

Making Monitoring Work for Managers

Thoughts on a conceptual framework for improved monitoring within large-scale ecosystem management efforts

Danny C. Lee¹ and Gay A. Bradshaw²

¹Pacific Southwest Research Station, Arcata, CA dclee@fs.fed.us

²formerly, Pacific Northwest Research Station, Corvallis, OR
Gay.Bradshaw@orst.edu

Making Monitoring Work for Managers	1
PREFACE	3
I. INTRODUCTION	3
II. THE CONVENTIONAL APPROACH	5
The Stressor-Response Model	5
Four Types of Monitoring.....	6
Sample-Based versus Model-Based Inference.....	7
Measuring Trends	7
Summary of the Conventional Approach	8
III. COMPLICATIONS	9
Baseline Monitoring.....	9
Implementation Monitoring.....	10
Effectiveness Monitoring.....	10
Validation Monitoring	12
Further Considerations.....	12
IV. A DECISION-BASED APPROACH.....	14
Statistical Evidence.....	14
Influence Diagrams.....	16
An Example using Wildfire and Fish.....	17
Benefits of the Influence Diagram.....	18
V. INTEGRATING THE CONVENTIONAL AND THE DECISION-ANALYTIC APPROACHES	27
Linkage to Adaptive Management.....	28
VI. IMPLICATIONS FOR BROAD-SCALE MANAGEMENT EFFORTS	31
Conclusions.....	32
VII. LITERATURE CITED.....	33
Web sites and Links.....	35

PREFACE

Monitoring has become a dominant theme among environmental scientists, land management, and policy makers alike. The number of publications and plans which propose to detect and identify system state and change continue to multiply, each suggesting alternative approaches and solutions. Despite considerable effort by various institutions and individuals, effective environmental monitoring remains an unanswered challenge. This is particularly the case for large-scale resource management projects such as the Interior Columbia Basin Ecosystem Management Project (ICBEMP), the Northwest Forest Plan (NWFP), and the Sierra Nevada forest plan amendment process.

The following report began in 1998 as a dialogue about an appropriate conceptual framework for organizing and developing a monitoring plan for broad-scale efforts. We prepare the original version of this report for a group drafting a monitoring charter for the ICBEMP. Because so much effort had already been invested in preparations for monitoring within the NWFP, it seemed logical to look first at those efforts. Our general impression is that the monitoring plans developed for the NWFP, while statistically sound, lacked an integrated strategy that allowed one to easily see why certain information is important and how such information might influence future decisions and investments. Thus, we believed that our comments offered applied equally well to the NWFP and ICBEMP.

I. INTRODUCTION

Monitoring has become a cause célèbre among scientists, managers, and policy makers and has occasioned a flurry of activity, discussion, and publications. Despite considerable attention in the literature on what, how, and when to monitor (see Spellerberg 1991 and *Ecological Applications*, May 1998), monitoring continues to remain a conundrum. Much of the difficulty with monitoring stems from the multiple purposes that it is intended to serve (Ringold et al. 1996). For some, monitoring is viewed as a way to resolve some of the uneasiness among constituents seeking resolution to environmental conflict. In this context, monitoring serves as a watchdog to detect adverse conditions in a timely manner that allows ameliorative steps. For example, monitoring might be expected to inform and evaluate the impact of implemented forest management on a given population of spotted owls or endangered salmon. Data must be of sufficient quality to provide statistically significant verification and sufficient accessibility to allow any management plan reformulation necessary to stop further negative impacts. In contrast, the scientific community may use monitoring to provide insight into ecosystem functions and data to inform their recommendations to the public and policymakers. Finally, policy makers benefit indirectly from the assurance that monitoring provides for their constituents and scientists. More directly, monitoring demonstrates active attention to a given environmental issue.

In this paper, we take the position that monitoring is first and foremost a tool for managers. That is, the principal role of monitoring is to illuminate decision making. Monitoring works best in three ways: (1) by providing an accurate assessment of the status of the resources being managed, (2) by validating that management decisions are

correctly interpreted and implemented, and that such decisions achieve desired consequences, and (3) by providing improved insight into how systems operate. Any monitoring system that achieves these three objectives will almost certainly meet the additional expectations placed upon it by interested publics.

Large-scale regional plans such as the Northwest Forest Plan ([NWFP](#)), Interior Columbia Basin Ecosystem Management Project ([ICBEMP](#)), and Sierra Nevada Forest Plan Amendment ([SNFPA](#)) create special monitoring problems. These ecosystem management projects cover relatively large, diverse areas and involve multiple social and ecological objectives. Their very nature suggests that monitoring designed to track the effectiveness of these plans inevitably will be complex and ambitious (Mulder et al. 1999, Manley et al. 2004). There is a tendency to identify a large number of potential ecological and socioeconomic indicators and try to monitor them all. A large number of indicators is not inherently bad—it may be unavoidable for complex systems. Difficulties can arise, however, in two ways: (1) without a integrated strategy for processing monitoring information, the multiple indicators deliver a cacophonous signal with no clear message, or (2) monitoring budgets are overstretched, leading to haphazard prioritization that drops important indicators in favor of those more easily measured, or worse, results in insufficient effort to monitor any single element effectively.

Some of these problems can be avoided by having a sound conceptual strategy guide the development of the monitoring plan. Herein, we offer our views on how to develop such a strategy using a decision analysis approach. Our report is organized along two major themes.

First, we examine the conceptual model that drives many conventional monitoring efforts and discuss some complicating issues that lead to confusion and debate for scientists and land managers alike. For example, poor articulation of scaling concepts and the use of ill-defined variables which are heavily value-laden (e.g., watershed integrity) can increase uncertainty. In the context of monitoring, such imprecision confounds appropriate quantification of ecological change. In response to some of these issues and complications, we propose an alternative approach for the design and context of monitoring plans in our second theme.

We propose the adoption of a formalized, quantitative process for the formulation of monitoring plans—in short, the design and use of a decision analysis approach explicitly linked to monitoring. We suggest that the lack of hypothesis formulation embedded in a decision-making framework creates the core source of the present monitoring quandary, and describe in general terms what such an approach might look like. This approach stresses the importance of clearly articulated hypotheses germane to the land-management decisions. Extant statistical and decision theories combined with a rigorous interpretation of adaptive management provide the appropriate tools to effect accurate translation of scientific information into inference, management decisions, and policy. The proposed approach offers several advantages, specifically: 1) it documents the decision making pathway (from concept to data, to management objective, to decision, to policy); 2) it elucidates what and how information is used in decision making; 3) it maps types and sources of uncertainty; 4) it provides an objective method for prioritization of monitoring targets; 5) it explicitly integrates research, management, and policy.

We begin with an overview of monitoring as it is commonly perceived.

II. THE CONVENTIONAL APPROACH

There is a steadily increasing literature on monitoring of environmental resources. Many of the ideas and techniques advanced in this literature share a common conceptual framework that we refer to as the *conventional approach*. The first principle of the conventional approach is that ecological monitoring is designed to measure and understand change. Many basic statistical concepts that were developed in other natural resource fields such as agriculture and biology have been incorporated into the vernacular of ecological monitoring. In the following sections, we step through some of the basic monitoring concepts in order to develop a common understanding and vocabulary that will be used throughout this paper.

The Stressor-Response Model

Underlying the conventional approach to monitoring is a simple conceptual model of how the world operates. Basically, this model divides ecosystem dynamics into two components--stressors and responses. Stressors are agents of change that operate through ecological processes to either directly or indirectly create a response. Both stressors and responses are measured through a set of surrogate indicators. For example, cattle grazing in riparian areas (a stressor) is believed to change the composition of streamside vegetation and reduce bank stability (responses). In a monitoring sense, cattle grazing could be indexed by tracking the number of animals grazing within a riparian area over a designated time period, or more simply by noting whether an area is grazed or not. Similarly, riparian vegetation could be tracked through various measurements of composition and structure, and bank stability might be indexed by looking at the frequency of slumping or the proportion of bank that is overhanging the stream channel. As this example illustrates, every stressor-response combination can be indexed in a variety of ways.

The conventional approach to monitoring next assumes that a statistical model can be constructed relating stressor indicators to response indicators in a manner that caricaturizes the relationships between ecological stressors and responses. The idea of caricature is important; no statistical model purports to perfectly describe the underlying ecological linkage. These statistical models can vary widely in their complexity, ranging from simple univariate tests between pre- and post-treatments in a given area, to complex multivariate models with nonlinear or interacting terms. The statistical models are vital to the monitoring effort: ultimately the power to measure change is wholly dependent upon the statistical properties of the models linking ecological indicators. Thus, the first rule of monitoring is to start with a sound statistical model. One should be extremely wary anytime that a monitoring plan is proposed without proper specification of and justification for an explicit statistical model of the elements in question.

Four Types of Monitoring

Noss and Cooperrider (1994) divide monitoring into four basic types: baseline, implementation, effectiveness, and validation. This classification has gained favor as a useful typology within the Forest Service (e.g., [ICBEMP DEIS 1996](#)), and we review them here.

Baseline monitoring is used to establish baseline reference conditions that can be used to quantify a change that might be due to management activities. Often, areas are included in the baseline set that are deemed to be minimally affected by management (e.g., Overton et al. 1993). In such cases, the underlying assumption is that differences in measurements taken in reference areas versus managed areas can be attributed to management.

