiley & Sons.

Pielou, E.C. 1984. The interpretation of ecological data: a primer on classification and ordination. New York: John Wiley & Sons.

Qinfeng, G., Rundel, P.W. 1997. Measuring dominance and diversity in ecological communities: choosing the right variables. Journal of Vegetation Science 8:405–408.

Ratliff, R.D., Mori, S.R. 1993. Squared Euclidean distance: a statistical test to evaluate plant community change. Research Note. PSW-RN-416. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station.

Schlâpfer, F., Schmid, B. 1999. Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. Ecological Applications 9:893–912.

Simon, T.P., Lyons, J. 1995. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. In: W.S. Davis, T.P. Simon (eds.), Biological assessment and criteria—tools for water resource planning and decision making. Boca Raton, FL: CRC Press.

Smith, E.P. 1998. Randomization methods and the analysis of multivariate ecological data. Environmentrics 9:37–51.

Smith, E.P., Pontasch, K.W., Cairns, J. 1990. Community similarity and the analysis of multispecies environmental data: a unified statistical approach. Water Research 24:507–514.

Smith, E.P., van Belle, G. 1984. Nonparametric estimation of species richness. Biometrics 40:119–129.

Solow, A R. 1993. A simple test for change in community structure. Ecology 62:191–193.

Solow, A.R. 1994. On the Baysian estimation of the number of species in a community. Ecology 75: 2139–2142.

Solow, A.R., Broadus, J.M., Tonring, N. 1993. On the measurement of biological diversity. Journal of Environmental and Economic Management 24:60–68.

Spies, T.A. 1998. Forest structure: a key to the ecosystem. Northwest Science 72:34–39.

Stroup, W.W., Stubbendieck, J. 1983. Multivariate statistical methods to determine changes in botanical composition. Journal of Range Management 36:208–212.

Swink, F., Wilhelm, G. 1994. Floristic quality assessment. In Plants of the Chicago Region, 4th ed. Indianapolis: Indiana Academy of Science.

Szaro, R.C. 1986. Guild management: an evaluation of avian guilds as a predictive tool. Environmental Management 10:681–688.

Tilman, D. 1999. The ecological consequences of changes in biodiversity: a search for general principles. Ecology 80:1455–1474.

Washington, H.G. 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. Water Resources 18:653–694.

Wilhelm, G., Ladd, D. 1988. Natural area assessment in the Chicago Region. Transactions of the North America Wildlife and Natural Resource Conference 53:361–375.

Wilhelm, G., Master, L. 1994. Floristic changes after five growing seasons in burned and unburned woodland. Erigenia 13:141-150.

Wilson, J.B. 1991. Methods for fitting dominance/diversity curves. Journal of Vegetation Science 2:35–46.

APPENDIX II
Sample Size Equations

Five different sample size equations are presented in this appendix for the following situations:

Equation #1: Determining the necessary sample size for estimating a single population mean or a single population total with a specified level of precision 300

Equation #2: Determining the necessary sample size for detecting differences between two means with temporary sampling units 305

Equation #3: Determining the necessary sample size for detecting differences between two means when using paired or permanent sampling units 307

Equation #4: Determining the necessary sample size for estimating a single population proportion with a specified level of precision 310

Equation #5: Determining the necessary sample size for detecting differences between two proportions with temporary sampling units 312

Equation #6: Determining the necessary sample size for detecting differences between two proportions with permanent sampling units 314

Each separate section is designed to stand alone from the others. Each section includes the sample size equation, a description of each term in the equation, a table of appropriate coefficients, and a worked out example based on a stated management and sampling objective.

The examples included in this appendix all refer to monitoring with a quadrat-based sampling procedure. The equations and calculations also work with other kinds of monitoring data such as measurements of plant height, number of flowers, or measures of cover.

The examples of management objectives included in this appendix for detecting changes between two means or two proportions could be evaluated with one-tailed significance tests (Chapter 9). The sampling objectives and worked-out examples show calculations for two-tailed significance tests. This implies an interest in being able to detect either *increases* or *decreases* over time, even though the management objectives specify a desire to achieve a change in only one direction or the other. If you are only interested in detecting changes in one direction, and you only plan on analyzing your monitoring results with one directional null hypotheses (e.g., H_o = density has not increased), then you should apply a sample modification to the simple size procedures. To change any sample size procedure to a one-tailed situation, simply double the false-change (Type I) error rate (α) and look up the new doubled-α value in the table of coefficients (e.g., use $\alpha = 0.20$ instead of $\alpha = 0.10$ for a one-tailed test with a false-change (Type I) error rate of $\alpha = 0.10$).

The coefficients used in all of the equations are from a standard normal distribution (Z_α and Z_β) instead of the t-distribution (t_α and t_β). These two distributions are nearly identical at large sample sizes but at small sample sizes ($n < 30$) the Z coefficients will slightly underestimate the number of sampling units needed. The correction procedure described for Equation #1 (using the

sample size correction table) already adjusts the sample size using the appropriate t-value. For the other equations, t_α and t_β values can be obtained from a t-table and used in place of the Z_α and Z_β coefficients that are included with the sample size equations. The appropriate t_α-coefficient for the false-change (Type I) error rate can be taken directly from the $\alpha(2)$ column of a t-table at the appropriate degrees of freedom (v). For example, for a false-change error rate of 0.10 use the $\alpha(2) = 0.10$ column. The appropriate t_β coefficient for a specified missed-change error level can be looked up by calculating 2(1-power) and looking up that value in the appropriate $\alpha(2)$ column. For example, for a power of 0.90, the calculations for t_β would be 2(1-.90) = 0.20. Use the $\alpha(2) = 0.20$ column at the appropriate degrees of freedom (v) to obtain the appropriate t_β value.

SAMPLE SIZE EQUATION #1: DETERMINING THE NECESSARY SAMPLE SIZE FOR ESTIMATING A SINGLE POPULATION MEAN OR A POPULATION TOTAL WITH A SPECIFIED LEVEL OF PRECISION.

Estimating a sample mean vs. total population size. The sample size needed to estimate confidence intervals that are within a given percentage of the estimated total population size is the same as the sample size needed to estimate confidence intervals that are within that percentage of the estimated mean value. The instructions below assume you are working with a sample mean.

Determining sample size for a single population mean or a single population total is a two- or three-step process.

(1) The first step is to use the equation provided below to calculate an uncorrected sample size estimate.

(2) The second step is to consult the Sample Size Correction Table (Table 1) appearing on pages 303–304 of these instructions to come up with the corrected sample size estimate. The use of the correction table is necessary because the equation below under-estimates the number of sampling units that will be needed to meet the specified level of precision. The use of the table to correct the underestimated sample size is simpler than using a more complex equation that does not require correction.

(3) The third step is to multiply the corrected sample size estimate by the finite population correction factor if more than 5% of the population area is being sampled.

1. Calculate an initial sample size using the following equation:

$$n = \frac{(Z_\alpha)^2(s)^2}{(B)^2}$$

Where:

n = The uncorrected sample size estimate.

Z_α = The standard normal coefficient from the table below.

s = The standard deviation.

B = The desired precision level expressed as half of the maximum acceptable confidence interval width. This needs to be specified in absolute terms rather than as a percentage. For example, if you wanted your confidence interval width to be within 30% of your sample mean and your sample mean = 10 plants/quadrat then B = (0.30 x 10) = 3.0.

Table of standard normal deviates (Z_α) for various confidence levels

Confidence level	Alpha (α) level	(Z_α)
80%	0.20	1.28
90%	0.10	1.64
95%	0.05	1.96
99%	0.01	2.58

2. To obtain the adjusted sample size estimate, consult Table 1 on page 303–304 of these instructions.

n = the uncorrected sample size value from the sample size equation.
n* = the corrected sample size value.

3. Additional correction for sampling finite populations.

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. If you are sampling more than 5% of the whole population then you should apply a correction to the sample size estimate that incorporates the finite population correction (FPC) factor. This will reduce the sample size.

The formula for correcting the sample size estimate with the FPC for confidence intervals is:

$$n' = \frac{n^*}{(1 + (n^*/N))}$$

Where:

n' = The new FPC-corrected sample size.
n^* = The corrected sample size from the sample size correction table (Table 1).
N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of one quadrat.

Example:

Management objective:

Restore the population of species Y in population Z to a density of at least 30 plants/quadrat by the year 2001.

Sampling objective:

Obtain estimates of the mean density and population size with 95% confidence intervals that are within 20% of the estimated true value.

Results of pilot sampling:

Mean ($\bar{x}$) = 25 plants/quadrat.
Standard deviation (s) = 7 plants.

Given:

The desired **confidence level** is 95% so the appropriate Z_α from the table above = 1.96.

The desired **confidence interval width** is 20% (0.20) of the estimated true value. Since the estimated true value is 25 plants/quadrat, the desired confidence interval (**B**) = 25 x 0.20 = 5 plants/quadrat.

Calculate an unadjusted estimate of the sample size needed by using the sample size formula:

$$n = \frac{(Z_\alpha)^2(s)^2}{(B)^2} \quad n = \frac{(1.96)^2(7)^2}{(5)^2} = 7.5$$

Round 7.5 plots up to 8 plots for the unadjusted sample size.

To adjust this preliminary estimate, go to Table 1 on pages 303–304 of these instructions and find n = 8 and the corresponding n* value in the 95% confidence level portion of the table. For n = 8, the corresponding n* value = 15.

The corrected estimated sample size needed to be 95% confident that the estimate of the population mean is within 20% (± 5 plants) of the true mean = **15 quadrats**.

If the pilot data described above was gathered using a 1m x 10m ($10m^2$) quadrat and the total population being sampled was located within a 20m x 50m macroplot ($1000m^2$) then N = $1000m^2/10m^2$ = 100. The corrected sample size would then be:

$$n' = \frac{n^*}{(1 + (n^*/N))} \quad n' = \frac{15}{(1 + (15/100))} = 13.0$$

The new, FPC-corrected, estimated sample size to be 95% confident that the estimate of the population mean is within 20% (± 5 plants) of the true mean = **13 quadrats**.

Sample size correction table for single parameter estimates, Part 1

80% confidence level						90% confidence level					
n	n*	n	n*	n	n*	n	n*	n	n*	n	n*
1	5	51	65	101	120	1	5	51	65	101	120
2	6	52	66	102	121	2	6	52	66	102	122
3	7	53	67	103	122	3	8	53	67	103	123
4	9	54	68	104	123	4	9	54	69	104	124
5	10	55	69	105	124	5	11	55	70	105	125
6	11	56	70	106	125	6	12	56	71	106	126
7	13	57	71	107	126	7	13	57	72	107	127
8	14	58	73	108	128	8	15	58	73	108	128
9	15	59	74	109	129	9	16	59	74	109	129
10	17	60	75	110	130	10	17	60	75	110	130
11	18	61	76	111	131	11	18	61	76	111	131
12	19	62	77	112	132	12	20	62	78	112	132
13	20	63	78	113	133	13	21	63	79	113	134
14	22	64	79	114	134	14	22	64	80	114	135
15	23	65	80	115	135	15	23	65	81	115	136
16	24	66	82	116	136	16	25	66	82	116	137
17	25	67	83	117	137	17	26	67	83	117	138
18	27	68	84	118	138	18	27	68	84	118	139
19	28	69	85	119	140	19	28	69	85	119	140
20	29	70	86	120	141	20	29	70	86	120	141
21	30	71	87	121	142	21	31	71	88	121	142
22	31	72	88	122	143	22	32	72	89	122	143
23	33	73	89	123	144	23	33	73	90	123	144
24	34	74	90	124	145	24	34	74	91	124	145
25	35	75	91	125	146	25	35	75	92	125	147
26	36	76	93	126	147	26	37	76	93	126	148
27	37	77	94	127	148	27	38	77	94	127	149
28	38	78	95	128	149	28	39	78	95	128	150
29	40	79	96	129	150	29	40	79	96	129	151
30	41	80	97	130	151	30	41	80	97	130	152
31	42	81	98	131	152	31	42	81	99	131	153
32	43	82	99	132	154	32	44	82	100	132	154
33	44	83	100	133	155	33	45	83	101	133	155
34	45	84	101	134	156	34	46	84	102	134	156
35	47	85	102	135	157	35	47	85	103	135	157
36	48	86	104	136	158	36	48	86	104	136	158
37	49	87	105	137	159	37	49	87	105	137	159
38	50	88	106	138	160	38	50	88	106	138	161
39	51	89	107	139	161	39	52	89	107	139	162
40	52	90	108	140	162	40	53	90	108	140	163
41	53	91	109	141	163	41	54	91	110	141	164
42	55	92	110	142	164	42	55	92	111	142	165
43	56	93	111	143	165	43	56	93	112	143	166
44	57	94	112	144	166	44	57	94	113	144	167
45	58	95	113	145	168	45	58	95	114	145	168
46	59	96	115	146	169	46	60	96	115	146	169
47	60	97	116	147	170	47	61	97	116	147	170
48	61	98	117	148	171	48	62	98	117	148	171
49	62	99	118	149	172	49	63	99	118	149	172
50	64	100	119	150	173	50	64	100	119	150	173

APPENDIX I—TABLE 1. Sample size correction table for adjusting "point-in-time" parameter estimates. n = the uncorrected sample size value from the sample size equation. n* = the corrected sample size value. This table was created using the algorithm reported by Kupper and Hafner (1989) for a one-sample tolerance probability of 0.90. For more information consult Kupper and Hafner (1989).

Sample size correction table for single parameter estimates, Part 1

80% confidence level						90% confidence level					
n	n*	n	n*	n	n*	n	n*	n	n*	n	n*
1	5	51	66	101	121	1	6	51	67	101	122
2	7	52	67	102	122	2	8	52	68	102	123
3	8	53	68	103	123	3	9	53	69	103	124
4	10	54	69	104	124	4	11	54	70	104	126
5	11	55	70	105	125	5	12	55	72	105	127
6	12	56	71	106	126	6	14	56	73	106	128
7	14	57	72	107	128	7	15	57	74	107	129
8	15	58	74	108	129	8	16	58	75	108	130
9	16	59	75	109	130	9	18	59	76	109	131
10	18	60	76	110	131	10	19	60	77	110	132
11	19	61	77	111	132	11	20	61	78	111	133
12	20	62	78	112	133	12	21	62	79	112	134
13	21	63	79	113	134	13	23	63	80	113	135
14	23	64	80	114	135	14	24	64	82	114	136
15	24	65	81	115	136	15	25	65	83	115	138
16	25	66	83	116	137	16	26	66	84	116	139
17	26	67	84	117	138	17	28	67	85	117	140
18	28	68	85	118	139	18	29	68	86	118	141
19	29	69	86	119	141	19	30	69	87	119	142
20	30	70	87	120	142	20	31	70	88	120	143
21	31	71	88	121	143	21	32	71	89	121	144
22	32	72	89	122	144	22	34	72	90	122	145
23	34	73	90	123	145	23	35	73	92	123	146
24	35	74	91	124	146	24	36	74	93	124	147
25	36	75	92	125	147	25	37	75	94	125	148
26	37	76	94	126	148	26	38	76	95	126	149
27	38	77	95	127	149	27	39	77	96	127	150
28	39	78	96	128	150	28	41	78	97	128	152
29	41	79	97	129	151	29	42	79	98	129	153
30	42	80	98	130	152	30	43	80	99	130	154
31	43	81	99	131	154	31	44	81	100	131	155
32	44	82	100	132	155	32	45	82	101	132	156
33	45	83	101	133	156	33	46	83	103	133	157
34	46	84	102	134	157	34	48	84	104	134	158
35	48	85	103	135	158	35	49	85	105	135	159
36	49	86	105	136	159	36	50	86	106	136	160
37	50	87	106	137	160	37	51	87	107	137	161
38	51	88	107	138	161	38	52	88	108	138	162
39	52	89	108	139	162	39	53	89	109	139	163
40	53	90	109	140	163	40	55	90	110	140	165
41	54	91	110	141	164	41	56	91	111	141	166
42	56	92	111	142	165	42	57	92	112	142	167
43	57	93	112	143	166	43	58	93	114	143	168
44	58	94	113	144	168	44	59	94	115	144	169
45	59	95	114	145	169	45	60	95	116	145	170
46	60	96	116	146	170	46	61	96	117	146	171
47	61	97	117	147	171	47	62	97	118	147	172
48	62	98	118	148	172	48	64	98	119	148	173
49	63	99	119	149	173	49	65	99	120	149	174
50	65	100	120	150	174	50	66	100	121	150	175

APPENDIX I—TABLE 1. (continued) Sample size correction table for adjusting "point-in-time" parameter estimates. n = the uncorrected sample size value from the sample size equation. n* = the corrected sample size value. This table was created using the algorithm reported by Kupper and Hafner (1989) for a one-sample tolerance probability of 0.90. For more information consult Kupper and Hafner (1989).

SAMPLE SIZE EQUATION #2: DETERMINING THE NECESSARY SAMPLE SIZE FOR DETECTING DIFFERENCES BETWEEN TWO MEANS WITH TEMPORARY SAMPLING UNITS.

The equation for determining the number of samples necessary to detect some "true" difference between two sample means is:

$$n = \frac{2(s)^2(Z_\alpha + Z_\beta)^2}{(MDC)^2}$$

Where:

s = sample standard deviation.

Z_α = Z-coefficient for the false-change (Type I) error rate from the table below.

Z_β = Z-coefficient for the missed-change (Type II) error rate from the table below.

MDC = Minimum detectable change size. This needs to be specified in absolute terms rather than as a percentage. For example, if you wanted to detect a 20% change in the sample mean from one year to the next and your first year sample mean = 10 plants/quadrat then MDC = (0.20 x 10) = 2 plants/quadrat.

Table of standard normal deviates for Z_α		Table of standard normal deviates for Z_β		
False-change (Type I) error rate (α)	Z_α	Missed-change (Type II) error rate (β)	Power	Z_β
0.40	0.84	0.40	0.60	0.25
0.20	1.28	0.20	0.80	0.84
0.10	1.64	0.10	0.90	1.28
0.05	1.96	0.05	0.95	1.64
0.01	2.58	0.01	0.99	2.33

Example:

Management objective:

Increase the density of species F at Site Y by 20% between 1999 and 2004.

Sampling objective

I want to be 90% certain of detecting a 20% change in mean plant density and I am willing to accept a 10% chance that I will make a false-change error (conclude that a change took place when it really did not).

Results from pilot sampling:

Mean ($\bar{X}$) = 25 plants/quadrat.
Standard deviation (s) = 7 plants.

Given:

The acceptable **False-change error rate (α) = 0.10** so the appropriate Z_a from the table = 1.64.

The desired **Power** is 90% (0.90) so the **Missed-change error rate (β) = 0.10** and the appropriate Z_β coefficient from the table = 1.28.

The **Minimum Detectable Change (MDC)** is 20% of the 1999 value or (0.20)(25) = 5 plants/quadrat.

Calculate the estimated necessary sample size using the equation provided above:

$$n = \frac{2(s)^2(Z_\alpha + Z_\beta)^2}{(MDC)^2} \qquad n = \frac{2(7)^2(1.64 + 1.28)^2}{(5)^2} = 33.4$$

Round up 33.4 to 34 plots.

Final estimated sample size needed to be 90% confident of detecting a change of 5 plants between 1999 and 2004 with a false-change error rate of 0.10 = **34 quadrats.** The sample size correction table is not needed for estimating sample sizes for detecting differences between two population means.

Correction for Sampling Finite Populations:

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. If you are sampling more than 5% of the whole population area then you should apply a correction to the sample size estimate that incorporates the finite population correction factor (FPC). This will reduce the sample size. The formula for correcting the sample size estimate is as follows:

$$n' = \frac{n}{(1 + (n/N))}$$

Where:

n′ = The new sample size based upon inclusion of the finite population correction factor.

n = The sample size from the equation above.

N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

Example:

If the pilot data described above was gathered using a 1m x 10m (10 m^2) quadrat and the total population being sampled was located within a 20m x 50m macroplot (1000 m^2) then N = $1000m^2/10m^2$ = 100. The corrected sample size would then be:

$$n' = \frac{n}{(1 + (n/N))} \qquad n' = \frac{34}{(1 + (34/100))} = 25.3$$

Round up 25.3 to 26.

The new, FPC-corrected estimated sample size needed to be 90% certain of detecting a change of 5 plants between 1999 and 2004 with a false-change error rate of 0.10 = **26 quadrats**.

