3 International Series on Applied Systems Analysis

Adaptive Environmental Assessment and Management

Edited by

C.S. Holling
Institute of Animal Resource Ecology
University of British Columbia

Sponsored by the United Nations Environmental Program

A Wiley–Interscience Publication
International Institute for Applied Systems Analysis

JOHN WILEY & SONS
Chichester–New York–Brisbane–Toronto

Library of Congress Cataloging in Publication Data:
Main entry under title:

Adaptive environmental assessment and management.

(Wiley IIASA international series on applied systems analysis; 3)
'A Wiley–Interscience publication.'
Bibliography: p.
Includes index.
1. Environmental impact analysis. 2. Economic development. 3. Environmental protection. 4. Ecology.
I. Holling, C. S.
TD194.6.A33 333.7 78–8523
ISBN 0 471 99632 7
Typeset at the Alden Press, Oxford, London and Northampton and printed by The Pitman Press, Bath.

1 Overview and Conclusions

Although the focus of this book is environmental assessment, its central message is that the process itself should be replaced. Environmental concerns are now often dealt with in a fixed review of an independently designed policy. We argue that this reactive approach will inhibit laudable economic enterprises as well as violate critical environmental constraints. We offer, as an alternative, the process of adaptive environmental management and policy design, which integrates environmental with economic and social understanding at the very beginning of the design process, in a sequence of steps during the design phase and after implementation. This argument is directed to senior administrators and policymakers who are responsible for the design of mechanisms and processes for dealing with developmental issues.

At the same time, however, we recognize that in many countries environmental assessment is practiced as a reactive review process. Even in that mode, the goal of environmental protection can be more validly and effectively achieved by the application of concepts, procedures, and techniques different from those commonly used. We describe these methods in some detail, directing our analysis to those persons with operational responsibility for doing environmental assessment and for communicating the results to senior administrators.

Because we are speaking to these two audiences, not all chapters will be of equal interest to all readers. Some concentrate on broad conceptual issues, some on fundamental procedures, and some on nontechnical but still detailed descriptions of techniques. The final chapters provide specific examples of five case studies.

This first chapter is designed for both audiences. It presents a broad overview and summary of the book – the issues, concepts, procedures, and techniques. Since it is written as an extended executive summary meant to stand largely alone, the themes and framework of analysis presented here will be repeated throughout the remaining chapters in greater detail.

In this summary we will treat five themes. The first is a brief encapsulation of present practice, presented in a rather exaggerated way for emphasis. The second provides a background that describes how present assessment practices have evolved. The third concerns the issue of uncertainty and the problem it now presents. The fourth offers a view of stability and resilience of systems, pointing to resilient or robust policy design criteria that differ from the traditional. The fifth and final topic reviews the processes and techniques that have emerged from our experience in dealing with specific problems of environmental policy design and assessment. Together, this set of issues, concepts, and techniques defines our approach.

MYTHS OF ENVIRONMENTAL MANAGEMENT AND ASSESSMENT

Perhaps the best way to introduce what adaptive environmental management and assessment is, is to indicate what it is not. Below we discuss twelve "myths" of present management and assessment. However much these appear to be straw men, they are still inherent in present practice. Most of us have subscribed to at least one or two at some time or another.

MYTHS OF ENVIRONMENTAL MANAGEMENT

The first set of myths concerns policy design and decisions.

Myth 1 The central goal for design is to produce policies and developments that result in stable social, economic, and environmental behavior.

Stability is a two-edged sword. If our knowledge of objectives and structure is complete, then design should indeed minimize the chance of the unexpected. But what we know of social, economic, and environmental behavior is much less than what we do not know. Therefore, the opportunity to benefit from change and the unexpected should be part of the design goal.

Myth 2 Development programs are fixed sets of actions that will not involve extensive modification, revision, or additional investment after the development occurs.

Program goals change, and unexpected impacts trigger corrective actions that result in progressively greater economic and political commitments to make further corrections if the initial ones are not successful. Thus, present decisions have *future decision consequences* as well as direct environmental ones, and these subsequent induced decisions often generate greater environmental impacts than seemed possible originally.

Myth 3 Policies should be designed on the basis of economic and social goals with environmental concerns added subsequently as constraints during a review process.

We must ride with ecological forces as much as with social and economic ones. Unless all are incorporated at the very beginning of the design, opportunities to achieve social goals are lost and subverted. The design will be more costly and the benefits too sensitive to the unexpected.

Myth 4 Environmental concerns can be dealt with appropriately only by changing institutional constraints.

This might ultimately be necessary, but constraints are more often perceived than real. Often, for example, one agency will have policy and management responsibility, and another, research or assessment responsibility. But the latter agency can hardly fulfill its research role without a policy perspective. That perspective can be developed internally if the goal is to design a number of alternative, but possible, policies. Each of these implies distinct or shared priorities for research that can be a powerful guide for research planning. At the same time, they provide an interface of communication between those responsible for the research and those responsible for decisions and management.

MYTHS OF ENVIRONMENTAL ASSESSMENT

This second set of myths concerns the details of how assessments are done.

