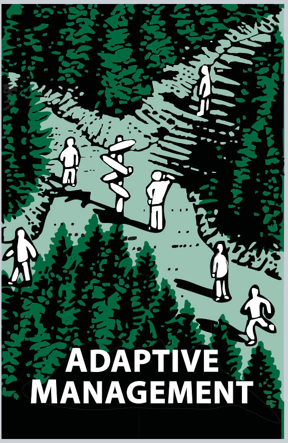
LAND MANAGEMENT HANDBOOK

42

Statistical Methods for Adaptive Management Studies



1998

British Columbia

Ministry of Forests
Research Program

Statistical Methods for Adaptive Management Studies

Vera Sit and Brenda Taylor (Editors)

BRITISH COLUMBIA

Ministry of Forests
Research Program

Citation
Sit, V. and B. Taylor (editors). 1998. Statistical Methods for Adaptive Management Studies. Res. Br., B.C. Min. For., Res. Br., Victoria, BC, Land Manage. Handb. No. 42.

Canadian Cataloguing in Publication Data
Main entry under title:
Statistical methods for adaptive management studies

(Land management handbook ; 42)

Includes bibliographical references: p.
ISBN 0–7726–3512–9

1. Forest management – Statistical methods.
2. Forest management – British Columbia. I. Sit, Vera.
II. Taylor, Brenda, 1962– . III. British Columbia. Ministry of Forests. Research Branch. IV. Series.

SD387.S73S72 1998 634.9'2'072 C98–960081–5

Prepared for:
B.C. Ministry of Forests
Research Branch
PO Box 9519, Stn Prov Govt
Victoria, BC V8W 9C2

Published by
B.C. Ministry of Forests
Forestry Division Services Branch
Production Resources
595 Pandora Avenue
Victoria, BC V8W 3E7

Copies of this and other Ministry of Forests titles are available from:
Crown Publications Inc.
521 Fort Street
Victoria, BC V8W 1E7

Send comments to: Vera Sit, B.C. Ministry of Forests, Research Branch,
PO Box 9519, Stn Prov Govt, Victoria, BC V8W 9C2

Ministry of Forests Publications Internet Catalogue: www.for.gov.bc.ca/hfd

9 SELECTING APPROPRIATE STATISTICAL PROCEDURES AND ASKING THE RIGHT QUESTIONS: A SYNTHESIS

BRUCE G. MARCOT

Abstract

In this chapter, I synthesize statistical approaches and considerations for adaptive management studies. I review approaches to learning from management actions and address questions of space and time. I also present a set of guidelines for asking the right questions about statistical reliability, for selecting the appropriate adaptive management study, and for guiding how different types of information can contribute at different stages in adaptive management. These guidelines are presented in a table, which can be used as a decision tree to determine the best kinds of studies for each step in the adaptive management process, and the most appropriate use of exisiting information.

9.1 Introduction

How should managers and researchers select an approach for designing an adaptive management study and analyzing the results? The chapters in this report provide some guidance; for example, Nemec (this volume, Chap. 2), summarizes principles of experimental design, and Schwarz (this volume, Chap. 3) lists types of nonexperimental and experimental designs. Other publications (e.g., Green 1979), while not specific to adaptive management as defined in this volume, also provide guidance on designing ecological studies. This chapter reviews issues to consider in designing adaptive management studies, synthesizes the methods discussed in preceding chapters of this report, and summarizes the roles different types of information can play in adaptive management.

Statistical approaches and study designs can be selected only when the management question is first well articulated. In the first section of this chapter, I review three types of monitoring, differentiated by the types of question they each address, and then address how the spatial and temporal elements of a management question can influence study design. In the second section, I review the characteristics of powerful studies and the principles of experimental design. The third section summarizes various types of information (including existing data, retrospective studies, and nonexperimental studies) and experimental studies, and how they can contribute to adaptive management. In the final section, I discuss some points to consider in interpreting and communicating the results from adaptive management studies, and in particular the difficulty in "unravelling the causal web." Throughout this chapter, I use the oversimplistic labels "researcher" and "manager," fully realizing that in the real world many resource professionals don both hats.

9.2 Types of Questions Addressed in Adaptive Management

A little experience often upsets a lot of theory.

– Cadman

The B.C. Ministry of Forests defines adaptive management as a formal process entailing problem assessment, study design, implementation, monitoring, evaluation, and feedback (B.C. Ministry of Forests 1996). In this approach, management activities are crafted as experiments to fill critical gaps in knowledge. The key questions are: (1) To what extent did the managment action lead to the measured outcome? and (2) Are our assumptions valid about how the system works?

Other institutions use the term "adaptive management" differently. For example, the USDA Forest Service incorporates the general concepts of adaptive management into its planning, but not as a formal process. Regardless of the definition of adaptive management and how it is institutionalized, monitoring activities and evaluation of data are key steps in adaptive management. The statistical approaches discussed in this report can help in both the design of monitoring activities and in the interpretation of data.

9.2.1 Types of monitoring

There are three types of monitoring: implementation monitoring, effectiveness monitoring, and validation monitoring. Each type of monitoring serves a unique function in an adaptive management study.

Implementation monitoring

Implementation monitoring (or compliance monitoring) essentially asks: Have the management guidelines been implemented correctly (Collopy et al. 1993)?

Correct implementation can be determined by a complete census of all activities or by sampling activities stratified by administrative unit or location. Obviously, asking more detailed questions of the effects and validity of particular management activities should proceed only when they have been correctly implemented. Implementation monitoring, however, does not teach us about effects of management actions. Thus, the focus of adaptive management is effectiveness and validation monitoring.

Effectiveness monitoring

Effectiveness monitoring asks: Are the management guidelines and activities producing the desired effects? Do the management activities really alter the biophysical conditions as expected? Many questions can be asked of the effects of management guidelines. Highest priority should be directed to potential effects that have the most serious economic, biological, or ecological ramifications, and those carrying the greatest uncertainty.

Validation monitoring

Validation monitoring, technically the most difficult of the three kinds of monitoring, asks: Are the ultimate expectations for the guidelines being met? Are the basic assumptions about how the biophysical system operates really correct, or does it operate in a very different way that would invalidate the selected management approach? If so, how?

Validation monitoring may be used to validate ecosystem models (Gentiol and Blake 1981), which is vital to ensuring the models' successful and appropriate use. In adaptive management, validation monitoring should focus on the ecosystem elements that have the greatest implications on the decision about the best course of action. Problem assessment—identifying which relationships to validate—is the first step of adaptive management.