Note on the Statistical Analysis for Two Sample Tests from Finite Populations

If you have sampled more than 5% of an entire population then you should also apply the finite population correction factor to the results of the statistical test. This procedure involves dividing the test statistic by the square root of the finite population factor (1-n/N). For example, if your t-statistic from a particular test turned out to be 1.645 and you sampled n = 26 quadrats out of a total N=100 possible quadrats, then your correction procedure would look like the following:

$$t' = \frac{t}{\sqrt{1-(n/N)}} \qquad t' = \frac{1.645}{\sqrt{1-(26/100)}} = 1.912$$

Where:

t = The *t*-statistic from a *t*-test.

t′ = The corrected *t*-statistic using the FPC.

n = The sample size from the equation above.

N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

You would need to look up the *p*-value of *t′* = 1.912 in a *t*-table at the appropriate degrees of freedom to obtain the correct *p*-value for this statistical test.

SAMPLE SIZE EQUATION #3: DETERMINING THE NECESSARY SAMPLE SIZE FOR DETECTING DIFFERENCES BETWEEN TWO MEANS WHEN USING PAIRED OR PERMANENT SAMPLING UNITS.

When paired sampling units are being compared or when data from permanent quadrats are being compared between two time periods, then sample size determination requires a different procedure than if samples are independent of one another. The equation for determining the number of samples necessary to detect some "true" difference between two sample means is:

$$n = \frac{(s)^2(Z_\alpha + Z_\beta)^2}{(MDC)^2}$$

Where:

s = Standard deviation of the differences between paired samples (see examples below).

Z_α = Z-coefficient for the false-change (Type I) error rate from the table below.

Z_β = Z-coefficient for the missed-change (Type II) error rate from the table below.

MDC = Minimum detectable change size. This needs to be specified in absolute terms rather than as a percentage. For example, if you wanted to detect a 20% change in the sample mean from one year to the next and your first year sample mean = 10 plants/quadrat then MDC = (0.20 x 10) = 2 plants/quadrat.

Table of standard normal deviates for Z_α

False-change (Type I) error rate (α)	Z_α
0.40	0.84
0.20	1.28
0.10	1.64
0.05	1.96
0.01	2.58

Table of standard normal deviates for Z_β

Missed-change (Type II) error rate (β)	Power	Z_β
0.40	0.60	0.25
0.20	0.80	0.84
0.10	0.90	1.28
0.05	0.95	1.64
0.01	0.99	2.33

If the objective is to track changes over time with permanent sampling units and only a single year of data is available, then you will not have a standard deviation of differences between the paired samples. If you have an estimate of the likely degree of correlation between the two years of data, and you assume that the among sampling units standard deviation is going to be the same in the second time period, then you can use the equation below to estimate the standard deviation of differences.

$$s_{diff} = (s_1)\left(\sqrt{(2\,(1 - corr_{diff}))}\right)$$

Where:

s_{diff} = Estimated standard deviation of the differences between paired samples.

s_1 = Sample standard deviation among sampling units at the first time period.

$corr_{diff}$ = Correlation coefficient between sampling unit values in the first time period and sampling unit values in the second time period.

Example #1:

Management objective:

Achieve at least a 20% higher density of species F at site Y in areas excluded from grazing as compared to grazed areas in 1999.

Sampling objective:

I want to be able to detect a 20% difference in mean plant density in areas excluded from grazing and adjacent paired grazed areas. I want to be 90% certain of detecting that difference, if it occurs, and I am willing to accept a 10% chance that I will make a false-change error (conclude that a difference exists when it really did not).

Results from pilot sampling:

Five paired quadrats were sampled where one member of the pair was excluded from grazing (with a small exclosure) and the other member of the pair was open to grazing.

Quadrat number	# of plants/quadrat		Difference between grazed and ungrazed	
	grazed	ungrazed		
1	2	3	1	
2	5	8	3	
3	4	9	5	
4	7	12	5	
5	3	7	4	
	$\bar{X}$ =4.20 s =1.92	$\bar{X}$ =7.80 s=3.27		
Summary statistics for the differences between the two sets of quadrats			$\bar{X}$ 3.60	s 1.67

Given:

The sampling objective specified a desired minimum detectable difference (i.e., equivalent to the MDC) of 20%. Taking the larger of the two mean values and multiplying by 20% leads to: (7.80) x (0.20) = **MDC = 1.56** plants quadrat

The appropriate **standard deviation** to use is **1.67**, the standard deviation of the differences between the pairs.

The acceptable **False-change error rate (α) = 0.10**, so the appropriate Z_α from the table = 1.64.

The desired Power is 90% (0.90), so the **Missed-change error rate (β) = 0.10** and the appropriate Z_b coefficient from the table = 1.28.

Calculate the estimated necessary sample size using the equation provided above:

$$n = \frac{(s)^2(Z_\alpha + Z_\beta)^2}{(MDC)^2} \qquad n = \frac{(1.67)^2(1.64 + 1.28)^2}{(1.56)^2} = 9.7$$

Round up 9.7 to 10 plots.

Final estimated sample size needed to be 90% certain of detecting a true difference of 1.56 plants/quadrat between the grazed and ungrazed quadrats with a false-change error rate of 0.10 = **10 quadrats**.

Example #2:

Management objective:

Increase the density of species F at Site Q by 20% between 1999 and 2002.

Sampling objective:
I want to be able to detect a 20% difference in mean plant density of species F at Site Q between 1999 and 2001. I want to be 90% certain of detecting that change, if it occurs, and I am willing to accept a 10% chance that I will make a false-change error (conclude that a change took place when it really did not).

The procedure for determining the necessary sample size for this example would be very similar to the previous example. Just replace "grazed" and "ungrazed" in the data table with "1999" and "2002" and the rest of the calculations would be the same. Because the sample size determination procedure needs the standard deviation of the difference between two samples, you will not have the necessary standard deviation term to plug into the equation until you have two years of data. The standard deviation of the difference can be estimated in the first year if some estimate of the correlation coefficient between sampling unit values in the first time period and the sampling unit values in the second time period is available (see the s_{diff} equation above).

Correction for Sampling Finite Populations:

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. If you are sampling more than 5% of the whole population area then you should apply a correction to the sample size estimate that incorporates the finite population correction factor (FPC). This will reduce the sample size. The formula for correcting the sample size estimate is as follows:

$$n' = \frac{n}{(1 + (n/N))}$$

Where:

n′ = The new sample size based upon inclusion of the finite population correction factor.

n = The sample size from the equation above.

N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

Example:
If the pilot data described above were gathered using a 1m x 10m ($10m^2$) quadrat and the total population being sampled was located within a 10m x 50m macroplot ($500m^2$) then N = $500m^2/10m^2$ = 50. The corrected sample size would then be:

$$n' = \frac{n}{(1 + (n/N))} \qquad n' = \frac{10}{(1 + (10/50))} = 8.3$$

Round up 8.3 to 9.

The new, FPC-corrected estimated sample size needed to be 90% confident of detecting a true difference of 1.56 plants/quadrat between the two years with a false-change error rate of 0.10 = **9 quadrats.**

Note on the Statistical Analysis for Two Sample Tests from Finite Populations

If you have sampled more than 5% of an entire population then you should also apply the finite population correction factor to the results of the statistical test. This procedure involves dividing the test statistic by the square root of (1-n/N). For example, if your *t*-statistic from a particular

test turned out to be 1.782 and you sampled n=9 quadrats out of a total N=50 possible quadrats, then your correction procedure would look like the following:

$$t' = \frac{t}{\sqrt{1-(n/N)}} \qquad t' = \frac{1.782}{\sqrt{1-(9/50)}} = 1.968$$

Where:

t = The *t*-statistic from a *t*-test.
t' = The corrected *t*-statistic using the FPC.
n = The sample size from the equation above.
N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

You would need to look up the *p*-value of t' = 1.968 in a *t*-table for the appropriate degrees of freedom to obtain the correct *p*-value for this statistical test.

SAMPLE SIZE EQUATION #4: DETERMINING THE NECESSARY SAMPLE SIZE FOR ESTIMATING A SINGLE POPULATION PROPORTION WITH A SPECIFIED LEVEL OF PRECISION.

The equation for determining the sample size for estimating a single proportion is:

$$n = \frac{(Z_\alpha)^2(p)(q)}{d^2}$$

Where:

n = Estimated necessary sample size.
Z_α = The coefficient from the table of standard normal deviates below.
p = The value of the proportion as a decimal percent (e.g., 0.45). If you don't have an estimate of the current proportion, use 0.50 as a conservative estimate.
q = 1 - p.
d = The desired precision level expressed as half of the maximum acceptable confidence interval width. This is also expressed as a decimal percent (e.g., 0.15) and this represents an *absolute* rather than a *relative* value. For example, if your proportion value is 30% and you want a precision level of ±10% this means you are targeting an interval width from 20% to 40%. Use 0.10 for the d-value and *not* 0.30 x 0.10 = 0.03.

Table of standard normal deviates (Z_α) for various confidence levels

Confidence level	Alpha (α) level	(Z_α)
80%	0.20	1.28
90%	0.10	1.64
95%	0.05	1.96
99%	0.01	2.58

Example:

Management objective:

Maintain at least a 40% frequency (in $1m^2$ quadrats) of species Y in population Z over the next 5 years.

Sampling objective:

Estimate percent frequency with 95% confidence intervals no wider than ± 10% of the estimated true value.

Results of pilot sampling:

The proportion of quadrats with species Z is estimated to be p = 65% (0.65).

Because q = (1-p), q = (1-0.65) = 0.35.

Given:

The desired **confidence level** is 95% so the appropriate Z_α from the table above = 1.96.

The desired **confidence interval width (d)** is specified as 10% (0.10).

Using the equation provided above:

$$n = \frac{(Z_\alpha)^2(p)(q)}{d^2} \qquad n = \frac{(1.96)^2(0.65)(0.35)}{0.10^2} = 87.4$$

Round up 87.4 to 88.

The estimated sample size needed to be 95% confident that the estimate of the population percent frequency is within 10% (± 0.10) of the true percent frequency = **88 quadrats.**

This sample size formula works well as long as the proportion is more than 0.20 and less than 0.80 (Zar 1999). If you suspect the population proportion is less than 0.20 or greater than 0.80, use 0.20 or 0.80, respectively, as a conservative estimate of the proportion.

Correction for Sampling Finite Populations:

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. If you are sampling more than 5% of the whole population area then you should apply a correction for your sample size estimate that incorporates the finite population correction factor (FPC). This will reduce the sample size estimate. The formula for correcting the sample size estimate is as follows:

$$n' = \frac{n}{(1 + (n/N))}$$

Where:

n' = The new sample size with the inclusion of the FPC factor.

n = The sample size estimate from the above equation.

N = The total number of possible quadrat locations in the population. To calculate N, divide the total population area by the size of the sampling unit.

Example:

If the pilot data described above was gathered using a 1m × 1m ($1m^2$) quadrat and the total population being sampled was located within a 25m × 25m macroplot ($625m^2$) then N = $625m^2/1m^2$ = 625. The corrected sample size would then be:

$$n' = \frac{n}{(1 + (n/N))} \qquad n' = \frac{88}{(1 + (88/625))} = 77.1$$

Round up 77.1 to 78.

The new, FPC-corrected, estimated sample size needed to be 95% confident that the estimate of the population percent frequency is within 10% (+/- 0.10) of the true percent frequency = 78 quadrats.

SAMPLE SIZE EQUATION #5: DETERMINING THE NECESSARY SAMPLE SIZE FOR DETECTING DIFFERENCES BETWEEN TWO PROPORTIONS WITH TEMPORARY SAMPLING UNITS.

The equation for determining the number of samples necessary to detect some "true" difference between two sample proportions is:

$$n = \frac{(Z_\alpha + Z_\beta)^2(p_1q_1 + p_2q_2)}{(p_2 - p_1)^2}$$

Where:

- n = Estimated necessary sample size.
- Z_a = Z-coefficient for the false-change (Type I) error rate from the table below.
- Z_b = Z-coefficient for the missed-change (Type II) error rate from the table below.
- p_1 = The value of the proportion for the first sample as a decimal (e.g., 0.65). If you don't have an estimate of the current proportion, use 0.50 as a conservative estimate.
- q_1 = $1 - p_1$.
- p_2 = The value of the proportion for the second sample as a decimal (e.g., 0.45). This is determined based on the magnitude of change you wish to detect (see example, below).
- q_2 = $1 - p_2$.

Table of standard normal deviates for Z_α

False-change (Type I) error rate (α)	Z_α
0.40	0.84
0.20	1.28
0.10	1.64
0.05	1.96
0.01	2.58

Table of standard normal deviates for Z_β

Missed-change (Type II) error rate (β)	Power	Z_β
0.40	0.60	0.25
0.20	0.80	0.84
0.10	0.90	1.28
0.05	0.95	1.64
0.01	0.99	2.33

Example:

Management objective:

Decrease the frequency of invasive weed F at Site G by 20% between 1999 and 2001.

Sampling objective:

I want to be 90% certain of detecting an absolute change of 20% frequency and I am willing to accept a 10% chance that I will make a false-change error (conclude that a change took place when it really did not).

Note that the magnitude of change for detecting change over time for proportion data is expressed in absolute terms rather than in relative terms (relative terms were used in earlier examples that dealt with sample means values). The reason absolute terms are used instead of relative terms relates to the type of data being gathered (percent frequency is already expressed as a relative measure). Think of taking your population area and dividing it into a grid where the size of each grid cell equals your quadrat size. When you estimate a percent frequency, you are estimating

the proportion of these grid cells occupied by a particular species. If 45% of all the grid cells in the population are occupied by a particular species then you hope that your sample values will be close to 45%. If over time the population changes so that now 65% of all the grid cells are occupied, then the true percent frequency has changed from 45% to 65%, representing a 20% absolute change.

Results from pilot sampling:

The proportion of quadrats with species Z in 1999 is estimated to be p_1 = 65% **(0.65)**.

Because $q_1 = (1-p_1)$, $q_1 = (1-0.65) =$ **0.35**.

Because we are interested in detecting a 20% shift in percent frequency, we will assign p_2 = 0.45. This represents a shift of 20% frequency from 1999 to 2001. A decline was selected instead of an increase (e.g., from 65% frequency to 85% frequency) because sample size requirements are higher at the mid-range of frequency values (i.e., closer to 50%) than they are closer to 0 or 100. Sticking closer to the mid-range gives us a more conservative sample size estimate.

Because $q_2 = (1-p_2)$, $q_2 = (1-0.45) = 0.55$.

Given:

The acceptable **False-change error rate (α) = 0.10** so the appropriate Z_α from the table = 1.64.

The desired **Power is 90% (0.90)** so the **Missed-change error rate (β) = 0.10** and the appropriate Z_b coefficient from the table = 1.28.

Using the equation provided above:

$$n = \frac{(Z_\alpha + Z_\beta)^2(p_1q_1 + p_2q_2)}{(p_2 - p_1)^2} \qquad n = \frac{(1.64 + 1.28)^2((0.65)(0.35)+(0.45)(0.55))}{(0.45 - 0.65)^2} = 101.3$$

Round up 101.3 to 102.

The estimated sample size needed to be 90% sure of detecting a shift of 20% frequency with a starting frequency of 65% and a false-change error rate of 0.10 = 102 quadrats.

Correction for Sampling Finite Populations:

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. If you are sampling more than 5% of the whole population area then you should apply a correction to the sample size estimate that incorporates the finite population correction factor (FPC). This will reduce the sample size. The formula for correcting the sample size estimate is as follows:

$$n' = \frac{n}{(1 + (n/N))}$$

Where:

n′ = The new sample size based upon inclusion of the finite population correction factor.

n = The sample size from the equation above.

N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

Example:

If the pilot data described above was gathered using a 1m x 1m ($1m^2$) quadrat and the total population being sampled was located within a 10m × 30m macroplot ($300m^2$) then N = $300m^2/1m^2$ = 300. The corrected sample size would then be:

$$n' = \frac{n}{(1 + (n/N))} \qquad n' = \frac{102}{(1 + (102/300))} = 76.1$$

Round up 76.1 to 77.

The new, FPC-corrected estimated sample size needed to be 90% sure of detecting an absolute shift of 20% frequency with a starting frequency of 65% and a false-change error rate of 0.10 = **77 quadrats.**

Note on the Statistical Analysis for Two Sample Tests from Finite Populations

If you have sampled more than 5% of an entire population then you should also apply the finite population correction factor to the results of the statistical test. For proportion data, this procedure involves dividing the test statistic by (1-n/N). For example, if your χ^2 -statistic from a particular test turned out to be 2.706 and you sampled n = 77 quadrats out of a total N = 300 possible quadrats, then your correction procedure would look like the following:

$$\chi^{2\prime} = \frac{\chi^2}{1-(n/N)} \qquad \chi^{2\prime} = \frac{2.706}{1-(77/300)} = 3.640$$

Where:

χ^2 = The χ^2 -statistic from a χ^2 -test.
$\chi^{2\prime}$ = The corrected χ^2 -statistic using the FPC.
n = The sample size from the equation above.
N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of each individual sampling unit.

You would need to look up the *p*-value of $\chi^{2\prime}$ = 3.640 in a χ^2 -table for the appropriate degrees of freedom to obtain the correct *p*-value for this statistical test.

SAMPLE SIZE EQUATION #6: DETERMINING THE NECESSARY SAMPLE SIZE FOR DETECTING DIFFERENCES BETWEEN TWO PROPORTIONS WITH PERMANENT SAMPLING UNITS.

$$n = \left(\frac{\left(Z_a \sqrt{AP + PA}\right) + \left(Z_b \sqrt{\frac{4(PA)(AP)}{AP+PA}}\right)}{PA - AP} \right)$$

Where:

n = Estimated necessary sample size.
Z_α = Z-coefficient for the false-change (Type I) error rate from the table below.
Z_β = Z-coefficient for missed-change (Type II) error rate from the table below.
PA = the proportion of quadrats in which the plant was present at the first measurement, but absent on the second, expressed as a decimal (e.g., 0.20).
AP = the proportion of quadrats in which the plant was absent at the first measurement, but present on the second, expressed as a decimal (as a decimal, e.g., 0.30)

Table of standard normal deviates for Z_α	
False-change (Type I) error rate (α)	Z_α
0.40	0.84
0.20	1.28
0.10	1.64
0.05	1.96
0.01	2.58

Table of standard normal deviates for Z_β		
Missed-change (Type II) error rate (β)	Power	Z_β
0.40	0.60	0.25
0.20	0.80	0.84
0.10	0.90	1.28
0.05	0.95	1.64
0.01	0.99	2.33

In a permanent frequency design, you track changes in presence/absence in each of the permanent quadrats. There are four possible quadrat transitions, as shown in the following matrix:

		Year 2	
		Present (P)	Absent (A)
Year 1	Present (P)	PP	PA
	Absent (A)	AP	AA

To calculate sample size, you need to know the fate of quadrats in the second year. How can you ensure that you establish an adequate sample size in the first year?

One approach to calculating permanent frequency quadrat sample sizes is to determine the desired minimum detectable change size and then create a range of hypothetical population changes that could lead to such a magnitude of change. Sample sizes can be calculated for this range of population changes and used as a guide to guessing the "right" sample size to use.

Consider the following example where we specify a minimum detectable change of 10%, with a starting frequency of 50%. What are some different ways that this frequency change could occur? For simplicity's sake, let's work through our example assuming a sampling design with 100 quadrats (the actual numbers don't matter, only the ratio between them). In the first year, we have 50 quadrats with plants and 50 quadrats without plants. In the second year, we want to be able to detect a shift to either 40 quadrats now having plants or to 60 quadrats now having plants. Let's stick with the 40 quadrats with plants scenario. What are some different ways that a 50% to 40% frequency change could occur?

We could simply lose plants from 10 of our previously occupied quadrats and not have any new plants show up in previously unoccupied quadrats. This would create the quadrat transitions shown below. These quadrat transitions can be used along with the 0.10 minimum detectable change and some specified power and false-change error rate (power = 0.90 and false change error rate = 0.10 in the following examples) in the STPLAN program. (The finite population correction factor was not applied to any of the sample sizes in the examples listed below.)

PP	PA	AP	AA
40	10	0	50

Permanent quadrat sample size	51
Temporary quadrat sample size	423

Alternatively, we could obtain a 10% change in frequency by losing plants from 15 of the originally occupied quadrats and gaining plants in 5 quadrats that did not previously have plants. Here are the transitions and sample sizes.

PP	PA	AP	AA
35	15	5	45

Permanent quadrat sample size	156
Temporary quadrat sample size	423

We could look at any number of these hypothetical population changes and compare permanent vs. temporary quadrat sample sizes. Table 2 lists 9 different population changes (including the two listed above) that all lead to 10% declines in percent frequency. Each new row shows the result of 5 additional quadrats losing plants that had plants in the original population.