Myth 5 Environmental assessment should consider *all* possible impacts of the proposed development.

The interesting question is rather: What does the fact that it is impossible to foresee all (or even most) of the impacts imply for the structure of the basic development plan and assessment research?

Myth 6 Each new assessment is unique. There are few relevant background principles, information, or even comparable past cases.

It is true that each environmental situation has some unique features (e.g., rare animal species, geological formations, settlement patterns). But most ecological systems face a variety of natural disturbances, and all organisms face some common problems. The field of ecology has accumulated a rich descriptive and functional literature that makes at least some kinds of studies redundant and some predictions possible. The same is true for economic, social, and physical aspects of the assessment.

Myth 7 Comprehensive "state of the system" surveys (species lists, soil conditions, and the like) are a necessary step in environmental assessment.

Survey studies are often extremely expensive yet produce nothing but masses of uninterpreted and descriptive data. Also, they seldom give any clues to natural changes that may be about to occur independently of development impacts. Environmental systems are not static entities, and they cannot be understood by simply finding out what is where over a short survey period.

Myth 8 Detailed descriptive studies of the present condition of system parts can be integrated by systems analysis to provide overall understanding and predictions of systems impacts.

The predictions from systems analysis are built up from an understanding of causal relationships between changing variables. Descriptive studies seldom give more than one point along each of the many curves that would normally be used to express such critical relationships. In short, what a complex system *is doing* seldom gives any indication of what it *would do* under changed conditions. Again, the interesting question is: What are the assessment, monitoring, and policy implications of the fact that even comprehensive systems models can make predictions only in sharply delimited situations?

Myth 9 Any good scientific study contributes to better decision making.

The interests of scientists are usually quite narrow and reflect the particular history of a discipline. There is thus no guarantee that in a scientific study the appropriate variables or processes will be measured, or that information will be collected on the proper spatial and temporal scales to address management questions. The research necessary for adaptive assessment and design must be focused through policy concerns.

Myth 10 Physical boundaries based on watershed areas or political jurisdictions can provide sensible limits for impact investigations.

Modern transportation systems alone produce environmental impacts in unexpected places. Transfers of impacts across political boundaries lead to a wide range of political and economic reactions from the other side. A narrow study that fails to recognize at least some of these impacts and reactions will provide inadequate and misleading information for the decision maker.

Myth 11 Systems analysis will allow effective selection of the best alternative from several proposed plans and programs.

This assertion would be incorrect even if systems models could produce reliable predictions. Comparison of alternative policies can occur only if someone places values on the results of each alternative. Rarely is this an explicit part of environmental assessment.

Myth 12 Ecological evaluation and impact assessment aim to eliminate uncertainty regarding the consequences of proposed developments.

Attempts to eliminate uncertainty are delusory and often counterproductive. The appropriate concept for both assessment and policy design is a recognition of the inevitability of uncertainties and the consequent selective risk-taking.

These shortcomings of present assessment practice are in part the consequence of the sudden and recent broad perception that environmental issues are important to the health of societies. The shortcomings reflect an urgent response to apparent crises, and before providing suggestions for an alternative, it is useful to explore this historical background.

DEVELOPMENT OF CONTEMPORARY ASSESSMENT PRACTICES

It is commonplace now to perceive limits – limits to growth, to resources, to climatic and environmental stability. Although the general perception of the importance of those limits is relatively new, mankind has always been confronted by them. There have always been problems of resource depletion, environmental contamination, and poverty. Moreover, industrial man's history, by and large, has been one of successful resolution of these problems, at least in the short term. In recent years, however, they seem to have taken the shape of crises, perhaps because the problems are ours and not our fathers'; more likely because our perceptions and methods, having once helped, now hinder.

The current approach to environmental concerns has been very much colored by a sudden shift of public awareness in the industrialized nations. What was once the concern of a minority became the concern of the public at large. The problems were not that qualitatively different from those of the past, but in the past they were largely local and often transient. Solutions were often found by simply waiting – next year's weather for crop production could well be better. And when this was not the case, there was often "somewhere else" that provided a way out – an unexploited resource, an unsettled piece of land, a new river to dam. In seeking elsewhere for solutions, the knowledge and technological devices needed could evolve at an easy pace. It required more innovation of spirit than innovation of technique for the Young Man To Go West.

With the "elsewheres" gradually becoming scarcer, however, alternatives had to be sought in new knowledge and technology rather than in new places. In seeking them, the scale and intensity of impact inevitably grew, eventually triggering that sharp shift of public awareness.

The past solutions however, provided little experience with ways of dealing with the environment. In most instances the goals of economic and social advance were most promptly achieved by subduing nature. The present protective response was therefore natural. In the face of limits now so suddenly perceived, time at least could be bought by protection of the environment and regulation of its use. The response is, therefore, largely reactive. Regional developments or policies are still designed within an economic context and reviewed only after the fact for their environmental consequences.

