United States
Department of
Agriculture

Forest Service

Pacific Northwest
Research Station

General
Technical Report
PNW-GTR-437
January 1999

The Strategy and Design of the Effectiveness Monitoring Program for the Northwest Forest Plan

Technical Coordinators

BARRY S. MULDER is a wildlife biologist, U.S. Department of the Interior, Fish and Wildlife Service, P.O. Box 3623, Portland, OR 97208; BARRY R. NOON and HARTWELL H. WELSH are research ecologists, U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, 1700 Bayview Dr., Arcata, CA 95521; THOMAS A. SPIES is a research forester and GORDON H. REEVES is a fisheries biologist, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3200 Jefferson Way, Corvallis, OR 97331; MARTIN G. RAPHAEL is a chief research wildlife biologist, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 3625 93rd Ave., Olympia, WA 98512-9193; CRAIG PALMER is a program director, Harry Reid Center for Environmental Studies, University of Nevada—Las Vegas, 505 South Maryland Parkway, Las Vegas, NV 89154-4009; and ANTHONY R. OLSEN is a mathematical statistician, U.S. Environmental Protection Agency, 200 SW 35th St., Corvallis, OR 97333.

The Strategy and Design of the Effectiveness Monitoring Program for the Northwest Forest Plan

Barry S. Mulder, Barry R. Noon, Thomas A. Spies,
Martin G. Raphael, Craig J. Palmer, Anthony R. Olsen,
Gordon H. Reeves, and Hartwell H. Welsh

Technical Coordinators

Published by:
U.S. Department of Agriculture
Forest Service
Pacific Northwest Research Station
Portland, Oregon
General Technical Report PNW-GTR-437
January 1999

In cooperation with:
U.S. Department of the Interior
Bureau of Land Management
Fish and Wildlife Service
Environmental Protection Agency

Abstract

Mulder, Barry S.; Noon, Barry R.; Spies, Thomas A.; Raphael, Martin G.; Palmer, Craig J.; Olsen, Anthony R.; Reeves, Gordon H.; Welsh, Hartwell H. 1999. The strategy and design of the effectiveness monitoring program for the Northwest Forest Plan. Gen. Tech. Rep. PNW-GTR-437. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 138 p.

This report describes the logic and design of an effectiveness monitoring program for the Northwest Forest Plan. The program is prospective, providing an early warning of environmental change before irreversible loss has occurred. Monitoring is focused at two resource levels: individual species and specific ecosystem types. Selection of prospective indicators for the status of species or ecosystems is based on the development of conceptual models relating resource change to reliable, early warning signals of change. Ecosystems, such as late seral stage forest communities, are monitored on the basis of critical structural and compositional elements that reflect the state of underlying ecological processes. The assumption is that systems retain their ecological integrity to the extent that key biotic and physical processes are sustained. For species of concern, the design integrates animal populations with their necessary habitat and projects changes in population status by monitoring significant changes in habitat at several spatial scales. Anticipatory forecasting of changes in population status assumes habitat to be a reliable surrogate for direct population measures. A surrogate-based approach requires an active period of model building that relates population to habitat variation to develop robust wildlife relation models. Essential components needed for program implementation, such as data collection, information management, report preparation, and feedback to management, are discussed. This discussion includes recommendations for staffing, funding, and establishing a long-term commitment for a large, interagency monitoring program.

Keywords: Northwest Forest Plan, ecological monitoring, effectiveness monitoring, adaptive management, regional scale, habitat basis, conceptual model, predictive model, integration, summary report, interpretive report, institutionalize.

Preface

Under the direction of an Intergovernmental Advisory Committee that oversees the implementation and management of the Northwest Forest Plan, interagency Federal teams have been developing a monitoring program to evaluate the success of the Forest Plan. This complex and challenging task has required a large commitment of time and agency expertise. This report represents another important step in implementing a comprehensive monitoring program for the Forest Plan, a program that eventually will cover the three types of monitoring required by the plan: implementation, effectiveness, and validation monitoring. Overall direction for monitoring under the Forest Plan began with the description of the key role of monitoring in adaptive management outlined in FEMAT (July 1993), defined in the interagency report entitled "Interagency Framework for Monitoring the President's Forest Ecosystem Plan" (March 1994), and summarized in the record of decision for the Forest Plan (April 1994).

This report is the second in a series that addresses effectiveness monitoring. The first report, "Effectiveness Monitoring: An Interagency Program for the Northwest Forest Plan" (July 1995), describes the general framework for effectiveness monitoring under the Forest Plan. The approach was accepted by the Intergovernmental Advisory Committee and approved as the appropriate direction for developing an effectiveness monitoring program. This second report, after taking into consideration all peer, agency, and other reviews, concludes development by the Effectiveness Monitoring Team of the overall strategy and guidance for effectiveness monitoring. The strategy described here has not changed from that reviewed and accepted by the Intergovernmental Advisory Committee; this document incorporates edits, clarifications, and further background material, as requested.

This report represents a further step in our understanding of and approach to ecosystem monitoring. It provides the scientific basis for the effectiveness monitoring program; separate reports or modules provide specific options for monitoring assigned priority resources: late-successional and old-growth forest, northern spotted owl, marbled murrelet, and aquatic and riparian ecosystems; and monitoring plans for survey-and-manage and other late-successional and aquatic species. Reports on other issues, such as socioeconomic and tribal, will be developed in the future. Our approach, described here, provides the template for designing these and other monitoring modules to help address the Forest Plan.

The effectiveness monitoring program will consist of many modules and their supporting guidance and plans. This report and the individual reports or modules for the individual resources respond to the assignment to identify a range of options for monitoring these issues from which the Federal agencies can select an appropriate approach (or approaches). Because of the complexity of the science related to this issue, the Federal research agencies (USDA Forest Service, Environmental Protection Agency, and U.S. Geological Survey—Biological Resources Division), at the request of the Intergovernmental Advisory Committee, took responsibility to establish the underlying scientific framework and develop monitoring options. Federal agency selection of a set of options will trigger the third and final stage in the development of an effectiveness monitoring program: assignment by the agencies of the plans, people, and funding to implement the program. The completed set of documents will function as integrated guidance to Forest Plan monitoring.

The Effectiveness Monitoring Team

Executive Summary

The Northwest Forest Plan is a large-scale ecosystem management plan for Federal lands in the Pacific Northwest, encompassing 24 million acres of federally managed forests over 18 National Forests and 7 Bureau of Land Management Districts in northern California, western Oregon, and western Washington. Three types of monitoring are mandated by the Forest Plan: implementation, effectiveness, and validation. The purpose of this report is to provide the strategy and design for effectiveness monitoring of priority resources identified in the Forest Plan.

The primary goals and objectives for the Forest Plan are both ecological and socioeconomic. In the context of the Forest Plan, the primary question that effectiveness monitoring is designed to answer is, "To what extent are the goals and objectives of the Forest Plan being achieved?" Following the goals and objectives of the Forest Plan, the basic scientific premise underlying the proposed program is to implement a predictive and integrated habitat-based approach to monitoring, intended to produce useful and timely results more efficiently and cost-effectively than past programs have.

The general approach for developing the effectiveness monitoring program has been to develop the scientific framework for monitoring; this document describes that approach. The approach has been used in developing monitoring strategies for specific priority resources identified by management, including late-successional and old-growth forests, northern spotted owls, marbled murrelets, and aquatic and riparian ecosystems. This document also provides the basis for designing future monitoring modules that may address other important resource issues (for example, socioeconomic, tribal, survey-and-manage species, or other species associated with late-successional or aquatic ecosystems).

The goals for effectiveness monitoring are to evaluate the success of the Forest Plan by assessing the status and trends of selected resources. These goals are consistent with emerging national and international frameworks for monitoring. The program outlined in this report is designed to build on and improve ongoing monitoring activities of regional as well as local forest management units to accomplish these goals. Its scope and complexity, however, mean that it will be significantly different from how agency activities have traditionally been monitored, a difference that will lead to a change in thinking about how to manage and operate a monitoring program.

Scientific Approach

The task of developing a monitoring system to detect and recognize significant change is complex because natural systems are inherently dynamic and spatially heterogeneous. Further, many changes in space and time are not a consequence of human-induced actions, and many are not amenable to management intervention. It is not surprising, therefore, that few examples exist of successful monitoring programs at the ecosystem scale. Environmental monitoring programs often are discussed in abstract terms, have little theoretical foundation, try to measure too many attributes, have vague objectives, and have no institutionalized connections to the decision process. In times of budget reductions, monitoring programs can be the first to be eliminated.

To be most meaningful, a monitoring program should provide insights into cause-and-effect relations between environmental stressors and anticipated ecosystem responses. Indicators should be chosen based on a conceptual model clearly linking stressors and indicators with pathways leading to effects on ecosystem structure and function. This process enables the monitoring program to investigate the relations between anticipated stressors and environmental consequences, and provides the opportunity to develop predictive models to anticipate trends instead of waiting until trends have been demonstrated.

The emphasis chosen for effectiveness monitoring of the Forest Plan may best be described as prospective monitoring. This approach incorporates causal relations between effects and stressors through the judicious selection of indicators. It starts with characterizing threats (stressors) to the ecological integrity and ecosystem functioning (effects) of the management unit. A conceptual model then outlines the pathways from the stressor(s) to the ecological effects. Attributes indicative of the anticipated changes in specific ecological conditions are then selected for measurement. The ultimate success of this approach depends on the validity of the assumed cause-effect relations between the stressor(s), their ecological effects, and the selected indicators of stress.

The essential steps, described in the scientific literature, that we followed in developing the approach to the effectiveness monitoring program for the Forest Plan were:

1. Specify goals and objectives

2. Characterize stressors and disturbances

3. Develop conceptual models—outlines the pathways from stressors to the ecological effects on one or more resources

4. Select indicators—detects stressors acting on resources

5. Determine detection limits for indicators—to guide sampling design

6. Establish “trigger points” for management intervention

7. Establish clear connections to the management decision process

Given the great diversity of species—plant and animal, vertebrate and invertebrate—monitoring of all biotic components of managed ecosystems is clearly impossible. Based solely on pragmatic considerations, only a few surrogate measures can be used that allow indirect (but reliable) inference to the integrity of the larger set of biological processes and components. A possible surrogate for the biota is to measure the pattern and dynamics of habitat structure.