9.2.2 Issues of space and time

Issues of space and time will in part determine the type of study design that is possible. For example, studies of large geographic areas may preclude replication, suggesting before-after-control-impact paired (BACI-P) study (Schwarz, this volume, Chap. 3). Similarly, long response times may suggest retrospective analysis of past actions to provide a preliminary assessment of the impact of a proposed action.

Issues of space

The five kinds of spatial effects to consider can influence the design of a study as well as the interpretation of its results.

1. What is the influence of on-site management activities on off-site conditions? That is, local management may influence remote conditions, both directly and indirectly (Loehle 1990). An example is the downstream effect of stream temperature or sedimentation on fish populations due to local reduction, removal, or restoration of riparian vegetation cover.

2. What is the relative influence of off-site management activities on on-site (desired) conditions? On-site conditions can be influenced by other off-site activities. For example, despite protection of old-growth forest groves, some arboreal lichens might nonetheless decline because of degraded air quality from industrial pollutants originating elsewhere in the airshed. The potential influence of downstream dams and fish harvesting on the abundance of local fish populations is another example.

3. To what *degree* do local management activities influence the on-site (desired) conditions? That is, to what extent do background noise and other environmental factors affect on-site conditions? Local management may influence only a portion of the total variation in local conditions. For example, providing local breeding habitat only partially succeeds in conserving populations of neotropical migratory birds, whose numbers may still decline due to pesticide loads or habitat loss encountered during wintering in the neotropics.

4. What is the relative influence of conditions and activities from different spatial scales, particularly the effects on local stand-level conditions from broader landscape-level factors? That is, desired conditions and management actions are best addressed at appropriate scales of geography. As examples, effects of forest management on abundance of coarse woody debris are best assessed at the stand level; effects of forest management on vegetation conditions that affect visual quality or goshawk (*Accipiter gentilis*) habitat are best assessed at the landscape level; and effects of overall management policy and ownership patterns on grizzly bear (*Ursus arctos*) populations are best assessed at subregional or regional levels.

5. What are the cumulative effects of stand-level treatments as they spread across the landscape? For example, wind fetch and thus wind speed may increase as clearcuts become wider with sequential, adjacent cuts. Thus, the windthrow hazard in one cutblock may increase as adjacent areas are cut, and the windthrow hazard in those cutblocks cannot simply be extrapolated from the hazard measured in a single cutblock surrounded by trees.

For each of these five kinds of spatial effects, adaptive management monitoring studies would be designed and implemented differently. Where this is not possible, spatial influences should at least be acknowledged as potential sources of variation and included in the analysis.

Issues of time

Answering questions about time effects can help distinguish true cause from non-causal correlation, and treatment effects from natural variation. Three typical time scale issues follow.

1. What are the response times of variables? For some variables, response may be apparent in a relatively short period of time; others may respond more slowly. Examples are the relatively short and quick response time of seedling survival compared with the long and slow response times associated with many biodiversity indices (e.g., changes in grizzly bear populations).
2. What are the lag times of variables? Some variables may not immediately respond to a treatment or may depend greatly on site history. For example, because acorn woodpeckers (*Melanerpes formicivorous*) show high fidelity to particular sites, a lag will exist before they respond to the removal of granary trees (Ligon and Stacey 1996). This lack of short-term response should not lead one to conclude that management actions—in this example, the reduction or removal of granary trees—have no effect. Sometimes these lags in response result when conditions from prior time periods overwhelm or influence responses from current actions. For example, the intensity of a fire will be influenced by site history, in addition to current management actions. Thus short-term changes in a response variable may reflect both the management action and past site history. Some time-lag effects can be quite variable and manifest as non-monotonic (up and down) trends over the long term. For example, annual non-monotonic variations in bird populations—both increases and decreases—may belie truer long-term declines in some population counts (Thomas and Martin 1996).
3. What are the cumulative effects of a variable over time? Some variables do not make a mark except over time or until a particular threshold has been exceeded. An example is the adverse effect of certain pesticides on wildlife reproduction. The detrimental effect may not be apparent until the pesticide concentrations reach a particular level of toxicity (Tiebout and Brugger 1995).

The design of adaptive management studies and selection of analysis methods are guided in part by these considerations of space and time. For example, replication is one major consideration in designing studies. Given a large geographic area, as tends to be the focus in ecosystem management, or a rare condition, such as a threatened species population, are spatial replicates possible? That is, can landscapes or threatened populations be replicated at all, or in adequate numbers? If the conditions cannot be replicated, then pseudoreplication (e.g., dividing a single area into smaller blocks) may be the only recourse (Hurlbert 1984). Alternatively, other kinds of studies (e.g., analytical surveys, expert testimony) might help in assessing the impact of the treatments, although they do not allow strong inference about cause. Similarly, long response times and time lags make temporal replication difficult. Retrospective studies (see Smith, this volume, Chap. 4) provide one alternative for gaining insight into the long-term effects of management actions. In cases where either spatial or temporal replication is severely limited, a higher probability of Type I and II errors might need to be tolerated (see Anderson, this volume, Chap. 6).

In some cases, a powerful adaptive management study may be possible but managers, decision-makers, industries, or other interested bodies may not be willing to bear the cost, duration, and tight controls on management activities. The consequences of not using an optimum study must be explicitly considered and understood by all.

9.3 Considerations in Designing an Adaptive Management Study

9.3.1 Characteristics of a powerful adaptive management study

To help in evaluating management actions and validating functional and causal relationships, an adaptive management study should be consistent (i.e., should represent the system of interest), accurate, precise, and unbiased (see Routledge, this volume, Chap. 5). Managers and researchers should work together in designing an adaptive management study that represents the real system and provides information within acceptable limits of Type I and Type II errors (Anderson, this volume, Chap. 6). They may also want to consider the trade-offs inherent in relaxing any of the conditions, such as accepting a lower but still acceptable level of precision in exchange for lower cost or more rapid results. The study design should also be independently reviewed to assess its capability to meet the desired (and often conflicting) criteria of high consistency, high accuracy, high precision, and low bias.

9.3.2 What managers need to ask of reliability

Managers should ask four general questions regarding the reliability of adaptive management studies and their results.