Quadrat transitions				Permanent quadrat sample size	Temporary quadrat sample size	Difference in sample size
PP	PA	AP	AA			
40	10	0	50	51	463	412
35	15	5	45	156	463	307
30	20	10	40	247	463	216
25	25	15	35	335	463	128
20	30	20	30	422	463	41
15	35	25	25	508	463	-45
10	40	30	20	595	463	-132
5	45	35	15	681	463	-218
0	50	40	10	767	463	-304

TABLE 2. Population changes that lead to a 50% to 40% change in frequency based upon 100 quadrats.

Table 3 shows the same type of information as Table 2 except now the magnitude of change is 20% and the frequency changes from 50% to 30%.

Quadrat transitions				Permanent quadrat sample size	Temporary quadrat sample size	Difference in sample size
PP	PA	AP	AA			
30	20	0	50	25	101	76
25	25	5	45	53	101	48
20	30	10	40	77	101	24
15	35	15	35	100	101	1
10	40	20	30	123	101	-22
5	45	25	25	145	101	-44
0	50	30	20	167	101	-66

TABLE 3. Population changes that lead to a 50% to 30% change in frequency based upon 100 quadrats.

With these sorts of tables in hand, we can evaluate the likelihood of various types of changes occurring and select a sample size that fits. For example, how likely is the following transition (taken from the middle line of Table 3)?

PP	PA	AP	AA
15	35	15	35

In this transition, the number of quadrats that had plants present in both years (15) is equal to the number of quadrats that were empty in the first year but gained at least one plant in the second year (15). This implies that quadrats having plants in them during the first year are no more likely to have plants during the second year than quadrats that did not have plants present in the first year. You could get this sort of result when the between-year correlation in quadrat counts is close to zero. How likely is this sort of result? For most plant species it is probably highly unlikely given typical patterns of sexual and asexual reproduction in plants (but perhaps it occurs in some situations where allelopathy is operating). Even many (most?) annual plants set a higher proportion of seed close to the parent plant location than at distances far from the parents.

The last transition listed in Table 3 (0-50-30-20) shows a permanent quadrat sample size penalty since it would take 167 permanent quadrats as compared to 101 temporary quadrats to detect this particular 20% change. In this transition, all 50 quadrats with plants in the first year lose their plants and 30 of the 50 quadrats that were previously empty gain plants. This would indicate a strongly negative correlation in plant counts between years and there would have to be a strong degree of allelopathy to create this sort of transition. Few, if any, plant species would be suspected of showing this type of response.

By developing an ecological model for your target plant or animal species that describes the types of transitions likely to occur between years, you can therefore arrive at an initial estimate of the sample size necessary to detect a particular level of change based on the first year's data. You can then calculate the sample size based on this model using either the formula or a computer program (see our web page). After you've collected the second year's data, you should recalculate sample size based on the actual transitions.

LITERATURE CITED

Kupper, L. L.; Hafner, K. B. 1989. How appropriate are popular sample size formulas? The American Statistician (43): 101-105.

Zar, J. H. 1999. Biostatistical analysis, 3rd edition. Upper Saddle River, New Jersey: Prentice Hall.

APPENDIX III
Confidence Interval Equations

A confidence interval is the interval within which a true parameter value lies with known probability. It is a measure of the reliability of our sample estimate of the parameter value. In this appendix, we provide examples illustrating calculation of confidence intervals for estimates of 1) a mean; 2) a population total; 3) a proportion; and 4) a median. These would be the kinds of estimates made for target/threshold types of objectives.

Confidence intervals can also be used to assess change/trend objectives. In Chapter 9 we describe the significance tests used to assess these types of objectives, but as briefly described there, confidence intervals can also be used when we wish to compare two quantities (such as mean density measured in two different years, or at two different sites). A confidence interval can be calculated for 1) the difference in two population means using independent sampling units, 2) the difference in two population means using paired (or permanent) sampling units; 3) the difference between two proportions; and 4) the difference between two paired proportions.

I. TARGET/THRESHOLD OBJECTIVES: CONFIDENCE INTERVALS FOR ESTIMATES OF A SINGLE VALUE

A. Confidence Interval for a Mean

We want to be able to specify the interval within which the true population mean most likely lies. In other words we want to be able to specify:

lower limit $< \mu <$ *upper limit*

We usually use a value from a table of the t distribution to determine a confidence interval. The formula for calculating a confidence interval is as follows:

$$\bar{X} - (t_{\alpha(2),v})(SE) < \mu < \bar{X} + (t_{\alpha(2),v})(SE)$$

Where:

$\bar{X}$ = sample estimate of the population mean.
SE = standard error, $s/\sqrt{n}$.
μ = the true population mean.
$t_{\alpha(2),v}$ = the critical value of t from a t table for a given level of α and sample size ($v = n - 1$).

The (2) indicates that we are using both tails of the t distribution (which will always be the case for calculating a confidence interval of α mean).

The α value we choose depends upon how certain we wish to be that μ lies within our confidence interval. If we want to be 95% confident of this we choose $\alpha = 0.05$. If we want to be 80% confident we choose $\alpha = 0.20$ and so on.

Another, more concise, way of expressing the confidence interval is:

$$\bar{X} \pm (t_{\alpha(2),v})(SE)$$

If you are sampling from a finite population and you have sampled more than 5% of the population, you should apply the finite population correction factor (FPC) to your estimate of the SE. You do this as follows:

$$SE' = (SE)\left(\sqrt{1-\frac{n}{N}}\right)$$

Where:

SE′ = Corrected standard error.
SE = Uncorrected standard error.
n = The sample size (the number of quadrats sampled).
N = The total number of possible quadrats in the population. To calculate N, determine the total area of the population and divide by the area of each individual quadrat.

You then insert the corrected standard error (SE′) into the equation for the confidence interval for the population mean.

For example, we wish to calculate the 90% confidence interval around our estimate of 20.5 plants/quadrat. The standard deviation (s) is 17.2 plants/quadrat. We sampled 40 out of 140 possible quadrats. Each quadrat is $2m^2$ in area.

First we calculate the standard error (SE):

$$SE = \frac{s}{\sqrt{n}} = \frac{17.2}{\sqrt{40}} = 2.72$$

Because we have sampled more than 5% of the population we apply the finite population correction factor to the standard error:

$$SE' = (SE)\left(\sqrt{1-\frac{n}{N}}\right) = (2.72)\left(\sqrt{1-\frac{40}{140}}\right) = 2.30$$

Then calculate the confidence interval:

$$\bar{X} - \left(t_{\alpha(2),\nu}\right)\left(SE'\right) < \mu < \bar{X} + \left(t_{\alpha(2),\nu}\right)\left(SE'\right)$$

Where t is derived from a table or computer program for $\alpha_{(2)} = 0.10$ (for the 90% confidence level) and $\nu = 39$ (for $n - 1$, or, in this example, $40 - 1$).

$$20.5 - (1.685)(2.30) < \mu < 20.5 + (1.685)(2.30)$$

The 90% confidence interval is $16.62 < \mu < 24.38$.

You can also express the 90% confidence interval as 20.50 ± 3.88 plants/quadrat or you can express the 90% confidence interval as 10.25 ± 1.94 plants/m^2 (remember each of our quadrats were $2m^2$ in area).

B. Confidence Interval for a Population Total

To calculate a confidence interval for a population total you must know the size (N) of the population you have sampled from. You then calculate your estimate of the population total as follows:

$$\tau = (N)(\bar{X})$$

Where:

τ = Estimate of population total.
N = The total number of possible quadrats in the population. To calculate N, determine the total area of the population and divide by the area of each individual quadrat.
$\bar{X}$ = Estimate of population mean.

The confidence interval around the estimate of the population total is then calculated as follows:

$\tau \pm (N)(\mathrm{CI})$

Where:

τ = Estimate of population total.
N = The total number of possible quadrats in the population. To calculate N, determine the total area of the population and divide by the area of each individual quadrat.
CI = Confidence interval calculated for population mean as described above.

Using our example from Section I.A., calculate the 90% confidence interval for the total population. Our estimate of the population mean was 20.5 plants/quadrat, the 90% confidence interval for the population mean was ± 3.88 plants/quadrat, and we sampled 40 of the 140 total possible quadrats.

First, calculate the total number of plants within the sampled area:

$\tau = N(\bar{X}) = 140(20.5) = 2870$ plants

Calculate the confidence interval:

$\tau = \pm(N)(CI) = \pm(140)(3.88) = 543.2$

The 90% confidence interval for the total is 2870 ± 543.2 plants within the sampled area.

C. Confidence Interval for a Proportion

We want to be able to specify the interval within which the true population proportion most likely lies. In other words we want to be able to specify:

lower limit < p < upper limit

There are several ways of calculating a confidence interval around a proportion. Krebs (1998:21, Figure 2.2) provides a graph that can be used to estimate the confidence interval. Zar (1999:527-529) gives an "exact" method that uses a relationship between the F distribution and the binomial distribution. To use the method you need access to an F table, which can be also be found in Zar (1999); alternatively, you can use computer programs to calculate the F values needed for the procedure (links available on our web page).

The following method, taken from Cochran (1977), approximates the confidence interval by using the normal distribution. It is accurate if the sample size is reasonably large, as shown in Table 1.

Sample Proportion ($\hat{p}$)	Number of sampling units in the smaller class	Total sample size (n)
0.5	15	30
0.4	20	50
0.3	24	80
0.2	40	200
0.1	60	600
0.05	70	1400

TABLE 1. Sample sizes needed to use the normal approximation to calculate confidence intervals for proportions (Krebs 1998; Cochran 1977). Do not use the normal approximation unless you have a sample size this large or larger.

Calculate a confidence interval around the estimate of the population proportion ($\hat{p}$) obtained from your sample:

$$\hat{p} \pm \left[\left((Z_\alpha) \left(\sqrt{1 - (n/N)} \right) \left(\sqrt{\frac{\hat{p}\hat{q}}{n-1}} \right) \right) + \frac{1}{2n} \right]$$

Where:

$\hat{p}$ = Estimated proportion.
Z_α = Standard normal deviate from Table 2.
$\hat{q}$ = 1 - $\hat{p}$.
n = Sample size.

Confidence level	Alpha (α) level	(Z_α)
80%	0.20	1.28
90%	0.10	1.64
95%	0.05	1.96
99%	0.01	2.58

TABLE 2. Table of standard normal deviates (Z_α) for various confidence levels

The value, 1 – (n/N), in the above equation is the finite population correction factor (FPC). If your population is finite (i.e., you used quadrats and not points) and you've sampled more than 5% of the population you should use the above equation. Otherwise, you can leave the FPC out of the equation, in which case the equation reduces to:

$$\hat{p} \pm \left[\left((Z_\alpha) \left(\sqrt{\frac{\hat{p}\hat{q}}{n-1}} \right) \right) + \frac{1}{2n} \right]$$

For example, we sample 200 frequency quadrats and find species X in 75 of the 200 quadrats. Our estimate, $\hat{p}$, of the population proportion is 75/200 = 0.375. There are 1000 possible quadrat positions in the population we sampled. The 95% confidence interval around this estimate is therefore:

$$0.375 \pm \left[\left((1.96) \left(\sqrt{1 - (200/1000)} \right) \left(\sqrt{\frac{(0.375)(0.625)}{200-1}} \right) \right) + \frac{1}{2(200)} \right] = 0.375 \pm 0.063$$

For frequency sampling, you can (and should) adjust your quadrat size so the proportion of quadrats with the species of interest is close to 0.50. Doing this will ensure that the normal approximation method will give good estimates of the confidence interval at reasonable sample sizes (Table 1). When you are using the point-intercept method of estimating cover, however, the proportion of "hits" on the species of interest depends entirely on the amount of cover of the species. For a species with low cover values, you will end up with a small proportion of "hits." In this situation you must pay heed to the sample size requirements of Table 1. If your sample size is less than that given in Table 1, you should not use the normal approximation method. Instead, you should use the exact method given by Zar (1999).

D. Confidence Interval for a Median

If sample sizes are large,[1] a normal approximation can be used to approximate the confidence interval around an estimated median. For small sample sizes, Zar (1999:543) provides an exact method based on the binomial distribution that is computationally simple, but requires the use of a table of critical values.

[1] *What constitutes a "large" sample is debatable, but the method seems to work fairly well for sample sizes of 30 or more. If in doubt, use the method described by Zar (1999).*

For the normal approximation, calculate:

$$i = \frac{n - \left(Z_\alpha \sqrt{n}\right)}{2}$$

(rounded to the nearest integer)

Where:

Z_α = the standard normal deviate for the selected confidence level (Table 2 above).
n = the sample size.
i = the ith placed data point from the smallest when the data are ranked from smallest to largest.

The lower confidence limit is the X_i value, the value that is the ith largest in the ranked list of data points. The upper confidence limit is X_{n-i+1}, the value that is the next larger value from the value that is ith from the largest in the ranked list. An example makes this clearer.

Heights were measured on 29 shrubs in a burn area (in centimeters). In rank order, these are the measured heights (from smallest to largest):

Rank number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Height (cm)	10	12	12	15	17	22	23	37	40	41	53	56	56	59	60
Rank number	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
Height (cm)	62	63	66	167	190	222	238	241	245	267	272	274	282	297	

The median is the 15th measurement, 60cm. The 95% confidence interval is calculated

$$i = \frac{29 - \left(1.96\sqrt{29}\right)}{2} = 9.22$$

Round this value to the nearest whole integer. The lower confidence limit is the X_i value, or the X_9 value, that is the 9th value in the ranked list. This value is 40cm.

The upper confidence limit is X_{n-i+1}, or X_{29-9+1}, or value X_{21}, which is 222cm.

In this example, the exact method gives the same answer.

II. CHANGE/TREND OBJECTIVES: CONFIDENCE INTERVALS FOR A DIFFERENCE BETWEEN TWO VALUES

A. The Difference in Two Population Means Using Independent Sampling Units

Several authors, particularly those in the behavioral sciences, have recently criticized the use of significance testing to determine whether two population means are different (see, for example, Cohen 1994). In its place they recommend calculating a confidence interval for the difference between two population means. This interval specifies:

lower limit $< \mu_1 - \mu_2 <$ *upper limit*

Just as for calculating the confidence interval around a single population mean, we resort to a value from the t distribution. Here is the formula:

$$\bar{X}_1 - \bar{X}_2 \pm (t_{\alpha(2),v})\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

Where:

$\bar{X}_1$ = The mean from the first sample.
$\bar{X}_2$ = The mean from the second sample.
s_1 = The standard deviation of the first sample.
s_2 = The standard deviation of the second sample.
n_1 = The sample size of the first sample.
vn_2 = The sample size of the second sample.
$t_{\alpha(2),v}$ = The t value from a t table for a given α value and degrees of freedom.
v = The degrees of freedom ($v = n_1 + n_2 - 2$).

The subscript (2) after the t indicates that we are using both tails of the t distribution.

Let us say we decide to calculate the 90 percent confidence interval for the difference in two population means. We take independent samples in two time periods and come up with the following information:

$\bar{X}_1$ = 10 plants
$\bar{X}_2$ = 5 plants
s_1 = 3.5 plants
s_2 = 3 plants
n_1 = 40 quadrats
n_2 = 40 quadrats

The 90 percent confidence interval for the difference between the means of the populations from which these two samples came is derived as follows:

$$10 - 5 \pm (1.665)\left(\sqrt{\frac{3.5^2}{40} + \frac{3^2}{40}}\right) = 5 \pm 1.21$$

Thus, we can be 90 percent confident that the true difference between the population means at times 1 and 2 falls within the interval 5 - 1.21 and 5 + 1.21 or between 3.79 and 6.21. Note that this interval does not include 0. If it did we would know that a significance test would yield a P value greater than 0.10 and we would conclude that the difference is not significant at the $\alpha = 0.10$ level. Because the interval is not even close to including 0 we can be very confident that the observed difference is real (this is not surprising since the difference between sample means is rather large and the estimates are rather precise). A significance test would yield a very low P value.

Consider the case, however, where either the difference between sample means is not so great and/or the estimates of the means are not very precise (s rather large compared to the means). Let's say we calculated a 90 percent confidence interval for the difference between two means and came up with the following interval:

-0.2 to 8

We can immediately determine two things from this. The first is that (because the interval contains 0) a significance test would yield a P value greater than 0.10. We would also note that the interval is rather large, meaning that our study design didn't have much power to detect change (the missed-change error rate would be high). The first thing we would have determined from a significance test (provided that the test provided exact P values). The second thing, however, is not as obvious. It is in providing this second important piece of information that the confidence interval between two population means is such a valuable statistical tool.

If you are sampling from a finite population and you have sampled more than 5% of the population, you should apply the finite population correction factor (FPC) to the formula for calculating the confidence interval as follows:

$$\bar{X}_1 - \bar{X}_2 \pm (t_{\alpha(2),v})\left(\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)\left(\sqrt{1-\frac{n}{N}}\right)$$

Where:

$\bar{X}_1$ = Sample mean for year one.
$\bar{X}_2$ = Sample mean for year two.
s_1^2 = Population variance for year one.
s_2^2 = Population variance for year two.
n = The sample size (the number of quadrats sampled in each year; note that you do not add the number of quadrats sampled the first year to the number of quadrats sampled in the second year).
N = The total number of possible quadrats in the population. To calculate N, determine the total area of the population and divide by the area of each individual quadrat.

Using our previous example, let us say that there were 500 possible quadrat locations in the population we sampled. Our sample size was 40 quadrats in each year. The confidence interval is therefore:

$$10 - 5 \pm (1.665)\left(\sqrt{\frac{3.5^2}{40}+\frac{3^2}{40}}\right)\left(\sqrt{1-\frac{40}{500}}\right) = 5 \pm 1.16$$

B. The Difference in Two Population Means Using Paired (or Permanent) Sampling Units

Paired sampling units produce a single value, the difference, for each pair of measurements. Thus, calculating the confidence interval is similar to calculating the confidence interval for a point estimate of the mean (Section I.A, above). The difference is you are calculating a confidence interval

Quadrat	1998	2000	Difference
1	12	15	3
2	6	12	6
3	21	25	4
4	13	18	5
5	5	2	-3
6	32	36	4
7	0	0	0
8	18	25	7
9	6	14	8
10	4	5	1
11	7	41	34

(*continued*)

Quadrat	1998	2000	Difference
12	41	39	-2
13	32	38	6
14	10	14	4
15	3	2	-1
16	56	67	11
17	12	14	2
18	7	22	15
19	37	43	6
20	23	20	-3

around the estimate of the mean difference. For example, here are the data for 20 permanent density quadrats measured twice. Within the sampled area there are 500 possible quadrats.

The mean difference (the mean of the values in the difference column) is 5.35 individuals/quadrat. The standard deviation is 8.15 individuals/quadrat.

Calculate the 90% confidence interval around this mean difference:

First calculate the standard error:

$$SE = \frac{s}{\sqrt{n}} = \frac{8.15}{\sqrt{20}} = 1.822$$

Because we have only sampled 20 out of 500 possible quadrats (4%), we do not apply the finite population correction factor.[2] Now calculate the confidence interval:

$$\overline{X} \pm (SE)(t_{(\alpha(2),\nu)}) = 5.35 \pm (1.822)(1.729) = 5.35 \pm 3.15 \text{ individuals/quadrat}$$

Where $(t_{(\alpha(2),\nu)})$ is the t value for the 90% confidence level ($\alpha(2) = 0.10$) and with $\nu = 20 - 1 = 19$.

C. The Difference between Two Proportions Using Independent Sampling Units

The frequency of a rare plant has been sampled with frequency quadrats in two different years. The locations of the quadrats were determined independently each year. Only the sampled area was permanently marked. Each year 200 quadrats were distributed throughout the area. In the first year, 66 (33%) of those quadrats contained the plant. In the second year 98 (49%) of the 200 quadrats contained the plant. The difference in frequency between the two years is 16% ($Freq_{diff}$). What is the 95% confidence interval around this difference? We can construct an approximate confidence interval based on the normal distribution.[3]

First, calculate the standard error:

$$SE_{(p1-p2)} = \sqrt{\left(\frac{p_1(100-p_1)}{n_1}\right) + \left(\frac{p_2(100-p_2)}{n_2}\right)}$$

---

[2] *You could apply the FPC, but when you have sampled less than 5% of the population, applying the FPC does not "reward" you very much, and is not worth the extra effort to calculate.*

[3] *The normal approximation is acceptably accurate given a reasonably large sample size. See discussion under Section I.C. for additional guidance.*

Where:

p_1 = proportion the first year.
p_2 = proportion the second year.
n_1 = sample size the first year.
n_2 = sample size the second year.

$$SE_{(p_1-p_2)} = \sqrt{\left(\frac{33(100-33)}{200}\right)+\left(\frac{49(100-49)}{200}\right)} = \sqrt{11.055 = 12.495} = 4.85$$

Then, calculate the 95% confidence interval:

$\text{Freq}_{\text{diff}} \pm (Z_\alpha)(SE_{(p_1-p_2)})$

$16\% \pm (1.96)(4.85)$ The 95% confidence interval is $16\% \pm 9.5\%$

Where $Z_\alpha = 1.96$, the standard normal deviate for the 95% confidence level. If you wish to use another confidence level: $Z_\alpha = 2.58$ at the 99% confidence level, $Z_\alpha = 1.64$ at the 90% confidence level, and $Z_\alpha = 1.28$ at the 80% confidence level.