The justification for using habitat structure as surrogate variables for predicting wildlife populations is based on both pragmatic and theoretical arguments. Habitat loss and fragmentation were the primary drivers or stressors behind creation of the Forest Plan. The theoretical argument is based on the belief that animals respond to habitat adaptively; that is, where an animal selects to live is believed to be an evolved behavioral response stimulated by structural and compositional features of the landscape. Predictive habitat suitability models will need to consider the relations between landscape pattern and life history characteristics of individual species and population-scale dynamics to provide a realistic portrayal of potential trends. The assessment strategy, which emphasizes both remotely sensed and ground-plot habitat data, should allow inferences about habitat quality at different spatial scales across a range of resource issues.

The foundation of our approach to effectiveness monitoring for the Forest Plan is to initiate a gradual transition from an intensive, individual species-resource focus to a more extensive, ecosystems approach. This transition assumes identifying and measuring surrogate variables that allow reliable inferences about the integrity of the primary resources. Such a fundamental shift means a movement away from the current crisis response to individual endangered species-resource issues, to a prospective evaluation of management decisions in an ecosystem context. The transition to a habitat-based monitoring program has several advantages:

- Monitoring vegetation change will be more cost-effective than directly monitoring populations of all the possible species for which agencies are responsible
- Existing forest inventory programs can be the foundation for monitoring programs
- A habitat focus is more in line with the mandates of the Forest Plan to manage vegetation communities (habitat), not species populations directly
- Estimating the trends in habitat structure and composition represents an anticipatory as opposed to a retrospective approach to ecological monitoring, and allows evaluation of alternative management strategies

Approach to Management

To be successful, a monitoring program must be able to collect data, summarize the data into useful information, and interpret that information to advance understanding and knowledge to improve management decisions. Key components of a structured monitoring program include data collection, information management, preparation of data summaries and interpretive reports, feedback to management, and program coordination and support.

Many inventory, monitoring, and research projects are currently collecting data of value to effectiveness monitoring in the region of the Forest Plan. Rather than duplicate these efforts, we recommend building as much as possible on ongoing data collection activities. Coordination among these programs will be encouraged through direct staff links, direct data links, and quality assurance systems.

Two types of reports are integral to the effectiveness monitoring program: data summaries and interpretive reports. Data summaries are brief, comprehensive reports of essential data collected for effectiveness monitoring and are to be produced annually for each resource being monitored. The key products of the effectiveness monitoring program will be periodic regionwide interpretive reports produced at 5-year intervals. The purpose of interpretive reports is to evaluate the ecological significance of status and trends emerging in the monitoring data in relation to the Forest Plan, and to provide statements of the implications of monitoring results, documented in the summary reports, to management; pertinent information from other sources or lands also would be considered. The resulting information is critical to adaptive management; it can be used to change plans, direction, or policies and contribute to budgetary and other decisions.

As the program develops, the challenges to success will expand because of the complexity of the data being collected. In addition to assisting in interpreting the monitoring data, research support will be needed to address emerging information needs, such as selecting new indicators and associated monitoring designs. Pilot or test studies also will offer important opportunities to test new methods and concepts, which will allow the monitoring program to be improved or adapted over time. The program must provide monitoring results that are legally defensible. Therefore, an information management and quality assurance system will be needed to assist in collecting, validating, storing, and retrieving data and in preparing reports.

Strategy for Implementation

Given the complexity of a monitoring program of this scale, magnitude, and importance, we propose that the initial goal be to develop the first regionwide interpretive monitoring report at the end of 1999. Not only will this product test the success of the program, but it also will provide the baseline for assessing future trends and offer an opportunity to adjust the program for future operation. The challenge to the effectiveness monitoring program will be to integrate all the critical components into an efficient and responsive program to meet this goal. This task is daunting, given the diversity of cooperating agencies, the number of resources being monitored, and the plethora of different monitoring groups. A primary concern has been to develop a strategy for integrating the assigned resources. Our approach has been to develop a scientific and management framework that fosters integration. The monitoring plans for each resource propose a common monitoring approach, conceptual framework, indicator-selection strategy and monitoring design, and data assessment and reporting process. Similarly, strategies to address research needs, pilot studies, data management, and quality assurance have been identified.

Because this program represents a step forward from how monitoring has been handled, specific steps will need to be taken to institutionalize all aspects of the program and to establish base funding to support program activities over the long-term. Assigning permanent monitoring staff and establishing core agency teams including program managers is critical to foster integration, management, and coordination for the monitoring program. If the approaches to staffing, data sharing, and quality assurance are followed, integrating all the monitoring efforts is likely and the information necessary for adaptive management of the Forest Plan will be available.

Contents

This page has been left blank intentionally.
Document continues on next page.

Chapter 2: Conceptual Basis for Designing an Effectiveness Monitoring Program

Barry R. Noon, Thomas A. Spies, and Martin G. Raphael

Scientific Basis for Monitoring

What Is Monitoring?

Monitoring is the "measurement of environmental characteristics over an extended period of time to determine status or trends in some aspect of environmental quality" (Suter 1993:505). The challenge in this definition, and the topic of this chapter, is to clearly understand why monitoring is an important activity, to decide which characteristics of the environment to measure, to determine what information these characteristics indicate about environmental quality, and to use that information to make better management decisions about the Forest Plan.

Monitoring is purpose oriented (Goldsmith 1991). In general, monitoring data are intended to detect long-term environmental change, provide insights to the ecological consequences of these changes, and to help decisionmakers determine if the observed changes dictate a correction to management practices. Monitoring is conducted at regular intervals to assess the current status and the time trend in various environmental attributes. By its very nature, monitoring is a dynamic exercise; that is, it is a continuing activity and its temporal span may be indefinite. The time frames for monitoring programs are frequently unspecified, because human behavior and continuing human population growth lead to ongoing environmental change with unexpected ecological events as unavoidable consequences.

In the following discussion, environmental attributes are broadly defined to include any biotic or abiotic feature of the environment that can be measured or estimated. The convention is to refer to the measured attributes as “indicators,” under the assumption that their values are somehow indicative of the quality, health, or integrity of the larger system to which they belong (refer to definitions in EMAP 1993, Hunsaker and Carpenter 1990).

The most common reason to monitor a specific indicator is to detect differences in its value among locations at a given moment (status), or changes in value across time at a given location (trend). Changes in the value of an indicator are useful and relevant to the extent that they provide an early warning of adverse changes to an ecosystem before unacceptable loss has occurred. Trend, viewed as the estimated time trajectory of a state variable (a variable that describes some fundamental attribute of the system), is particularly relevant because even if the value of an indicator is currently acceptable, a declining (or increasing) trend may indicate a trajectory towards system degradation, or an undesired state.

The task of detecting and recognizing meaningful change is complex because natural systems are inherently dynamic and spatially heterogeneous. Further, many changes in space and time are not a consequence of human-induced effects, and many are not amenable to management intervention. For example, at least three kinds of change are intrinsic to natural systems: stochastic variation, successional trends after natural disturbance, and cyclic variation. Assuming that sustained ecosystems maintain these dynamic variations with predictable bounds of variation (Chapin et al. 1996), management intervention may be appropriate even when change is not human induced. For example, developing an underlying structural model that predicts the expected magnitude of change in state variables arising from natural variation may be possible. Values of indicators could then be viewed in the context of deviations from expectations based on the structural models. Repeated observations of indicator variables whose values appeared “out of range” could trigger a management response.

Extrinsically driven changes to biological indicators that arise as a consequence of some human action are of most interest to environmental monitoring programs. Concern arises when extrinsic factors, acting singly or in combination with intrinsic factors, drive ecosystems outside the bounds of sustainable variation. Thus, one key goal of a monitoring program is to discriminate between extrinsic and intrinsic drivers of change; that is, a mechanism is needed to filter out the effects of expected intrinsic variation or cycles (noise) from the effects of additive, human-induced patterns of change (signal).

Intrinsic and extrinsic drivers of change can, for the most part, be collectively referred to as “disturbances events.” Disturbances alter processes or act as physiological disruptors that elicit a response from the biota; they generate a change in the value of the state variables that characterize an organism or an ecosystem. The term “stressor” is used to refer to disturbance events that result in significant ecological effects. These effects can be either positive or negative; however, our focus is usually on those resulting in undesired outcomes. In this context, stressors are considered as the proximate causes of adverse effects on an organism or system. This terminology is consistent with the literature on monitoring (for example, Suter 1993). The focus is on stressors arising from human activities because they are amenable to management intervention and changes in policy. Further, we focus on stressors that cannot be incorporated within the natural disturbance dynamics of a system, exceed the resilience of the system, and drive an ecosystem to new state.

Stressor effects are evaluated in the context of induced changes in one or more indicators. The magnitude of indicator change that could generate a management response, however, is difficult to determine *a priori*. This uncertainty arises primarily from an incomplete understanding of the dynamics of ecosystems and the bounds of variation to which they are resilient. Interpretation of the significance of changes in the value of an indicator is also complicated by nonlinear, cause-effect relations between the indicator and its stressor(s). The assumption of linearity implies that marginal increases in the magnitude of the stressor generate fixed, marginal changes in the value of the indicator. Such assumptions fail to recognize the fundamental nonlinearity of most ecological systems (see Jones and Lawton 1994 for examples).

The real danger for monitoring programs, however, is that assumptions of linearity fail to acknowledge the possible existence of thresholds. Thresholds are regions of change in the value of a stressor that generate precipitous declines in the value of the indicator or, more seriously, the larger ecosystem. A familiar analogy is an acid-base titration in analytical chemistry. Increasing acidity (the stressor) is indicated by changes in color (the indicator) of the liquid, but the change in color is not uniform with marginal increases in acidity; rather, the change is precipitous when the buffering capacity (threshold) of the liquid has been surpassed.

A lesser known, but extremely relevant example in public land management, considers the effects of habitat loss and fragmentation on the extinction process. Loss of some area from relatively continuous habitat may have no effect for some time. But, at some point, landscape connectivity is lost, and populations become isolated and vulnerable to stochastic processes (Opdam et al. 1993). Computer simulations of these scenarios suggest that critical threshold amounts and distribution of habitat exist, below which species populations rapidly decline (Lamberson et al. 1992, Lande 1987).

Why Monitor?