1. What *confidence* can I have in the results of this adaptive management study, particularly for avoiding false positives? Statistically, this question can be answered by calculating the probability of a Type I error (Anderson, this volume, Chap. 6).

2. What *power* do the results provide for avoiding false negatives (Anderson, this volume, Chap. 6)? Statistically, this can be answered by calculating the probability of a Type II error (although Bayesian approaches differ significantly in not dealing with questions of confidence and power). Type I and Type II errors hold different implications for managers (Marcot 1986; Anderson, this volume, Chap. 6). For example, if the adaptive management study is aimed at determining adverse effects of some management activity on a wildlife species that is threatened, then managers may be more tolerant of a Type I error than of a Type II error. However, if the species is not threatened and the activity results in important commodity production and economic return, then they may be more tolerant of a Type II error.

3. What is the *relevance* of the results? How representative is the study of other sites or conditions? Some studies may reveal only local conditions and the chance effects of unique site histories, rather than overall effects, or they may pertain to only one vegetation type or climatic condition. The manager should know the contexts under which results apply. For example, results of a forest thinning operation may apply to only a particular initial stand density or forest type.

4. Were the effects truly a result of the *management activity*? This question cuts to the heart of separating cause from noise, and determining what really influenced the outcome. The experimental studies that are central to adaptive management are designed to determine causality. Researchers and managers should not assume that demonstration of pattern and correlation constitutes valid evidence of causation.

9.3.3 Principles of experimental design

To help ensure success in evaluating management actions, researchers should review adaptive management studies for the four main principles of experimentation: randomization, replication, blocking, and representation (see Nemec, this volume, Chap. 2). *Randomization* reduces bias. *Replication* allows an estimation of variance, which is vital for confirming observed differences. *Blocking* increases precision and reduces cost and sample size. *Representation* helps to ensure study of the correct universe of interest.

In the real world, these four principles cannot always be met and compromises are necessary. It is often impossible to fully randomly allocate treatments, such as forest clearcuts or fire locations. In such cases, study sites may be randomly selected from existing clearcuts or fire locations, resulting in nonexperimental studies (e.g., observational studies, analytical surveys, retrospective studies, or impact studies; see Schwarz, this volume, Chap. 3). When interpreting study results, researchers should account for the site-specific characteristics leading to the initial nonrandom assignment of the treatment. Furthermore, the researcher should recognize that the altered study can no longer provide reliable knowledge of cause, but only generates hypotheses for validation when future management actions are implemented.

When replication is not possible, suspected causal effects can be masked by confounding hidden causes or by spurious correlations. Researchers may be tempted to resolve the problem by taking multiple samples as pseudoreplications. The drawback of this solution is that study results apply to study areas only and cannot be generalized to the entire system of interest.

When blocking is not feasible, precision suffers. Larger sample sizes, hence increased cost, are necessary to achieve desired levels of confidence and power. Finally, when a study considers only a portion of the system of interest (due to lack of randomization, replication, or funding), generalization of the results to the entire system could be inappropriate and misleading. In this case, researchers and managers together must re-evaluate the study objectives and scope.

Even though researchers are responsible for designing studies, managers and decision-makers should be aware of these issues and possible limitations. Other useful aspects of measurement errors are reviewed by Routledge (this volume, Chap. 5), who presents a useful set of criteria for selecting indices.

9.4 Types of Information and Study Designs

Study the past if you would divine the future.
– Confucius

Information from sources other than management experiments can play important roles in adaptive management. For example, expert judgement, anecdotes, existing data, and literature can help in building simulation models used to explore alternative scenarios and identify key uncertainties. Information from these sources can also provide supporting evidence, which becomes important when real world limitations prevent the design of "ideal" management experiments. Each source of information provides different levels of reliability.

9.4.1 Learning from existing data, expertise, and expert testimony

Using existing data and literature

In the initial stages of adaptive management, existing data and literature can be used to evaluate scenarios, project effects, or devise guidelines. However, the ability to determine treatment effects from existing data is often limited because such data may not cover the range of environments or treatments proposed, or may be knitted from disparate databases. In addition, the spatial, temporal, or ecological scope and the degree of reliability of such data may be poorly documented. Perhaps a good reminder of the potential weaknesses of using existing information is to remember the acronym for "best available data." When existing data are used, how well they can address the critical management question should be assessed honestly and accurately.

Gathering expertise and expert testimony

Another source of information is expert judgement, review, and testimony. Broad-scale assessments of wildlife population viability conducted recently by federal resource management agencies of the western United States have relied on panels of experts and contracted technical reports to fill in many gaps left by existing databases and publications (e.g., Schuster et al. 1985). In my own work using expert-panel approaches, I have modified[1] the Delphi technique (Zuboy 1981; Richey, Horner, and Mar 1985) to collect expert knowledge and judgement (Marcot et al. 1997). However, expert judgement cannot replace statistically sound experiments.

9.4.2 Learning from management actions

Probably the most reliable means of gathering information for assessing the impact of management actions is to conduct field studies. But, like publications and expert opinion, empirical evidence comes in many forms and levels of usefulness. A few key sources of evidence for the manager to know about—listed here in increasing order of reliability—include anecdotes and expert judgement, retrospective studies, nonexperimental (observational) studies, and experimental manipulation.

Anecdotes and expert judgement

The results of management actions are often evaluated informally by simple observations with no measurements. Such opportunistic observations are a two-edged foil: while the collective expertise from field experts can constitute a valuable and irreplaceable pool of wisdom, individual anecdotes can prove strikingly misleading. As a whole, anecdotal information should be used with a great deal of caution—or at least with rigorous peer review—to help avoid problems such as motivational bias (Marcot et al. 1997).

1 Modifications addressed the need to adhere to the U.S. *Federal Advisory Committee Act*, by polling individual experts for basic ecological information and not reaching group consensus on specific management actions.

Anecdotes and expert judgement alone are not recommended for evaluating management actions because of their low reliability and unknown bias. In the BC Forest Service, use of this source of information *alone* to evaluate management actions is not considered adaptive management.

Retrospective studies

Sometimes the results of management actions are provided by measuring the outcomes of future actions taken in the past. Retrospective studies (evaluating the outcomes of actions taken in the past) are valuable for helping to predict the outcomes of future actions. These studies can provide some insights to support or refute proposed hypotheses, and are particularly valuable for problems where some indicators take a long time to respond. However, because the treatments might not have been randomly assigned, and the initial conditions and the details of the treatments are often unknown, teasing out causal factors may be challenging at best and misleading at worst.