Implementation monitoring, sometimes called compliance monitoring, is simply tracking to see if management direction is accurately interpreted and followed. Conceptually, it should be straightforward to monitor implementation, assuming that the initial direction is clear. Where there is imprecision in the direction itself, it can be difficult to tell who is at fault for poor implementation success.

Effectiveness monitoring is intended to measure whether progress is being made towards an objective. Every management decision is intended to achieve a given set of future conditions. Effectiveness monitoring can be used to compare existing conditions to both past conditions and the desired future conditions to describe the overall progress or success of the management activities.

Validation monitoring seeks to verify the assumed causal linkages between cause and effect. Generally, validation monitoring requires a more intensive effort than the other forms, albeit at a more limited number of sampling sites. As its name implies, validation monitoring is intended to validate the basic assumptions under which the management direction was developed.

It is important to note that validation monitoring and effectiveness monitoring reach their full potential only when used in combination. That is, knowing that the overall goals are or are not being met (effectiveness monitoring) is rather pointless without some assurance that the observed effect is due to the management activities (validation monitoring). Similarly, while it is nice to know that one's understanding of the world is correct (validation monitoring), that is of little utility in determining whether overall goals are being met (effectiveness monitoring). These two types of monitoring can be highly complementary if fully integrated.

There also is a misconception by some that effectiveness monitoring is somehow easier or less rigorous than validation monitoring. This has led to suggestions that effectiveness monitoring can be readily handled by the National Forest System and the BLM districts, while validation monitoring is shifted to research institutions. The fallacy in this case is (1) assuming that effectiveness monitoring is necessarily easy, and (2) decoupling validation and effectiveness monitoring into separate and perhaps disconnected exercises.

Sample-Based versus Model-Based Inference

A major topic in the monitoring literature is the difference between design-based and model-based inference. In simplest terms, design-based approaches derive their utility from the strength of the sampling design used to gather the information. The observed sample is randomly drawn from a specific target population; inferences from the sample to the population at large follow naturally. In contrast, model-based approaches do not rely on random sampling. Rather, an attempt is made to identify representative sites, sometimes called sentinel sites, that are intensively studied for the purposes of constructing a more detailed model of the ecological process in question. This model is then applied more widely to similar sites or locations.

As an example, consider a plan to monitor eutrophication in small, high-elevation lakes through time. In a design-based approach, a relatively large number of lakes would be randomly sampled across the landscape, with every lake having the same probability of being included in the sample. Because of their potential remoteness, each lake might be sampled periodically for a limited set of water quality indicators. Tracking these indicators through time would indicate water quality changes in the population of lakes in question. In contrast, a model-based approach would intensely study eutrophication in a small number of high-elevation lakes in order to develop a more thorough understanding of eutrophication and use the results of that study to make some predictions about the level of eutrophication expected in lakes throughout the region.

The advantages of design-based approaches are that there should be no question about the populations for which the findings apply, assuming that the designs are correctly applied. The drawback with design-based approaches is that they can require large sample sizes and the probabilistic sampling called for in the design can increase the cost of collecting data substantially. Careful attention to probabilistic sampling and the use of advanced sampling designs can help alleviate some of these problems.

Model-based approaches can be more efficient in terms of collecting and using a variety of information, but it can be very hard to find truly representative sites. Using a small number of sites as indicators for the population at large, the so-called sentinel site approach, is especially problematic. Jassby (1998) identifies three demanding conditions that must be fulfilled for a network of sites to serve as sentinel sites for a particular stressor: (1) some subset of the network must encounter the stressor; (2) at some sites must be responsive to the stressor; and (3) the background noise at the sites must not disguise the response to the stressor. Jassby cautions that, "reliable extrapolation from sentinel-site networks to regional trends appears to be beyond our ecological understanding at the present time." Most statisticians tend to agree that model-based approaches such as sentinel sites are best used for understanding processes rather than measuring regional trends (Olsen et al. 1997).

Measuring Trends

In general terms, trend refers to a unidirectional change in an indicator variable across a second dimensional variable (e.g., space or time). In an ecological monitoring context, trends are often expressed across time, where time is viewed discretely as in annual

increments. The statistical literature is rich in methodologies for analyzing trends in individual time series. In the general case, these time series analyses require only the time trace of the indicator variable. The underlying assumption is that future observations can be modeled solely as a function of past observations; causality is not addressed. Unfortunately, robust results from time series analysis require series considerably longer series than are generally available for ecological indicators.

Short time series raise all kinds of problems for ecological inference. For example, we know that natural systems exhibit a variety of cyclic patterns with alternating periods of increases and decreases. Many times there are even multiple cyclic patterns with different frequencies and amplitudes superimposed on each other. This creates the obvious problem of interpreting a relatively short increase or decrease in an indicator. Is the observed increase/decrease a part of the natural cycling or is it due to some fundamental shift in ecological process?

The case of a regional trend analysis that simultaneously assesses multiple sites that may display contrasting trends is an especially complex and challenging case. Urquhart et al. (1998) provide an example of regional trend analysis using water quality data collected by EPA. Urquhart et al. (1998) simply their analysis by using strictly linear models as a first approximation. Their analysis suggests that as concordant variation--the variation of all sites together around a common trend--increases, the ability to detect regional trends decreases. The best sampling designs required 10-15 years of sampling data to detect moderate trends. Urquhart et al. (1998) used the arithmetic average of the trends observed across sites as an indicator of regional trend. This raises the question, however, of whether a situation where half the sites are increasing and half are decreasing is truly a "no change" scenario ecologically.

Summary of the Conventional Approach

We summarize key points regarding the conventional approach to ecological monitoring as follows:

- Monitoring begins with a conceptual model of the ecosystem as a system of stressors and responses represented by measurable indicator variables.
- Sound statistical models are the key to detecting and understanding change. If you don't see the model, don't buy into the plan.
- A comprehensive plan requires a mix of the four types of monitoring (baseline, implementation, effectiveness, and validation). These elements should be fully integrated for optimal performance.
- Design-based approaches are better suited for assessing regional changes, while model-based approaches better support understanding of process. This suggests that design-based methods more naturally fall into the realm of effectiveness monitoring while model-based approaches are more useful in validation monitoring.

- Trend analysis using conventional time series analysis may be of limited utility in effectiveness monitoring, but are unlikely to provide ecological understanding without long time series. Measuring regional trends across numerous sites is especially problematic.

III. COMPLICATIONS

While the conventional approach to monitoring has enjoyed moderate success (Spellerberg 1991), monitoring plan development for large, complex landscapes soon encounter difficulties. In large part, the potential disparity between expectations of monitoring and its success derives from a desire for certainty at spatial scales that are much greater than before. Not only are larger scales involved, but large landscapes are notoriously open systems where change can result from unexpected quarters. Managers have little to no control over unintended or unexpected factors such as flooding events or unknown influences on the landscape behavior from past disturbances. Finally, both management agencies and the public desire high levels of certainty despite the intrinsic complexity and non-linear behavior of the environment.

The emergence of ecological theories (i.e., hierarchy theory, patch dynamics, landscape ecology, metapopulations), all which constitute the theoretical foundations for ecosystem management, involve spatially and temporally explicit consideration of ecological processes and patterns. As a result, large-scale monitoring has shifted to include not only the description of populations distributed across large spatial extents, but also to the description of large-scale *processes*: those ecological phenomena that encompass and occupy large, contiguous areas (e.g., watershed functions, metapopulations). In contrast, while ecological knowledge and scientific confidence are confined mostly to fine scales, great uncertainties about large-scale ecological processes and how to manage them abound (see papers in Ricklefs and Schulter [1993] and Peterson and Parker [1998]). We examine how the shift to an ecosystem and large-scale perspective challenges implementation of the four basic monitoring types.

Baseline Monitoring

A fundamental assumption in baseline monitoring (namely that measurement differences in reference versus managed areas can be attributed to management) must be carefully reassessed in the process of monitoring plan formulation for large-scale systems. The presence of multiple variables in an open system cause landscape ecologists to wrestle with such questions: What constitutes reference conditions and reference sites in the landscape context? What constitutes a replication at the watershed scale, and what constitutes a reasonable "population" of watersheds?

The concept of a reference or replicate landscape is difficult. First, very often it is difficult to distinguish the effects of either past human or natural disturbances on the present landscape pattern. While at first glance two landscapes may appear quite similar, their subsequent behavior may be quite different. Given the subtleties of ecological interactions over time and space, it is problematic to designate an unmanaged or control

area and its supposed complement for treatment. Differences in the order of disturbance, or differing combinations of various processes may create the same landscape pattern relative to our levels of detection. Thus, if we observe a difference in the response between the control and treated watershed, can we conclude with certainty that the difference derives from the treatment? Perhaps the difference in observed responses is the result of differing histories and the management treatment under question is relatively incidental. These issues require careful planning and problem definition to achieve a statistically sound experiment.

Implementation Monitoring

In practice, implementation monitoring should be the most straightforward of the monitoring types, provided that management direction is sufficiently clear. While scaling to larger frames may create logistical challenges, other issues should be minimal. The main challenge is to provide clear direction, yet maintain the flexibility to accommodate local conditions and needs. It is easy to tell if an agency has followed the letter of the law, but more difficult to tell if the intent was properly followed. Problems arise with insufficient articulation, poor training, and a lack of the right combination of specificity and flexibility to address the spectrum of objectives across scales.