If you have sampled more than 5% of the population you should adjust the confidence interval by multiplying it (9.5% in this example) by the FPC. This example assumes that less than 5% of the population has been sampled.

D. The Difference between Two Proportions Using Paired Sampling Units

Now, let us place 100 permanent frequency quadrats in the same plant population. This time, we will place 10 quadrats along 10 permanently marked transects. The quadrats are placed in the same position in both years. In the first year, 43 quadrats (43%) contained the plant. In the second year, 3 of those quadrats that contained the plant in the first year no longer contained it, but 12 quadrats that lacked the plant in the first year now contained it. A total of 52 quadrats (52%) contained the plant in year 2. What is the 95% confidence interval around the difference in frequency ($\text{Freq}_{\text{diff}}$) of 9%? Again, we will use a normal approximation.[4] Here are the fates of the quadrats:

	Present Year 2	Absent Year 2
Present Year 1	40 (PP)	3 (PA)
Absent Year 1	12 (AP)	45 (AA)

First, calculate the standard error of the difference:

$$SE_{diff} = 100\left(\frac{1}{n}\right)\left(\sqrt{(AP+PA)-\frac{(AP-PA)^2}{n}}\right) = 100\left(\frac{1}{100}\right)\left(\sqrt{(12+3)-\frac{(12-3)^2}{100}}\right)$$

$$= 100\left(\frac{1}{100}\right)\left(\sqrt{15-\frac{(9)^2}{100}}\right) = 3.77$$

Now calculate the 95% confidence interval:

$\text{Freq}_{\text{diff}} \pm (Z_\alpha)(SE_{diff})$ $\qquad 9\% \pm (1.96)(3.77)$

The 95% confidence interval of the difference in frequency between years 1 and 2 is 9% ± 7.39%.

[4] *The normal approximation provides acceptable results if sample sizes are reasonably large. See Section I.C. for additional discussion and guidance.*

Where $Z_\alpha = 1.96$, the standard normal deviate for the 95% confidence level. If you wish to use another confidence level: $Z_\alpha = 2.58$ at the 99% confidence level, $Z_\alpha = 1.64$ at the 90% confidence level, and $Z_\alpha = 1.28$ at the 80% confidence level.

If you have sampled more than 5% of the population you should adjust the confidence interval by multiplying it (7.39% in this example) by the FPC. This example assumes that less than 5% of the population has been sampled.

LITERATURE CITED

Cochran, W. G. 1977. Sampling techniques, 3rd ed. New York, NY: John Wiley & Sons.

Cohen, J. 1994. The earth is round ($p < 0.05$). American Psychologist 49:997–1003.

Krebs, C. J. 1998. Ecological methodology. New York, NY: Harper & Row.

Zar, J. H. 1999. Biostatistical analysis, 4th ed. Upper Saddle River, NJ: Prentice Hall.

APPENDIX IV
Sample Size and Confidence Intervals for Complex Sampling Designs

The examples presented in this appendix have been adapted from: Platts, W. S.; Armour, C.; Booth, G. D.; Bryant, M.; Bufford, J. L.; Cuplin, P.; Jensen, S.; Lienkaemper, G. W.; Minshall, G. W.; Monsen, S. B.; Nelson, R. L.; Sedell, J. R.; Tuhy, J. S. 1987. Methods for evaluating riparian habitats with applications to management. General Technical Report INT-221. Ogden, Utah: USDA Forest Service, Intermountain Research Station.

STRATIFIED RANDOM SAMPLING

If the population of interest falls naturally into several subdivisions, or strata, stratified random sampling is found to be substantially more efficient than simple random sampling. For example, if the number of shrubs is a management concern in a riparian zone that extends through several homogeneous vegetation types (such as sagebrush, sagebrush-grass, and ponderosa pine-Idaho fescue), this method of sampling is suitable. This procedures requires that the investigator clearly identify each stratum in advance of sampling. Then a simple random sample (SRS) is taken independently within each stratum.

In addition to being more efficient in estimating the overall population mean or total, stratified random sampling provides separate estimates for each stratum. This feature alone might be reason enough for using this method over SRS.

Example 1. Assuming that the following information is collected from three strata, what are the mean number of shrubs per acre and the 95 percent confidence interval for the population mean (m)? Sample means and variances were calculated for each stratum. Approximately 13 percent of the acres were sampled in each stratum. This is a finite population with three strata such that $N_1 = 155$, $N_2 = 62$, and $N_3 = 93$.

Stratum	Total acres/ stratum (N_h)	Total acres sampled (n_h)	Sample statum mean $\bar{X}_h$	Total shrubs $N_h\bar{X}_h$	Straum variance s_h^2	$N_h s_h^2$
1 Sagebrush	155	20	33.900	5,254.500	35.358	5,480.49
2 Sagebrush-grass	62	8	25.125	1,557.750	232.411	14,409.48
3 Ponderosa pine-Idaho Fescue	93	12	19.000	1,767.000	87.636	8,150.15
	310	40		8,578.750		28,040.12

$N = \Sigma N_h = 310$ $\quad n = \Sigma n_h = 40$ $\quad T = \Sigma N_h\bar{X}_h = 8,578.750$ $\quad s^2 = \Sigma N_h s_h^2 = 28,040.12$

Step 1 - Calculate sample mean

$$\bar{X}_{st} = \frac{T}{N}$$

$$= \frac{8{,}578.750}{310} = 27.673$$

= sample estimate of μ, the population mean number of shrubs per acre

Step 2 - Calculate an estimate of the variance of $\bar{X}_{st}$

$$\hat{V}(\bar{X}_{st}) = \frac{1}{N^2} \Sigma \left[N_h^2 \left(\frac{N_h - n_h}{N_h} \right) \left(\frac{s_h^2}{n_h} \right) \right]$$

$$= \frac{1}{(310)^2} \left[(155)^2 \left(\frac{155 - 20}{155} \right) \left(\frac{35.358}{20} \right) + (62)^2 \left(\frac{62 - 8}{62} \right) \left(\frac{(232.411)}{8} \right) + (93)^2 \left(\frac{93 - 12}{93} \right) \left(\frac{(87.636)}{12} \right) \right]$$

$$= \frac{1}{(310)^2} (36{,}993.308 + 97{,}264.004 + 55{,}013.499)$$

$$= \frac{189{,}270.81}{96{,}100} = 1.970$$

Step 3 - Calculate the 95 percent confidence interval for the population mean (m) number of shrubs per acre.

The interval is calculated as:

Lower limit: $\bar{X}_{st} - (Z)\sqrt{\hat{V}(\bar{X}_{st})} = 27.673 - (1.96)\sqrt{1.970} = 27.673 - 2.751 = 24.922$ shrubs per acre

Upper limit: $\bar{X}_{st} + (Z)\sqrt{\hat{V}(\bar{X}_{st})} = 27.673 + (1.96)\sqrt{1.970} = 27.673 + 2.751 = 30.424$ shrubs per acre

Where Z = 1.96, the standard normal deviate for the 95% confidence level. If you wish to use another confidence level: Z = 2.58 at the 99% confidence level, Z = 1.64 at the 90% confidence level, and Z = 1.28 at the 80% confidence level.

Example 2. What should the sample size be for each stratum if we want to be 95 percent confident that the confidence interval width, (B), is no larger than ±2.0 shrubs/acre?

Step 1 - Calculate the denominator for stratum weights

$$\text{Denominator} = \Sigma N_h S_h$$

$$= (155)\sqrt{35.358} + (62)\sqrt{232.411} + (93)\sqrt{87.636}$$

$$= 921.67 + 945.19 + 870.61$$

$$= 2{,}737.47$$

Step 2 - Calculate the stratum weights

$$w_h = \frac{N_h s_h}{\Sigma N_h s_h}$$

= the proportion of the total sample size, n, that will come from stratum h.

$$w_1 = \frac{921.67}{2{,}737.47} = 0.337$$

$$w_2 = \frac{945.19}{2{,}737.393} = 0.345$$

$$w_3 = \frac{870.573}{2{,}737.393} = 0.318$$

Notice that the weights over all three strata add up to 1.000. To determine the size of sample required from stratum h, multiply the total sample size by w_h. Therefore,

$n_h = w_h n.$

We still need to determine the overall sample size, n.

Step 3 - Calculate the numerator for the n' equation.

$$\text{Numerator} = \sum \frac{N_h^2 s_h^2}{w_h}$$

$$= \frac{(155)^2(35.358)}{0.337} + \frac{(62)^2(232.411)}{0.345} + \frac{(93)^2(87.636)}{0.318}$$

$$= 2{,}520{,}700.148 + 2{,}589{,}530.099 + 2{,}383{,}533.849$$

$$= 7{,}493{,}764.096$$

Step 4 Calculate n'

$$D = \frac{B^2}{Z^2} = \frac{(2.0)^2}{(1.96)^2}$$

= 1.041, where Z = 1.96, the standard normal deviate for the 95% confidence level (see Example 1, above, for values for other confidence levels).

$$\text{Finally,} \quad n' = \frac{\text{Numerator}}{N^2 D + s^2}$$

$$= \frac{7{,}493{,}764.096}{(310)^2(1.041) + 28{,}040.12} = \frac{7{,}493{,}764.096}{100{,}040.10 + 28{,}040.12}$$

$$= \frac{7{,}493{,}764.096}{128{,}080.22} = 58.508 \text{ or } 59$$

Therefore, an overall sample of n = 59 should give the investigator high probability of obtaining an estimate that is no more than 2.0 shrubs per acre from the population mean being estimated.

Step 5 - Calculate sample size for each stratum

$n_1 = w_1 n' = (0.337)(59) = 19.883$ or 20
$n_2 = w_2 n' = (0.345)(59) = 20.355$ or 20
$n_3 = w_3 n' = (0.318)(59) = 18.762$ or 19
Total 59

NOTE: The weights, w_h, were determined in such a way that the variance of $\bar{X}_{st}$ is minimized for a fixed value of n. Therefore, once we determined an estimate of n, say n', we applied the weights to it to obtain the sample size in each stratum.

Example 3. Using the results of example 2, what is the estimate of the total number of shrubs in the three strata, the 95 percent confidence interval for the estimate, and the estimated number of samples that would have to be collected for B, the confidence interval width, not to exceed ±400 shrubs?

Step 1 - Calculate the value for $\hat{\tau}$, the estimate of the population total number of shrubs

$$\hat{\tau} = N\bar{X}_{st}$$

$$= (310)(27.673)$$

$$= 8{,}578.630 \text{ shrubs}$$

Step 2 - Calculate the estimated variance of $\hat{t}$

$$\hat{V}(N\bar{X}_{st}) = N^2\hat{V}(\bar{X}_{st})$$
$$= (310)^2(1.970)$$
$$= 189{,}317$$

Step 3 - Calculate the 95 percent confidence interval for the total number of shrubs in the population.

The interval is computed as:

Lower limit: $\hat{t}_{st} - Z\sqrt{\hat{V}(N\bar{X}_{st})} = 8{,}578.63 - 1.96\sqrt{189{,}317} = 7{,}725.82$

Upper limit: $\hat{t}_{st} + Z\sqrt{\hat{V}(N\bar{X}_{st})} = 8{,}578.63 + 1.96\sqrt{189{,}317} = 9{,}431.44$

Step 4 - Calculate n', the estimated sample size for B not to exceed ±400 shrubs.

The only difference between this case and the estimation of μ in example 2 is in the computation of D. We now have

$$D = \frac{B^2}{Z^2N^2} = \frac{(400)^2}{(1.96)^2(310)^2} = 0.433$$

Where Z is from a table of the normal distribution for 95 percent confidence level.

$$n' = \frac{\text{Numerator}}{N^2D + s^2} = \frac{7{,}493{,}764.096}{(310)^2(0.433) + 28{,}040.12}$$
$$= \frac{7{,}509{,}992.786}{69{,}651.420}$$
$$= 107.59 \text{ or } 108 \text{ rounded up}$$

We can apply the weights from example 2 to obtain the sample sizes for each stratum. We get

$n_1 = (0.337)\ 108 = 36.40$ or 36
$n_2 = (0.345)\ 108 = 37.26$ or 37
$n_3 = (0.318)\ 108 = 34.34$ or 34

CLUSTER SAMPLING

Cluster sampling should not be confused with cluster analysis, which is a classification and taxonomic technique. Here, cluster sampling refers to a method of collecting a sample when the individual elements cannot be identified in advance. Instead, we are only able to identify groups or clusters of these elements. A sample of the clusters is then obtained, and every element in each cluster is measured.

For example, we may wish to take measurements on individual trees in a riparian area but are only able to identify 1-acre plots along the stream. Each plot can contain a different number of trees, and the individual trees cannot be identified before taking the sample. Cluster sampling allows us to select a sample of clusters, instead of individual trees. We would then measure every tree within each cluster.

Cluster sampling is convenient and inexpensive with regard to travel costs. To gain maximum advantage of this method, elements within a cluster should be close to each other geographically.

If we compare cluster sampling with either simple random sampling or stratified random sampling, we find one major advantage of the cluster method: the cost per element sampled is lower than for the other two methods. Unfortunately, two disadvantages of cluster sampling are:

(1) the variance among elements sampled tends to be higher, and (2) the computations required to analyze the results of the sample are more extensive. Therefore, cluster sampling is preferable to the other methods if the cost benefits exceed the disadvantages.

If we have only a few clusters, each quite large, we minimize our costs—especially of travel. However, samples with only a few clusters produce estimates with low precision (that is, high variance). On the other hand, if we increase the number of clusters (making each cluster smaller), the variance is reduced while the cost is increased. The user must find a compromise.

Whether sampling 40 clusters of 0.5 acre each is better than 20 clusters of a full acre each is not clear, although approximately the same number of trees may be measured with either sample. There would be a larger number of the smaller clusters, and therefore they would be dispersed more evenly over the population. The estimates produced would have lower variability than those from fewer but larger clusters. However, the sampler would have to travel to twice as many sites, thus increasing costs. Knowledge of the variability and costs involved would be the key to planning such a study effectively.

Example 4. Suppose that we have 30 clusters of 1 acre each ($N = 30$) in a riparian area. Calculate the average number of cavities per snag tree, the bound on the error of estimation (B), and the 95 percent confidence interval for the population mean (μ). Five clusters (n) are selected for sampling and data are collected for all snag trees in each cluster. Sampling data are tabulated below:

Cluster	Number of snag trees (m_i)	Total cavities (X_i)
1	8	5
2	9	7
3	4	8
4	5	9
5	6	10
	$\Sigma m_i = 32$	$\Sigma X_i = 39$

Step 1 - Calculate an estimate of μ, the population mean, for cavities per snag tree

$$\bar{X} = \frac{\Sigma X_i}{\Sigma m_i} = \frac{39}{32} = 1.22 \text{ cavities per snag tree}$$

Step 2 - Calculate $\bar{m}$, the average cluster size for the sample

$$\bar{m} = \frac{\Sigma m_i}{n} = \frac{32}{5} = 6.4 \text{ snag trees per cluster}$$

An estimate of the total number of snag trees in the 30 clusters is $N\bar{m} = (30)(6.4) = 192.0$ trees.

Step 3 - Calculate sum of squares

Cluster	m_i	X_i	$\bar{X}m_i$	$(X_i - \bar{X}m_i)^2$
1	8	5	9.76	22.66
2	9	7	10.98	15.84
3	4	8	4.88	9.73
4	5	9	6.10	8.41
5	6	10	7.32	7.18
			Total	63.82

where $\bar{X}$ came from step 1.

Step 4 - Calculate $(\hat{X})$ = estimated variance for $\bar{X}$

$$\hat{V}\bar{X} = \left(\frac{N - n}{(N)(n)(m)^2}\right)\left(\frac{\Sigma(X_i - \bar{X}m_i)^2}{n - 1}\right)$$

$$= \left(\frac{30 - 5}{(30)(5)(6.4)^2}\right)\left(\frac{63.82}{4}\right)$$

$$= (0.004)(15.955) = 0.0649$$

Step 5 - Calculate the 95 percent confidence interval for the population mean number of cavities per snag tree:

Lower limit: $\bar{X} - \left(Z\sqrt{\hat{V}(\bar{X})}\right) = 1.22 - 0.4994 = 0.7206$

Upper limit: $\bar{X} + \left(Z\sqrt{\hat{V}(\bar{X})}\right) = 1.22 + 0.4994 = 1.7194.$

Where $Z = 1.96$, the standard normal deviate for the 95% confidence level. If you wish to use another confidence level: $Z = 2.58$ at the 99% confidence level, $Z = 1.64$ at the 90% confidence level, and $Z = 1.28$ at the 80% confidence level.

Example 5. Assuming that information for example 4 is preliminary, how can we determine the number of clusters to sample if we want the confidence interval width, (B), to be within ±0.1?

Step 1 - Calculate s_c^2= estimate of the population variance among clusters

$$s_c^2 = \frac{\Sigma(X_i - \bar{X}m_i)^2}{n - 1}$$

$$= \frac{63.82}{4} = 15.955$$

Step 2 - Calculate

$$D = \frac{B^2\bar{m}^2}{Z^2} = \frac{(0.1)^2(6.4)^2}{(1.96)^2} = 0.1066$$

where:

1.96 is the Z value from the normal distribution for 95 percent confidence level.

Step 3 - Calculate n′ = total number of clusters to sample

$$n' = \frac{(N)(s_c^2)}{ND + s_c^2} = \frac{(30)(15.955)}{(30)(0.1066) + 15.955}$$

$$= \frac{(30)(15.955)}{19.153} = 24.99 \text{ or 25 clusters rounded up}$$

TWO-STAGE SAMPLING

Suppose we have clusters with so many elements in them that it is prohibitive to measure all elements in the cluster. It is natural to think of sampling elements within each cluster—that is, to measure only part of the elements within each cluster. This situation is a common one and is referred to as two-stage sampling.

Another common use of two-stage sampling is when it is apparent that even though there are many elements within a cluster, all elements are so nearly the same that to sample all of them would provide little additional information. The reasonable thing to do might be to measure only a part of the elements available within the cluster.

Two-stage sampling introduces a high degree of flexibility in defining clusters and sampling within them. The give and take between the number of clusters and the number of elements to be sampled within each cluster has been studied in some detail. Unfortunately, the results are complicated and beyond the scope of this publication. Interested readers are referred to one of the more extensive books on sampling (Cochran 1963; Kish 1965).

The following examples serve to give the reader a brief introduction to the concepts of two-stage sampling.

Example 6. Suppose that there are $N = 90$ clusters in a riparian zone and we can sample 10 clusters (n = 10) and 20 percent of the pools in each cluster. Estimate the mean depth of pools in the population, the bounds on the error of estimation (B), and the 95 percent confidence interval for the population mean (m). Assume that there is a total of M = 4,500 pools in the 90 clusters. Data for each cluster have been used to calculate the cluster means ($\bar{X}_i$), and variances (s_i^2).