The ultimate rationale for monitoring arises from the fact that long-term human welfare and environmental integrity are inseparable. Monitoring is usually justified in the context of a more immediate goal or mandate; however, on multiple-use public lands, management actions are subject to many environmental standards. The public demands information about whether these standards are being realized and resources sustained; for example, monitoring is mandated on National Forest lands to ascertain the degree of compliance with the population viability requirement of National Forest Management Act (NFMA) and with minimum water quality standards of the Clean Water Act of 1972, as amended. Even for lands reserved from resource extraction and multiple use, such as the National Parks, compliance with the broad mandate to sustain “wild” resources for the enjoyment of future human generations must be assessed. In this document, we are developing a monitoring program for the Forest Plan to determine whether the goals and objectives of that plan are being met.

Determining compliance with a monitoring goal requires a predetermined standard or norm for comparison. The degree of deviation of the indicator from its desired value serves as a signal of noncompliance or a measure of environmental degradation. Standard or benchmark values for indicators are particularly important when monitoring is part of a large restoration project. In highly degraded ecosystems (for example, coastal watersheds in the Pacific Northwest [Bisson et al. 1997, Reeves et al. 1995]), some time may elapse before indicator values begin to approach the standard, but evidence that the indicator is changing in the direction of the benchmark value is evidence that the restoration effort is working.

One way to establish the benchmark value of an environmental indicator is to refer to documented historical values or to conduct preliminary, baseline monitoring of a nonaffected ("pristine") system. Given the scarcity of truly pristine systems, however, benchmarks may have to be based on some concept of a "desired condition" (see discussion of this concept in Bisson et al. 1997). Therefore, in the absence of reference systems, some other method must be used to generate expected values or time trajectories of indicator variables.

In addition to assessing compliance, environmental monitoring programs have great value as early warning systems. By providing measures, in the early stages of decline, of those attributes indicative of ecological change, monitoring can result in prompt intervention before unacceptable environmental losses occur. Note, however, that compliance monitoring and early warning monitoring can lead to selection of very different indicators. A simple example will demonstrate this difference. On a parcel of public land, the ESA may require compliance monitoring for a top-level, vertebrate predator, such as the northern spotted owl. The life history of this species (long lived, high survival rate, low fecundity, high site fidelity) may introduce lags in its response to environmental change, however, and thus make it a particularly poor choice as an early warning indicator of all but large-scale changes in old-growth ecosystems.

Thus monitoring, whether for compliance or early warning, is undertaken to ascertain whether the current state of the system matches the expected norm or lies within some acceptable confidence region about the norm. If monitoring results indicate that conditions lie outside the acceptance region, then some specific attribute of land management practice or resource policy should be changed. Alternatively, the information from monitoring can be used to investigate the response of the system to specific management actions. This information will allow the question, "Is the system responding as predicted?" to be addressed.

What Can a Monitoring Program Tell Us?

Before a monitoring program can be developed for the Forest Plan or any given management unit (for example, National Wildlife Refuge, state park, National Forest), understanding and agreement on what environmental monitoring is must be reached. What are realistic goals for a monitoring program? What biological insights can and cannot be inferred from a monitoring program. How are the costs of a monitoring program justified?

To determine if a monitoring program is, by itself, adequate to assess attainment of management objectives, what can and cannot be legitimately inferred from the results of monitoring must be understood. A logical first step, often referred to as “implementation monitoring,” is to determine if the management guidelines or environmental regulations have been implemented. Given implementation, a monitoring program can help to evaluate the effectiveness of current management practices, develop a predictive understanding (in the form of one or more testable hypotheses) of why an environmental indicator is changing, and decide when more active management or intervention is required.

If the purpose of the monitoring program is to provide an early warning of ecosystem decline (or signs of improvement), then its success depends on having selected an appropriate indicator or indicators, and knowledge of how much change in the value of the indicator signals a significant biological change. By itself, however, a monitoring program cannot unambiguously determine the cause of a change; help decide on how much change is acceptable—that is, whether the observed change is still within the range of acceptable variation; decide on threshold values of the indicator that trigger specific management actions; or avoid false alarms—that is, concluding the state of the system has changed significantly when no meaningful change has occurred.

Because changes in the status of an indicator are of limited value without evidence of causation, cause-effect relations are best established by concurrent assessment of suspected ecosystem stressors. The second and third limitations are largely research problems; minimizing these deficiencies clearly demonstrates the complementary nature of research and monitoring programs. The last limitation is, to varying degrees, unavoidable. For a fixed sampling effort, limiting false positives (type I errors) occurs at the cost of increasing the likelihood of type II errors (that is, failing to detect a significant biological effect). The tradeoff between these risks is determined by which error is considered most important to avoid.

The Legacy of Environmental Monitoring Programs

Monitoring to estimate the viability of an individual species or a group of related species is more straightforward than assessing the integrity of an entire ecosystem. Therefore, it is not surprising that the few examples of successful, long-term monitoring programs that exist have a narrow focus on specific taxa. Arguably, the best example of such a program is the North American waterfowl monitoring program administered by the U.S. Fish and Wildlife Service (Nichols et al. 1995).

Despite the obvious value of holistic environmental monitoring, few examples exist of successful monitoring programs at the ecosystem scale. Unfortunately, little evidence supports the idea that such programs have contributed to informed management decisions, or proved valuable in averting biological crises (NRC 1990, U.S. GAO 1988). In fact, the most ambitious (and expensive) monitoring program to date, EMAP, has little tangible evidence of success and has been heavily criticized both scientifically and technically (NRC 1990, 1994a, 1994b, 1995). Given the obvious importance of knowledge of the status and trends of the Nation’s natural resources, and the integrity of the ecosystems that provide these resources, why have monitoring programs contributed so little to environmental decisions or policy formulation at the ecosystem scale?

One fundamental reason for consistent failure is that monitoring costs are perceived by managers and the public to be prohibitively high, so there is reluctance to commit to implementation. In addition, environmental monitoring programs often are discussed in abstract terms, have little theoretical foundation, try to measure too many attributes, have vague objectives, and have no institutionalized connections to the decision process. The result has been a shallow comprehension of the need for, and components of, an effective monitoring program. Further, almost all previous programs have been given low priority, seldom have been fully implemented, and have been insufficiently funded. In times of budget reductions, monitoring programs often are the first to be eliminated.

The limited investment in environmental monitoring by most public land management agencies demonstrates its low priority and lack of appreciation. One example is in the U.S. Department of Agriculture. To assess whether resource management practices are maintaining biological diversity on National Forest lands, environmental monitoring is required under NFMA; however, a review of existing Forest Service monitoring programs indicates that they often exist in name only and are funded at a fraction of programs for resource extraction. And what monitoring has been done is often ad hoc, has little foundation in ecological theory, or fails to follow the fundamental statistical principles of sampling and estimation. Those few National Forests that have implemented scientifically defensible monitoring programs have not developed a formal mechanism to link the results of monitoring to management decisions (Morrison and Marcot 1995). The primary reasons for the failure of monitoring programs are:

- Minimal foundation in ecological theory or knowledge
- Little logic to support selection of indicators
- No necessary understanding of causation
- Trigger points not identified
- No connection to decisionmaking

To gain institutional support, the concept of environmental monitoring must become less abstract, its purposes more relevant, and its contributions more apparent. At a minimum, a defensible monitoring program should do the following:

1. Clearly state management goals and objectives, emphasizing how periodic information about the status of the resources is needed for informed management decisions.

2. Provide a clear statement of why the monitoring program has value, what information it will provide, and how the interpretation of that information will lead to a more responsible management response.

3. Establish the relation between those factors that may compromise the management goals and their ecological expression. This action is best accomplished by developing a conceptual model of how the system works and how it will be affected by external stresses.

4. Provide a clear exposition of the logic and rationale underlying the selection of the environmental attributes (indicators) to be measured. Recognizing that every species or physical or biological process of interest cannot (and need not) be measured, on what basis should attributes to be monitored be selected from among all possible candidates? Inherent in this step is the need to select indicators that can be measured simply and cost-effectively.

5. Outline the sampling design and methods of measurement to estimate the value of the indicator variable. This element includes, but should not be limited to, the sampling and measurement protocols.

6. Ensure statistical precision of the measurement protocols. For example, the sampling design must address the necessary precision of indicator estimation to detect a given magnitude of change, and the likelihood of detecting this change should it occur (for a good example, see Zielinski and Stauffer 1996).

7. Include those procedures that connect the monitoring results to the decision process. For example, determine what magnitude of change in a given indicator should trigger a management response, and what the response or responses should be?

Most existing monitoring programs frequently omit the first, second, third, sixth, and seventh elements or address them only superficially. Most attention has been given to the fifth element, and even here, the focus has been narrow, often restricted to an exhaustive discussion of the sampling and measurement protocols. It is not unusual to discover that great thought and deliberation have gone into how, when, and where to measure a given indicator, but little discussion of why that particular attribute is being measured or what magnitude of change needs to be detected (that is, issues of monitoring design and management decisionmaking).

Prospective (Predictive) or Retrospective Monitoring?

To be most meaningful, a monitoring program should provide insights into cause-and-effect relations between environmental stressors and anticipated ecosystem responses; that is, prior scientific knowledge and an understanding of the factors likely to stress ecosystem functions should be incorporated into the selection of variables to measure and the sampling design (NRC 1995). Indicators should be chosen based on a conceptual model clearly linking stressors and indicators with pathways that lead to effects on ecosystem structure and function (NRC 1995). This process enables the monitoring program to investigate the relations between anticipated stressors and environmental consequences and provides the opportunity to develop predictive models.

Prospective and retrospective studies focus on determining if a cause-and-effect relation exists as postulated. In epidemiology, a prospective study begins by selecting cases with and without a suspected antecedent cause and following cases to determine if the anticipated effect is associated with the antecedent cause. Conversely, a retrospective study begins by selecting cases with and without an effect and tracing back the cases to determine whether the effect is associated with the suspected antecedent cause. Both approaches have their foundation in identifying a supposed causal relation between an antecedent cause and its expected effect. The two perspectives differ only about whether the study begins with a set of cases with or without a suspected antecedent (stressor) or with a set of cases with or without an anticipated effect.

The NRC report (1995) states that “retrospective or effects-oriented monitoring is monitoring that seeks to find effects by detecting changes in status or condition of some organism, population, or community,” and “predictive or stress-oriented monitoring is monitoring that seeks to detect the known or suspected cause of an undesirable effect (a stressor) before the effect has had a chance to occur or to become serious.” Effects-oriented monitoring does not require knowing a cause-effect relation, but if stressors and effects are *both* included in the monitoring, then the program permits analyses directed at establishing cause-effect relations. Stress-oriented monitoring assumes that a cause-effect relation is known. See Thornton et al. (1994) and Suter (1993) for additional discussions. A specific effort must be made to gather cause-effect data; this effort was not part of the current assignment for effectiveness monitoring of the Forest Plan (FEMAT 1993, USDA and USDI 1994) but rather left for future research.