Nonexperimental (observational) studies

Nonexperimental studies (called observational studies by some authors) are the most common kind of field studies reported in wildlife journals. Like retrospective studies, nonexperimental studies are not based on experimental manipulations. Although it may be debatable whether nonexperimental studies should entail hypothesis testing, they should nonetheless meet statistical assumptions, including adequacy of samples sizes and selection of study sites, to ensure reliable results. Much can be learned from taking advantage of existing conditions and unplanned disturbances (Carpenter 1990; Schwarz, this volume, Chap. 3).

Nonexperimental studies usually entail analysis of correlations among environmental and organism parameters, such as studying the correlations between clearcutting and wildlife response. Causes are inferred and corroborated through repeated observations under different conditions. Because results may be confounded by uncontrolled (and unknown) factors, nonexperimental studies are best interpreted as providing only insights to cause. These insights can be valuable in predicting outcomes of actions, but again, the veracity of such predictions and the effects of management actions are best evaluated through controlled experiments (McKinlay 1975, 1985). Of nonexperimental studies, BACI-P designs allow the strongest inferences about causes (Schwarz, this volume, Chap. 3).

Inventories and *surveys* are not the same as nonexperimental studies; they display patterns but do not reveal correlates. Nevertheless, inventories and surveys can be useful in adaptive management. They provide information from which to select random samples, or a baseline of conditions from which to monitor changes over time. Inventories and surveys should still adhere to strict sampling protocols and can use more advanced statistical methods to streamline efficiencies (Schwarz, this volume, Chap. 3). For example, Max et al. (1990) presented an inventory method of random sampling of Northern Spotted Owl (*Strix occidentalis caurina*) territories with partial, annual replacement of samples to increase accuracy and reduce bias in estimates of site occupancy.

One particularly terse version of inventories is *rapid assessment procedure* (RAP) or *rapid survey*, used by some conservation groups "running ahead of the bulldozers" to survey biota of tropical forests (Oliver and Beattie 1993, 1996). Rapid surveys may prove useful in some temperate ecosystems as well, but should be used only to provide quick, initial, mostly qualitative or categorical information from which to design more formal adaptive management studies.

Experimental manipulation

Management actions can best be evaluated through experimentation (Nemec, this volume, Chap. 2). Experimental manipulations can be used to quantify the contributions from each suspected causal factor, and ultimately to develop, refine, and validate prediction models. The kind of experimentation referred to here involves deliberate, planned alterations of one or more sites, one of which may be an unaltered control.

Finally, *demonstrations* are not adaptive management per se, but often appear in the adaptive management literature (e.g., Yaffee et al. 1996). Demonstrations are designed to showcase the execution of specific management activities such as silvicultural techniques but they do not provide the evidence that controlled, replicated experiments do. When faced with a proposal for a demonstration "study," the manager might first ask if they need evidence of cause and effect, and, if so, if a management experiment with controls and replicated treatments would better provide evidence as well as the opportunity to demonstrate the activities.

9.4.3 Information for improving study designs

Study designs can be improved by using prior knowledge of the system of interest gained through retrospective analysis of past events, existing literature, and expert testimony. This information can aid in *blocking* samples to increase study efficiency, and in ensuring correct spatial and temporal representation of samples.

Study design can also benefit from *initial field sampling*. This sampling can provide preliminary estimates of variance of parameters that can be used to calculate sample size necessary to meet desired levels of precision. Initial field sampling also gives information on stratification or blocking strategy and helps to reveal conditions not originally considered in a study.

The relative merit of alternative study designs can be assessed using the tools of quantitative decision analysis, including Bayesian statistics (see Bergerud and Reed, this volume, Chap. 7; Peterman and Peters, this volume, Chap. 8). Such analysis may suggest, for example, the sampling period, sampling frequency, and sample size necessary for providing reliable information in a suitable time frame and at an acceptable cost.

The past several sections have discussed characteristics of AM study designs and use of information sources. I turn next to the topic of integrating study results into statements of risk. The topic of risk is also addressed by Peterman and Peters (this volume, Chap. 8).

9.5 Risk Analysis and Risk Management

Lots of folks confuse bad management with destiny.
– Kin Hubbard

9.5.1 Risk: speaking the same language between analysis and management

The concept of risk has pervaded much of the adaptive management literature and much of land management planning. However, researchers and managers often use the term "risk" in vastly different ways. This use can lead to, at best, confusion in interpreting results, or, at worst, misrepresentation of study results. For adaptive management, risk is defined as the expected value of adverse outcomes of a management action.

It is useful in adaptive management to differentiate risk analysis from risk management. In *risk analysis*, the researcher lists possible outcomes, estimates their likelihoods under one or more alternative future scenarios, and calculates their individual "utilities" by weighting outcome likelihoods by outcome values. These values are usually derived by managers and may pertain to social, economic, or political interests, as well as to legal regulations and objectives for resource management. Weighting outcome values with outcome likelihoods helps the manager to determine the overall risk of a management action. Then, in *risk management*, the manager defines and applies their risk attitude (their degree of risk-seeking, risk-neutral, or risk-avoidance behaviour) and then decides on the best course of action. In separating risk analysis from risk management, the onus of articulating outcome values, describing personal attitudes to risk, and defining personal decision criteria is correctly placed on the manager, not the researcher.

Formal decision analysis (Peterman and Peters, this volume, Chap. 8) is a method for assessing the risk of alternative outcomes of actions, taking uncertainty into account. Most managers do weigh the relative values or outcomes, their likelihoods, and a host of other factors that limit the decision space, such as political acceptability, effects on career, and effects on potential future decisions. However, decision analysis "in your head" is a poor substitute for quantitative decision analysis. At a minimum, managers should explicitly reveal their own outcome values, risk attidudes, and decision criteria.

9.5.2 Expressing uncertainties and unknowns

Uncertainty is a hallmark of science. However, managers—as well as politicians, the media, the public, and courts—typically view the issue of uncertainty differently than do researchers. To the researcher, uncertainty in adaptive management may represent error of measurement, confounding effects of natural variation, or other unstudied causes; such uncertainty is to be expected and results are to be treated with due care (Kodrick-Brown and Brown 1993). In some sense, the researcher may be certain of a particular level of variance, and may still view adaptive management study results as strong evidence of some effect of a management activity within some range of outcome. To the manager and others, however, such variance may be seen as lack of evidence of effects, or even as strong evidence of little or no effect, if the researcher cannot be "certain" of the outcome.