Effectiveness Monitoring

Judging management success (i.e., the goal of much of effectiveness monitoring) first requires a definition of success and specific criteria by which the plan may be assessed. For example, in the case of a single species such as the Northern Spotted Owl (*Strix occidentalis caurina*), this may mean that the population has reached some predetermined level or composition. The identification and establishment of success criteria implicitly assumes the use of a working hypothesis (i.e., model) which articulates the relationship between the stressor (e.g., management) and response (e.g., population status). In the case of regional level assessments then, a specific model needs to be identified to allow for scaling of information collected at the fine scale (e.g., district) to the large scale (i.e., region). In other words, we need to articulate what assumptions are made when we use data collected at one set of scales (and under one set of assumptions) and use them at another scale (under a different set of assumptions). Model development needs to address both conceptual and technical issues involved. For example, the scaling process requires addressing of such questions as: Does the simple sum of individual numbers collected in the districts constitute an adequate representation of ecological condition for the region scale? Since the spatial distribution of responses and stressors involved are likely to vary, should individual reports be weighted by area? Has management direction succeeded if there are increases in some areas and decreases in others, leading to a zero net change for the region? Should a mechanistically or statistically based model be used for the region?

Formulation of effectiveness measures for the region and relating them to data throughout the region requires the development of a conceptual framework which allows not only for both scaling from the bottom up but also the top down; it is necessary to scale down from the basin and regional levels to local levels, in the process of stepping down policy directives to district procedure. This entails the ability to concretely relate broad-brush

policies all the way down to specific sampling designs. For example, in the NWFP this type of challenge begins with translating the Presidential mandate to "maintain biological diversity" and the series of decisions which relate this directive to a given sampling design for the Northern Spotted Owl at the district level.

Large scales imply not only larger spatial scales but also larger temporal frames. General ecological theory suggests that as the spatial scale of processes increase, the temporal scales over which significant change occurs also increase. Climatic responses and ecological scales of variation occur over decades to centuries and even millennia. From the sampling perspective, this means that time scales of 20 to hundreds of years may be needed to detect meaningful change in the system. However, while many of the target organisms and processes have life histories of decades to centuries, monitoring and decisions based on these data are scheduled annually or at most a few years; the institutions which oversee monitoring long-term processes and organisms are governed at very different, and much shorter, time scales. The scientific process and time scales associated with many ecological phenomena are infrequently commensurate with the scales at which the public and politicians function (Bradshaw 1998; Bradshaw and Fortin 2000). Ecological theory demands consideration of large-scales; policy requires statistically rigorous information for decision-making tractable thus far only at much finer scales (Bradshaw and Fortin 2000; Ludwig et al. 1993). As a result, it is often not possible to statistically distinguish human-derived from natural change given such short periods for detection (Innes 1998). Inference strength used to judge management success depends on the appropriateness of both the spatial and temporal sampling frame used for a given variable or attribute. This entails not only extending the temporal extent (e.g., allowing longer periods of time for monitoring before a decision is made) but also allowing flexibility in institutional response times (see adaptive management: Walters 1986; Gunderson et al. 1995). Many of these issues are central to some of the limitations of trend analysis discussed above.

The levels, magnitude and types of change which define management success or effectiveness should be specified prior to monitoring to inform how, when, where, and what variables need to be sampled. That is, it must be decided what constitutes significant change for the region. Ecologically important change can occur in many forms: changes in trend, frequency, timing, variability, etc. Subsequently, all criteria for effectiveness may not be expected to be explained or supported by trend analysis. Ecosystems are notoriously non-linear in behavior. Because we have incomplete knowledge of system history, including land-use and natural disturbances, and the nature of multi-scale interactions, the significance of a simple increase or decrease is difficult to interpret. Ascertaining cause and effect is difficult. A change in slope over time may actually reflect a longer period response to climatic shifts rather than a specific shift in management practices. Specific responses to management may be hidden by other environmental signals during the sampling period. Complex and multiple interactions across scales impede ready identification of cause and effect and attenuate our abilities to detect meaningful change. The existence of cumulative effects, non-linear, and threshold behaviors argue strongly for the coupling of validation and effectiveness monitoring.

Once models and success criteria are selected, it is necessary to design a sampling scheme. This task is often frustrating. As discussed above, the identification of the appropriate sampling units, frames, and population for large-scale, multi-scale, and multi-variate systems is difficult. Additionally, once these attributes are identified and sampling implemented, the results are often disappointing. Statistics and ecology do not always seem to match. For example, in some cases the biologist has a firmly held belief, either from extant literature from studies elsewhere, experience etc., that a given stressor has an important effect on a given process (e.g., sedimentation on fish health). However, when data comes back and is analyzed, the statistics indicate differently. The apparent mismatch between what is perceived as ecologically important change and its detection using statistical criteria occurs because of the aforementioned spatial and temporal scaling problems, but also because it is sometimes difficult to accurately measure critical, discrete quantities which will capture the dynamic and connected behavior of a complex system.

To this end, the use of indicators or indices (e.g., watershed integrity) have been suggested to help address the problem of handling multiple stressors and responses for large-scale behaviors. These integrative measures are intended to capture the essence of the ecosystem behavior when single or clusters of variables seem to fail. However, this introduces another problem. Often these value-laden indices rely on a qualitative or non-rigorous way of quantification. As a result, they are only relative terms with limited scope of inference beyond a certain set of ecological conditions, or are too subjective and introduce significant observer error. Thus, using such indices introduces another source of uncertainty in the inference and interpretation of monitoring results. In part, some monitoring can be directed to understanding these scale-related issues; such monitoring lies at the interface of research and validation monitoring.

Understanding knowledge gaps and sources and types of uncertainty, and articulation of the conceptual underpinnings of the monitoring plan are essential for appropriate assessment of effectiveness.

Validation Monitoring

Many of the questions raised for effectiveness monitoring hold for validation monitoring. One of the most difficult, unresolved questions is: given the uniqueness of each watershed and landscape, how valid is it to generalize intensive, sentinel experiments to other sites with distinct environmental and land-use histories, let alone extrapolate to the region? Coordination with extant long-term sites and inter-site comparison coupled with modeling efforts can help in some cases.

Further Considerations

Generally, larger spatial extents imply increases in sample sizes. Increased sample size means more time and effort are required for data collection and quality assurance. However, finite resources oblige management and planning teams operating at the local level to be parsimonious in the design of monitoring plans. Adequate sampling in space, time, and number of variables generally exceed local resource capacities. Logistics and

funding constrain many studies which seek to use landscapes or watersheds as sampling units. As a result, managers must decide what variables receive priority. Fiscal constraints may be such that even the way in which a variable is sampled has to be altered to be tractable. The combination of multiple goals, large scales, and fiscal constraints put a tremendous pressure on resources and leave the question of prioritization up to the local manager. A given district or national forest is subject to multiple pressures and requirements ranging from local issues to national policies and regulations. Without thoughtful anticipation, insufficient or inadequate data for regional assessment may result. Prioritization of management objectives becomes a critical step to optimize resources and information.

Summary Points

- The expansion to larger scales and increased number of variables of concern coupled with a need to provide high levels of certainty creates a significant challenge for the formulation of effective monitoring plans.
- A number of conceptual and technical issues still need to be researched and be resolved before the conventional approach to monitoring can be successfully applied to large-scale management efforts.
- Fundamental questions such as "what constitutes significant change at the regional scale" have not been answered, yet this question forms the basis for judging the success of any given effectiveness monitoring plan.
- The formulation of effectiveness measures for the region and relating them to data throughout the region requires the development of a conceptual framework which will be able to both scale from the top down as well as the bottom up. The conceptual framework not only identifies relationship between variables and processes across scales, but also will reveal where there are knowledge gaps or uncertainty. Uncertainty is inescapable, as is the pressure to act under conditions of uncertainty. Thus, the best safeguard and use of scientific knowledge is an explicit conceptual framework and the identification of knowledge gaps.
- Another source of difficulty for designing large-scale monitoring plans derives from the fact that temporal scales of human institutions and ecological systems are not matched; detection of change and statistically significant results are not always possible under short (e.g., annual to five year) sampling windows. Efforts to match human institutions and ecological scales will improve the ability to design effective monitoring.
- Variables should be selected carefully, avoiding ill-defined indices that offer little more than convenience.
- Strong coupling between validation and effectiveness monitoring will help address these complications.

IV. A DECISION-BASED APPROACH

As we have discussed above, the conventional approach offers a useful foundation from which to start developing a monitoring plan, but the approach can be thwarted by some of the complexities involved in monitoring a large-scale, complex set of activities. Two obvious shortcomings of the conventional approach are (1) the monitoring effort is not directly tied to the decision process itself, rather leaving this important feedback loop to a more *ad hoc* process, and (2) the conventional approach does not explicitly address how one prioritizes among the multiple possible indicators and questions that might be monitored.

In this section, we reach beyond the conventional approach to bring in concepts of quantitative decision analysis and adaptive management. Our purpose is to introduce a more rigorous framework for tying monitoring to information needs for decision making. Specifically, we propose a formalized method for hypothesis formulation and testing embedded in a decision analysis framework. The methods that we introduce are not novel or untested, enjoying extensive application throughout industry and government, but seemingly have not seen widespread exposure or use in natural resource management. Because the ideas are new to many, our discussions are relatively elementary and the examples simple. This does not suggest that common resource issues are too complex to yield to similar analyses, although the effort to complete such analyses can be demanding.