The emphases chosen for effectiveness monitoring of the Forest Plan may best be described as anticipatory monitoring and predictive effects monitoring. Each incorporates supposed causal relations between effects and stressors through the judicious selection of indicators. Anticipatory monitoring starts with a characterization of threats (stressors) to the ecological integrity and ecosystem functioning (effects) of the management unit. A conceptual model then outlines the pathways from the stressor(s) to the supposed ecological effects. Attributes that indicate the anticipated changes in specific ecological conditions are then selected for measurement. The ultimate success of this approach depends on the validity of the assumed cause-effect relations among the stressor(s), their ecological effects, and the selected indicators of stress. Anticipatory monitoring does not require monitoring ecological condition or assessment endpoints of interest. It attempts to detect effects as they are occurring by measuring anticipatory indicators, rather than describing effects after they have occurred. An advantage of this approach at the local and regional scales is that the emphasis on anticipated cause-effect relations allows an earlier and more focused management response to environmental change. Given that all potential stressors cannot be identified, complete reliance on this approach is not without some risk. A possibility exists of failing to detect the ecological effects of significant but unanticipated stressors.

Predictive effects monitoring incorporates the basis of anticipatory monitoring and extends it to predicting ecological effects. Not only does this extension require the assumption of an assumed cause-effect relation but it also requires developing a predictive model for the relation. As an example, an anticipatory monitoring program could be established to measure the vegetation characteristics necessary to support northern spotted owls. Based on these characteristics, a model is developed to predict the probable distribution or population status of spotted owls. The model may assume the vegetation characteristics remain as measured, hence predicting presence under steady-state conditions, or it may predict future vegetation characteristics under natural growth or harvest assumptions to allow a prediction of population trend. In this case, predictive monitoring focuses on estimating the future effects of changes in habitat. Initial phases of a predictive monitoring program would include additional monitoring of the effect of interest (that is, population response) to construct the predictive models and establish their reliability. Subsequently, the direct monitoring of owl populations would be conducted periodically as required for model validation and model refinement.

Challenges of Monitoring Ecological Systems

Ecosystems are poorly understood, complex systems subject to stochastic variation and unpredictable behaviors. In addition, the process of ecosystem adaptation and accommodation to stress is not well explored scientifically (Rapport and Reiger 1995). Given this reality, it is not surprising that the task of monitoring ecosystems, and drawing reliable inferences to system integrity before irreversible degradation, has proved such a daunting task. Incomplete understanding of ecosystem process and function, and limited ability to predict system response to stress will remain for the foreseeable future. As a consequence, research and monitoring are inextricably entwined and mutually dependent; a successful monitoring program will require a parallel research program.

Despite the complexity of ecosystems and the limited knowledge of their functions, to begin monitoring, we must first simplify our view of the system. The usual method has been to take a species-centric approach, focusing on a few high-profile species; that is, those of economic, social, or legal interest. Because of the current wide (and justified) interest in all components of biological diversity, however, the species-centric approach is no longer sufficient. This wide interest creates a conundrum; we acknowledge the need to simplify our view of ecosystems to begin the process of monitoring, and at the same time we recognize that monitoring needs to be broadened beyond its usual focus to consider additional ecosystem components.

To address this dilemma, we need information about a small number of surrogate variables whose status and trend provide insights to the integrity of the larger system. This is the logical basis for the indicator variable concept. But, no body of ecological theory or empiricism that will unambiguously tell us what to measure currently exists. To develop a step-down process to move towards a solution requires that we begin to build on experience and existing ecological knowledge and theory.

One step toward a comprehensive but simplified approach to ecosystem monitoring is to focus on the structural and composition elements of the landscape that express underlying process and function (fig. 7). Applying this logic to managing public lands, such as through the Forest Plan, suggests an emphasis on living and nonliving elements that collectively define the habitat of a species. Thus, an assessment of the status and trend of habitat types and key habitat elements may be a useful surrogate set of variables to substitute for the direct monitoring of numerous biotic populations. Indicators may vary, however, depending on the class or classes of organisms being addressed.

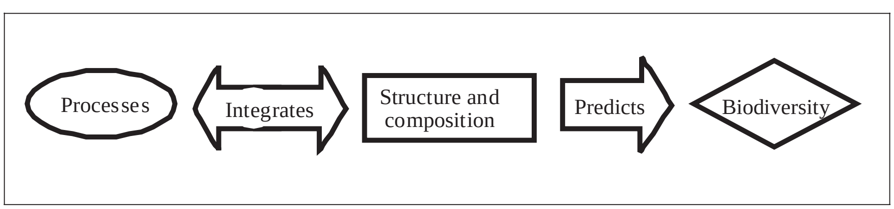


Figure 7—The conceptual model that is the basis for identifying indicators from structural and compositional landscape elements. We assumed that these elements reflect underlying ecological processes and allow predictions of the biodiversity response.

Table 2—A sequential list of key issues to address in the design of a prospective monitoring program

Steps	Design topics
1	Specify goals and objectives
2	Characterize stressors and disturbances
3	Develop conceptual models—outlines the pathways from stressors to the ecological effects on one or more resources
4	Select indicators—detects stressors acting on resources
5	Determine detection limits for indicators—to guide sampling design
6	Establish "trigger points" for management intervention
7	Establish clear connections to the management decision process

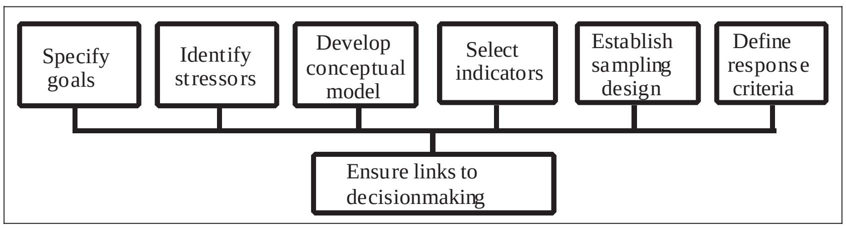


Figure 8—Steps in the design of a monitoring program.

Key Steps in Designing the Monitoring Program

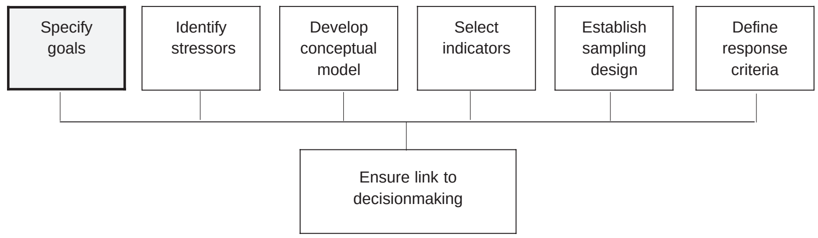
A key premise of our effort is that improving the framework for monitoring, specifically grounding the process in ecological theory and empiricism, will result in a defensible and useful monitoring program for the Forest Plan. The value of monitoring will become more apparent because the process will become less abstract and better focused; the relevance of the indicators to the integrity of the larger system will be more obvious; and a better theoretical framework should prove to be more cost-effective. We therefore sought to apply the concepts presented here by following a logical process to creating a monitoring program. The steps we followed are summarized in table 2.

These represent, in a step-down fashion, the key components in the design of the effectiveness monitoring program for the Forest Plan. Figure 8 is a model to guide the reader through each step discussed in the following sections.

The principles and concepts discussed in this chapter are general, and they can be applied to environmental monitoring programs regardless of spatial scale: local, regional, or national. The process and recommendations expressed here, however, are targeted to regional-scale monitoring programs, as required for the Forest Plan. Monitoring programs to evaluate the effects of specific local projects are more amenable to the experimental designs of environmental impact studies (see Schmitt and Osenberg 1996).

Our intent here is to provide a brief overview of the scientific basis underlying development of the monitoring proposals by the team; a more thorough understanding can be gained from the literature we applied to this effort. Because monitoring is an ongoing, active process, however, implementing these components will never be completely finished. These components must constantly be revisited and revised as scientific knowledge is acquired and as the threats to the integrity of ecosystem functions change.

State the Goals of the Monitoring Program



No universal set of goals characterizes a “quality” environment, assures the maintenance of biological diversity, and applies to all ecosystems experiencing a diversity of stresses. No single benchmark condition applies to all ecosystems. The concept of ecological integrity (Karr 1991), however, serves as a broad unifying concept and provides a universal set of goals for ecosystem management. Ecological integrity has been defined as the capacity to support and maintain a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitats of the region (Karr 1987, 1991, 1996). The key aspect of this definition is that it ties ecological integrity to evolution—the ability of the biota to persist by way of adaptive responses to environmental variation. In this broad context, the goal of monitoring of the Forest Plan is to provide the information needed to answer the question, “Are current management practices maintaining the ecological integrity of the ecosystem, including human uses of these resources for needed goods and services?”

A relevant example of a human-induced disturbance leading to a loss of biological integrity is silviculture as practiced in the forests of the Pacific Northwest from 1950 to 1990. The conversion of old-growth coniferous forests to intensively managed forests has resulted in significant changes in forest structure, decreased biological diversity, and a loss of resilience to natural disturbance events such as fire and windthrow (Spies 1991; Spies and Franklin 1991, 1995). The ecological integrity of these forests has been compromised.

Invoking the concept of ecological integrity puts the problem in the context of an ecological system composed of integrated biological components (individual organisms, populations, species, and communities) connected by exchanges of matter and energy. This model represents the traditional notion of an ecological hierarchy, and it will be a comfortable starting point for most ecologists, though it may not be a good starting point for decisionmakers responsive to societal, not necessarily biological, values. A connection, therefore, must be made between measured biological and physical attributes and what society values. This link requires a conceptual framework identifying the relations between societal values (the ultimate assessment endpoints for an environmental monitoring program) and biotic integrity.