Step 1 - Tabulate data as follows:

cluster	total pools (M_i)	Pools sampled (m_i)	mean depth $\bar{X}_i$	$(M_i)(\bar{X}_i)$	$(M_i\bar{X}_i - \bar{M}\bar{X})^{2*}$
1	50	10	5.40	270.00	900.00
2	65	13	4.00	260.00	400.00
3	45	9	5.67	255.15	229.52
4	48	10	4.80	230.40	92.16
5	52	10	4.30	223.60	268.96
6	58	12	3.83	222.14	318.98
7	42	8	5.00	210.00	900.00
8	66	13	3.85	254.10	198.81
9	40	8	4.88	195.20	2,007.04
10	56	11	5.00	280.00	1,600.00
	$\Sigma M_i = 522$			$\Sigma(M_i\bar{X}_i) = 2{,}400.59$	$\Sigma(M_i\bar{X}_i - \bar{M}\bar{X})^2 = 6{,}915.47$

*Calculated $\bar{M}$ and $\bar{X}$ from Step 2 and Step 3 below

cluster	s_i^2	$M_i(M_i - m_i) = A_i$	$s_i^2/m_i = B_i$	$(A_i)(B_i)$
1	11.38	2,000	1.138	2,276.00
2	10.67	3,380	0.821	2,774.98
3	16.75	1,620	1.861	3,014.82
4	13.29	1,824	1.329	2,424.10
5	11.12	2,184	1.112	2,428.61
6	14.88	2,668	1.240	3,308.32
7	5.14	1,428	0.643	918.20
8	4.31	3,498	0.332	1,161.34
9	6.13	1,280	0.766	980.48
10	11.80	2,520	1.073	2,703.96
				$\Sigma M_i(M_i - m_i)\frac{s_i^2}{m_i} = 21{,}990.81$

Step 2 - Calculate $\bar{M}$ = average number of elements (pools) in each cluster

$$\bar{M} = \frac{M}{N} = \frac{4{,}500}{90} = 50 \text{ pools}$$

Step 3 - Calculate $\bar{X}$ = the estimated population mean depth for pools

$$\bar{X} = \frac{N}{(M)(n)} \Sigma M_i \bar{X}_i$$

$$= \frac{90}{(4{,}500)(10)} (2{,}400.59) = 4.8012 \text{ ft deep}$$

Step 4 - Calculate the estimated variance for $\bar{X}$

A. Calculate:

$$s_b^2 = \frac{1}{n-1} \Sigma(M_i\bar{X}_i - \bar{M}\bar{X})^2 = \frac{1}{10-1}(6{,}915.47)$$

$$= \frac{6{,}915.47}{9} = 768.4;$$

B. and calculate:

$$\hat{V}(\bar{X}) = \left[\left(\frac{N-n}{N}\right)\left(\frac{1}{n\bar{M}^2}\right)(s_b^2)\right] + \left(\frac{1}{nN\bar{M}^2}\right)\left[\Sigma M_i (M_i - m_i)\left(\frac{s_i^2}{m_i}\right)\right]$$

$$= \left[\left(\frac{90-10}{90}\right)\left(\frac{1}{(10)(50)^2}\right)(768.4)\right] + \left[\frac{1}{(10)(90)(50)^2}\right](21{,}990.81)$$

$$= 0.037095$$

Step 5 - Calculate the 95 percent confidence interval for the population mean pool depth (μ), which is:

Lower limit: $\bar{X} - Z\sqrt{\hat{V}(\bar{X})} = 4.8012 - 1.96\sqrt{0.037095} = 4.42$

Upper limit: $\bar{X} + Z\sqrt{\hat{V}(\bar{X})} = 4.8012 + 1.96\sqrt{0.037095} = 5.18$

Where $Z = 1.96$, the standard normal deviate for the 95% confidence level. If you wish to use another confidence level: $Z = 2.58$ at the 99% confidence level, $Z = 1.64$ at the 90% confidence level, and $Z = 1.28$ at the 80% confidence level.

Example 7. If M is unknown in example 6, calculate the estimate of the population mean depth of pools, and the 95 percent confidence interval for the population mean depth of pools.

Step 1 - Estimate μ = ratio estimate of the population mean μ

$$\bar{X}_r = \frac{\Sigma M_i \bar{X}_i}{\Sigma M_i} = \frac{2{,}400.59}{522} = 4.599 \text{ ft}$$

Step 2 - Complete tabulations for extension of table for example 6

$M_i^2\bar{X}_i$	$(M_i\bar{X}_i)^2$	M_i^2
13,500.00	72,900.00	2,500
16,900.00	67,600.00	4,225
11,481.75	65,101.52	2,025
11,059.20	53,084.16	2,304
11,627.20	49,996.96	2,704
12,884.12	49,346.18	3,364
8,820.00	44,100.00	1,764
16,770.60	64,566.81	4,356
7,808.00	38,103.04	1,600
15,680.00	78,400.00	3,136
$\Sigma M_i^2\bar{X}_i$ = 126,530.87	$\Sigma(M_i\bar{X}_i)2$ = 583,198.67	ΣM_i^2 = 27,978

Step 3 - Calculate $\bar{M}$ = estimate of average number of pools per cluster

$$\bar{M} = \frac{\Sigma M_i}{n} = \frac{522}{10} = 52.2 \text{ pools per cluster}$$

Step 4 - Calculate estimated variance for μ
A. Calculate s_r^2:

$$s_r^2 = \frac{1}{n-1} \Sigma M_i^2 (X_i - \bar{X}_r)^2$$

$$= \frac{1}{n-1} \left[\Sigma (M_i \bar{X}_i)^2 - 2\bar{X}_r \Sigma M_i^2 \bar{X}_i + (\bar{X}_r)^2 \Sigma M_i^2 \right]$$

$$= \frac{583{,}198.67 - 2(4.599)(126{,}530.87) + (4.599)^2 (27{,}978)}{9}$$

$$= \frac{583{,}198.67 - 1{,}163{,}830.94 + 591{,}757.11}{9}$$

$$= \frac{11{,}124.84}{9} = 1{,}236.09;$$

B. and calculate $\hat{V}(\bar{X}_r)$, the estimated variance of $\bar{X}_r$

$$\hat{V}(\hat{\mu}) = \left(\frac{N-n}{N}\right)\left(\frac{1}{n\bar{M}^2}\right)(s_r^2) + \left(\frac{1}{nN\bar{M}^2}\right) \Sigma M_i (M_i - m_i)\left(\frac{s_i^2}{m_i}\right)$$

$$= \left(\frac{90-10}{90}\right)\left(\frac{1}{(10)(52.2)^2}\right)(1{,}236.09) + \left(\frac{1}{(10)(90)(52.2)^2}\right)(21{,}990.81)$$

$$= \left(\frac{80}{90}\right)\left(\frac{1}{(10)(2{,}724.84)}\right)(1{,}236.09) + \left(\frac{1}{(10)(90)(52.2)^2}\right)(21{,}990.81)$$

$$= \left(\frac{80}{2{,}452{,}356}\right)(1{,}236.09) + \left(\frac{1}{2{,}452{,}356}\right)(21{,}990.81)$$

$$= 0.0403 + 0.0090 = 0.0493$$

Step 5 - Calculate the 95 percent confidence interval for the population mean (μ) for pool depth, which is:

Lower limit: $\bar{X}_r - \left(Z\sqrt{\hat{V}(\hat{\mu})}\right) = 4.599 - 0.435 = 4.164$ ft.

Upper limit: $\bar{X}_r + \left(Z\sqrt{\hat{V}(\hat{\mu})}\right) = 4.599 + 0.435 = 5.034$ ft.

LITERATURE CITED

Cochran, W. G. 1963. Sampling techniques, 2nd ed. New York, NY: John Wiley & Sons.

Kish, L. 1965. Survey sampling. New York, NY: John Wiley & Sons.

LITERATURE CITED

Ahmed, J., Bonham, C.D., Laycock, W.A. 1983. Comparison of techniques used for adjusting biomass estimates by double sampling. Journal of Range Management 36:217–221.

American Ornithologists' Union. 1998. Check-list of North American birds: the species of birds of North America from the Arctic through Panama, including the West Indies and Hawaiian Islands (7th ed.) (p. 829). Washington, DC.

Andren, H. 1996. Population responses to habitat fragmentation: statistical power and the random sample hypothesis. Oikos 76:235–242.

Ausden, M. 1996. Invertebrates. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Avery, T.E., Berlin, G.L. 1992. Fundamentals of remote sensing and airphoto interpretation. New York: MacMillan.

Avery, T.E., Burkhart, H.E. 1994. Forest measurements (4th ed.). New York: John Wiley & Sons.

Awbrey, R.T. 1977. Locating random points in the field. Journal of Range Management 30:157–158.

Bailey, A.W., Poulton, C.E. 1968. Plant communities and environmental relationships in a portion of the Tillamook Burn, northwest Oregon. Ecology 49:1–14.

Baldwin, J. 1996. [Personal communication]. Albany, CA: U.S. Forest Service, Pacific Southwest Research Station.

Barcikowski, R.S., Robey, R.R. 1984. Decisions in single group repeated measures analysis: statistical tests and three computer packages. American Statistician 38:149–150.

Bartmann, R.M., White, G.C., Carpenter, L.H., Garrott, R.A. 1987. Aerial mark-recapture estimates of confined mule deer in pinyon-juniper woodland. Journal of Wildlife Management 51:41–46.

Bartz, K.L., Kershner, J.L., Ramsey, R.D., Neale, C.M.U. 1993. Delineating riparian cover types using multispectral, airborne videography. Proceedings of the 14th biennial workshop on color aerial photography and videography for resource monitoring (pp. 58–67). Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Becker, D.A., Crockett, J.J. 1973. Evaluation of sampling techniques on tall grass prairie. Journal of Range Management 26:61–67.

Becker, E.F. 1991. A terrestrial furbearer estimator based on probability sampling. Journal of Wildlife Management 55:730–737.

Beier, P., Cunningham, S.C. 1996. Power of track surveys to detect changes in cougar populations. Wildlife Society Bulletin 24:540–546.

Benoit, D.L., Kenkel, N.C., Cavers, P.B. 1989. Factors influencing the precision of soil seed bank estimates. Canadian Journal of Botany 67:2833–2840.

Bernstein, B.B., Zalinski, J. 1983. An optimal sampling design and power tests for environmental biologists. Journal of Environmental Sciences 16:35–43.

Bibby, C.J., Burgess, N.D., Hill, D.A. 1992. Bird census techniques. London: Academic Press.

Blackwood, L.G. 1991. Assurance levels of standard sample size formulas. Environmental Science and Technology 25:1366–1367.

Blomberg, S., Shine, R. 1996. Reptiles. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Blyth, K. 1982. On robust distance-based density estimators. Biometrics 38:127–135.

Bonham, C.D. 1989. Measurements for terrestrial vegetation. New York: John Wiley & Sons.

Bowden, D.C., Kufeld, R.C. 1995. Generalized mark-sight population size estimation applied to Colorado moose. Journal of Wildlife Management 59:840–851.

Brady, W.W., Mitchell, J.E., Bonham, C.D., Cook, J.W. 1995. Assessing the power of the point-line transect to monitor changes in basal plant cover. Journal of Range Management 48:187–190.

Braun-Blanquet, J. 1965. Plant sociology: the study of plant communities. London: Hafner.

Brewer, J.K. 1972. On the power of statistical tests in the American Educational Research Journal. American Educational Research Journal 9:391–401.

Brewer, L., Berrier, D. 1984. Photographic techniques for monitoring resource change at backcountry sites (p. 13). General Technical Report NE-86. Radnor, PA: USDA Forest Service. Northeastern Forest Experiment Station.

Brown, B.W., Brauner, C., Chan, A., Gutierrez, D., Herson, J., Lovato, J., Polsley, J. 1996. STPLAN, Version 4.1. Houston, TX: University of Texas, M.D. Anderson Cancer Center, Department of Biomath-ematics.

Brun, J.M., Box, W.T. 1963. Comparison of line intercepts and random point frames for sampling desert shrub vegetation. Journal of Range Management 16:21–25.

Buckland, S.T.,Anderson, D.R., Burnham, K.P., Laake, J.L. 1993. Distance sampling: estimating abundance of biological populations. New York: Chapman & Hall.

Buell, M.F., Cantlon, J.E. 1950. A study of two communities of New Jersey pine barrens and a comparison of methods. Ecology 31:567–586.

Bullock, D.J., Evans, P.G.H. 1990. The distribution, density and biomass of terrestrial reptiles in Dominica, West Indies. Journal of Zoology 222:421–443.

Burgman, M.A., Ferson, S., Akcakaya, H.R. 1993. Risk assessment in conservation biology (pp. 314). London: Chapman & Hall.

Burleson, W.H. 1975. A method of mounting plant specimens in the field. Journal of Range Management 28:240–241.

Burton, T.M., Likens, G.E. 1975. Salamander populations and biomass in the Hubbard Brook Experimental Forest, New Hampshire. Copeia 1975:541–546.

Catana, A.J. 1963. The wandering quarter method of estimating population density. Ecology 44:349–360.

Catchpole, W.R., Wheeler, C.J. 1992. Estimating plant biomass: A review of the techniques. Australian Journal of Ecology 17:121–131

Caughley, G. 1977. Analysis of vertebrate populations. London: Wiley.

Chambers, J.C., Brown, R.W. 1983. Methods for vegetation sampling and analysis on revegetated mined lands. Ogden, UT: U.S. Forest Service Intermountain Research Station, General Technical Report INT-151.

Chapman, D.G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publications in Statistics 1:131–160.

Chernick, M.R. 1999. Bootstrap methods: a practitioner's guide. New York: Wiley-Interscience.

Clymo, R.S. 1980. Preliminary survey of the Peat-Bog Knowe Moss using various numerical methods. Vegetation 42:129–148.

Cochran, W.G. 1977. Sampling techniques (3rd ed.). New York: John Wiley & Sons.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. 1994. The earth is round ($p < 0.05$). American Psychologist 49:997–1003.

Collins, J.T. 1997. Standard common and current scientific names for North American amphibians and reptiles (4th ed.) (p. 40). Society for the Study of Amphibians and Reptiles. Herpetol. Circ. No. 25.

Crouse, D.T., Crowder L.B., Caswell, H. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology 68:1412–1423.

D'Agostino, R.B., Belanger, A., D'Agostino, R.B., Jr. 1990. A suggestion for using powerful and informative tests of normality. American Statistician 44:316–321.

Dahl, T.E., Johnson, C.E. 1991. Status and trends of wetlands in the conterminous United States, mid-1970's to mid-1980's. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service.

Daniel, W.W. 1990. Applied nonparametric statistics (2nd ed.). Boston, MA: PWS-KENT Publishing.

Daubenmire, R.F. 1959. A canopy-coverage method. Northwest Science 33:43–64.

Davidson, A.C., Hinkley, D.V. 1997. Bootstrap methods and their application. London: Cambridge University Press.

Day, R.W., Quinn, G.P. 1989. Comparison of treatments after an analysis of variance in ecology. Ecological Monographs 59:433–463.

Dennis, A. 1993. Sampling methods for vegetation description and monitoring. Unpublished draft on file at USDA Forest Service, Pacific Southwest Station/Region 5, Research Natural Areas Program, Berkeley, CA.

DeSante, D.F. 1992. Monitoring avian productivity and survivorship (MAPS): a sharp, rather than blunt, tool for monitoring and assessing bird populations. In: D.R. McCullough, R.H. Barrett (eds.), Wildlife 2001: populations (pp. 511–521). Essex, UK: Elsevier Applied Science.

Dethier, M.N., Graham, E.S., Cohen, S., Tear, L.M. 1993. Visual versus random-point percent cover estimations: "objective" is not always better. Marine Ecology Progress Series 96:93–100.

DeVries, P.G. 1979. Line intersect sampling—statistical theory, applications, and suggestions for extended use in ecological inventory. In: R.M. Cormack, G.P. Patil, D.S. Robson (eds.), Sampling biological populations: Vol. 1. Statistical ecology (pp. 1–70). Fairland, MD: International Cooperative Publishing House.

Digby, P.G.N., Kempton, R.A. 1987. Mutivariate analysis of ecological communities. London: Chapman & Hall.

Diggle, P.J. 1975. Robust density estimation using distance methods. Biometrika 62:39–48.

Dilworth, J.R. 1989. Log scaling and timber cruising. Corvallis, OR: Oregon State University.

Dilworth, J.R., Bell, J.F. 1973. Variable probability sampling: variable plot and 3P. Corvalis OR: Oregon State University.

Dixon, W.J., Massey, F.J., Jr. 1983. Introduction to statistical analysis (4th ed.). New York: McGraw-Hill.

Doak, D.F., Mills, L.S. 1994. A useful role for theory in conservation. Ecology 75:615–626.

Droege, S., Cyr, A., Larivee, J. 1998. Checklists: an under-used tool for the inventory and monitoring of plants and animals. Conservation Biology 12:1134–1138.

Efron, B., Tibshirani, R. 1991. Statistical data analysis in the computer age. Science 253:390–395.

Efron, B., Tibshirani, R. 1993. An introduction to the bootstrap. New York: Chapman & Hall.

Ellison, A.M. 1993. Exploratory data analysis and graphic display. In: S.M. Scheiner, J. Gurevitch (eds.), Design and analysis of ecological experiments. New York: Chapman & Hall.

Ellison, A.M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6:1036–1046.

Elzinga, C. 1997. Habitat conservation assessment and conservation strategy for *Primula alcalina*. Unpublished paper on file at U.S. Department of Interior, Bureau of Land Management, Salmon District, Salmon, ID.

Elzinga, C., Salzer, D., Willoughby, J. 1998. Measuring and monitoring plant populations. Technical Reference 1730-1. Denver, CO: Bureau of Land Management, National Business Center.

Engeman, R.M., Sugihara, R.T, Pank, L.F., Dusenberry, W.E. 1994. A comparison of plotless density estimators using Monte Carlo simulation. Ecology 75:1769–1779.

Eschmeyer, W. N. 1998. Catalog of fishes. San Francisco: California Academy of Sciences.

Evans, R.A., Love, R.M. 1957. The step-point method of sampling: a practical tool in range research. Journal of Range Management 10:208–212.

Fairweather, P.G. 1991. Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research 42:555–567.

Faith, D.P., Dostine, P.L., Humphrey, C.L. 1995. Detection of mining impacts on aquatic macroinvertebrate communities: results of a disturbance experiment and the design of a multivariate BACIP monitoring program at Coronation Hill, Northern Territory. Australian Journal of Ecology 20:167–180.

Fellers, G.M., Drost, C.A. 1994. Sampling with artificial cover. In: W.R. Heyer, M.A. Donnelly, R.W. McDiarmid, L.-A.C. Hayek, M.S. Foster (eds.), Measuring and monitoring biological diversity: standard methods for amphibians (pp. 146–150). Washington, DC: Smithsonian Institution Press.

Fisser, H.G., VanDyne, G.M. 1966. Influence of number and spacing of points on accuracy and precision of basal cover estimates. Journal of Range Management 19:205–211.

Fitch, H.S. 1987. Collecting and life-history techniques. In: R.A. Seigel, J.T. Collins, S.S. Novak (eds.), Snakes: ecology and evolutionary biology (pp. 143–164). New York: Macmillan.

Fitch, H.S. 1992. Methods of sampling snake populations and their relative success. Herpetological Review 23:17–19.

Floyd, D.A., Anderson, J.E. 1982. A new point frame for estimating cover of vegetation. Vegetation 50:185–186.

Floyd, D.A., Anderson J.E. 1987. A comparison of three methods for estimating plant cover. Journal of Ecology 75:221–228.

Foresman, K.R., Pearson, D.E. 1998. Comparison of proposed survey procedures for detection of forest carnivores. Journal of Wildlife Management 62:1217–1226.

Fosberg, F.R., Satchet, M. 1965. Manual for tropical herbaria. Utrecht, Netherlands: International Association of Plant Taxonomy and Nomenclature. Regnum Vegetabile 39:1–132.

Foster, M.S., Harrold, C., Hardin, D.D. 1991. Point versus photo quadrat estimates of the cover of sessile marine organisms. Journal of Experimental Marine Biology and Ecology 146:193–203.

Fowler, J., Cohen, L. 1998. Practical statistics for field biology. Open University Press. Philadelphia, PA: Milton Keynes.

Frank, N., Ramus, E. 1995. A complete guide to scientific and common names of reptiles and amphibians of the world. Pottsville: Ramus/NG Publishing.

Freddy, D.J., Bowden, D.C. 1983. Sampling mule deer pellet-group densities in juniper-pinyon woodland. Journal of Wildlife Management 47:476–485.

Frohn, R.C. 1998. Remote sensing for landscape ecology: new metric indicators for monitoring, modeling and assessment of ecosystems. Lewis Publishers.

Frost, Darrel R. 1999. Amphibian species of the world: an online reference. V2.1.

Fuller, T. K. 1992. Do pellet counts index white-tailed deer numbers and population change?: a reply. Journal of Wildlife Management 56:613.

Fulton, M.R. 1996. The digital stopwatch as a source of random numbers. Bulletin of the Ecological Society of America 77:217–218.

Garner, D.L., Underwood, H.B., Porter, W.F. 1995. Use of modern infrared thermography for wildlife population surveys. Environmental Management 19:233–238.

Gauch, H.G. 1982. Multivariate analysis in community ecology. Cambridge, England: Cambridge University Press.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68:1364–1372

Gibbons, D.W., Hill D.A., Sutherland, W.J. 1996. Birds. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Gibbons, J.W., Semlitsch, R.D. 1981. Terrestrial drift fences with pitfall traps: an effective technique for quantitative sampling of animal populations. Brimleyana 7:1–16.