Scientific unknowns should be treated as a qualitatively different beast than scientific uncertainty. For the researcher, uncertain outcomes can be quantified as some measure of variation (such as variance or confidence interval), but unknowns cannot be quantified at all. The influence of unknowns may be deterministic, stochastic, strong, weak, or nonexistent; the researcher often simply cannot say. Again, however, the manager might erroneously view unknowns as lack of evidence of effect and thus as justification to proceed unless some contrary "proof" is provided.

Managers also need to understand how to interpret results of adaptive management studies, particularly in the context of a risk analysis. If adaptive management studies are designed as good statistical investigations, then results can serve to either falsify, or fail to falsify, the null hypothesis being tested; results can never "prove" a hypothesis.[2] Failing to falsify the null hypothesis of no effect lends only incremental credence to the management hypothesis. One of the ways to lend greater credence is through replicate findings that would further corroborate results.

Therefore, researchers and managers (as well as courts, media, and the public) must come to a common understanding of the concepts and implications of scientific uncertainty, unknowns, risk and associated concepts of proof, errors, and statistical falsification. Otherwise, results of adaptive management studies can be severely misrepresented, misunderstood, and misapplied.

9.5.3 Unravelling the causal web: when is it our fault and what can be done?

One of the main reasons for conducting adaptive management studies of resource use or ecosystem elements is to determine not just patterns and trends but also their causes. The manager should ask: What is the true cause? Do our management activities directly affect the outcome, or merely set the stage for other, more direct factors? To what degree do our management activities influence the outcome?

Untangling the causal web in field situations can be a great challenge. Seldom are causal factors affecting ecosystems single, simple, or easily quantified. Most often, factors interact in complex ways, such as with indirect and secondary effects, and through feedback relations (Figure 9.1). Even in the simplest model (Figure 9.1a), the relative contributions of known and unknown causes must be estimated. In simple models, the contribution from linear associations—which may or may not be causal—is indicated by the value of the coefficient of determination R^2 (or adjusted R^2), with the contribution from unknown associations being $1–R^2$. In more complex models, (Figures 9.1b, c, d), estimating relative contributions can be more involved. In real-world cases, it is not always evident which factors act as proximate causes, which act as less direct causes, which are mere correlates with no causal relation, and which participate in obligate feedback relations.

Of course, some relations are obvious, such as removal of forest canopy causing the local elimination of obligate canopy-dwelling organisms. But less obvious effects or gradations, though difficult to unravel, may be of great interest to the manager. For example, what degree of effect does partial removal of the forest canopy have on local plant or animal populations that are only facultatively associated with canopy environments? Might there be compounding, cumulative effects that exacerbate or ameliorate such effects, such as wider regional loss of forest canopies, or restoration of canopy conditions in adjacent stands?

To determine the relative influence of specific management activities, the researcher may turn to statistical techniques using estimation of partial correlations. These methods help determine the contribution of one factor, such as a management activity, given the effects of all other factors (e.g., other activities, natural changes in environments, unknown causes). Traditional analyses such as step-wise multiple regression help identify such partial influences. Other, less well-known techniques such as regression trees and path regression analysis (e.g., Schemske and Horvitz 1988) can also be used.

Determining the relative influence of management actions is vital for setting realistic expectations for management results. For example, determining that fragmentation of local forests affects breeding habitat for migrating songbirds (Wilcove 1985) is only part of the puzzle; loss of habitat on neotropical wintering grounds is also a significant cause of declines in songbird populations. Therefore, changing local management to reduce fragmentation should be expected to have only a partial impact on songbird populations.

2 Some authors suggest that Bayesian analyses also can be interpreted as the testing of null hypotheses, that is, the prior probabilities.