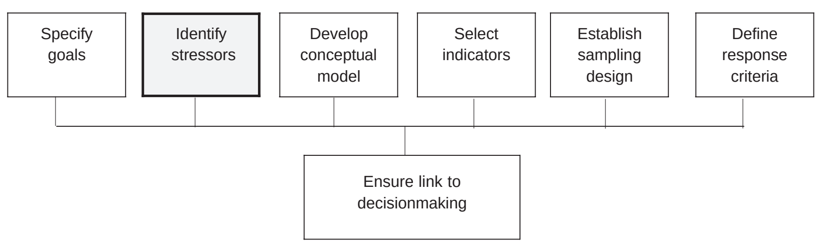
The Forest Plan (FEMAT 1993, USDA and USDI 1994) provides general direction for management in the form of standards and guidelines, and makes qualitative predictions of anticipated changes in the forest ecosystem, given their implementation. To develop a monitoring program for the Forest Plan, the general management objectives and predictions of the Forest Plan must first be refined into a set of specific monitoring questions. The monitoring program is designed to answer these questions, restated as parameters to be estimated or formal hypotheses to be tested with monitoring data.

A key ecological resource identified in the Forest Plan, and mandated for monitoring in subsequent legal decisions (USDA and USDI 1994), is late-successional and old-growth forests (Hemstrom et al., in press [see footnote 2, Chapter 1]). Monitoring goals, stated in the form of questions that can be addressed with monitoring data from late-successional forests, include:

- What are the amounts and distribution of forest age classes (including LSOG) at the landscape scale?
- What are the patch size distribution, patch interior area distribution, and interpatch distance distribution for LSOG at the landscape scale?
- Based on stand sample data, what changes have been produced by stressors in the amount and distribution of forest age classes, beginning with data collected for the 1993 FEMAT analysis?
- What are the effects of silvicultural treatment and salvage logging on LSOG structure and composition at the stand scale?
- Are the standards and guidelines leading to an increase in the amount and distribution of late-successional forest?

These questions refine the monitoring goals and suggest attributes (indicators) to measure. Measured attributes are those components of late-successional forests assumed to be indicative of the successful implementation of the standards and guidelines of the Forest Plan.

Identify Stressors Relating to Management Goals



This step usually will take the form of identifying the anticipated extrinsic environmental stressors that may compromise the integrity of the ecosystem and its component species and resources. From previous studies of disturbed ecosystems (for example, Delcourt et al. 1983), we know if the effects of an extrinsic stressor exceed the resilience or adaptational limits of the ecosystem, change occurs, the ecosystem moves to a new state, and the management goal may be compromised. Stressors, as envisioned here, can be both human-induced and "natural." Examples include (see Barber 1994):

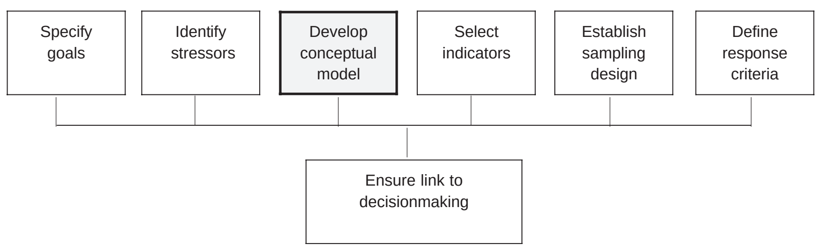
- Loss of late-successional habitat by fire
- Alterations of hydrologic cycles because of dams or water diversions
- Reduction, loss, or fragmentation of critical habitat
- Increased sediment loads to streams after storm events

- Overharvest of game species
- Changes in the transport of minerals and nutrients resulting from road construction
- Increased pollution from point sources or diffuse input of toxins

To retain the possibility of establishing cause-effect relations from the monitoring program, the status of the stressor also must be periodically estimated; that is, to infer causation from an observed change in the value of an indicator requires concurrent estimates of the status of the indicator and the magnitude of the supposed stressor.

To aid the process of indicator selection, identifying the ecological resource(s) likely to be affected by a given stressor, is important. A resource is broadly defined as an ecological entity subject to stressor effects. In practice, a resource is usually a key component of the larger ecosystem or management unit. Examples include freshwater lakes and montane meadows in National Forests. A resource can be either discrete or extensive (EMAP 1993). Examples from the Forest Plan of extensive resources include late-successional forests and aquatic-riparian ecosystems; discrete resources include the northern spotted owl and marbled murrelet. Establishing the functional relations between stressors (natural or human-induced) and resources is an essential first step in developing the conceptual model.

Develop a Conceptual Model Linking Relevant Ecosystem Components



To select indicators that reflect underlying ecological structure and function requires well-developed conceptual models of the resources of concern (Barber 1994; NRC 1990, 1995). The conceptual model outlines the interconnections among ecosystem resources (key system components), the strength and direction of those links, and the attributes that characterize the state of the resources. The model should demonstrate how the system works, with particular emphasis on anticipated system responses to stressor input. The model also should indicate the pathways by which the system accommodates natural disturbances and how the system may acquire resilience to disturbance. These processes could be portrayed by illustrating the acceptable bounds of variation of system components, and normal patterns of variation in input and output among the model elements.

As a general goal, management will strive to maintain ecological processes. These functions, however, are often difficult or impossible to measure directly. Conceptual models should identify structural and compositional elements of the resources affected by, and affecting, the underlying processes. A heuristic device to guide the model development would link process and function to measurable aspects of structure and composition. These elements, in turn, can be used to make predictions of expected biological response (see fig. 7).

Ecological hierarchy; components	Biotic consequence	Measurable attributes	Scale of measurement	Sampling methods
Landscape: Function-process Structure-composition				
Community-ecosystem: Function-process Structure-composition				
Population-species: Function-process Structure-composition				
Genetic: Function-process Structure-composition				

Figure 9—The stressor-specific worksheet used to identify the biotic consequences of stressor action at several scales of the ecological hierarchy. The attributes that reflect the biotic consequence (that is, indicators) and their measurement also are listed.

Measurements and inferences from biological systems are affected by the scale of observation. Therefore, to determine the appropriate scale for measuring an indicator, the temporal and spatial scales at which processes operate and resources respond must be estimated (at least to a first approximation) and clearly identified in the conceptual model. As a result, the most useful conceptual models will have a hierarchial structure; that is, a given structural-compositional resource in the model will reflect processes operative at smaller temporal and spatial scales, and indicate the constraints operating at larger scales (Allen and Hoekstra 1992, Allen and Starr 1982).

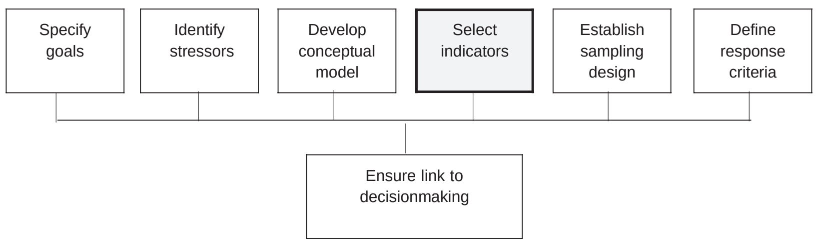
To make the process of scale an explicit component of the conceptual models developed for the Forest Plan, we developed a worksheet to characterize stressors and their anticipated effects on the ecosystem and its components (fig. 9). The purpose of this exercise is to assist with the development of the conceptual models leading to the selection of indicators for measurement. Scale was considered by allocating the effects of specific stressors to various levels in the ecological hierarchy: landscape, community-ecosystem, population-species, or genetic (see Noss 1990). Formalizing of the conceptual model required identifying the scale associated with each model component (fig. 9). As a result, insights to both the resolution and the range of the measured indicators become apparent in the conceptual model.

To illustrate the use of the worksheet in developing a model, consider the addition of roads as a stressor. The biotic consequences at the landscape scale for road building could be a disruption of landscape connectivity for plants and animals (function-process) leading to the isolation of habitats or species (structure-composition). A consequence at the community-ecosystem level could be changes in the dynamics of predator-prey systems resulting in changes to the species abundance distribution. At the population-species scale, a decrease in connectivity among individuals within a population may result in inbreeding depression. At the genetic scale, gene flow is altered via the barriers to dispersal and migration, thereby resulting in a change to the distribution of genotypes.

Developing the conceptual model should highlight that the links between stressors and biotic responses may be indirect. The building of roads, for example, may lead to an increase in erosion resulting in excess fine particle sedimentation in streams and associated biotic responses. The preliminary effects of stressors on the physical-chemical components of the ecosystem also should be considered hierarchically during development of the conceptual models (Ulrich 1994).

The indicators arising from the conceptual model are the attributes that characterize structural and compositional resources of the system. Their values indicate the current state of those resources. The indicators subsequently selected for measurement are those best reflecting known or suspected cause-effect relations among system components as identified in the model. Resources occupying central positions in the model should receive increased weight when the indicators are selected. As a result, in terms of contemporary ecological principles and theory, the model justifies the indicator or indicators selected for monitoring, and demonstrates how knowledge of the status and trend of the indicator reflects underlying process and function and will meet the goal of the monitoring program. Usually, modeling a restricted, but relevant component of the system will be sufficient. Thus, a complete model of an ecosystem is seldom necessary before proceeding with a reliable monitoring program.

Identify Candidate Indicators Responsive to Environmental Stressors



On the basis of the conceptual model and characterization of its central components, indicators are proposed for monitoring and subsequent field testing. At this point, the primary criteria for selecting indicators are that they reflect underlying ecological processes **and** changes in stressor levels, represent the larger resource of which they are a structural or compositional component, and are measurable. We begin with candidate indicators because our knowledge of the stressors affecting the system is limited. Thus, we identify a set of indicators that, based on our current knowledge, best meets our needs, but with the understanding that these may change as the program is implemented and new knowledge is gained.

Before field or simulation testing, the list of candidate indicators can be narrowed by focusing on those with the following properties:

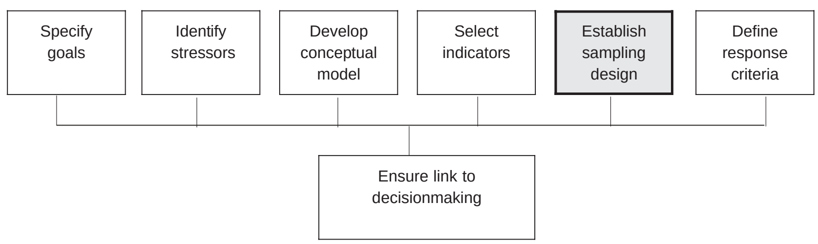
- Their dynamics parallel those of the larger environmental component or system of ultimate interest.
- They each show a short-term but persistent response to change in the status of the environment.
- They can be accurately and precisely estimated (that is, a high signal-to-noise ratio).
- The likelihood of detecting a change in their magnitude is high, given a change in the status of the system being monitored.
- Each demonstrates low natural variability, or additive variation, and changes in their values can readily be distinguished from background variation.
- The costs of measurement are not prohibitive.