Statistical Evidence

We begin with a discussion of the nature of statistical evidence, that is, how should one interpret data. In this task we will rely heavily on the concepts of likelihood, following the excellent description by Richard Royall (1997).

In the preface to his book, Statistical Evidence: A likelihood paradigm, Royall (1997) states,

Science looks to statistics for help in interpreting data. Statistics is supposed to provide objective methods for representing scientific data as evidence and measuring the strength of that evidence. Statistics serves science in other ways as well, contributing to both efficiency and objectivity through theories for the design of experiments and decision-making, for example. But its most important task is to provide objective quantitative alternatives to personal judgment for interpreting the evidence produced by experiments and observational studies. In this role statistics has made fundamental contributions to science.

Royall proceeds from there to note that all is not well in statistical science and builds his case for adopting the likelihood paradigm for statistical inquiry. We will not go into the details of the likelihood paradigm here, nor describe how it differs from classical statistical techniques, but it is worth noting that most of the statistical techniques used in ecology and elsewhere do not follow directly from the likelihood paradigm. Rather, they

follow from statistical theory advocated by Neyman and Pearson (1933) (e.g., hypothesis testing, Type 1 and Type 2 error probabilities), or alternatively, Fisher (1959) (e.g., significance tests). There is ongoing debate among the statistical community as to whether the likelihood paradigm should play a more central role. What is not contested, however, is that the likelihood paradigm is fully consistent with a statistical decision making framework.

What is the likelihood paradigm, and what role can it play in linking monitoring and decision making? We start, following Royall, with three basic questions that one can answer when confronted with data: (1) what do the data say with regard to the relative truthfulness of a given hypothesis or set of hypotheses; (2) what should one believe given the data; and (3) what action should one take.

The first question is one that falls solidly within the realm of science. That is, there should be formal methods for determining which hypothesis is best supported by observed data, and further quantifying the strength of that support. The law of likelihood quantifies this evidentiary support. In simplest terms, the law of likelihood states that the hypothesis which is best supported by the observed data is the hypothesis which has the highest likelihood of having produced that data. Furthermore, the strength of evidence supporting a given hypothesis vis-à-vis a competing hypothesis is determined by the ratio of their likelihoods. For example, if the likelihood of observing $X = 3$ is 0.2 under hypothesis A and 0.8 under hypothesis B, then B is 4 times more likely to be correct than A, given $X = 3$ is observed.

The second question ventures beyond the data themselves and into the realm of personal belief. What one believes given new data will depend in part on prior beliefs. For example, if a tossed coin lands on heads, is that sufficient evidence for one to believe that every toss will be heads? What if three tosses in succession land on heads? What if three hundred tosses land on heads? In each case, the most likely hypothesis given only the observed data is that all tosses land on heads. Yet, our prior belief that it is a fair coin leads us to continue to believe that a toss landing on tails is equally likely, despite the occurrence of one heads or three heads in succession. Even the most die-hard believer in a fair coin, however, would have her belief shaken by 300 successive heads.

This intuitive sense of how evidence affects beliefs was first articulated mathematically by Reverend Thomas Bayes in 1763 and has given rise to a field of statistics known as Bayesian statistics. Bayesian statistics allows prior beliefs to be combined with empirical evidence in order to update beliefs regarding parameters (Press 1989), taking advantage of Bayes theorem and the likelihood principle. Bayesian approaches are common in the decision analysis literature and have been suggested for ecological analyses (Ellison 1996), but remain heresy to many empirically based statisticians because of the subjective nature of prior beliefs (see Dennis 1996). (Other classical methods have subjective elements in them as well, but users don't readily admit to it.)

The last question of what one should do moves beyond evidence and belief and into the arena of risk management and values. Consider the coin tossing example again. In this case, imagine that someone extracts a coin from their pocket and tosses it five times in succession, and each time it lands on heads. Before tossing it a sixth time, however, a bet

is proposed. If the coin lands on tails, the tosser will pay you \$1000; if it lands on heads, you pay him \$100 dollars. Do you take the bet? The answer, of course, depends on a variety of factors, among them the character of the person tossing the coin, i.e., do you believe that the coin tosser is intentionally trying to deceive you. Assuming for the moment that the character of the tosser can be ignored, foremost in one's consideration of the bet is a personal attitude towards risk. For some, \$100 is valued too highly to risk it in all but the surest cases (i.e., they are risk adverse). For others, \$100 seems a small sum to risk for a possible 10-fold payoff (i.e., they are risk-neutral); there are even others that will be excited by the gamble and take the bet irregardless of the odds (i.e., risk seeking; casinos love these individuals). Both risk-adverse and risk-neutral individuals will want to carefully assess their belief that the coin is fair; each group, however, will accept or reject the bet at a different odds ratio.

Examining evidence, beliefs, potential payoffs, as well as attitudes about risk falls within the domain of statistical decision theory (Clemen 1996; Hamburg 1987). In the following sections, we present a more formal representation of these ideas and show how decision analysis can be used to both inform decisions and structure a monitoring program. By *decision analysis* we mean a rational, formal analysis of options available in the course of making any decision, and the potential consequences of those options. For a brief overview of the lexicon of decision analysis, see the paper by Tom Spradlin (1997) on the Web Pages of the Decision Analysis Society (<http://www.fuqua.duke.edu/faculty/daweb/lexicon.htm>).

Influence Diagrams

The next step in building a framework for linking monitoring to decisions is to introduce the concept of influence diagrams (Howard and Matheson 1981; Clemen 1996). The purpose of an influence diagram is to represent the decision process in a way that explicitly recognizes the uncertainty in consequences or outcomes of the decision. Influence diagrams consist of nodes or variables connected by directed arrows. There are three kinds of nodes: (1) *decision nodes* representing alternative actions that can be taken by decision makers; (2) *chance nodes* representing events or system variables that are outcomes of the decision or other chance variables; and (3) *value or utility nodes*, variables that summarize the final outcome of a decision. In business decisions, *value* nodes are often expressed in monetary units. For other kinds of outcomes, the relative satisfaction offered by a particular outcome is summarized by its *utility*, a non-dimensional metric that allow comparisons of dissimilar elements (e.g., fish versus timber). Relations between outcomes and utility are expressed as utility or *preference* functions; such functions reflect both comparative value and attitudes about risk (Keeney and Raiffa 1976). The numeric values of the preference function generally is less important in ranking alternatives than the function's shape. This feature helps focus discussion of preference functions, which can be a contentious matter among interested parties with very different value systems. While it is possible to analyze decisions without explicitly assigning values or utilities to outcomes, the act of choosing one outcome or the other as preferable implicitly reveals a preference function.

The arrows connecting the nodes in an influence diagram represent conditional dependence, which is related to the conventional notion of causality (Pearl 1988; Shipley 2000). That is, the likelihood of the receiving node taking on a particular value depends on the decision or the realized values of donor chance nodes. Utility nodes are strictly receiver nodes; they assume particular values based solely on the realized value(s) of chance nodes, i.e., the outcome. In the parlance of decision analysis, the utility values are attributed back to the decision. Thus, the objective of the influence diagram is to identify the decision option with the highest expected utility. As we will discuss below, influence diagrams serve other purposes as well.

The key to making an influence diagram more than simply a schematic representation of the interaction of decisions and chance variables is in the mathematical representation of causal dependencies. The strength of these dependencies is captured in conditional probability matrices that link chance nodes to decisions or other chance nodes. The underlying framework follows probability theory and the likelihood principle. Influences are propagated through the network using Bayes theorem. This mathematical framework provides a strong conceptual link between influence diagrams and statistics that is absent in other modeling approaches (e.g., simulation modeling, fuzzy logic). It also provides a natural link back to how one uses data to update or verify the diagrams.

An Example using Wildfire and Fish

We illustrate the use of an influence diagram with an example involving wildfire and fish populations. At issue is a decision regarding the level of fuels treatment (thinning, prescribed fire, etc.) that might be called for within a watershed with unnaturally high fuel loads. The fuel loads suggest a relatively high likelihood of a stand-replacing wildfire occurring within the next 50 years. The decision is complicated by the fact that the watershed supports a sensitive fish population. There are two competing lines of reasoning regarding the fuels treatment. The first argues that a high intensity, stand-replacing fire leads to loss of scenic, recreational, and other values, and furthermore could lead to increased flooding and erosion, and thus degradation of instream fish habitats. The competing argument is that while fire effects are uncertain, the fuel treatment itself would surely increase erosion to some extent, which might also affect fish. The worst possible scenario would be fuel treatment that was ineffective in preventing a high intensity fire, but which aggravated erosion and habitat degradation even more once a fire occurred.

To examine this decision, we constructed the influence diagram shown in Figure 4.1. The diagram consists of a single decision node, six chance nodes, and a utility node. At the top of the diagram is the decision node, *fuel treatment*. Two possible fuels treatment levels are proposed, treat the entire watershed or treat the upslopes only and leave the riparian areas untreated; a third decision option is to do no fuels treatment (Figure 4.2). A chance node, *fuel ignition*, represents the frequency of natural fire ignitions and is the only independent chance node in the diagram, i.e., its value does not depend on other nodes but is determined exogenously. Fuel treatment and fire ignition combine to affect the likelihood of wildfires of different intensity within the watershed. Wildfire and fuel treatment in turn affect fire in the riparian zone; wildfire also affects flooding regime.