Gibbs, J.P. 2000. Monitoring populations. In: L. Boitani, T.K. Fuller, (eds.), Research techniques in animal ecology: controversies and consequences (pp.213–252). New York: Columbia University Press.

Gibbs, J.P., Droege, S., Eagle, P. 1998. Monitoring populations of plants and animals. BioScience 48: 935–940.

Gibbs, J.P., Stanton, E.J. 2000. Habitat fragmentation and arthropod community change: carrion beetles, phoretic mites, and flies. Ecological Applications.

Gilbert, R.O. 1987. Statistical methods for environmental pollution monitoring. New York: Van Nostrand Reinhold.

Glantz, S.A. 1997. Primer of biostatistics (4th ed.). New York: McGraw-Hill.

Glassberg, J. 1993. Butterflies through binoculars: a field guide to the butterflies of the Boston–New York–Washington region. Oxford: Oxford University Press.

Goebel, C.J., DeBano, L.F., Lloyd, R.D. 1958. A new method of determining forage cover and production on desert shrub vegetation. Journal of Range Management 11:244–246.

Goldsmith, F.B. (ed.). 1991. Monitoring for conservation and ecology. London: Chapman & Hall.

Goldsmith, F.B., Harrison, C.M., Morton, A.J. 1986. Description and analysis of vegetation. In: P.D. Moore, S.B. Chapman (eds.), Methods in plant ecology (2nd ed.) (pp. 437–524). Palo Alto, CA: Blackwell Scientific Publications.

Good, P. 1999. Resampling methods: a practical guide to data analysis. New York: Springer-Verlag.

Good, P. 2000. Permutation tests—a practical guide to resampling methods for testing hypotheses. New York: Springer Verlag.

Goodall, D.W. 1952. Some considerations in the use of point quadrats for the analysis of vegetation. Australian Journal of Scientific Research (Series B) 5:1–41.

Gray, J.S., Jensen, K. 1993. Feedback monitoring: a new way of protecting the environment. Trends in Ecology and Evolution 8:267–268.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. New York: John Wiley & Sons.

Green, R.H., Young, R.C. 1993. Sampling to detect rare species. Ecological Applications 3:351–356.

Greene, S. 1984. Botanical baseline monitoring in research natural areas in Oregon and Washington. In: J.L. Johnson, J.F. Franklin, R.G. Krebill (coordinators), Research natural areas: baseline monitoring and management: proceedings of a symposium (pp. 6–10), March 21, 1984, Missoula, MT. General Technical Report INT-173. Ogden, UT: USDA Forest Service, Intermountain Research Station.

Greene, S. 1993. RNA Monitoring protocols installation instructions. Unpublished draft (July 1993) on file at: USDA Forest Service, Pacific Northwest Station, Research Natural Areas Program, Corvallis, OR.

Greenwood, J.D., Jr. 1996. Basic techniques. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Greig-Smith, P. 1983. Quantitative plant ecology (3rd ed.). Berkeley, CA: University of California Press.

Gross, K.L. 1990. A comparison of methods for estimating seed numbers in the soil. Journal of Ecology 78:1079–1093.

Gunderson, L. 1999. Resilience, flexibility and adaptive management—antidotes for spurious certitude? Conservation Ecology 3:7. [online] URL:http://www/consecol.org/vol3/iss1/art7.

Hahn, G.J., Meeker, W.Q. 1991. Statistical intervals: a guide for practitioners. New York: John Wiley & Sons.

Hairston, N.G. 1989. Ecological experiments: purpose, design and execution. Cambridge, England: Cambridge University Press.

Hankin, D.G., Reeves, G.H. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Hanley, T.A. 1978. A comparison of the line interception and quadrat estimation methods of determining shrub canopy coverage. Journal of Range Management 31:60–62.

Hanski, I. 1989. Metapopulation dynamics: does it help to have more of the same? Trends in Ecology and Evolution 4:113–114.

Harper, J.L. 1977. Population biology of plants. London: Academic Press.

Hart, R.H., Laycock, W.A. 1996. Repeat photography of range and forest lands in the western United States. Journal of Range Management 49:60–67.

Hatfield, J.S., Gould IV, W.R., Hoover, B.A., Fuller, M.R., Lindquist, E.L. 1996. Detecting trends in raptor counts: power and type I error rates of various statistical tests. Wildlife Society Bulletin 24(3):505–515.

Hatton, T.J., West, N.E., Johnson, P.S. 1986. Relationship of the error associated with ocular estimation and actual total cover. Journal of Range Management 39:91–92.

Hayes, J.P. 1987. The positive approach to negative results in toxicology studies. Ecotoxicology and Environmental Safety 14:73–77.

Hayes, J.P., Steidl, R.J. 1997. Statistical power analysis and amphibian population trends. Conservation Biology 11:273–275.

Heady, H.F., Gibbens, R.P., Powell, R.W. 1959. A comparison of the charting, line intercept, and line point methods of sampling shrub types of vegetation. Journal of Range Management 12:180–188.

Heyer, W.R., Donnelly, M.A., McDiarmid, R.W., Hayek, L.C., Foster, M.S. 1994. Measuring and monitoring biological diversity, standard methods for amphibians. Washington, DC: Smithsonian Institution Press.

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

Hintze, J. 1996. NCSS 6.0.21 Jr., Help version 3.10.425. Kaysville, UT: NCSS.

Hirst, S.M. 1983. Ecological and institutional bases for long-term monitoring of fish and wildlife populations. In: J.F. Bell, T. Atterbury (eds.), Renewable resource inventories for monitoring changes and trends: proceedings of an international conference (pp. 175–178), August 15, 1983, Corvallis, OR. Corvallis, OR: Oregon State University, College of Forestry.

Hoaglin, D.C., Mosteller, F., Tukey, J.W. 1983. Understanding robust and exploratory data analysis. New York: John Wiley & Sons.

Hollander, M., Wolfe, D.A. 1999. Nonparametrical statistical methods (2nd ed.). New York: John Wiley & Sons.

Holling, C.S. (ed.). 1978. Adaptive environmental assessment and management. New York: John Wiley & Sons.

Holmgren, P, Thuresson, T. 1998. Satellite remote sensing for forestry planning—a review. Scandinavian Journal of Forest Research 13:90–110.

Hope-Simpson, J.F. 1940. On the errors in the ordinary use of subjective frequency estimations in grasslands. Journal of Ecology 28:193–209.

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

Husband, B.C., Barrett, S.C.H. 1996. A metapopulation perspective in plant population biology. Journal of Ecology 84:461–469.

Husch, B., Miller, C.I., Beers, T.W. 1982. Forest mensuration. New York: John Wiley & Sons.

James, F.C., McCulloch, C.E., Weidenfeld, D.A. 1996. New approaches to the analysis of population trends in birds. Ecology 77:13–27.

Jensen, J.R., Burkhalter, S.G., Althausen, J.D., Narumalani, S., Mackey, H.E., Jr. 1993. Integration of historical aerial photography and a geographic information system to evaluate the impact of human activities in a cypress-tupelo swamp. Proceedings of the 14th biennial workshop on color aerial photography and videography for resource monitoring (pp. 125–131). Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Jensen, M.E., Hann, W., Keane, R.E., Caratti, J., Bourgeron, P.S. 1994. ECODATA—a multiresource database and analysis system for ecosystem description and analysis. In: M.E. Jensen, P.S. Bourgeron (eds.), Eastside forest ecosystem health assessment: Vol. 2. Ecosystem management: principles and applications (pp. 203–216). General Technical Report GTR-PNW-318. Portland, OR: U.S. Department of Agriculture, Forest Service.

Jernvall, J, Wright, P.C. 1998. Diversity components of impending primate extinctions. Proceedings of the National Academy of Sciences USA 95:11279–11283.

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

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

Johnson, R. 1993. What does it all mean? Environmental Management and Assessment 26:307–312.

Jones, C., McShea, W.J., Conroy, M.J., Kunz, T. 1996. Capturing mammals. In: D.E. Wilson, F.R. Cole, J.D. Nichols, R. Rudran, M.S. Foster (eds.), Measuring and monitoring biological diversity, standard methods for mammals (pp. 115–155). Washington, DC: Smithsonian Institution Press.

Jongman, R.H.G.H., Van Tongeren, P. (eds.). 1995. Data analysis in community and landscape ecology. Cambridge: Cambridge University Press.

Kendall, K.C., Metzgar, L.H., Patterson, D.A., Steele, B.M. 1992. Power of sign surveys to monitor population trends. Ecological Applications 2(4):422–430.

Kennedy, K.A., Addison, P.A. 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. Journal of Ecology 75:151–157.

Kerans, B.L., Karr, J.R., Ahlstedt, S.A. 1992. Aquatic invertebrate assemblages: spatial and temporal differences among sampling protocols. Journal of the North American Benthological Society 11:377–390.

Klemm, D.J., Lewis, P.A., Fulk, F., Lazorchak, J.M. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA/600/4-90/030. Washington, DC: USEPA

Knapp, P.A., Warren, P.L., Hutchinson, C.F. 1990. The use of large-scale aerial photography to inventory and monitor arid rangeland vegetation. Journal of Environmental Management 31:29–38.

Koch, G.S., Link, R.F. 1970. Statistical analysis of geological data. New York: John Wiley & Sons.

Krebs, C.J. 1998. Ecological methodology (2nd ed.). New York: Harper & Row.

Kremen, C. 1994. Biological inventory using target taxa: a case study of the butterflies of Madagascar. Ecological Applications 4:407–422.

Kupper, L.L., Hafner, K.B. 1989. How appropriate are popular sample size formulas? The American Statistician 43:101–105.

Laferriere, J.E. 1987. A central location method for selecting random plots for vegetation surveys. Vegetatio 71:75–77.

Lancia, R.A., Nichols, J.D., Pollock, K.H. 1994. Estimating numbers of animals in wildlife populations. In: T.A. Bookout (ed.), Research and management techniques for wildlife and habitats. Bethesda, MD: The Wildlife Society.

Leck, M.A., Parker, V.T., Simpson, R.L. 1989. Ecology of soil seed banks. New York: Academic Press.

Lee, K.N. 1993. Compass and gyroscope: integrating science and politics for the environment. Washington DC: Island Press.

Lee, K.N. 1999. Appraising adaptive management. Conservation Ecology 3:3 [online] URL:http://www/consecol.org/vol3/iss2/art3.

Leonard, G.H., Clark, R.P. 1993. Point quadrat versus video transects estimates of the cover of benthic red algae. Marine Ecology Progress Series 101:203–208.

Leps, J., Hadicova V. 1992. How reliable are our vegetation analyses? Journal of Vegetation Science 3:119–124.

Lesica, P., Steele, B.M. 1996. A method for monitoring long-term population trends: an example using rare arctic-alpine plants. Ecological Applications 6:879–887.

Levy, E.B., Madden, E.A. 1933. The point method for pasture analysis. New Zealand Journal of Agriculture 46:267–279.

Li, J.C.R. 1964. Statistical inference (Vol. 1). Ann Arbor, MI: Edward Brothers.

Lillesand, T.M., Kiefer, R.W. 1999. Remote sensing and image interpretation (4th ed.). New York: John Wiley & Sons.

Long, B.G., Dennis, D.M., Skewes, T.D., Poiner, I.R. 1996. Detecting an environmental impact of dredging on seagrass beds with a BACIR sampling design. Aquatic Botany 53:235–243.

Lucas, H.A., Seber, G.A.F. 1977. Estimating coverage and particle density using the line intercept method. Biometricka 64:618–622.

Lunneborg, C.E. 2000. Data analysis by resampling: concepts and applications. North Scituate, MA: Duxbury Press.

Luque, S.S., Lathrop, R.G., Bognar, J.A. 1994. Temporal and spatial changes in an area of the New Jersey pine barrens landscape. Landscape Ecology 9:287–300.

Lyon, J.G., McCarthy, J. (eds.). 1995. Wetland and environmental applications of GIS. Boca Raton, FL: Lewis Publishers.

Lyon, L.J. 1968. An evaluation of density sampling methods in a shrub community. Journal of Range Management 21:16–20.

MacDonald, L.H., Smart, A.W. 1993. Beyond the guidelines: practical lessons for monitoring. Environmental Monitoring and Assessment 26:203–218.

MacDonald, L.H., Smart, A.W., Wissmar, R.C. 1991. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. EPA/910/9-91-001. Seattle, WA: United States Environmental Protection Agency, Water Division.

Malone, C.R. 1968. Determination of peak standing crop biomass of herbaceous shoots by the harvest method. American Midland Naturalist 79:429–435.

Manly B.F.J. 1992. The design and analysis of research studies. Cambridge, England: Cambridge University Press.

Manly, B.F.J. 1997. Randomization, bootstrap, and Monte Carlo methods in biology. New York and London: Chapman & Hall.

Mattson, D.E. 1984. Statistics: difficult concepts, understandable explanations. St. Louis, MO: C.V. Mosby.

McCall, C.H., Jr. 1982. Sampling and statistics handbook for research. Ames, IA: Iowa State University Press.

McGarigal, K., Cushman, S., Stafford, S.G. 2000. Multivariate statistics for wildife ecology and research. New York: Springer-Verlag.

McNeill, L., Kelly, R.D., Barnes, D.L. 1977. The use of quadrat and plotless methods in the analysis of the tree and shrub component of woodland vegetation. Proceedings of The Grassland Society of Southern Africa 12:109–113.

Meagher, T.R., Antonovics, J., Primack, R. 1978. Experimental ecological genetics in Plantago. III. Genetic variation and demography in relation to survival of *Plantago cordata*, a rare species. Biological Conservation 14:243–257.

Meese, R.J., Tomich, P.A. 1992. Dots on the rocks: a comparison of percent cover estimation methods. Journal of Experimental Marine Biology and Ecology 165:59–73.

Menges, E.S. 1986. Predicting the future of rare plant populations: demographic monitoring and modeling. Natural Areas Journal 6:13–25.

Menges, E.S. 1990. Population viability analysis for an endangered plant. Conservation Biology 4:52–62.

Merritt, R.W., Cummins, K.W. 1996. An introduction to the aquatic insects of North America. Dubuque, IA: Kendall/Hunt Publishing.

Milne, A. 1959. The centric systematic area-sample treated as a random sample. Biometrics 15:270–297.

Moore, N.W., Corbet, P.S. 1990. Guidelines for monitoring dragonfly populations. Journal of the British Dragonfly Society 6:21–23.

Morrison, R.G., Yarranton, G.A. 1970. An instrument for rapid and precise sampling of vegetation. Canadian Journal of Botany 48:293–297.

Mozingo, H.N., Williams, M. 1980. Threatened and endangered plants of Nevada. Reno: U.S. Department of Interior, Bureau of Land Management, Nevada State Office.

Muir, P.S., Moseley, R.K. 1994. Responses of *Primula alcalina*, a threatened species of alkaline seeps, to site and grazing. Natural Areas Journal 14:269–279.

Munro, B.H., Page, E.P. 1997. Statistical methods for health care research (3rd ed.). Philadelphia, PA: J.B. Lippincott.

Norton-Griffiths, M. 1973. Counting the Serengeti migratory wildebeest using two-stage sampling. East African Wildlife Journal 11:135–149.

Nowling, S., Tueller, P.T. 1993. A low-cost multispectral airborne video image system for vegetation monitoring on range and forest lands. Proceedings of the 14th biennial workshop on color aerial photography and videography for resource monitoring (pp. 1–8). Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Nyberg, J.B. 1998. Statistics and the practice of adaptive management. In: V. Sit, B. Taylor (eds.), Statistical methods for adaptive management studies. Land management handbook No. 42 (pp. 1–7). Victoria, British Columbia: Research Branch, B.C. Ministry of Forests.

Osenberg, C.W., Schmitt, R.J., Holbrook, S.J., Abu-Saba, K.E., Flegal, A.R. 1994. Detection of environmental impacts: natural variability, effect size, and power analysis. Ecological Applications 4:16–30.

Otis, D.L., Burnham, K.P., White, G.C., Anderson, D.R. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:1–135.

Palmer, M.E. 1986. A survey of rare plant monitoring: programs, regions and species priority. Natural Areas Journal 6:27–42.

Palmer, M.E. 1987. A critical look at rare plant monitoring in the United States. Biological Conservation 39:113–127.

Pavlik, B.M. 1988. Nutlet production and germination of *Amsinckia grandiflora*. I. Measurements from cultivated populations. Sacramento, CA: California Department of Fish and Game, Endangered Plants Program.

Pavlik, B.M. 1993. Demographic monitoring and recovery of endangered plants. In: M. Bowles, C. Whelen (eds.), Recovery and restoration of endangered species. Cambridge, England: Cambridge University Press.

Perrow, M.R., Cote, I.M., Evans, M. 1996. Fish. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Perry, J.A., Schaeffer, D.J., Herricks, E.E. 1987. Innovative designs for water quality monitoring: are we asking the questions before the data are collected? In: T.P. Boyle (ed.), New approaches to monitoring aquatic ecosystems, ASTM STP 940 (pp. 28–39). Philadelphia, PA: American Society for Testing and Materials.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2–15.

Peterman, R.M., M'Gognigle, M. 1992. Statistical power analysis and the precautionary principle. Marine Pollution Bulletin 24:231–234.

Peterson, C.R., Dorcas, M.E. 1994. Automated data acquisition. In: W.R. Heyer, M.A. Donnelly, R.W. McDiarmid, L.C. Hayek, M.S. Foster (eds.), Measuring and monitoring biological diversity, standard methods for amphibians (pp. 47–57). Washington, DC: Smithsonian Institution Press.

Peterson, R.T. 1990. A field guide to bird songs. Compact disc. ISBN 0-395-50257-8. Ithaca, NY: Cornell Laboratory of Ornithology.

Pierce, W.R., Eddleman, L.E. 1970. A field stereo-photographic technique for range vegetation analysis. Journal of Range Management 23:218–220.

Platts, W.S., Armour, C., Booth, G.D., Bryant, M., Bufford, J.L., Cuplin, P., Jensen, S., Lienkaemper, G.W., Minshall, G.W., Monsen, S.B., Nelson, R.L., Sedell, J.R., Tuhy, J.S. 1987. Methods for evaluating riparian habitats with applications to management. General Technical Report INT-221. Ogden, UT: U.S. Forest Service Intermountain Research Station.

Podani, J. 1994. Multivariate analysis in ecology and systematics. Champaign, IL: Balogh Scientific Books.

Poissonet, P.S., Poissonet, J.A., Bodron, M.P., Long, G.A. 1973. A comparison of sampling methods in dense herbaceous pasture. Journal of Range Management 26:65–67.

Pollard, E., Yates, T.J. 1993. Monitoring butterflies for ecology and conservation. London: Chapman & Hall.

Pollock, K., Nichols, J., Brownie, C., Hines, J. 1990. Statistical inference from capture-recapture experiments. Wildlife Monograph 107:1–97.

Prescott, G.W. 1980. How to know the aquatic plants. 2nd edition. Dubuque, IA: Wm.C. Brown Co. Publishers.

Ralph, J.C., Geupel, G.R., Pyle, P., Martin, T.E., DeSante, D.F. 1993. Handbook of field methods for monitoring landbirds. General Technical Report PSW-GTR-144. Albany, CA: U.S.Department of Agriculture, Forest Service, Pacific Southwest Research Station.

Ratliff, R.D., Westfall, S.E. 1973. A simple stereophotographic technique for analyzing small plots. Journal of Range Management 26:147–148.

Redd, T.H., Neale, C.M.U., Hardy, T.B. 1993. Use of airborne multispectral videography for the classification and delineation of riparian vegetation. Proceedings of the 14th biennial workshop on color aerial photography and videography for resource monitoring (pp. 202–211). Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Reed, J.M., Blaustein, A.R. 1995. Assessment of "nondeclining" amphibian populations using power analysis. Conservation Biology 9:1299–1300.

Reichhardt, C.S., Gollob, H.F. 1997. When confidence intervals should be used instead of significance tests, and vice versa. In: L.L. Harlow, S.A. Mulaik, J.H. Steiger (eds.), What if there were no significance tests? (pp. 259–284). Mahwah, NJ: Lawrence Erlbaum Associates Publishing.

Resh, V.H. 1979. Sampling variability and life history features: basic considerations in the design of aquatic insect studies. Journal of the Fisheries Research Board of Canada 36:290–311.

Rexstad, E.A., Burnham, K.P. 1992. User's guide for interactive program CAPTURE. Colorado Cooperative Fish and Wildlife Research Unit. Fort Collins, CO: Colorado State University.

Rice, W.R., Gaines, S.D. 1994. "Heads I win, tails you lose": testing directional alternative hypotheses in ecological and evolutionary research. TREE 9:235–237.