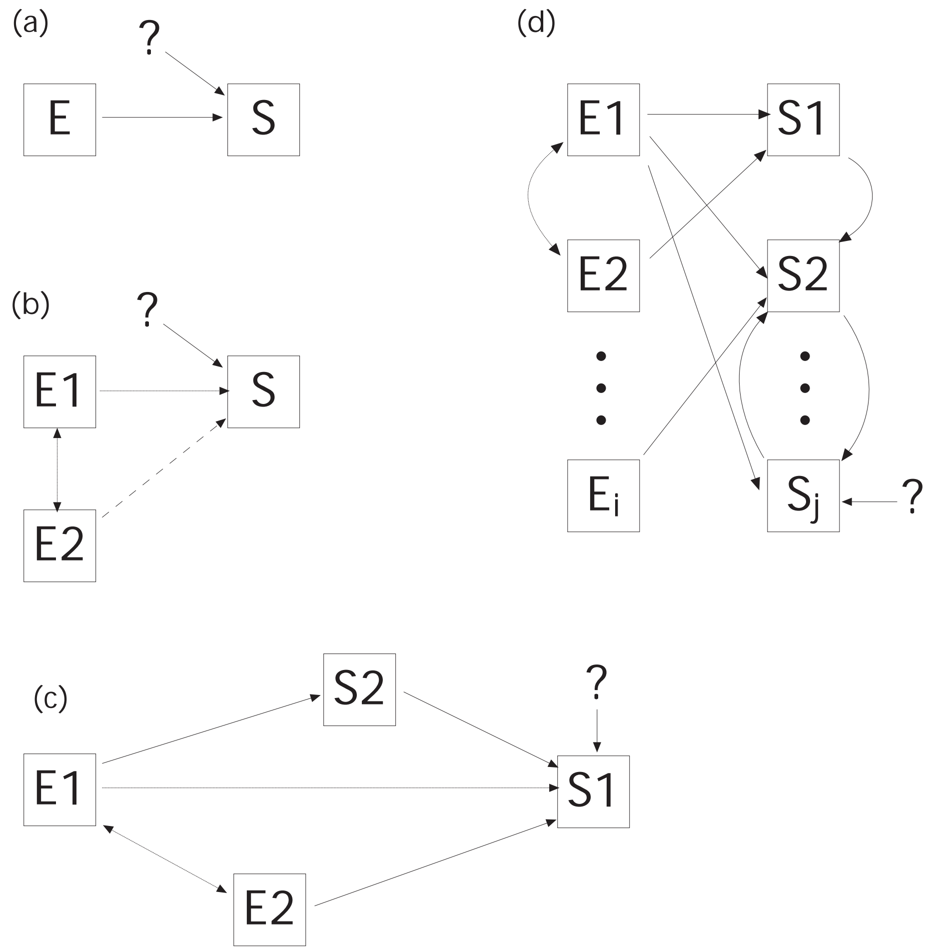


FIGURE 9.1 *Causes and correlates: four examples. In all figures, S = wildlife species response; ? = unexplained variation due to measurement error, experimental error, or effects of other environmental or species factors; solid arrows = causal relations; dotted arrows = correlational relations that may or may not be causal. (a) In this simplest case, some wildlife species response S, such as population presence or abundance, is assumed to be explained and caused by some environmental factor E. (b) In a more complex case, we may be measuring one environmental factor E1 when the real cause is another environmental factor E2. (c) Getting closer to the real world, a second species response S2 may be part of the cause. (d) Most like the real world, with feedback relations among the dependent (response) variables S. (Adapted from Morrison et al. 1998, Fig. 10.2.)*

9.5.4 A dilemma for managers: when samples are few and crises are many

One bane of adaptive management is that, in many cases, the unique field conditions make it difficult to correctly design statistical studies to identify and quantify causes. Especially when studying landscapes, ecosystems, rare or threatened species, and infrequent events, the major problems in the design of such studies are small sample size and inability to replicate conditions. In such circumstances, what can the researcher do, and how should the manager interpret results? The answer may be found in several courses of action: selecting correct indicators, merging disparate lines of evidence, and using statistical procedures that take advantage of prior knowledge or that function adequately with small sample sizes.

Selecting correct indicators

Indicators that are objective, repeatable measurements, whose quality is documented quantitatively should be selected. For adaptive management studies, an indicator should (1) respond rapidly to changes, (2) signal changes in other variables of interest, (3) be monitored efficiently, and (4) be causally linked to changes in stressors. Most "ecological indicators" purported to fit these criteria usually fail (Block et al. 1987; Patton 1987; Landres et al. 1988). For example, the Northern Spotted Owl, often selected by USDA Forest Service as an "old-growth indicator," may serve criterion (4), but fails with the other three criteria: spotted owls have low reproductive rates and long life spans, so they respond slowly to changes; changes in their populations may not necessarily correlate well with other desired facets of old-growth forests (e.g., habitat for anadromous fish); and their population trends are terribly costly to monitor.

Indicators that do meet these criteria include soil arthropods as indicators of soil productivity (McIver et al. 1990; Moldenke and Lattin 1990; Pankhurst et al. [editors] 1997); butterfly diversity as an indicator of overall ecosystem diversity (Kremen 1994); and some arboreal lichens as indicators of air quality (Stolte et al. 1993; Geiser et al. 1994) or persistence of old forests (Tibell 1992). See Murtaugh (1996) for a review of the statistical basis of ecological indicators.

Merging disparate lines of evidence

Merging different study results is a second tactic that can help in identifying causal relations when good experimental design is impossible or impractical. In statistics, this process is called "combining information" (CI). Draper et al. (1992) provide a useful overview of various CI techniques, including methods of meta-analysis (Hedges and Olkin 1985) that can be useful in conservation research (Fernandez-Duque and Valeggia 1994). For example, meta-analysis was used by Burnham et al. (1996) to determine overall trends of Northern Spotted Owls by combining results from individual population demography studies.

CI is not a panacea, as it can be fraught with difficulties such as matching consistency and representativeness among studies designed for different initial objectives. Still, the researcher may wish to use CI methods to merge lines of evidence taken from available information. This available information could include anecdotes and local experience, retrospective studies, observational studies, experimental manipulations, and demonstrations. The reliability of each source for inferring causes should be judged very carefully.

In contrast to formal meta-analysis, simply pooling data from different studies could lead to spurious and misleading conclusions. For example, to assess the impact of clearcutting on grizzly bear populations, the data from several studies might be combined into an overall regression. This regression might suggest a significant correlation between clearcutting and grizzly bear populations. However, grizzly bears within individual study areas might respond differently to clearcutting because they come from different geographic areas, latitudes, or forest types. Thus the correlation may reflect these differences between populations, rather than any treatment effect. The incorrect conclusion of correlation would arise because such an analysis violates an assumption underlying regression: that the data come from the same statistical population with the same causal mechanisms. On the other hand, a formal meta-analysis approach would analyze results from each study with differences among studies as an explanatory factor. CI has great utility, especially where powerful experimental studies are difficult. However, managers and researchers must be careful in its use, ensuring that studies are truly from the same causal web.

Using statistical procedures that take advantage of prior knowledge

Bayesian statistics were developed specifically for using prior knowledge and incrementally gathered field data (Ver Hoef 1996). Bayesian statistical techniques include empirical Bayes and sequential Bayes procedures, in which initial estimates of the likelihood of conditions become incrementally adjusted and refined over time as new evidence is gathered (e.g., Gazey and Staley 1986; Link and Hahn 1996). Expert opinion, existing literature and data, retrospective studies, and non-experimental studies can all be used to establish preliminary values of prior probabilities in a Bayesian analysis. Bayesian methods were reviewed by Bergerud and Reed (this volume, Chap. 7), who advocate their use to incorporate accumulated knowledge of experts.

9.6 Conclusions and Recommendations

Knowledge is the small part of ignorance that we arrange and classify.

– Ambrose Bierce

9.6.1 A decision tree for managers

The six stages of adaptive management and sources of information appropriate for each stage are presented in Table 9.1. This table can be used by managers as a decision tree to guide the choice of

TABLE 9.1 *Stages of an adaptive management (AM) project and sources of information appropriate for each stage. This table can be used by managers as a decision tree to guide (1) the choice of study for each AM stage (reading across rows), and (2) the use of existing information (reading down columns).* ✔ *= recommended sources;* ✘ *= not recommended;* [✔] *= most recommended for a given project stage; – = does not apply; (pilot) = use for pilot study only. See Chapter 1 for full descriptions of AM stages.*