Instream fish habitat is the next link in the network, and is affected by the combination of fuel treatment, fire in the riparian area, and the flooding regime. The terminal chance node, fish population, is affected only by instream habitat. Utility is expressed as a function of wildfire, instream habitat, and fish population. The reason for including fish habitat in the utility function is the recognition that fish populations can be affected by factors other than habitat that are not explicitly included in the diagram. Thus, some utility is given to an outcome of improved or protected habitat even when the fish populations fail to follow suit.

Each of the chance nodes is divided into two to four discrete levels. The relationships among them are quantified by the conditional probability matrices (CPM). The values in these matrices were contrived for this example and should not be taken seriously (Table 4.1). Table 4.1 also contains the utility values assigned to various outcomes. For example, one possible outcome would be no fires in the watershed, no fire in the riparian zone, no change in flooding frequency, no change in fish habitat, and a stable fish population. When a decision is specified, its "influence" is propagated through the network of nodes, changing the probabilities of all possible outcomes. Overall expected utility is calculated by summing the utilities associated with each possible outcome, weighted by their respective probabilities, given a decision.

Based on the influence diagram as constructed, we can examine the relative utilities of different fuel treatments (Table 4.2). The results suggest that the optimal decision depends on the natural fire ignition frequency. If all possible frequencies (rare, infrequent, or frequent) are equally likely, the suggested preferable decision is to treat the entire watershed. Now suppose that fire ignition frequency can be accurately estimated; then the decision can be evaluated based on known ignition frequency. In this case we find that the preferable decision depends on the ignition frequency. If fires are ignited either rarely or frequently in the watershed, the preferable decision is to treat only the upslope areas. If fires are infrequently ignited, the preferable decision is to treat fuels in the entire watershed. Note that in our contrived example, the cost of treatment is not factored into the expected utility. In cases with many possible decision options, one can plot the expected utility versus costs and define a cost-effectiveness frontier, i.e., the set of alternatives that maximize utility for a given cost.

For further exploration of this example, we suggest you open the file, fish_n_fire.dne, using the Netica software package that is available from Norsys Inc. (<http://www.norsys.com>) or convert the CPM for use in other programs (e.g., [Hugin](#), [others](#)). This example file is small enough to run on the evaluation copy of Netica that is available free of charge. Further information on influence diagrams also is available on the various websites¹. (Note:)

Benefits of the Influence Diagram

As described above, influence diagrams are generally constructed for the purpose of identifying preferable decisions in an uncertain environment. Examples of such use can

¹ This information is provided for information purposes only and is not an official endorsement of any commercial product by the USDA Forest Service.

be found in Lee and Reiman (1997), Reiman et al. (2001), and Marcot et al. (2001). This type of pre-decisional analysis, however, is but one benefit of having a formal influence diagram. Other benefits that are germane to the issue of developing a monitoring plan include the following.

1. The influence diagram is an explicit, formal representation of all knowledge of the system that is relevant to the decision. There is no vague logic trail behind the decision with competing world views arguing this or that is the best decision. All assumptions and preferences are fully revealed. In a sense, the influence diagram represents a formal, quantified, composite hypothesis of how the world operates, including how the decision makers value potential outcomes. Hypotheses are embedded in the influence diagram in two ways: (1) in the structure of the diagram itself; and (2) in the conditional probabilities assigned to each chance node. In a more complete analysis, one can construct alternative influence diagrams that reflect different views or hypotheses of how the world operates. The decision analyst can then seek out a decision that is optimal across all competing hypotheses. From a monitoring perspective, the hypotheses embedded in the influence diagram can be examined in a scientifically rigorous manner using empirical data.

2. The mathematical structure underlying the influence diagram lends itself to quantitative sensitivity analysis and formal testing of alternative model structures. Thus, one can readily determine where uncertainty in chance nodes is most influential in affecting outcomes, and the degree to which the conditional probabilities are supported by data. Furthermore, recent methods developed by Pearl (1988, 2000) allow for alternative causal pathways to be hypothesized and examined for their consistency with observational data (see Ripley 2000). One also can readily discover the sensitivity to the utility function. These mix of insights are extremely valuable. Such analyses point to areas of needed ecological research or priorities for ecological monitoring to either reduce the uncertainty in a chance node (if possible) or refine the node descriptions, causal structure, and conditional probabilities through data collection and analysis. The influence of the utility function suggests an active role for socioeconomic research to accurately determine how the public and decision makers value resources and regard risk.

3. Influence diagrams are a natural mechanism for integrating scientific information across disciplines. For example, if the fire and fish diagram that we illustrated above were to be constructed with the best available information, the services of a fire ecologist, a hydrologist, a fish biologist, and a social scientist would be required to build and parameterize the model. Add to that a decision analyst if none of the others have the prerequisite skill to put the influence diagram together. These sorts of interdisciplinary teams are not new to resource management agencies; what is new is that there rarely seems to be a common framework in which to express interdisciplinary understanding of the system of interest.

4. The modular structure of the influence diagrams allows them to be combined or partitioned in different ways depending on the scale and complexity of the decision in question. Furthermore, relations between nodes can also be collapsed or expanded depending on the detail required. Going back to our fire and fish example, the effects of riparian fire and flooding regime on instream habitat are complex. Our example simplifies this relationship tremendously in order to capture no more than is necessary to

inform the decision at hand. One might want to expand this segment of the model in order to more fully articulate relationships. Such expansion might lead to improved understanding that could be used to update the simpler version, or to create an alternative structure that better distinguishes conditions under which fire in the riparian zone has positive or negative benefits.

5. The influence diagram approach cleanly distinguishes how one addresses the three question posed above from Royall (1997). While the objective of the model is to answer the question, "what does one do," there are more specific questions as well. Scientists can play an objective role in structuring and parameterizing the diagrams in a way that is most consistent with scientific understanding and data, answering the question of which hypothesis is best supported by data. Similarly there are formal statistical techniques for updating these diagrams as monitoring data are collected, using Bayes theorem and the likelihood principle. The decision analyst can decide, in combination with the decision makers, if there is sufficient empirical data to build the diagrams or if more subjective professional opinion is required, quantifying beliefs in the process. If necessary, there are formal methods for eliciting such information (Morgan and Henrion 1990). Both decision analysts and social scientists will be engaged in building utility functions.

6. Influence diagrams are a natural fit to the concept of adaptive management. Building models such as influence diagrams that explicitly acknowledge uncertainty is a central step in adaptive management (Walters 1986). There are two kinds of uncertainties contained in an influence diagram: those arising because of random events in nature, and those arising from failure to fully understand how the world operates (ignorance). By examining these models, one can tell not only which uncertainties are likely to impede management objectives, but also identify management decisions that if properly monitored can reduce uncertainties based on ignorance. Such insights can lead to management actions for the sake of learning. That learning experience can be formally captured in decision models, which in turn improves management efficiency.

Table 4.1. Node descriptions and conditional probability tables for the example fire and fish influence diagram. These are contrived values and should not be taken seriously.

Decision Node: Fuel treatment

Levels (3): no treatment, treat entire watershed, treat upslopes outside riparian areas

Chance Node 1: Natural Fire Ignition

Levels (3): rare, infrequent, frequent

Chance Node 2: Wildfire in Watershed

Levels (3): no fires, low intensity fire, high intensity fire

Conditional Probability Matrix:

Fuel Treatment	Fire Ignition	Fire in Watershed		
		no fires	low intensity	high intensity
none	rare	0.90	0.02	0.08
	infrequent	0.50	0.20	0.30
	frequent	0.10	0.60	0.30
entire watershed	rare	0.92	0.06	0.02
	infrequent	0.60	0.30	0.10
	frequent	0.28	0.58	0.14
upslopes only	rare	0.92	0.05	0.03
	infrequent	0.13	0.52	0.35
	frequent	0.28	0.54	0.18

Chance Node 3: Wildfire in Riparian Area

Levels (3): no fires, low intensity fire, high intensity fire

Conditional Probability Matrix:

Fuel Treatment	Fire in Watershed	Fire in Riparian Area		
		no fires	low intensity	high intensity
none	none	1	0	0
	low intensity	0.40	0.60	0
	high intensity	0	0.10	0.90
entire watershed	none	1	0	0
	low intensity	0.40	0.60	0
	high intensity	0	0.40	0.60
upslopes only	none	1	0	0
	low intensity	0.40	0.48	0.12
	high intensity	0	0.20	0.80

Table 41. Continued.

Chance Node 4: Flooding Regime
 Levels (2): no change, increased flooding

Conditional Probability Matrix:

Fire in Watershed	Flooding regime	
	no change	increased frequency
none	1	0
low intensity	0.8	0.2
high intensity	0.3	0.7

Chance Node 5: Instream Fish Habitat
 Levels (4): improving, no change, minor loss of habitat, major loss of habitat

Conditional Probability Matrix:

Flooding	Fire in Riparian	Fuel Treatment	Instream Fish Habitat			
			improving	no change	minor loss	major loss
no change	none	none	0.10	0.80	0.09	0.01
		watershed	0.05	0.60	0.30	0.05
		upslopes	0.10	0.70	0.18	0.02
	low intensity	none	0.10	0.70	0.17	0.03
		watershed	0.05	0.53	0.37	0.05
		upslopes	0.05	0.60	0.33	0.02
	high intensity	none	0.05	0.50	0.30	0.15
		watershed	0	0.40	0.35	0.25
		upslopes	0	0.40	0.40	0.20
increased frequency	none	none	0.05	0.70	0.20	0.05
		watershed	0	0.50	0.40	0.10
		upslopes	0	0.60	0.35	0.05
	low intensity	none	0.05	0.50	0.40	0.05
		watershed	0	0.30	0.50	0.20
		upslopes	0	0.30	0.60	0.10
	high intensity	none	0	0.15	0.50	0.35
		watershed	0	0.05	0.45	0.50
		upslopes	0	0.10	0.50	0.40

Table 41. Continued.