Ringold, P.L., Alegria, J., Czaplewski, R.L., Mulder, B.S., Tolle, T., Burnett, K. 1996. Adaptive monitoring design for ecosystem management. Ecological Applications 6(3):745–747.

Rogers, G.F., Malde, H.E., Turner, R.M. 1984. Bibliography of repeat photography for evaluating landscape change. Salt Lake City, UT: University of Utah Press.

Rosenburg, D.K., Overton, W.S., Anthony, R.G. 1995. Estimation of animal abundance when capture probabilities are low and heterogeneous. Journal of Wildlife Management 59:252–261.

Rotenberry, J.T., Wiens J.A. 1985. Statistical power analysis and community-wide patterns. American Naturalist 125:164–168.

Routledge, R.D. 1999. Measurements and estimates. In: V. Sit, B. Taylor (eds.), Statistical methods for adaptive management studies (pp. 55–68). Victoria, BC: British Colombia Ministry of Forests, Research Branch.

Rudran R., Kunz, T.H., Southwell, C., Jarman, P., Smith, A.P. 1996. Observational techniques for non-volant mammals. In: D.E. Wilson, F.R. Cole, J.D. Nichols, R. Rudran, M.S. Foster (eds.), Measuring and monitoring biological diversity, standard methods for mammals (pp. 8–104). Washington, DC: Smithsonian Institution Press.

Ruyle, G.B. (ed.). 1991. Some methods for monitoring rangelands and other natural area vegetation. Extension Report 9043. Tucson, AZ: University of Arizona.

Rykken, J.J., Capen, D.E., Mahabir, S.P. 1997. Ground beetles as indicators of land type diversity in the Green Mountains of Vermont. Conservation Biology 11:522–530.

Salafsky, N., Margoluis, R. 1999. Threat reduction assessment: a practical and cost-effective approach to evaluating conservation and development projects. Conservation Biology 13:830–841.

Sample, V.E. (ed.). 1994. Remote sensing and GIS in ecosystem management. Washington, DC: Island Press.

Sandland, R.L., Alexander, J.C., Haydock, K.P. 1982. A statistical assessment of the dry-weight rank method of pasture sampling. Grass Forage Science 37:263–272.

Schaeffer, R.L., Mendenhall, W., Ott, L. 1979. Elementary survey sampling. North Scituate, MA: Duxbury Press.

Schreuder, H.T., Geissler, P.H. 1999. Plot designs for ecological monitoring of forest and range. In: C. Aguirre-Bravo, C.R. Franco (comp.), Toward a unified framework for inventorying and monitoring forest ecosytem resources (pp. 180–184) . Guadalajara, Mexico (November 2–6, 1998). Proceedings RMRS-P-12. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Schreuder, H.T., Gregoire, T.G., Wood, G.B. 1993. Sampling methods for multiresource forest inventory. New York: John Wiley & Sons.

Schwarz, C.J. 1998. Studies of uncontrolled events. In: V. Sit, B. Taylor (eds.), Statistical methods for adaptive management studies. Land management handbook No. 42 (pp. 19–40). Victoria, British Columbia: Research Branch, B.C. Ministry of Forests.

Schwegman, J. 1986. Two types of plots for monitoring individual herbaceous plants over time. Natural Areas Journal 6:64–66.

Scott, C.T., Tyrrell, L.E., Smith, M., Funk, D.T. 1999. A monitoring system for research natural areas in northeastern and midwestern United States. In: C. Aguirre-Bravo, C.R. Franco (comp.), Toward a unified framework for inventorying and monitoring forest ecosytem resources (pp. 315–318). Guadalajara, Mexico (November 2–6, 1998). Proceedings RMRS-P-12. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. London: Charles Griffen.

Sharp, L.A., Sanders, K., Rimbey, N. 1990. Forty years of change in a shadscale stand in Idaho. Rangelands 12:313–328.

Shindler, B., Cheek, K.A. 1999. Integrating citizens in adaptive management: a prepositional analysis. Conservation Ecology 3:9. [online] URL:http://www.consecol.org/vol3/iss1/art9.

Shivers, B.D., Borders, B.E. 1996. Sampling techniques for forest resource inventory. New York: John Wiley & Sons.

Sibley, C.G., Monroe, B.L., Jr. 1990. Distribution and taxonomy of birds of the world. New Haven, CT: Yale University Press.

Sibley, C.G., Monroe, B.L., Jr. 1993. A supplement to distribution and taxonomy of birds of the world. New Haven, CT: Yale University Press.

Silvertown, J.W., Lovett-Doust, J. 1993. Introduction to plant population biology. Oxford: Blackwell Scientific.

Simberloff, D. 1990. Hypotheses, errors, and statistical assumptions. Herpetelogica 46:351–357.

Snedecor, G.W., Cochran, W.G. 1989. Statistical methods (8th ed.). Ames, IA: Iowa State University Press.

Sokal, R.R., Rohlf, F.J. 1994. Biometry (3rd ed.). New York: W.H. Freeman.

Southwood, T.R.E. 1978. Ecological methods: with particular reference to the study of insect populations (2nd ed.). New York: Chapman & Hall.

Stafford, S.G. 1993. Data, data everywhere but not a byte to read: managing monitoring information. Environmental Monitoring and Assessment 26:125–141.

Steel, R.G.D., Torrie, J.H. 1980. Principles and procedures of statistics: a biometrical approach (2nd ed.). New York: McGraw-Hill.

Stehman, S.V., Overton, W.S. 1994. Environmental sampling and monitoring. In: G.P. Patil, C.R. Rao (eds.), Environmental statistics (pp. 263–306). Handbook of statistics 12. New York: North-Holland.

Stine, P.A., Davis, F.W., Csuti, B., Scott, J.M. 1996. Comparative utility of vegetation maps of different resolutions for conservation planning. In: R.C. Szaro, D.W. Johnson (eds.), Biodiversity in managed landscapes (pp. 210–220). New York: Oxford University Press.

Stohlgren, T.J. 1999. Measuring and monitoring biodiversity in nature reserves, forests and grasslands in the United States. In: C. Aguirre-Bravo, C.R. Franco (comp.). Toward a unified framework for inventorying and monitoring forest ecosytem resources (pp. 248–255). Guadalajara, Mexico (November 2–6, 1998). Proceedings RMRS-P-12. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Stohlgren, T.J., Bull, K.A., Otsuki, Y. 1998. Comparison of rangeland vegetation sampling techniques in the central grasslands. Journal of Range Management 51:164–172.

Sutherland, W.J. 1996. Mammals. In: W.J. Sutherland (ed.), Ecological census techniques. London: Cambridge University Press.

Sykes, J.M., Horril, A.D., Mountford, M.D. 1983. Use of visual cover assessments as quantitative estimators of some British woodland taxa. Journal of Ecology 71:437–450.

Taylor, B.L., Gerrodette, T. 1993. The uses of statistical power in conservation biology: the vaquita and northern spotted owl. Conservation Biology 7:489–500.

Thomas, J.M., McKenzie, D.H., Eberhardt, L.L. 1981. Some limitations of biological monitoring. Environment International 5:3–10.

Thomas, L. 1996. Monitoring long-term population change: why do we need so many methods? Ecology 77:49–51.

Thomas, L., Krebs, C.J. 1997. A review of statistical power analysis software. Bulletin of the Ecological Society of America 78:128–139.

Thomas, L., Martin, K. 1996. The importance of analysis method for breeding bird survey population trend estimates. Conservation Biology 10:479–490.

Thompson, S.K. 1992. Sampling. New York: John Wiley & Sons.

Thompson, S.K., Seber, G.A.F., Seber, G.F. 1996. Adaptive sampling. New York: John Wiley & Sons.

Thompson, W.L., White, G.C., Gowan, C. 1998. Monitoring vertebrate populations. San Diego, CA: Academic Press.

Titus, K. 1990. Trends in counts of scissor-tailed flycatchers based on a nonparametric rank-trend analysis. In: J.R. Sauer, S. Droege (eds.), Survey designs and statistical methods for the estimation of avian population trends (pp. 164–166). Biological Report 90(1). Washington, DC: U.S. Fish and Wildlife Service.

Titus, K., Fuller M., Jacobs D. 1990. Detecting trends in hawk migration count data. In: J.R. Sauer, S. Droege (eds.), Survey designs and statistical methods for the estimation of avian population trends (pp. 105–113). Biological Report 90(1). Washington, DC: U.S. Fish and Wildlife Service.

Todd, J.E. 1982. Recording changes: a field guide to establishing and maintaining permanent camera points. R6-10-095-1982. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Region.

Toft, C.A., Shea, P.J. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. American Naturalist 122:618–625.

Tukey, J.W. 1977. Exploratory data analysis. Reading, MA: Addison-Wesley.

Turner, R.M. 1990. Long-term vegetation change at a fully protected Sonoran desert site. Ecology 7:464–477.

Underwood, A.J. 1997. Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge, England: Cambridge University Press.

Usher, M.B. 1991. Scientific requirements of a monitoring programme. In: F.B. Goldsmith (ed.), Monitoring for conservation and ecology (pp. 15–32). New York: Chapman & Hall.

Van Strien, A.J., Van dePavert, R., Moss, D., Yates, T.J., VanSwaay, C.A.M., Vos, P. 1997. The statistical power of two butterfly monitoring schemes to detect trends. Journal of Applied Ecology 34:817–828.

Verbyla, D.L. 1995. Satellite remote sensing of natural resources. Boca Raton, FL: Lewis Publishers.

Walker, B.H. 1970. An evaluation of eight methods of botanical analysis on grasslands in Rhodesia. Journal of Applied Ecology 7:403–416.

Walters, C., Green, R. 1997. Valuation of experimental management options for ecological systems. Journal of Wildlife Management 61:9878–1006.

Walters, C., Holling, C.S. 1990. Large-scale experiments and learning by doing. Ecology 71:2060–2068.

Warren-Wilson, J. 1963. Errors resulting from thickness of point quadrats. Australian Journal of Botany 11:178–188.

Wells, K.F. 1971. Measuring vegetation changes on fixed quadrats by vertical ground stereophotography. Journal of Range Management 24:233–236.

Whitcomb, R.F., Robbins, C.S., Lynch, J.F., Whitcomb, B.L., Klimkiewicz, M.K., Bystrak, D. 1981. Effects of forest fragmentation on avifauna of the eastern deciduous forest. In: R.L. Burgess, D.M. Sharpe (eds.), Forest island dynamics in man-dominated landscapes (pp.125–201). New York: Springer-Verlag.

White, G.C. 1993. Evaluation of radio tagging, marking and sighting estimators of population size using Monte Carlo simulations. In: J.D. Lebreton, P.M. North (eds.), Marked individuals in the study of bird population (pp. 91–103). Basel, Switzerland: Birkhäuser Verlag.

White, G.C. 1996a. NOREMARK: population estimation from mark-resighting surveys. Wildlife Society Bulletin 24:50–52.

White, G.C. 1996b. Program NOREMARK Software Reference Manual. Boulder, CO: Colorado State University.

White, G.C., Anderson, D.R., Burnham, K.P., Otis, D.L. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos, NM: Los Alamos National Laboratory LA-8787-NERP.

Whitman, W.C., Siggeirsson, E.J. 1954. Comparison of line interception and point contact methods in the analysis of mixed grass range vegetation. Ecology 35:431–435.

Whorff, J.S., Griffing, L. 1992. A video recording and analysis system used to sample intertidal communities. Journal of Experimental Marine Biology and Ecology 160:1–12.

Wiegert, R.G. 1962. The selection of an optimum quadrat size for sampling the standing crop of grasses and forbs. Ecology 43:125–129.

Wilke, D.S., Finn, J.T. 1996. Remote sensing imagery for natural resource monitoring. A guide for first time users. New York: Columbia University Press.

Wilkinson, L. 1991. SYSTAT: the system for statistics, version 5.03. Evanston, IL: SYSTAT.

Williams, B. 1978. A sampler on sampling. New York: John Wiley & Sons.

Wilson, B., Hammond, P.S., Thompson, P.M. 1999. Estimating size and assessing trends in a coastal bottlenose dolphin population. Ecological Applications 9:288–300.

Wilson, D.E., Cole, F.R., Nichols, J.D., Rudran, R., Foster, M.S. 1996. Measuring and monitoring biological diversity: standard methods for mammals (p. 409). Washington, DC: Smithsonian Institution Press.

Wilson, D.E., Reeder, D.M. (eds) 1993. Mammal species of the world (p. 1206). Smithsonian Institution Press.

Wimbush, D.J., Barrow, M.D., Costin, A.B. 1967. Color stereo-photography for the measurement of vegetation. Ecology 48:150–152.

Windas, J.L. 1986. Photo-quadrat and compass-mapping tools. Natural Areas Journal 6:66–67.

Winkworth, R.E., Goodall, D.W. 1962. A crosswire sighting tube for point quadrat analysis. Ecology 43:342–343.

Winkworth, R.E., Perry, R.W., Rosetti, C.O. 1962. A comparison of methods of estimating plant cover in an arid grassland community. Journal of Range Management 15:194–196.

Wynne, R.H., Carter, D.B. 1997. Will remote sensing live up to its promise for forest management? Journal of Forestry 95:23–26.

Yaffee, S.L., Wondolleck, J. 1997. Building bridges across agency boundaries. In: K.A. Kohm, J.F. Franklin (eds.), Creating a forestry for the 21st century (pp. 381–396). Washington, DC: Island Press.

Zar, J.H. 1999. Biostatistical analysis (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Zielinski, W.J., Stauffer, H.B. 1996. Monitoring *Martes* populations in California: survey design and power analysis. Ecological Applications 6:1254–1267.

Zimmerman, B.L. 1994. Standard techniques in inventory and monitoring. III. Audio strip surveys. In: W.R. Heyer, M.A. Donnelly, R.W. McDiarmid, L.C. Hayek, M.S. Foster (eds.), Measuring and monitoring biological diversity. Standard methods for amphibians (p. 364). Washington, DC: Smithsonian Institution Press.

INDEX

A

Absence/presence techniques, 35, 38, 237
Abundance indices, 14, 237–38, 289
Accuracy vs. precision, 78–79
Action, in management objectives, 248, 252–53, 264
Adaptive management cycle
  completing, 278–79
  described, 2–3, **3,** 10
  overview of development, **12,** 12–20, **14, 16, 19**
  role of management objectives, 248
Aerial cover, defined, 218
Aerial photography, 41, 47–48, 289
Alidades/plane tables, 54
Amphibian collecting/monitoring, 63, 241–43
Analysis of variance (ANOVA), 167–69, 173–74, 178–79, 191
Animal collecting, 62–63
Animal monitoring
  amphibians/reptiles, 241–43
  aquatic invertebrates, 239–40
  birds, 243–44
  census method, 233
  difficulties, 6, 232, **232**
  distance sampling, 236–37
  double sampling, 132
  fish, 240–41
  indices, 237–38
  insects, 238–39
  mammals, 244–45
  management objective examples, **259**
  mark-recapture methods, 233–36
  photography-based, 47, 48
  systematic random sampling benefits, 125
  t-tests, 164
Annual species, monitoring difficulties, 6, 228
Aquatic invertebrate monitoring, 239–40
Attributes
  animal monitoring, 245, **251**
  indices of abundance, 14, 237–38
  in management objectives, 248, 264
  relationship to sample unit, 106, **107**
  selecting, 14, 227-30, 250–51, **251–52**
  and stratified random sampling, 122–23
Auditory surveys, 242

B

Background tasks, **12,** 12–13
Back-up systems
  data files, 73
  of data forms, 19
  field data, 67
  field notebooks, 52–53, 64
  monumenting, 55, 56–57, 204
  species information, 30
  study area location, 38, 57
Bar charts, **181,** 181–82, **182**
Bartlett's test, 157
Basal cover, defined, 218
Baseline studies, monitoring compared, 9–10
Basking traps, turtle monitoring, 243
Beetle monitoring, 238–39
Bias avoidance
  boundary decisions, 109, 207–8
  density estimations, 211
  investigator impact, 110, 136
  line intercept measures, 221–22
  nonsampling/sampling types compared, 79, **80,** 81
  permanent sampling units, 136
  qualitative monitoring, 15, 38–39, 41
  scale selection, 34–35
  visual cover estimations, 219–20
  See also random sampling
Binoculars, 52
Biological populations, defining, 102, **103**
Biological significance
  in census counts, 36
  minimum detectable change and, 88–89
  in objective development, 14
  power analysis for, 161, 180
Biomass estimations
  defined, 206
  double sampling, 132–33
  quadrat efficiencies, 116, 227
Biotic integrity indices, 290–91
Bird collecting/monitoring, 63, 243–44
Bonferri correction, 169, 172, 174
Bootstrapping, 158
Boundary decisions
  birds monitoring, 243–44

Boundary decisions (*cont.*)
  frequency estimations, 217
  overview, **207,** 207–9
  visual cover estimations, 219–20
Boundary mapping, 36, 41, 289
Box plots, 152, **153,** 154, 182, **183**
Butterfly monitoring/collecting, 62, 238

C

Camera equipment. *See* photography-based techniques
Canopy cover, defined, 218
Capture methods, 233–36, 240, 245
CAPTURE program, 235
Carabid monitoring, 238–39
Census monitoring
  advantages, 36, 150
  animals, 233
  methodology overview, 16, **16,** 17–18, 206
Change measurement, monitoring compared, 9–10
Change specifications, identifying, 14, 88–89, 253, 264
Change/trend objectives
  confidence interval calculations, 323–28
  described, 254–55, **258–59**
  sampling objectives for, 267–70
  significance test purposes, 154–55, **156, 179,** 179–80
Chi-square tests, 156, 164, 169–72, 178, 189, **190**
Class data, defined, 156
Clinometers, 54
Clipboards, 53
Closed population methods (animals), 234–35, 236
Clumped-gradient populations, quadrat efficiencies, 108–10, **111–13,** 113–14
Cluster sampling
  described, **118,** 127–29, **128,** 134
  formulas for, 332–34
Cochran's formula, 159–60
Codes for data forms, 69–70
Coefficient of variation, **142**
Collecting techniques
  animals, 62–63
  plants, 59–62, **62**
Communication guidelines
  monitoring plan, 275, **276**
  project design stage, **272–73,** 272–74
  results of monitoring, 20, 275, 279, **280**
Communities, monitoring
  difficulties, 283, 285–86
  indices for, 289–91
  multivariate analysis, **284–85,** 291–94
  qualitative methods, 289
  structure characteristics approach, 287–89
  univariate methods, 286–89
Compass, 52
Computers for data collection, 66–67
Computer simulations
  quadrat size/shape, 110, **111–13,** 113–14
  random sampling comparisons, 120, 125, 127
Confidence intervals
  calculating, 83–85, **85,** 319–28
  defined/described, **82,** 82–83, **83, 84,** 154, **154**
  error bars, **181,** 181–82, **182, 183**
  finite population correction, 86, **86**
  normality assumption and, 159
  in sample size calculations, 141, 159–60
  in target/threshold management objectives, 162–64, **163**
Correlation coefficient, 146, 148
Counting units, census, 16, 206
Cover estimations
  classes of, 219
  line intercept measures, **221,** 221–22
  methods compared, 226
  overview, 206, **218,** 218–19, **219**
  permanent sampling units, 136–37, 225–26
  photoplots, 42
  point intercept methods, **222,** 222–24
  sample units compared, 106, **107,** 108
  transects as sampling units, 116, 131–32, 222
  vigor indicators, 226–27
  visual methods, 219–21
Cover sheets, 67, **69**

D

Data analysis, purpose/timeliness, 18–20
Database programs, 71–72
Data collection
  back-up guidelines, 19, 64, 67, 73
  electronic, 66–67
  form design efficiencies, 17, 67, **68,** 69–71
  office entry/storage, 71–73, **73**
Data forms
  cover sheets, 67, **69**
  design efficiencies, 17, 67, **68,** 69–71
  field notebooks, 52–53
  photo logs, 46
  plant collections, 61
  population estimations, 39, **40**
  presence/absence monitoring, 38
  qualitative monitoring, 38, 39, **40,** 41
  site condition assessment, 41
Data loggers, 66–67
Decimal method, 201–2
Demographic distribution, estimating, 35
Demographic monitoring, 36
Density estimations
  advantages/disadvantages, 209–10, **210**
  boundary decisions, **207,** 207–9
  defined, 206
  distance measures, **107, 212,** 212–13, **213**
  permanent sampling units, 136–37
  photoplots, 42
  power analysis example, 89–97, **90, 93, 94, 95, 96, 97**
  quadrat size/shape efficiencies, 108–10, **111–13,** 113–14, **115,** 210–12
Density plots, 151–52, **153,** 154
Direction of change, specifying, 14
Distance measures
  animal monitoring, 236–37

density estimations, **212,** 212–13, **213**
as sampling units, **107**
Distribution shapes
defined/described, 81, 89–90
evaluating, 81–85, **83**
statistics selection considerations and, **157,** 157–61
See also graphs for statistical analysis; power analysis
Dit plots, 152, **153,** 154
Diversity indices, 289–90
Double sampling, **118,** 132–33
Draft monitoring plans, 18, 275
Drift fences, 242, 243
Duncan multiple-range test, 169