Additional evaluative criteria for screening candidate indicators are in NRC (1990) and Barber (1994).

Even if a monitoring program is fully funded and implemented for many years, it will fail if the wrong indicators were selected. Thus, *the ultimate success or failure of the program may be determined by this one step*. The likelihood of choosing appropriate indicators is greatly improved if the conceptual model thoroughly characterizes the dynamics of the system, and accurately reflects stressor inputs. (A review of the effort by EPA to produce a strategy for developing indicators for EMAP [for example, Barber 1994] and subsequent criticism [NRC 1995] clearly shows the difficulty of this task.)

We find the following a useful analogy for the process of indicator selection. Imagine a funnel-shaped filter into which are poured all possible attributes of an ecological system that can possibly be measured. The fabric of filter is composed of scientific, political, and social threads. Our goal is to design the scientific fibers of the filter so that only those attributes that allow the most comprehensive and reliable inferences to the status of the ecosystem, constrained by cost functions, remain in the filter. Those attributes retained by the filter become the indicators.

Estimate the Status and Trend of the Indicator



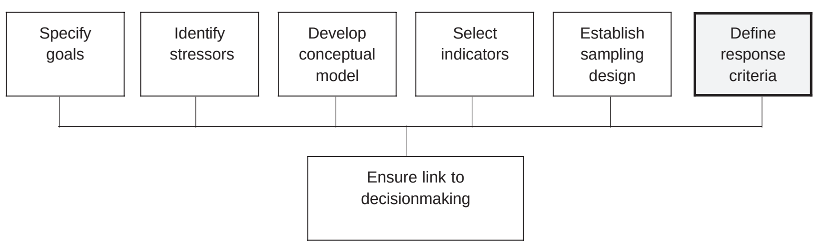
In general, determining the status of an indicator is a problem in estimating the value of an unknown parameter (that is, state variable) within some specified bounds of precision. Estimates of trend address the pattern of change over time in the status of the indicator. These problems are to be addressed by statisticians using the tools of survey and sample design (for example, Cochran 1977). As indicators change, the sampling design may change, so continual effort is needed to ensure that the design meets our monitoring needs. This topic is broad, and proper design requires substantial statistical expertise. Fortunately, a large body of statistical literature exists on parameter estimation, hypothesis testing, and trend estimation that is relevant to this problem (for example, Larsen et al. 1994, Overton and Stehman 1995, Sauer and Droege 1990, Stevens 1994).

Debate exists over the correct statistical framework for monitoring: parameter estimation or hypothesis testing (for example, Stewart-Oaten 1996). For the moment, we frame the monitoring question in terms of a statistical null hypothesis of no difference between the estimated value of the indicator and its hypothesized baseline value. The choice of significance level (α) for tests of the null hypothesis of no difference in the status of an indicator must be balanced against the likelihood of failing to detect a significant biological difference. Determining the α risk-level is a burden borne by decisionmakers. The β risk-level, in contrast, is a burden borne by those charged with maintaining ecological integrity.

The managers' responsibility is to implement an environmental monitoring program with sufficient statistical power (that is, an acceptable value of 1-β) to detect meaningful changes in the values of the indicators. For the monitoring design and analyses to be meaningful, statistical power must be considered *a priori* when determining sample sizes, and post hoc to interpret the result of statistical tests that failed to reject the null hypothesis (Skalski 1995, Zielinski and Stauffer 1996). In practice, to address questions of statistical power requires that the minimal magnitude of change in the indicator variable that is of biological significance be stated (this critical value must be estimated by some defensible process). Given this information, practical sampling issues, such as number of samples and resampling interval, can be addressed.

One of the most difficult challenges is to determine the value of an indicator, or the magnitude of change in its value over some interval, that indicates a significant biological effect. In statistical terms, this amount is referred to as the effect size (Δ) or magnitude of change in the value of the indicator that the monitoring program should be able to detect. Initial estimates of an appropriate effect size can be based on the spatial or temporal variation in the indicator (σ^2) under baseline or reference conditions (Skalski 1995). In sum, specification of acceptable levels for type I and II errors (α and β), natural variability of the indicator (σ), and the sensitivity of the test (Δ), determine the sampling effort for a given effect size. A comprehensive discussion of statistical power, and its relevance to decisionmaking in the context of responsible management of natural resources, is found in Peterman (1990).

Expected Values and Trends



An essential component of a monitoring program is the generation of expected values or expected time trends of the indicator variables; that is, the system is observed at time i and its state projected at time $i + \Delta$ given some management action. Only by comparing observed with expected values or trends can a determination be made about the effectiveness of management practices. The close approach to, or the passing of, an expected value is the threshold point that triggers a change in management practices. Estimating expected values (that is, benchmark conditions), however, is difficult and imprecise for five reasons: (1) the limited availability of pristine, undisturbed ecosystems to provide insights to benchmark conditions; (2) an incomplete understanding of the relation between the value of an indicator and the desired ecosystem state(s); (3) inadequate knowledge of the expected variability, over time and space, of the indicator of ecosystem state (or species status); (4) the nonlinear relations between indicator values and ecosystem processes (including the existence of sharp threshold regions); and (5) the fact that indicator benchmarks may be best represented by probability distributions rather than single target values.

Expected values and thresholds implicitly assume that an ecosystem will evolve to (or was historically at) a steady state of ecosystem integrity. This concept, often referred to as “the balance of nature,” has been replaced by one that recognizes the dynamic nature of ecosystems (Pickett et al. 1992). Therefore, when evaluated for periods ranging from decades to hundreds of years, the assumption of a steady state is clearly false. The dynamic nature of ecosystems argues for specifying a probability distribution of values rather than an expected value at a single moment. A second aspect of the nonequilibrium paradigm concerns the predicted time trajectory for an indicator (by trajectory, we mean how the value of the state variables change through time). Given the long delay between management actions and the response of the ecosystem to those actions, a monitoring program needs to be designed to predict the future, expected trajectory of the indicator. This prediction will require developing a mechanistic model that simulates the system response to management and whose state variables reflect both current and future ecological conditions.

Determining threshold values first requires the selection of a spatial scale to observe the ecosystem. If the spatial scale is a point in space, for example, when stream temperature is measured at a single location, an indicator threshold may be specified as a single value. An example would be a maximum water temperature beyond which conditions become lethal for cutthroat trout (temperature >22 ^{O}C). If the spatial scale includes a complete watershed, or the range of the species, however, then expecting the water temperature of all stream reaches within this area to be ≤ 22 ^{O}C may be unreasonable. Specifying an expected distribution of temperature values over the area would be more appropriate. Thus, two different categories of indicators may be described: those that lend themselves to threshold values (for example, water temperature for some fish and amphibians), and those best categorized by a target distribution (for example, number of snags and logs per acre). In practice, few indicators will be characterized by a single target value.

In addition, because the physical and biological processes and structural-compositional elements that characterize ecosystems differ in space and time, most indicators are best considered random variables; that is, when integrated across space, at a given moment, a specific process or landscape element is characterized by a dynamic distribution. To illustrate, assume that we have selected “forest stand age” as our measured indicator. We know that under a natural disturbance regime a dynamic distribution of stand ages would differ according to the spatial scale of aggregation of forest stands. If the goal of management is to mimic natural disturbance processes, then the scientific challenge is to estimate the benchmark distribution of stand ages that management should aspire to achieve. This distribution, however, depends on spatial scale. The age distribution would change as it is estimated for different-sized areas. As a consequence, a threshold value or an objective distribution cannot be specified without having some idea of the “correct” spatial and temporal scale for measuring the indicator, and the “correct” spatial scale for aggregating the measurements.

Once the scale of observation has been determined, indicator values can be aggregated into a frequency distribution. For a given moment, the observed distribution of indicator values would be compared to the expected distribution to detect both the magnitude and pattern of deviation from desired conditions. The concept of a spatial distribution of indicator values as the appropriate evaluative statistic is critical to the monitoring of ecological systems.

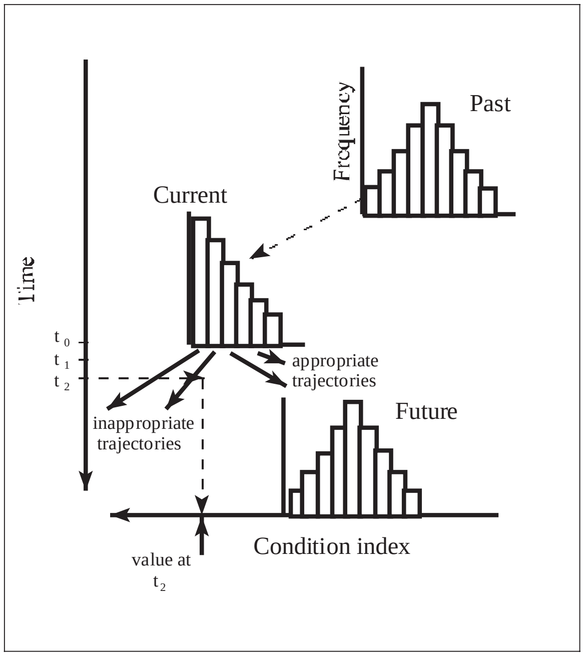


Figure 10—Frequency distribution of indicator values showing environmental condition at various points on the landscape. The past benchmark distribution, current distribution, and future, desired distribution are shown as they change through time. Current-to-figure changes are a consequence of management intervention.

Given the inherent dynamic nature of ecosystems, the value of a given ecosystem component (for example, process or landscape element) will follow a probability distribution. Based on this understanding, a monitoring program must address two distinct questions: Is the observed value of the process (or its indicator) at a specific area on the landscape, or at some moment in a time series, within acceptable bounds of the expected probability distribution? and When the observed value of the indicator, at a given time and space, is considered in the context of neighboring locations on the landscape, or in the context of a longer time series, does the expected distribution of indicator values result? For a given resource on the landscape (for example, a segment of stream, a forest stand, a riparian corridor), establishing a target value for a given indicator may be appropriate. When deviation from the desired ecosystem state at the landscape scale is evaluated, however, inferences drawn from the indicator's value at a site are of limited use without considering that signal in the broader context of values from neighboring landscape sites.