Chance Node 6: Fish Population

Levels (3): increasing, stable, decreasing

Conditional Probability Matrix:

Instream Fish Habitat	Fish Population		
	increasing	stable	decreasing
improving	0.60	0.35	0.05
no change	0.15	0.70	0.15
minor loss	0.10	0.50	0.40
major loss	0.02	0.30	0.68

Utility Node: Utility is the sum of utilities for Wildfire in Watershed, Instream Fish Habitat, and Fish Population. Individual utilities as follows:

Node	Outcome	Utility
Wildfire in Watershed	no fires	10
	low intensity	3
	high intensity	-10
Instream Fish Habitat	improving	10
	no change	0
	minor loss	-5
	major loss	-10
Fish Population	increasing	10
	stable	5
	decreasing	-10

Table 4.2. Calculated expected utilities for different fuel treatment options, conditional on fire ignition frequency

Fuel Ignition	Fuel Treatment		
	none	entire watershed	upslopes only
Unknown	5.08	5.69	3.91
Rare	11.40	9.53	11.53
Infrequent	3.56	5.44	-2.27
Frequent	0.27	2.09	2.47

Figure 4.1. An example influence diagram linking fuel treatment options to potential effects on sensitive fish populations.

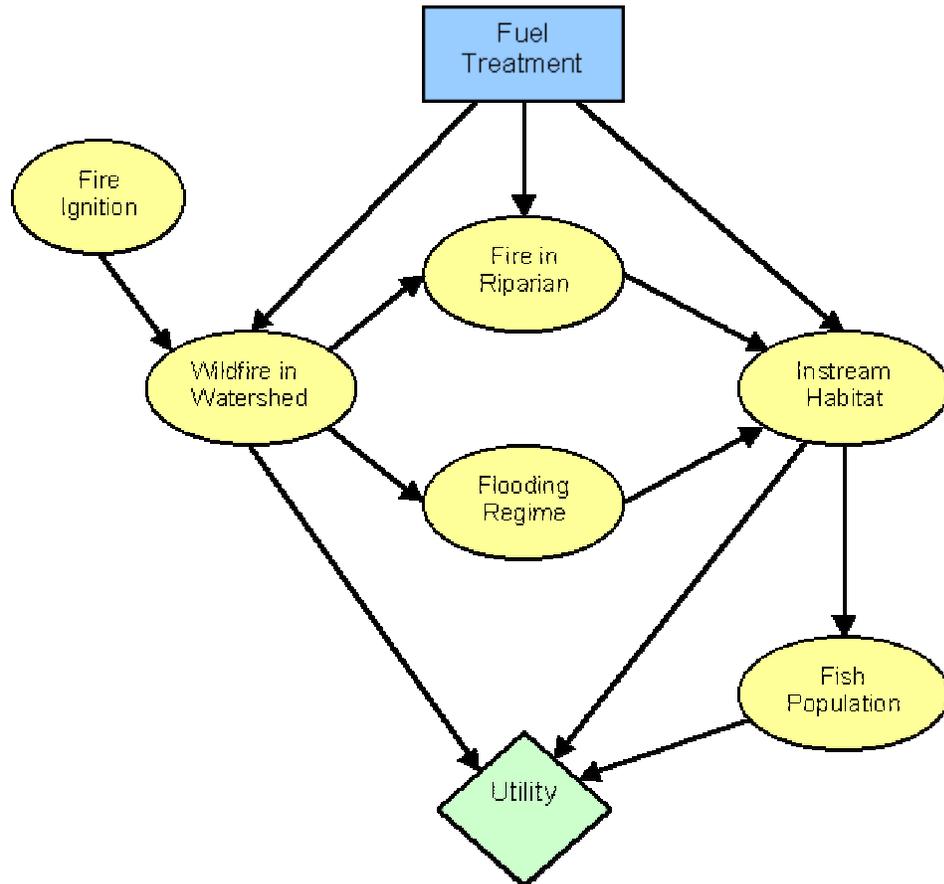
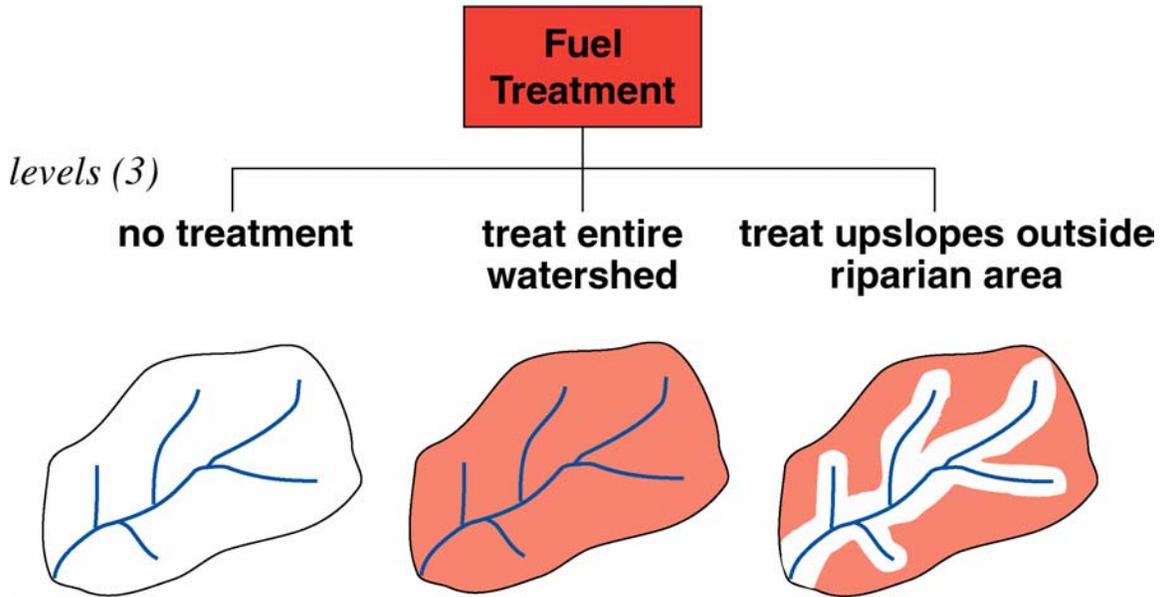


Figure 4.2 Schematic Representation of Different Fuel Treatments.

Decision Node: Fuel Treatment



V. INTEGRATING THE CONVENTIONAL AND THE DECISION-ANALYTIC APPROACHES

At this point, it is useful to note that the decision analysis framework that we are promoting is not mutually exclusive with many of the concepts articulated in the conventional approach. Rather, the decision analysis framework works hand-in-hand with the different types of monitoring and sampling designs to complete the informational loop among goals, decisions, and actions.

Figure 5.1 illustrates this informational loop and the essential role played by decision analysis. The process originates with general goals, objectives, and direction at multiple scales. Goals and direction may be articulated in regional directives such as the Northwest Forest Plan, local resource management plans, or more local concerns. Building an objectives hierarchy (Keeney 1992) can be useful in focusing the analysis on metrics directly related to management objectives and social values. Decision analysis, which can be subdivided into more discrete steps, allows further articulation of the linkages among objectives, actions, and values.

The first steps in decision analysis are to identify the relevant ecological processes at appropriate scales, and identify the range of management options that are available. These two component feed into the more formal analysis of options, which included use of the influence diagrams as discussed in previous sections. The formal analysis explicitly includes a sensitivity analysis to identify key uncertainties and dependencies, suggesting important indicators for monitoring. A potential list of indicators can be further screened for technical or administrative constraints (i.e., do we know how and can we afford to measure each indicator). A rough estimate of the feasibility and cost of monitoring a particular set of actions can help in choosing among alternatives.

The formal analysis should indicate the relative preferability of different options, given the overall goals. This leads directly to implementation decisions. These decisions initiate an entire set of consequences not shown on the diagram, that is, the actual implementation of management activities. Making an implementation decision also leads back to the formal analysis, but for a different purpose. Now the analysis is conditioned on the decisions that have been made, and the monitoring design begins in earnest. Using the influence diagrams and screening indicators for feasibility and affordability, a key set of management and ecological indicators are identified. This set of indicators is passed to the sampling design phase wherein the previous discussion about model-based versus sample-based designs comes to the fore.

Once a suitable sampling design is chosen, data collection and analysis soon follows. The results of this analysis can lead back to the formal analysis step or to an evaluation phase. Depending on the results of the evaluation of progress, feedback to the implementation decision step may occur.