AM stages	Literature review	Expert judgement	Demonstration	Anecdote	Retrospective study	Nonexperimental study	Experimental study
1. Assess problem Identify potential impacts of management actions and the potential reasons for them.							
Identify patterns and trends	✔	✔	✘	✔	✔	✘	✘
Identify correlates	✔	✔	✘	✔	✔	✘	✘
Identify potential causes of suspected impact [a]	✔	✔	✘	✔	✔	✘	✘
2. Design project Determine treatments to implement, sample size, effect size, power, etc.	✔	✔	✘	✘	✔	[✔][b] (pilot)	[✔][b] (pilot)
3. Implement	—	—	—	—	—	✔	[✔]
4. Monitor	Monitoring design is determined at the "Design project" stage.						
5. Evaluate Interpretation	✔[c]	✔[c]	✘	✘	✔[c]	✔	[✔]
6. Adjust management action	—	—	—	—	—	—	—

a Experimental and nonexperimental studies can provide information on patterns, correlates, etc., but typically these studies will not be done by taking advantage of management actions, but rather as part of applied research.

b All else being equal, if the cost of conducting a nonexperimental study is significantly less than that of an experimental study, choose the former.

c In the "Evaluation" stage, existing information based on literature, expert judgement, and retrospective analysis is updated using data collected from the management experiment to assess the effect or outcome of an action. It can also be used to determine the relative plausibility of suspected causes, and to estimate the prior probabilities in a Bayesian analysis.

study for each stage of adaptive management, as well as to guide the use of existing information.

At the problem assessment stage, existing information is valuable for identifying *potential* impacts of management actions. At the project design stage, pilot studies (experimental or nonexperimental) are recommended for fine tuning the study methodology. Pilots can be used to estimate variability in the response variables; these estimates can then be used to determine sample size, effect size, and power for the study. Controlled experiments allow the strongest inference about the actual impacts of management actions. Once a study has been implemented, relevant data are collected through a monitoring process. The data are then analyzed using appropriate statistical methods to answer questions set out at the beginning of the adaptive management project. In the evaluation stage, existing knowledge (based on literature, expert judgement, and retrospective analysis) is updated using data collected from the management experiment to assess the effect or outcome of an action. Using Bayesian analysis, existing knowledge together with collected data can also be used to determine the relative plausibility of suspected causes. Management actions are then adjusted based on this updated knowledge. During the course of the management experiment, new questions may arise that then lead to further problem assessment, project design, implementation, and so on, in a continuous cycle of learning.

9.6.2 Is there a "best" statistical approach to adaptive management?

The answer to this question is an unqualified "yes." The best approach for answering the questions "Did this action have the desired effect?" and "Are the basic assumptions underlying management decisions correct?" is to use controlled, randomized experiments with sufficient sample sizes and duration. This approach provides the best understanding of causal relations and the best basis for validating predictive models—assuming that the models can provide testable predictions.[3] Nevertheless, the informativeness of a statistical approach must also be weighed against its costs (ecological, social, and economic). Ultimately, designing management actions as controlled, randomized experiments will provide the best evidence for managers who face the difficult task of making management decisions and defending such decisions legally, politically, and scientifically.

Short of this ideal, both researchers and managers have their work cut out for them. They should maximize the use of available information, but not draw undue conclusions about causes. It may be useful to explicitly array the various available lines of evidence and to articulate the confidence in identifying causes from each. Managers and researchers must look for similarities and disparities among lines of evidence and investigate reasons for the differences. Moreover, they should seek peer review to ensure appropriate and rigorous use of various sources of information. Repeatability of findings and suspected causes is the basis for true scientific understanding and predictability.

Real-world adaptive management problems are often complicated by time exigencies or finite funding so that powerful experiments are not possible. If the study design must be compromised, then the ramifications of drawing incorrect conclusions should be thought out. When turning to field experts for their professional opinions, managers should be aware of potential problems such as motivational and personal bias.

Managers should weigh the benefits and cost of a more or less rigorous approach. For, in the end, reliable knowledge is a hard-won but essential commodity for ensuring successful conservation practices for future generations.

Acknowledgements

My thanks to Brian Nyberg and Brenda Taylor for suggesting topics to cover in this chapter and for their technical reviews of the manuscript. Roger Green, Rick Page, Martin Raphael, and Ian Thompson also provided technical reviews. Thanks to Vera Sit and Brenda Taylor for their fine edits. My gratitude to Tim Max for his technical review and for thoughtful discussions on the role of statistics in adaptive management.

3 As expressed by one statistician, if predictive models are so complex that they become essentially untestable, then they are nothing more than belief structures and their relation to science is questionable at best (T. Max, pers. comm., 1997).

References

Anderson, J.L. [n.d.]. Errors of inference. This volume.

Block, W.M., L.A. Brennan, and R.J. Gutierrez. 1987. Evaluation of guild-indicator species for use in resource management. Environ. Manage. 11:265–9.

British Columbia Ministry of Forests. 1996. Adaptive management: learning from our forests. Brochure. Integrated Resources Policy Branch, Victoria, B.C.

Burnham, K.P., D.R. Anderson, and G.C. White. 1996. Meta-analysis of vital rates of the northern spotted owl. Stud. Avian Biol. 17:92–101.

Carpenter, S.R. 1990. Large-scale perturbations: opportunities for innovation. Ecology 71:2038–43.

Collopy, M., J. Cannell, S. Daniels, D. Dippon, E. Gaar, G. Grant, B. Mulder, C. Philpot, J. Steffenson, and F. Swanson. 1993. Implementation and adaptive management. *In* Forest ecosystem management: an ecological, economic, and social assessment. Forest Ecosystem Management Assessment Team. U.S. Dep. Agric. For. Serv., Washington, D.C.

Draper, D., D.P. Gaver, Jr., P.K. Goel, J.B. Greenhouse, L.V. Hedges, C.N. Morris, J.R. Tucker, and C.M. Waternaux. 1992. Combining information: statistical issues and opportunities for research. National Academic Press, Washington, D.C. Contemporary statistics No. 1.

Fernandez-Duque, E. and C. Valeggia. 1994. Meta-analysis: a valuable tool in conservation research. Cons. Biol. 8:555–61.

Gazey, W.J. and M.J. Staley. 1986. Population estimation from mark-recapture experiments using a sequential Bayes algorithm. Ecology 67:941–51.

Geiser, L.H., C.C. Derr, and K.L. Dillman. 1994. Air quality monitoring on the Tongass National Forest: methods and baselines using lichens. U.S. Dep. Agric. For. Serv., Alaska Region. Petersburg, Alaska. R10-TB-46.

Gentiol, S. and G. Blake. 1981. Validation of complex ecosystem models. Ecol. Modelling 14:21–38.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. J. Wiley, New York, N.Y.