The concept of the distribution of indicator values as a collective index of ecosystem state at the landscape scale is illustrated in figure 10. This figure shows a historical distribution used as a benchmark, the current distribution of indicator values, and the future targeted distribution. In recognition of the impossibility of returning to preindustrial, pristine conditions, the target distribution is not identical to the historical distribution. Despite the need to establish benchmark distributions, the process of establishing such benchmarks is subject to some degree of arbitrariness. For example, the

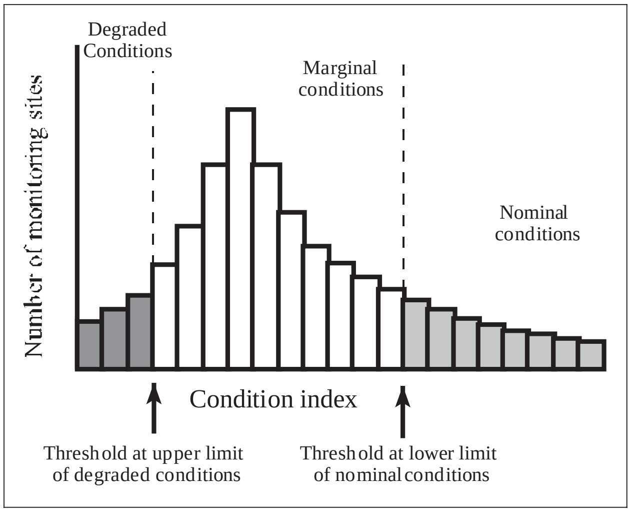


Figure 11—The distribution of indicator scores showing ecological condition at various points on the landscape. Index values relate to ecological conditions considered degraded, marginal, and nominal.

appropriate temporal reference point and uncertainty on how benchmark conditions are to be estimated from historical data are being debated. No clear guidance has been given on how far back in time to go to find an appropriate point of reference. Finally, how the concept of benchmarks can be reconciled with the dynamic nature of landscapes is unclear, especially when viewed over long time scales. For the time being, any evaluation of the ecological consequences of human activity will inescapably depend on value judgments.

In the interim, benchmark distributions, and the critical values that separate degraded from nominal conditions (fig. 11), will be based on best available information. Evaluating local conditions relative to these threshold points will be the basis of management decisions even though the location of threshold points is subject to change as ecological understanding increases. In the absence of decision thresholds or explicit objectives that management seeks to achieve, monitoring will be disconnected from management and policy formulation. Because of the complexity of this issue, the EMT did not believe this could be adequately addressed in this planning effort. This is an area that needs further work to improve sampling designs and make the program more responsive as we implement the program (see Chapter 4, "Research Support").

Most natural systems and resources recover slowly and will be slow to respond to changes in management practices. In the interim, while ecological resources are moving in the direction of a more desired ecosystem state, it is useful to identify appropriate trajectories of change in indicator values that, if continued, would lead to the target distribution (fig. 10). Thus, periodic estimates of the direction and magnitude of indicator change provide an ongoing evaluation of the appropriateness of the management strategy.

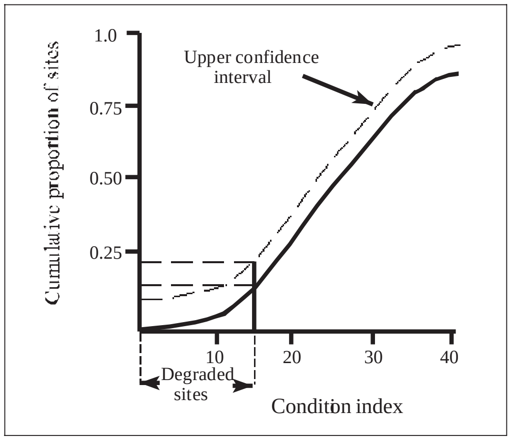
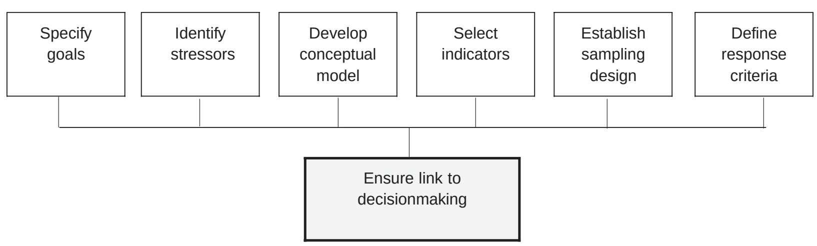


Figure 12—Cumulative distribution of indicator scores showing the collective ecological condition of many locations on the landscape. The figure illustrates that about 20 percent of the sample locations show degraded conditions.

A useful way to evaluate the integrity of a given ecological resource, concurrently across many locations, is to compute the cumulative distribution of indicator scores (fig. 12; Barber 1994). This distribution allows computing the proportion of sites below (or above) a given indicator value (that is, the lower or upper acceptable value of the indicator). In addition, the observed and expected distributions can be compared by using statistical tests (that is, Kolmogorov-Smirnov goodness-of-fit test; Zar 1984) to evaluate the deviation of the current distribution from the target distribution. Unacceptable test statistics for accepting goodness-of-fit, for example, can be used as pseudo-threshold points to trigger a change in management policy.

Linking Monitoring Results to Decisionmaking



Monitoring programs do not end with the collection of data, its analysis and synthesis, or even with summary reports. The results of monitoring programs are of value to the extent that they provide information for management decisions, and provide early warnings of ecosystem degradation. *The link between monitoring and decisionmaking begins with the formulation of and agreement on the monitoring questions*. The “correct” questions allow monitoring to be directed at areas where management requires information to adjust activities to mitigate unplanned and undesirable outcomes. Because the behaviors of complex systems are frequently unpredictable (Smith 1997), the link between decisionmaking and monitoring is essential.

Table 3—A sequential list of the steps to follow to make an optimal decision in the context of uncertainty and incomplete information

Steps	Decisions
1	Determine the bounds of the management decision space
2	Provide a range of possible management responses to the monitoring data
3	Estimate the probabilities associated with each possible interpretation of the monitoring data
4	Estimate the utilities associated with each possible combination of decision and monitoring data interpretation (that is, the costs of wrong decisions and misinterpretation of the monitoring signal)
5	Determine the decision that maximizes utility

Decisionmakers begin by asking questions such as, "Is the Forest Plan achieving its objectives for late-successional and old-growth forests?" A simple yes or no answer is not necessarily useful. A process must be instituted to connect the decisionmakers' questions to the analysis and summary of the monitoring data. One formalization uses principles of statistical decision theory (Lindley 1985).

Statistical decision theory involves determining the potential alternative ecological outcomes, assessing the probability each of these outcomes is valid, describing the management decisions under consideration, and associating a "utility" with each combination of decision and outcome (table 3). A few examples applying decision theory to the management of natural resources exist: Maguire and Boiney (1994) used decision analysis in conjunction with dispute-resolution techniques to resolve a public policy dispute in Zaire over the best policy for managing an endangered species; and Conroy and Noon (1996) applied decision theory to the question of reserve selection and species conservation.

A simple, but nonetheless relevant, example illustrates the value of monitoring data and how it is integrated with decision analysis to improve decisionmaking under the Forest Plan. Assume the plan is responsible for conserving a species listed as threatened under the ESA. In their simplest form, the possible management decisions are to take no action at all (the status quo decision) or to institute conservation measures. Based on available data, particularly the monitoring data relevant to status and trend, we estimate some probability that the species is stable or increasing ($p(\theta_1)$), or in decline ($p(\theta_2)$). (Note: Because of the impossibility of knowing the "true" status of a population, all we can do is to estimate the likelihood of the different status categories based on the best available data. The combination of alternative decisions by possible states of the population are presented in a two-way decision table [fig. 13].)

Decision (d_i)	Population status - likelihood (θ)	
	θ_1: increasing/stable	θ_2: declining
No action (d_1)	1.0	0.0
Conservation (d_2)	0.5	0.75
Probabilities	$p(\theta_1)$	$p(\theta_2)$

Figure 13—Hypothetical utility table illustrating the likelihood of different population states, the possible management decisions, and the utilities associated with the combinations of states and decisions.

The task now is to assign values to each combination of decisions and population states (fig. 13). These values are the utilities, $u(d_i, \theta_j)$, associated with the various outcomes. Utilities are scaled to the unit interval, with u = 1 "best" and u = 0 worst; $u(d_i)$ is the expected utility for decision i, over the probability space of the possible outcomes (Conroy and Noon 1996). Although utility is arguably subjective in many instances, certain outcomes (for example, the species goes extinct) are unequivocally the worst possible [$u(d_1, \theta_2) = 0$], and others (for example, the species persists with no economic costs) are the best [$u(d_1, \theta_1) = 1$]. The other outcomes have intermediate utilities. In this example, taking conservation action when none was needed (that is, the population was not declining) was assigned a lower utility because of the economic costs (for example, opportunity costs) that accompany most conservation actions.

Once the elements of the table are complete (d_i, θ_j, and $u_{i,j}$; fig. 13), the management decision is chosen that maximizes the expected (average) utility:

$$\bar{u}(d_i) = \sum_{i=1}^{u} (d_{i,}\theta_j)p(\theta_j)$$

That is, the decision (d_i) with the largest $u(d_i)$ is chosen.

As new data become available (for example, through monitoring the behavior of the ecosystem) the probabilities associated with the possible states of the system (the $p(\theta_i)$ values) are recomputed. The decision process is then revisited to determine if a different decision now maximizes overall utility. This iterative process is the substance of adaptive management.

In summary, application of decision analysis under uncertainty involves specifying management objectives and criteria for measuring success in achieving them; identifying alternatives to achieving the objectives; describing the uncertain events in the ecological and sociopolitical environment that influence the outcome of actions taken; assessing the outcome of each combination of management alternative and uncertain events in terms of the decision criteria; estimating the likelihood, or probability, of each uncertain event; calculating the expected values of the decision criteria for each alternative; resolving any tradeoffs among conflicting criteria; and reexamining the "optimal" decision by analyzing its sensitivity to changes in input parameters (Maguire et al. 1988). This is an area of current research and one not carried out by the EMT. We expect this to be addressed as we gather monitoring information for each resource issue and begin to study how to make the results useful to management (see Chapter 4, "Research Support").