The data collection and analysis step, and its feedback to the formal analysis of options step encompasses all of the monitoring types identified in the conceptual model (Figure 5.2). Implementation monitoring, for example, ensures that decisions were implemented as planned. Baseline monitoring is necessary to set up initial conditions in the network and parameterize some of the conditional dependencies. Validation monitoring provides the information necessary to examine the overall diagram structure for its agreement with observed data, and updates conditional probabilities matrices based on new observations. Finally, the overall linkage between decisions and outcomes provides the linkage necessary for effectiveness monitoring, in the sense of measuring progress towards an objective in a manner that the true effect of management actions is distinguishable.

Linkage to Adaptive Management

One of the insights that are sure to arise from the formal analysis of options is that there are certain management options that lead to outcomes with potentially high value, but also high uncertainty. That is, the most likely outcome given the option may be highly favorable, but the chance of unfavorable outcomes is also high. In some cases, this uncertainty is an inherent property of the natural system and cannot be avoided. In such instances, the management option is inherently risky, and there is no apparent way to avoid the risk.

Many times, however, the reason an option has high uncertainty is because of ignorance about how the system operates. It may be that the particular option has never been tried under a given set of circumstances, so one cannot say with certainty what the result might be. If the potential gain from the action is high, it might seem rational to try the option on a limited scope and carefully observe the outcome. This type of active probing of the system to increase knowledge of system behavior is central to the idea of adaptive management. Such management experiments must be carefully designed such that the results can be accurately attributed to the correct causal factors. Such designs cannot be achieved without something akin to the influence diagram approach.

Figure 5.1. The informational linkages among goals and direction, decision analysis, implementation, and monitoring.

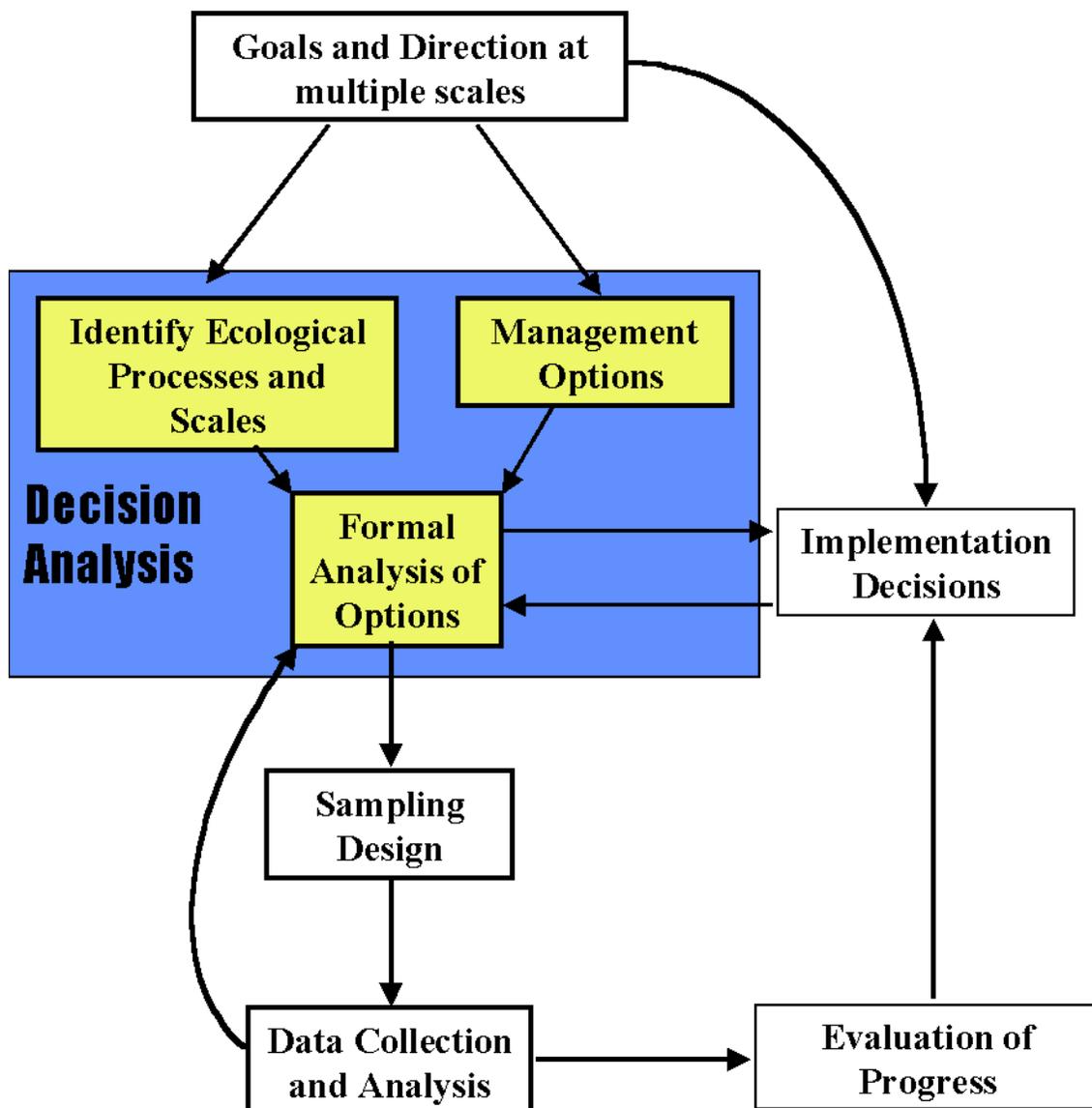
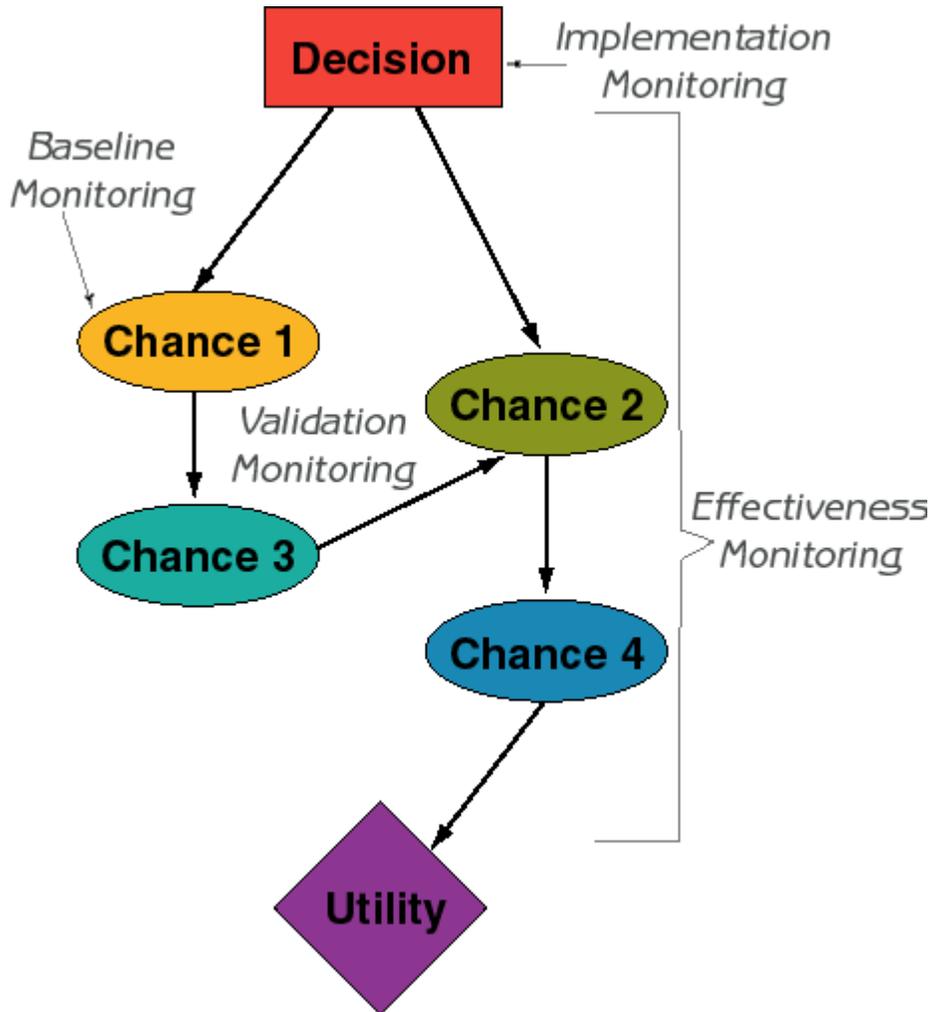


Figure 5.2. Schematic diagram of how the different types of monitoring interface with an influence diagram.



VI. IMPLICATIONS FOR BROAD-SCALE MANAGEMENT EFFORTS

In the preceding sections, we provide an overview of the conventional approach to monitoring, discuss some of the complications in applying the conventional approach to broad regional plans, and introduce decision analysis techniques as one inroad to explicitly bringing monitoring and decision making together. In our presentation, we have intentionally omitted much of the technical detail for the sake of brevity. By doing so, we run the risk of violating our own standards of rigor. Many of the ideas that we broach are areas of active research within the statistical and decision analysis arenas. Managing complex systems under uncertainty is not the exclusive purview of natural resource management agencies, nor is monitoring. Our central point is that there are techniques and approaches in the literature and under development in the larger research community that resource management agencies can use.