E

Ecological models, 13, 260, **261, 262**
Edge effects, 109, 227
Electronic Distance Measurer (EDM), 54, 55
Electroshock for fish monitoring, 240
Error bars, **181,** 181–82, **182, 183**

F

False-change errors
consequences, 86–88, **87,** 268–69
defined, 79
in power analysis example, 92–94, **93, 94,** 96–97
in power level equation, 89
in prior-power analysis, 99
in sample size calculations, 141
in sampling objectives, 267–68
Field equipment, **50,** 50–55, 63–64
Field notebooks, 52–53, 56, 61, 64
Field time
estimating, 33
sample size considerations, 108
Filenames for data storage, 72, **73**
Finite population correction, 85–86, **86,** 139, 177–79
Fish and Wildlife Service, species ranking, **23**
Fish collecting/monitoring, 62, 240–41
Fisher's measure of skewness, 159
Flagging materials, 52, 55
Frequency data analysis
analysis of variance, 167–69, 173–74
Chi-square test, 169–72
confidence interval calculations, 162–63
finite population correction, 177–79
McNemar's test, 174–76
significance test options, 155–56, **156**
statistics selection considerations, 157–61
t-tests, 164–67, 172–73
Frequency estimations
overview, 206, 214–15
pairing of quadrats, 174–75
permanent sampling units, 137
sampling design efficiencies, 215–18
significance test options, 156, **156,** 164, 167, 169–70
transects as sampling units, 131–32, 217
Frog monitoring, 242
F-test statistic, 167–69, 178–79
Functional group approach, 287–89

G

Gaussian distributions, 150–51, **151, 152**
Global Positioning System (GPS), 54–55, 203–4
Graphs for statistical analysis
benefits for monitoring, 150
density plots, 151–52, **153,** 154
normal distribution plots, 150–51, **151, 152**
summary statistics, **181,** 181–82, **182, 183,** 184
See also power analysis
Grasses, attribute selection, 227–28
Grid-cell method, 120, **120**
Guild approach, 287–89

H

Habitat monitoring, as species surrogate, 6–7, 14
Hinges, in box plots, 152, **153**
Hip chains, 53
Histograms, 151, **153**
Homogeneity assumption, described, 157
Hspread, in box plots, 152
Hypothesis test. *See* significance tests

I

Implementation monitoring, 8
Independence of sampling units, importance, 117
Index of Biotic Integrity (IBI), 290–91
Indicators
community monitoring, 286–89
in management objectives, 250, 263
monitoring benefits, 6–7, 250
selecting, 14
Indices, species, 14, 237–38, 289–91
Insects, collecting/monitoring, 62, 238–39
Intensity of monitoring, 13, 35–36, 123
Interspersion of sampling units
importance, 117
restricted sampling benefits, 127
systematic sampling benefits, 125–26
Interval data, 156
Inventories, monitoring compared, 8
Invertebrate collecting/monitoring, 62, 239–40

J

Jolly-Seber method, 235

K

Kick/dip methods, aquatic monitoring, 239

L

Labels for plant collections, 61, **62**
Landmarks & monumenting, 56–57
Landscape scale, defining, 33–34

Lincoln method, 234
Linear regression analysis, **186,** 186–88, **187, 188**
Line intercept measures. *See* transects
Lines. *See* transects
Local scale, defining, 34–35
Location, in management objectives, 248, 250, 264
Long-lived species, monitoring difficulties, 6, 228–29
Long-term ecological studies, 10

M

Macroplots
  defined, **77,** 103
  placement considerations, 34–35, 103–6, **104, 105**
  population size estimation, 39
Mammal collecting/monitoring, 63, 244–45
Management goals, identifying, 14
Management objectives
  community monitoring, 287
  components of, 248–54, **249–50, 251–52**
  example of developing, 263–65
  overview of developing, 13–15, **14**
  purpose, 248
  resources for developing, 260–62
  role in adaptive management cycle, 2–3, **3,** 10, 278–79
  sampling objectives and, 265–70
  types, 254–55, **255–59**
Management response, specifying, 15, 265, 278–79
Mann-Kendall test, 189–90, **191**
Maps, 38, 57
Mark-recapture methods, 233–36, 240
Matted plants, attribute selection, 228
McNemar's test, 156, 174–76
MDC. *See* minimum detectable change (MDC)
Mean values
  analysis of variance, 167–69
  confidence intervals for, 154, 162, 319–20, 323–26
  defined/described, 77–78, **77–78**
  in permanent sampling units, 135, **135, 136**
  in power analysis example, 89–97, **90, 93, 94, 95, 96, 97**
  in sample size calculations, 140–43, **142,** 299–310
  t-tests, 164–67
Measurement data analysis
  analysis of variance, 167–69, 173–74
  confidence interval calculations, 162–63
  finite population correction, 177–79
  significance test options, 155–56, **156**
  statistics selection considerations, 157–61
  t-tests, 164–67, 172–73
Measurement frequency, determining, 18
Measuring tapes, 51–52
Median values
  in box plots, 152, **153,** 182, **183**
  confidence interval calculations, 322–23
  defined, **152**
Metapopulations, monitoring difficulties, 229–30
Minimum detectable change (MDC)
  defined/described, 88–89
  in post hoc power analysis, 180, **181**
  in power analysis example, 93–94, 95–97, **98–99**
  in prior-power analysis, 99
  in sampling objectives, 267–68
Missed-change errors
  consequences, 86–88, **87,** 268–69
  described, 79
  in power analysis example, 92–93
  in power level equation, 89
  in sampling objectives, 267–68
Mist-netting, 244
Monitoring (overview)
  background tasks, **12,** 12–13
  communication guidelines, **272–73,** 272–75
  compared to other data-gathering, **4,** 4–6, 8–10
  definition/purpose, 2–3, 10, 272
  failure factors, **5,** 6, 274
  implementation/reporting, **19,** 19–20, 279, **280**
  indicators as species surrogates, 6–7
  methodology design, 15–18, **16**
  objective development, 13–15, **14**
  pilot studies, 18–19, **19**
  power analysis benefits, 86–89, **87,** 97, 99–100
  project development, **12,** 12–20, **276**
  role in adaptive management cycle, 2–3, **3,** 10, 278–79
MONITOR software, 191–93
Monumenting
  photoplots/photopoints, 42, 46
  procedural guidelines, 52, 55–59, 63, 204, 225
Multivariate analysis, **284–85,** 291–94

N

National High Altitude Photography Program, 47
Natural Heritage Program, **23–24,** 30
Natural history studies, monitoring compared, 8
Nature Conservancy, The, **23–24,** 30
Nearest individual measure, **212,** 213
Nearest neighbor measure, **212,** 213
Nested quadrat designs, 216–17
Nets for fish monitoring, 240
Nonparametric statistics
  selection considerations, 157–61
  test options, 155–56, **156**
Nonparametric trend analysis, 189–91, **190**
Nonsampling errors, minimizing, 79, 81
NOREMARK program, 236
Normal distributions, 150–51, **151, 152**
Normality assumption, described, 157
Normal probability plots, 150–51, **151, 152**
Notched box plots, 182, **183**
Null hypothesis, defined/described, **154,** 154–55

O

Objectives. *See* management objectives; sampling objectives, developing
Observational monitoring. *See* monitoring (overview)
Observer variability
  boundaries, 207–8
  communities, 285–86

nonsampling errors from, 79, 81
qualitative monitoring, 15, 38–39, 41
vigor measures, 227
visual cover estimations, 219–20
Office time, estimating, 33
One-tailed *t*-tests, 165, 166–67
Open population methods (animals), 235–36
Ordinal data, 156
Outliers, in box plots, 152, **153**

P

Pacing for random sample location, 203
Paint for monumenting, 52, 56, 63
Parameter, defined, **150**
Parametric statistics
defined, 150
selection considerations, 157–61
test options, 155–56
Partners in Flight, 22–23
Permanent sampling units
advantages/disadvantages, 58, 134–38, **137, 138**
cover estimations, 225–26
density estimations, 136–37, 211–12
frequency estimations, 217–18
monumenting considerations, 58–59
and sample size calculations, 140–41
standard deviation estimating, 146, 148
transects, 131–32
Permutation testing, 158
Petersen method, 234
Photography-based techniques
aerial, 41, 47–48, 289
historical photos, 262
in qualitative monitoring, 35, 42–48, **43**
Photoplots, 35, 42–46, **43**
Photopoints, 35, 44, 289
Pilot sampling
purpose, 79, 139, 141–42
quadrat size/shape efficiencies, 114, **115,** 116
Pilot studies
fish monitoring, 241
graphing data from, 150–52, **151, 152, 153,** 154
overview, 18–19, **19, 102,** 275, 277–78
sampling objectives, 269–70
Pin flags, 64, 203
Pitfall traps, beetle monitoring, 238–39
Plane tables/alidades, 54
Planning documents, 12, 260, 263
Plant collecting, 59–62, **62**
Plant database sources, 30
Plotless methods, **107**
Pocket electronic distance measures (PEDMs), 55
Pocket stereoscope, 53
Point-centered quarter measure, **212,** 213
Point counts, bird monitoring, 243–44
Point frames as sampling units, **107**
Point graphs, 182, **182, 183**
Points
intercept methods for cover estimations, **222,** 222–24, 226
permanency considerations, 137–38
as sampling units, **107**
Population condition, estimating, 35, 39, **40**
Population parameters, defined, 77–78
Population prioritization. *See* species/population prioritization
Population size, estimating, 35, 38–39
Population types, compared, 102–3, **103,** 105–6
Portable invertebrate box samplers (PIBS), 239
Post hoc power analysis, 180, **181**
Power analysis
benefits for monitoring, **87,** 87–89, 97, 99–100
graphical comparisons, 89–97, **90, 93, 94, 95, 96, 97, 98–99**
post hoc, 180, **181**
Power curve graphs, 95–97, **98–99**
Power levels. *See* missed-change errors
Precautionary principle, 269
Precision vs. accuracy, 78–79
Presence/absence techniques, 35, 38, 237
Primary sampling units, 116, 129, **130,** 131, 134
Priority setting
criteria for, 25, **26,** 27–29
importance, 22
information review benefits, 29–30, **31–32**
intensity selection, 35–36
resource assessment, 32–33
scale selection, 33–35
teamwork, 25
upper-level guidance, 22–23, **23–25**
Prior power analysis, 97, 99–100
Proportion values
Chi-square test, 169–72, 178
confidence interval calculations, 82–83, 321–22, 326–28
McNemar's test, 174–76
in sample size calculations, 140–41, 299–300, 310–17
in target/threshold objectives, 154, 162–63
P value
defined, **155**
in power analysis example, 91–92
significance of, 87, 155, 179
PVC frames, **43,** 43–44, 53

Q

Quadrats
defined, **77, 103**
pairing for frequency measurements, 174–75
permanency considerations, 137–38
as sampling units, **107**
Quadrat size/shape
computer simulations, 110, **111–13,** 113–14
cover estimations, 220–21, 226
density estimations, 108–14, **111–13, 115,** 116, 210–12
efficiency considerations, 108–10, 113–14, **115**
frequency estimations, 215–18
vigor estimations, 227

Qualitative monitoring
advantages, 35, 38
communities, 289
methodology overview, 15–16, **16,** 17–18
photography-based techniques, 35, 42–48, **43**
types of, 35–36, 38–39, **40,** 41
Quantitative monitoring
methodology overview, **16,** 17–18
types, 36
Quantity of change, specifying, 14, 248, 253, 264

R

Randomization testing, 158
Random pairs measure, **212**
Random sample selection
"calculatorless," 202
decimal method, 201–3
field location methods, 203–4
overview, 196
whole digit method, **197,** 197–99, **198, 200,** 201
Random sampling
cluster, **118,** 127–29, **128,** 134
double, **118,** 132–33
importance, 116
restricted, **118,** 126–27, **127**
simple, **117, 119,** 119–20, **121, 122, 133,** 133–34
stratified, **117,** 120, **122,** 122–23, **123**
systematic, **117,** 123–26, **124, 125, 126,** 134
two-stage, **118,** 129, **130,** 131–32, 134
types compared, **117–18**
Rank-based trend analysis, 189–91
Ranked data, 155–56
Ranked data analysis
analysis of variance, 167–69, 173–74
confidence interval calculations, 162–63
finite population correction, 177–79
significance test options, 155–56, **156**
statistics selection considerations, 157–61
t-tests, 164–67, 172–73
Ranking systems, species, 23, **23–25**
Red flag objectives, 255
Reference sites, 260–61
Regression analysis
linear, **186,** 186–88, **187,** 188
route, 188–89, **189**
Reinhardt Red-Mapper, 54
Relational databases, 71
Relocating the study area, 38, 57
Remote sensing techniques, 46–48
Removal method of animal monitoring, 236
Report schedules, 20, 279
Reptile collecting/monitoring, 63, 241–43
Resampling methods, 158
Research, compared to monitoring, **4,** 4–6
Research Natural Areas, 29
Resight method of animal monitoring, 235–36
Resource allocation, assessing, 13, 18, 19, 32–33, 277
Restricted random sampling, **118,** 126–27, **127**
Review process/intervals
background tasks, 13
at data collection intervals, 20
draft monitoring plan, 18
management objective/response, 15
pilot study results, 19
Rhizomatous species, attribute selection, 227–28
Rock picks, 53
Route regression, 188–89, **189**

S

Salamander monitoring, 242
Sample, described/defined, 36, 77
Sampled population, defined, 103, **103,** 105–6
Sampling objectives, developing, 17, 265–70
Sampling (overview)
definition/purpose, 76
distribution shapes, 81–85, **83,** 89–90
error potential, 79, **80,** 81
finite population correction, 86, **86**
long-term monitoring, 191–93, **192**
power analysis benefits, 86–89, 97, 99–100
precision levels, 78–79
statistics for calculations, **76,** 77–78
terms defined, 76–77
Sampling units, design decisions
boundaries, **207,** 207–9
number of, 138–40, **140**
overview, 17, 34, 102
permanency, 134–38, **138**
placement, 116–17
population of interest, 102–6, **103, 104–5**
size/shape, 108–10, **111–13,** 113–14, **115,** 116
type, 106, **107,** 108, 227
See also animal monitoring; random sampling; size of sample, calculating
Sampling universe, defined, 34, 76–77
Satellite imagery, 47, 289
Scale of interest, selecting, 13, 33–35
Scheffe test, 169
Screwdrivers, 64
Secondary sampling units, 116, 129, **130,** 131, 134
Seed banks, study difficulties, 228
Sequential sampling, 141–42, **142, 144, 146, 147**
Short-lived species, monitoring difficulties, 6
Significance tests
analysis of variance, 167–69, 173–74
Chi-square, 169–72
finite population correction factor, 177–79
interpreting, **179,** 179–80, **181**
limitations, 161
McNemar's, 156, 174–76
overview, 154–56, **156**
statistics selection considerations, 157–61
t-tests, 164–67, 172–73, 174
Silphid monitoring, 238–39
Simple random sampling
examples, **119, 121, 122**
individual plants/animals, **133,** 133–34
overview, **117,** 119–20
Site condition assessment, 35, 39, 41, 289

Size of sample, calculating
considerations, 138–40, **140**
equations for, 299–317
with pilot sampling data, 141–43, **142, 144, 145, 146, 147**
required information, 140–41
standard deviation estimating methods, 143–46, **144,** 148, **148**
standard error reduction, 82
See also power analysis; quadrat size/shape
Skewness measure (Fisher's), 159
Snake monitoring, 242–43
Software programs
animal monitoring, 235, 236
data management, 71–72
trend analysis, 191–93
Species
indicators for, 6–7, 14, 248, 250
listing systems, 23, **23–25**
in management objectives, 248, 261, 263
Species codes, 69–70
Species/population prioritization
criteria for, 25, **26,** 27–29
importance/overview, 12, 22
information review benefits, 29–30, **31–32**
stakeholder participation, 25
upper-level guidance, 13, 22–23, **23–25**
Spreadsheet programs, 71–72
Stage classes, monitoring benefits, 211, 218
Stakeholders
design stage, **272–73,** 272–74
priority setting and, 25
Stakes/T-posts, 55–56
Standard deviation values
defined/described, 77–78, **77–78**
normality assumption and, 159
power analysis example, 94, 97, **98–99**
in power level equation, 89
prior-power analysis, 99
in sample size calculations, 141–46, **142, 144, 145, 146, 147,** 148, **148**
sequential sampling for, 141–42, **142**
in standard error reduction, 81–82, **82**
two-stage sampling, 131
Standard error
in confidence intervals, 82–85
defined, 81, **82**
in finite population correction, 85–86
minimizing, 81–82
in permanent sampling units, 135, **135, 136**
two-stage sampling, 131
Statistical analysis methods
analysis of variance, 167–69, 173–74
Chi-square test, 169–72
confidence intervals, 154, 162–64, **163**
finite population correction, 177–79
graphing pre-analysis data, 150–52, **151, 152, 153,** 154
graphing summary statistics, **181,** 181–82, **182, 183,** 184
McNemar's, 156, 174–76
t-tests, 164–67, 172–73, 174
See also significance tests
Statistical population, defined, 103, **103,** 108
Statistical power. *See* power analysis
Statistics
defined, **150**
population vs. sample, 77–78, **77–78**
role in monitoring, 150
See also statistical analysis methods
Stereoscopes, 53
Stopwatch, random number generation, 202
Stratified random sampling
described, **117,** 120, **122, 123**
fish monitoring, 241
formulas for, 329–32
guidelines, 120, 122–23
Structure characteristics approach, 287–89
Student-Neuman-Keuls test, 169
Study area, identifying, 38, 57
Substrate sampling methods, 239–40
Surveillance studies, 9–10
Survey routes, butterfly monitoring, 238
Systematic sampling, **117,** 123–26, **124, 125, 126,** 134

T

Tags, 53
Tape measures, 51–52
Tape recorders, 66
Target population
defining, 102–3, **103**
sample unit placement, 103–6, **104, 105**
Target/threshold objectives
confidence interval calculations, 154, 162–63, 319–23
developing, 254–55, **255–57, 259,** 266–67
interpreting statistical data for, **163,** 163–64
Telephone book, random number generation, 202
The Nature Conservancy, **23–24**
Threat monitoring, as species surrogate, 6–7, 14
Threshold objectives. *See* target/threshold objectives
Time frame
estimating, 33
in management objectives, 14, 248, 253–54, 264–65
Tools, field, **50,** 50–55
T-posts/stakes, 55–56
Transects
amphibian/reptile monitoring, 241–43
cover estimations, 220–24, **221, 222,** 226
frequency estimations, 217–18
mammal monitoring, 244–45
monumenting, 58–59
as sampling units, **107,** 108, 116, 131–32
Trees for monumenting, 55, 56
Trend analysis
ANOVA-based approaches, 191
linear regression, **186,** 186–88, **187, 188**
nonparametric tests, 189–91, **190**
program planning considerations, **182,** 191–93
route regression, 188–89, **189**
TREND software, 191–93

Trend studies, monitoring compared, 9
T-square measure, 213
T-tests
  finite population correction factor, 177–78
  for independent samples, 164–67
  for paired samples, 172–73
  permanent sampling units, 136–37
Tukey test, 169
Turtle monitoring, 243
Two-stage sampling
  boundary decisions, 207, 209
  described, 116, **118,** 128, 129, **130**
  formulas, 334–37
  guidelines, 129, 131–32
  for individual plants/animals, 234
Two-tailed *t*-tests, 165–66

U

Units for sampling. *See* sampling units, design decisions
Univariate methods, 286–89
User groups
  design stage, **272–73,** 272–74
  priority setting and, 25

V

Variance (statistical), defined/described, 77–78, **77–78, 157**
Vests, field, 64
Video photography, 46
Vigor estimations, 206, 226–27
Visual cover estimations, 219–21

W

Wandering quarter measure, 212–13, **213**
Whiskers, in box plots, 152
Whole digit method, **197,** 197–99, **198,** 201
World Conservation Union, species ranking system, **24–25**

Y

Yates correction for continuity, 171