Summary of Key Points

The purpose of this chapter has been to explain our scientific framework for effectiveness monitoring. The definition and purpose of monitoring were presented along with an attempt to address the challenges and experiences encountered by others when developing monitoring programs for complex ecological systems. The concepts of prospective and retrospective monitoring were introduced and the reason for our selection of the former approach was given. Seven steps for developing a prospective monitoring program were explained, including specifying goals or monitoring questions, identifying stressors, developing conceptual models relating stressors to ecological responses, selecting indicators, establishing sample designs, defining response criteria, and linking monitoring results to decisionmaking. These seven steps have been used as guidance for developing the modules for each of the resources to be monitored under the Forest Plan. They are intended to be used as the template for developing future modules.

Acknowledgments

This chapter includes contributions from Anthony Olsen and Hart Welsh.

References

Allen, T.F.H.; Hoekstra, T.W. 1992. Toward a unified ecology. New York: Columbia University Press.

Allen, T.F.H.; Starr, T.B. 1982. Hierarchy: perspectives for ecological complexity. Chicago: University of Chicago Press.

Barber, M.C., ed. 1994. Environmental monitoring and assessment program indicator development strategy. EPA/620/R-94/022. Athens, GA: Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory.

Bisson, P.A.; Reeves, G.H.; Bilby, R.E.; Naiman, R.J. 1997. Watershed management and desired future conditions. In: Stouder, D.J.; Bisson, P.A.; Naiman, R.J., eds. Pacific salmon and their ecosystems: status and future options. New York: Chapman and Hall: 447-474.

Chapin, F.S., III; Torn, M.S.; Tateno, M. 1996. Principles of ecosystem sustainability. American Naturalist. 148: 1016-1037.

Cochran, W.G. 1977. Sampling techniques. New York: Wiley.

Conroy, M.J.; Noon, B.R. 1996. Mapping of species richness as a tool for the conservation of biological diversity: conceptual and methodological issues. Ecological Applications. 6: 763-773.

DelCourt, H.R.; DelCourt, P.A.; Webb, T. 1983. Dynamic plant ecology: the spectrum of vegetational change in space and time. Quaternary Science Review. 1: 153-175.

Environmental Monitoring and Assessment Program [EMAP]. 1993. Environmental monitoring and assessment program: master glossary. EPA/620/R-93/013. Research Triangle Park, NC: Environmental Protection Agency, Office of Research and Development.

Forest Ecosystem Management Assessment Team [FEMAT]. 1993. Forest ecosystem management: an ecological, economic, and social assessment. Portland, OR: U.S. Department of Agriculture; U.S. Department of the Interior [and others]. [Irregular pagination].

Goldsmith, F.B., ed. 1991. Monitoring for conservation and ecology. New York: Chapman and Hall.

Hemstrom, M.; Spies, T.; Palmer, C. [and others]. [In press]. Late-successional and old-growth forest effectiveness monitoring plan for the Northwest Forest Plan. Gen. Tech. Rep. PNW-GTR-438. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 37 p.

Hunsaker, C.T.; Carpenter, D.E., eds. 1990. Environmental monitoring and assessment program: ecological indicators. EPA/600/3-90/060, BP91-14196. Research Triangle Park, NC: Environmental Protection Agency.

Jones, C.G.; Lawton, J.H., eds. 1994. Linking species and ecosystems. New York: Chapman and Hall. 387 p.

Karr, J.R. 1987. Biological monitoring and environmental assessment: a conceptual framework. Environmental Management. 11: 249-256.

Karr, J.R. 1991. Biological integrity: a long neglected aspect of water resources management. Ecological Applications. 1: 66-84.

Karr, J.R. 1996. Ecological integrity and ecological health are not the same. In: Shultz, P.C., ed. Engineering within ecological constraints. Washington, DC: National Academy Press: 97-109.

Lamberson, R.H.; McKelvey, R; Noon, B.R.; Voss, C. 1992. A dynamic analysis of northern spotted owl viability in a fragmented forest landscape. Conservation Biology. 6: 505-512.

Lande, R. 1987. Extinction thresholds in demographic models of territorial populations. American Naturalist. 130: 624-635.

Larsen, D.P.; Thorton, K.W.; Urquhart, N.S.; Paulsen, S.G. 1994. The role of sample surveys for monitoring the condition of the nation's lakes. Environmental Monitoring and Assessment. 32: 101-134.

Lindley, D.V. 1985. Making decisions. New York: Wiley.

Maguire, L.A.; Boiney, L.G. 1994. Resolving environmental disputes: a framework incorporating decision analysis and dispute resolution techniques. Journal of Environmental Management. 42: 31-48.

Maguire, L.A.; Clark, T.W.; Crete, R. [and others]. 1988. Black-footed ferret recovery in Montana: a decision analysis. Wildlife Society Bulletin. 16: 111-120.

Morrison, M.L.; Marcot, B.G. 1995. An evaluation of resource inventory and monitoring program used in National Forest Planning. Environmental Management. 19: 147-156.

National Research Council [NRC]. 1990. Managing troubled waters: role of marine environmental monitoring. Washington, DC: National Academy Press.

National Research Council [NRC]. 1994a. Review of EPA's environmental monitoring and assessment program: forests and estuaries. Washington, DC: National Academy Press.

National Research Council [NRC]. 1994b. Review of EPA's environmental monitoring and assessment program: surface waters. Washington, DC: National Academy Press.

National Research Council [NRC]. 1995. Review of EPA's environmental monitoring and assessment program: overall evaluation. Washington, DC: National Academy Press.

Nichols, J.D.; Johnson, F.A.; Williams, B.K. 1995. Managing North American waterfowl in the face of uncertainty. Annual Reviews Ecology and Systematics. 26: 177-199.

Noss, R. 1990. Indicators for monitoring biodiversity—a hierarchial approach. Conservation Biology. 4: 355-364.

Opdam, P.; Van Apeldoorn, R.; Schotman, A.; Kalkhoven, J. 1993. Population responses to landscape fragmentation. In: Vos, C.C.; Opdam, P., eds. Landscape ecology of a stressed environment. London: Chapman and Hall: 147-171.

Overton, W.S.; Stehman, S.V. 1995. Design implications of anticipated data uses for comprehensive environmental monitoring programs. Environmental and Ecological Statistics. 2(4): 287-303.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Science. 47: 2-15.

Pickett, S.T.A.; Parker, V.T.; Fiedler, P.T. 1992. The new paradigm in ecology: implications for conservation biology above the species level. In: Fiedler, P.L.; Jain, S.K., eds. Conservation biology: the theory and practice of nature conservation, preservation, and management. New York: Chapman and Hall: 65-88.

Rapport, D.J.; Reiger, H.A. 1995. Disturbance and stress effects on ecological systems. In: Patten, B.C.; Jorgensen, S.E., eds. Complex ecology: the part-whole relation in ecosystems. [City unknown], NJ: Prentice Hall: 397-414.

Reeves, G.H.; Benda, L.E.; Burnett, K.M. [and others]. 1995. A disturbance-based ecosystem approach to maintaining and restoring freshwater habitats of evolutionary significant units of anadromous salmonids in the Pacific Northwest. American Fisheries Society Symposium. 17: 334-349.

Sauer, J.R.; Droege, S. 1990. Survey designs and statistical methods for the estimation of avian population trends. Biol. Rep. 90(1). Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service. 166 p.

Schmitt, R.J.; Osenberg, C.W., eds. 1996. Detecting ecological impacts: concepts and applications in coastal habitats. New York: Academic Press.

Skalski, J.R. 1995. Statistical considerations in the design and analysis of environmental damage assessment studies. Journal of Environmental Management. 43: 67-85.

Smith, G.R. 1997. Making decisions in a complex and dynamic world. In: Kohn, K.A.; Franklin, J.F., eds. Creating a forestry for the 21st century. Washington, DC: Island Press.

Spies, T.A. 1991. Plant species diversity and occurrence in young, mature, and old-growth Douglas-fir stands in western Washington and Oregon. In: Ruggiero, L.F.; Aubry, K.B.; Carey, A.B.; Huff, M.H., eds. Wildlife and vegetation of unmanaged Douglas-fir forests. Gen. Tech. Rep. PNW-GTR-285. Portland, OR: U.S. Department of Agriculture, Forest Service.

Spies, T.A.; Franklin, J.F. 1991. The structural of natural, young, mature, and old-growth forest in Washington and Oregon. In: Ruggiero, L.F.; Aubry, K.B.; Carey, A.B.; Huff, M.H., eds. Wildlife and vegetation of unmanaged Douglas-fir forests. Gen. Tech. Rep. PNW-GTR-285. Portland, OR: U.S. Department of Agriculture, Forest Service.

Spies, T.A.; Franklin, J.F. 1995. The diversity and maintenance of old-growth forests. In: Szaro, R.C.; Johnson, D.W., eds. Biodiversity in managed landscapes. New York: Oxford University Press.

Stevens, D.L. 1994. Implementation of a national environmental monitoring program. Journal of Environmental Management. 42: 1-29.

Stewart-Oaten, A. 1996. Problems in the analysis of environmental monitoring data. In: Schmitt, R.J.; Osenberg, G.W., eds. Detecting ecological impacts: concepts and applications in coastal habitats. New York: Academic Press: 109-132.

Suter, G.W. 1993. Ecological risk assessment. [Place of publication unknown]: Lewis Publishers.

Thornton, K.W.; Saul, G.E.; Hyatt, D.E. 1994. Environmental monitoring and assessment program assessment framework. EPA/620/R-94/016: Research Triangle Park, NC: Environmental Protection Agency, Office of Research and Development, Environmental Monitoring and Assessment Program, Research and Assessment Center.

Ulrich, B. 1994. Process hierarchy in forest ecosystems: an integrative ecosystem theory. In: Godbold, D.L.; Huttermann, A., eds. Effects of acid rain on forest processes. New York: Wiley: 353-397.

U.S. Department of Agriculture, Forest Service; U.S. Department of the Interior, Bureau of Land Management [USDA and USDI]. 1994. Record of decision for amendments to Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl. [Place of publication unknown]. 74 p.

U.S. General Accounting Office [U.S. GAO]. 1988. Environmental Protection Agency: protecting human health and the environment. GAO/RCED-880-101. Washington, DC.

Zar, J.H. 1984. Biostatistical analysis. 2d ed. [City unknown], 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.