Broad-scale regional plans create an environment in which multiple decisions are called for at a variety of scales. Monitoring serves a vital interest in determining whether progress is made towards overall goals, and in improving understanding such that sound, informed decisions are made, regardless of actual outcomes. The difficulties associated with moving and relating information across scales call for a formalized process of hypothesis formulation coupled with and embedded in a decision-making framework where science, management, and policy-making are linked. Too often monitoring fails when scientific and management objectives are poorly matched. Mismatches evolve when the perception of specific monitoring objectives (i.e., hypotheses) and the scales of concern differ among scientists, managers, and policy makers. While many monitoring plans contain both large-scale and to varying degrees, fine-scale goals, what is generally lacking is an explicit delineation of the path and precise steps taken in this scaling down process.

Failure may be defined in several ways but most commonly means that monitoring has failed to provide the information needed for judicious and defensible decision-making. In this context, failure will occur if data are collected at the local level and are unusable for the intended policy objectives mandated at the regional and even national levels (Edvardsson 2004). At worst, such an informal articulation leaves the science and logic vulnerable to criticism that it is arbitrary and subjective--and therefore indefensible. At best, the lack of rigorous articulation hides critical areas of uncertainty and leaves decision-making decoupled from science. A careful mapping of management objectives (Keeney 1992; Slocombe 1998) to monitoring objectives from start (i.e., policy or public expectations) to finish (i.e., data collection and analysis) is necessary to ascertain whether specific, testable monitoring objectives executed at the ground level match broad expectations of both policy and monitoring. This remains a challenge for large and multiscale, multivariate, open systems such as the interior Columbia River Basin. As we discuss above, much of the uncertainty can be identified and addressed, if not resolved, within an explicit decision-making framework.

Conclusions

In conclusion, we propose that current monitoring plans based solely on the conventional approach lack two critical features: (1) an explicit decision-making structure linked with the monitoring plan; and (2) no formalized prioritization procedure. Without a decision analysis approach, a framework is lacking in which monitoring results may be used to judge effectiveness and directly linked to management decisions. This leaves monitoring results decoupled from the decision-making process. A lack of articulation between monitoring and decision-analysis can hide areas of uncertainty and the source of uncertainty. In an articulated decision analysis framework, decision-makers at all levels can see the array of possible paths and weigh the relative risk or desirability. We suggest that monitoring plans can be written from a decision analysis perspective and suggest that this approach will help facilitate the step-down process.

There are several benefits to using a formal decision analysis approach:

- This approach allows integration across scales, between research and management, across disciplines, and provides a common channel of communication.
- Because it requires quantification and rigor, the decision analysis framework functions as a permanent repository of scientific knowledge which is accessible to all and future projects.
- The modularity of the decision analysis allows for flexibility linking across scales and geographic locations. As such, it provides a viable tool for comparative studies.
- Assumptions are transparent and risk is quantifiable, leading to more defensible decisions.
- Its explicit articulation helps identify areas of uncertainty and promising research directions.
- The inference diagram method of modeling is theoretically and practically integrated with existing statistical techniques, providing a coherent strategy for data interpretation.
- The decision analysis approach facilitates rigorous adaptive management.

We recognize that our vision of decision-based monitoring requires further testing and refinement in real-world applications. Only through experience will some of the more mechanistic details of implementing a decision analysis framework for monitoring emerge.

VII. LITERATURE CITED

- Bradshaw, G.A., 1998, What constitutes ecologically relevant pattern in the process of scaling up? in *Ecological Scale: Theory and Applications* (D. Petersen and V.T. Parker, eds). Columbia University Press.
- Bradshaw, G.A. and M.-J. Fortin. 2000. Landscape heterogeneity effects on scaling and monitoring large areas using remote sensing data. *Geographic Information Sciences*. Vol 6. No. 1, pp. 61-68.
- Clemen, R. T. 1996. *Making hard decisions: An introduction to decision analysis*. 2nd edition. Duxbury Press, Belmont, CA.
- Dennis, B. 1996. Discussion: Should ecologists become Bayesians? *Ecological Applications* 6(4):1095-1103.
- Edvardsson, K. 2004. Using Goals In Environmental Management: The Swedish System of Environmental Objectives. *Environmental Management* 34(2): 170-180.
- Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* 6(4):1036-1046.
- Fisher, R. A. 1959. *Statistical methods and scientific inference*, 2nd edition. Hafner, New York.
- Gunderson, L.H., C. S. Holling and S. S. Light, 1995. *Barriers and bridges to the renewal of ecosystems and institutions*. Columbia Press. New York. 593 pp.
- Hamburg, M. 1987. *Statistical analysis for decision making*, 4th edition. Harcourt Brace Jovanovich, San Diego.
- Howard, R., and J. Matheson. 1981. Influence diagrams. Pages 721-762 in Howard, R., and J. Matheson (editors). *Readings on the principles and applications of decision analysis*, Volume II. Strategic Decisions Group, Menlo Park, CA.
- Innes, J. L. 1988. Measuring environmental change. Pages 429-457 in Petersen, D. and V. T. Parker, editors. *Ecological Scale: Theory and applications*. Columbia University Press.
- Jassby, A. D. 1998. interannual variability at three inland water sites: implications for sentinel ecosystems. *Ecological Applications* 8(2):277-287.
- Jensen, F.V., 2001. *Bayesian Networks and Decision Graphs*. Springer Verlag, New York, NY.
- Keeney, R.L. 1992. *Value-Focused Thinking. A Path to Creative Decisionmaking*. Harvard University Press. 416 p.
- Keeney, R. L., and H. Raiffa. 1976. *Decisions with multiple objectives: Preferences and value tradeoffs*. John Wiley & Sons, New York.

- Lee, D.C., Rieman, B.E., 1997. Population viability assessment of salmonids by using probabilistic networks. *North American Journal of Fisheries Management* 17:1144–1157.
- Ludwig, D., R. Hilborn, and C.J. Walters, 1993. Uncertainty, resource exploitation, and conservation: lessons from history. *Science* 260: 17, 36.
- Manley, P.N., W.J. Zielinski, M.D. Schlesinger, and S.R. Mori. 2004. Evaluation of a multiple-species approach to monitoring species at the ecoregional scale. *Ecological Applications* 14(1):296-310.
- Marcot, B.G., Holhausen, R.S., Raphael, M.G., Rowland, M.M., Wisdon, M.J., 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest Ecology and Management* 153:29–42.
- Mulder, Barry S., Barry R. Noon, Thomas A. Spies, Martin G. Raphael, Craig J. Palmer, Anthony R. Olsen, Gordon H. Reeves, and Hartwell H. Welsh. 1999. The strategy and design of the effectiveness monitoring plan 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.
- Morgan, M. G., and M. Henrion. 1990. *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge Press.
- Neyman, J., and E. S. Pearson. 1933. On the problem of the most efficient test of statistical hypotheses. *Philosophical Transactions of the Royal Society, Series A*, 231:289-337.
- Noss, R. F., and A. Y. Cooperrider. 1994. *Saving nature's legacy: Protecting and restoring biodiversity*. Island Press, Washington, DC.
- Olsen, T. [and others]. 1997. Ecological resource monitoring: Change and trend detection workshop report. *Bulletin of the Ecological Society of America* 78(1):11-13.
- Overton, C. K, [and 4 others]. 1995. User's guide to fish habitat: Descriptions that represent natural conditions in the Salmon River Basin, Idaho. General technical report INT-GTR-322. USDA Forest Service, Intermountain Research Station, Ogden, UT.
- Pearl, J., 1988. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*, second edition. Morgan Kaufman, San Mateo, CA.
- Pearl, J. 2000. *Causality*. Cambridge University Press.
- Peterson, D. L., and V. T. Parker, editors. 1998. *Ecological Scale*. Columbia University Press, New York. 1998. *Ecological Scale*. Columbia University Press, New York.
- Press, S. J. 1989. *Bayesian statistics*. John Wiley & Sons, New York.

- Rieman, B., Peterson, J.T., Clayton, J., Howell, P., Thurow, R., Thompson, W., Lee, D. 2001. Evaluation of potential effects of federal land management alternatives on trends of salmonids and their habitats in the interior Columbia River basin. *Forest Ecology and Management* 153:43–62.
- Ricklefs, R. E., and D. Schuller, editors. 1993. *Species diversity in ecological communities*. University of Chicago Press.
- Ringold, P. L. [and 5 others]. 1996. Adaptive monitoring design for ecosystem management. *Ecological Applications* 6(3):745-747.
- Royall, R. 1997. *Statistical evidence: A likelihood paradigm*. Chapman & Hall, London.
- Shipley, B. 2000. *Cause and Correlation in Biology*. Cambridge University Press.
- Slocombe, D. S. 1998. Defining Goals and Criteria for Ecosystem-Based Management. *Environmental Management* 22(4): 483-493.
- Spellerberg, I. F. 1991. *Monitoring ecological change*. Cambridge University Press.
- Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. *Ecological Applications* 8(2):246-257.
- Walters, C. 1986. *Adaptive management of renewable resources*. Macmillian, New York.

Web sites and Links

Interior Columbia River Basin Ecosystem Management Project: <http://www.icbemp.gov/>

Northwest Forest Plan: <http://www.fs.fed.us/r6/nwfp.htm>

University of Oregon Library System: Northwest Forest Plan: Authoritative Documents: <http://libweb.uoregon.edu/govdocs/forestpl.html>

Sierra Nevada Forest Plan Amendment: <http://www.fs.fed.us/r5/snfpa/>