Hedges, L.V. and I. Olkin. 1985. Statistical methods for meta-analysis. Academic Press, New York, N.Y.

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

Kodric-Brown, A. and J. H. Brown. 1993. Incomplete data sets in community ecology and biogeography: a cautionary tale. Ecol. Applic. 43:736–42.

Kremen, C. 1994. Biological inventory using target taxa: a case study of the butterflies of Madagascar. Ecol. Applic. 4:407–22.

Landres, P.B., J. Verner, and J.W. Thomas. 1988. Ecological uses of vertebrate indicator species: a critique. Cons. Biol. 2:316–28.

Ligon, J.D. and P.B. Stacey. 1996. Land use, lag times and the detection of demographic change: the case of the acorn woodpecker. Cons. Biol. 10:840–6.

Link, W.A. and D.C. Hahn. 1996. Empirical Bayes estimation of proportions with application to cowbird parasitism rates. Ecology 77:2528–37.

Loehle, C. 1990. Indirect effects: a critique and alternate methods. Ecology 71:2382–6.

McIver, J.D., A.R. Moldenke, and G.L. Parsons. 1990. Litter spiders as bio-indicators of recovery after clearcutting in a western coniferous forest. NW Environ. J. 6:410–2.

McKinlay, S.M. 1975. Comprehensive review on design and analysis of observational studies. J. Am. Statist. Assoc. 70:503–20.

______. 1985. Observational studies. *In* Encyclopedia of statistical sciences. Vol. 6. J. Wiley, New York, N.Y. pp. 397–401.

Marcot, B.G. 1986. Summary: biometric approaches to modeling: the manager's viewpoint. *In* Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. J. Verner, M.L. Morrison, and C.J. Ralph (editors). Univ. Wis. Press, Madison, Wis. pp. 203–4.

Marcot, B.G., M.A. Castellano, J.A. Christy, L.K. Croft, J.F. Lehmkuhl, R.H. Naney, R.E. Rosentreter, R. E. Sandquist, and E. Zieroth. 1997. Terrestrial ecology assessment. *In* An assessment of ecosystem components in the interior Columbia Basin and portions of the Klamath and Great Basins. T.M. Quigley and S.J. Arbelbide (editors). Pac. N. Res. Sta., Portland, Oreg. U.S. Dep. Agric., For. Serv. Gen. Tech. Rep. PNW-GTR-405. pp. 1497–1713.

Max, T.A., R.A. Souter, and K.A. O'Halloran. 1990. Statistical estimators for monitoring spotted owls in Oregon and Washington in 1987. U.S. Dep. Agric. For. Serv., Portland Oreg. Res. Pap. PNW-RP-420.

Moldenke, A.R. and J.D. Lattin. 1990. Density and diversity of soil arthropods as "biological probes" of complex soil phenomena. NW Environ. J. 6:409–10.

Morrison, M.L., B.G. Marcot, and R.W. Mannan [1998]. Wildlife-habitat relationships: concepts and applications. 2nd ed. Univ. Wis. Press, Madison, Wis. In press.

Murtaugh, P.A. 1996. The statistical evaluation of ecological indicators. Ecol. Applic. 6:132–9.

Nemec, A.F.L. [n.d.]. Design of experiments. This volume.

Oliver, I. and A.J. Beattie. 1993. A possible method for the rapid assessment of biodiversity. Cons. Biol. 7:562–8.

______. 1996. Designing a cost-effective invertebrate survey: a test of methods for rapid assessment of biodiversity. Ecol. Applic. 6:594–607.

Pankhurst, C., B. Doube, and V. Gupta (editors). 1997. Biological indicators of soil health. Oxford Univ. Press, New York, N.Y.

Patton, D.R. 1987. Is the use of "management indicator species" feasible? W. J. For. 2:33–4.

Peterman, R.M. and C. Peters. [n.d.]. Decision analysis: taking uncertainties into account in forest resource management. This volume.

Richey, J.S., R.R. Horner, and B.W. Mar. 1985. The Delphi technique in environmental assessment. II. Consensus on critical issues in environmental monitoring program design. J. Environ. Manage. 21:147–59.

Richey, J.S., B.W. Mar, and R.R. Horner. 1985. The Delphi technique in environmental assessment. I. Implementation and effectiveness. J. Environ. Manage. 21:135–46.

Routledge, R.D. [n.d.]. Measurements and estimates. This volume.

Schemske, D.W. and C.C. Horvitz. 1988. Plant-animal interactions and fruit production in a neotropical herb: a path analysis. Ecology 69:1128–38.

Schuster, E.G., S.S. Frissell, E.E. Baker, and R.S. Loveless. 1985. The Delphi method: application to elk habitat quality. U.S. Dep. Agric. For. Serv. Res. Pap. INT-353.

Schwarz, C.J. [n.d.]. Studies of uncontrolled events. This volume.

Smith, G.J. [n.d.]. Retrospective studies. This volume.

Stolte, K., D. Mangis, R. Doty, and K. Tonnessen. 1993. Lichens as bioindicators of air quality. Rocky Mtn. For. Range Exp. Sta., Fort Collins Colo. U.S. Dep. Agric. For. Serv. Gen. Tech. Rep. RM-224.

Thomas, L. and K. Martin. 1996. The importance of analysis method for breeding bird survey population trend estimates. Cons. Biol. 10:479–90.

Tibell, L. 1992. Crustose lichens as indicators of forest continuity in boreal coniferous forests. Nor. J. Bot. 12:427–50.

Tiebout, H.M. III and K.E. Brugger. 1995. Ecological risk assessment of pesticides for terrestrial vertebrates: evaluation and application of the U.S. Environmental Protection Agency's quotient model. Cons. Biol. 9:1605–18.

Ver Hoef, J.M. 1996. Parametric empirical Bayes methods for ecological applications. Ecol. Applic. 6:1047–55.

Wilcove, D.S. 1985. Forest fragmentation and the decline of migratory songbirds. Ph.D. thesis, Princeton Univ., Princeton, N.J.

Yaffee, S.L., A.F. Phillips, I.C. Frentz, P.W. Hardy, S.M. Maleki, and B.E. Thorpe. 1996. Ecosystem management in the United States. Island Press, Covelo, Calif.

Zuboy, J.R. 1981. A new tool for fishery managers: the Delphi technique. N. Am. J. Fish. Manage. 1:55–9.