There has now been enough experience with this approach to suggest two major difficulties. First, the fundamental properties of any development or policy are set very early in the design stage. If problems arise because the original context was too narrow, any fundamental redesign is extremely difficult unless there is extraordinary pressure. Confrontation is guaranteed as different groups identify clear conflicts with their own interests. Confrontation and public debate are essential dimensions of the development of policies, but if the issues emerge only because the design phase was unnecessarily limited, economic enterprises offering legitimate social benefits can be halted and opportunities for husbanding and enhancing man's natural endowment can be subverted.

The second major difficulty with the present protective and reactive response is that it makes the practice of environmental assessment arbitrary, inflexible, and unfocused. Each issue is often dealt with as if it were unique, as if the environmental consequences could be separated from the social and economic ones. And yet the major environmental impact of a pipeline, for example, often occurs not along the route itself but at sites remote from it, as human settlements experience an acceleration of economic and population pressures. Such environmental effects induced through social forces are rarely considered. And the reverse is true. Deleterious social and economic impacts can be induced through ecological forces that, if recognized early, could at times be turned to man's benefit rather than simply suppressed and ignored.

The result of simple reactive assessment is therefore intolerable. How can we know what to measure for base-line information or assessment if the detailed character of the policy or development is not revealed until it has largely crystallized? The tendency is to measure everything, hence producing the indigestible tomes typical of many environmental impact statements. More time and effort are spent in measuring what is, rather than in projecting what is likely to be or could be made to be. Static and confused description replaces anticipation and clear prescription of alternatives.

But enough experience has now accumulated to allow a start to be made in developing and implementing an alternative approach. Systems ecology, in partnership with the physical sciences, has now matured enough to be capable of producing succinct representations of key elements of ecological and environmental systems. The resulting models mimic not simply static properties, but the dynamic ones that shift and change because of natural and man-induced influences. They can serve, alone or combined with similar economic representations, as a kind of laboratory world for the development of alternative policies and for the exploration of their impact.

The systems sciences have evolved methods of optimization that, if used with care, can point toward general policies that better achieve objectives by working with, rather than against, the rhythm of ecological and economic forces. There are techniques to deal with uncertain information, with mobilizing available data on partially known processes, and with the formulation of objectives that are less sensitive to the unexpected. All these lie at the heart of developing policies that recognize and benefit from both economic and environmental realities. Finally, decision theory provides a few theoretical hints and some practical experience in ways to explore decisions in the face of uncertainty and conflicting objectives.

This set of descriptive and prescriptive techniques provides the skeleton for policy design that can integrate economic, ecological, and environmental understanding. What's more, this integration can commence at the very beginning of the design process. But techniques alone are not enough. The best of techniques, unless guided by a clear vision of the fundamental issues and by a concept that gives them form, can turn solutions into larger problems. We argue that the fundamental challenge is not simply to better mobilize known information. Rather, it is to cope with the uncertain and the unexpected. How, in short, to plan in the face of the unknown. It is to that generic issue that we now turn.

THE ISSUE OF UNCERTAINTY

The design of policies or economic developments implies knowledge – knowledge to develop alternative policies, and knowledge to evaluate their respective consequences. And indeed a significant part of the contents of this book is concerned with how to deal with qualitative and quantitative data, how to use this knowledge of fundamental processes to construct models that can serve as "laboratory worlds" for the testing and evaluation of intrusions, developments, and policies. How, in short, to better reduce uncertainty. But however intensively and extensively data are collected, however much we know of how the system functions, the domain of our knowledge of specific ecological and social systems is small when compared to that of our ignorance.

Thus, one key issue for design and evaluation of policies is how to cope with the uncertain, the unexpected, and the unknown. It seems a common plea that too little is known of the structure and behavior of ecological systems. That can lead to the syndromes of living dangerously ("who cares how birds and bugs are affected – jobs and income are more important") or living safely (" nothing must be done until we know more"). But man has always molded the environment and been molded by it, and we will argue that more is known of ecological systems than is generally appreciated or used. Nevertheless, there is still uncertainty.

At the same time, there is growing unease about the economic systems with which ecological systems are linked. The unexpected increases in oil prices that have touched so many aspects of national economies have the same flavor as the unexpected appearance of a new crop pest after successful control of the original pests with insecticide. There is sufficient knowledge to anticipate both events, but both come as surprises. And, being unexpected, they are ignored in the original design of policies.

Even the ultimate objectives of environmental policies and developments are uncertain. A renewable-resource industry might have as an initial high-priority objective stabilized employment over the short term, which then shifts to a major concern for environmental standards, then to diversity of opportunity, and then to simple economic objectives. A design that assumes that objectives are immutable can rapidly foreclose options if those objectives shift.

Man has always lived in a sea of the unknown and yet has prospered. His customary method of dealing with the unknown has been trial-and-error. Existing information is used to set up a trial. Any errors provide additional information to modify subsequent efforts. Such "failures" create the experience and information upon which new knowledge is built. Both prehistoric man's exploration of fire and the modern scientist's development of hypotheses and experiments are in this tradition. The success of this time-honored method, however, depends on some minimum conditions. The experiment should not, ideally, destroy the experimenter – or at least someone must be left to learn from it. Nor should the experiment cause irreversible changes in the environment. The experimenter should be able to start again, having been humbled and enlightened by a "failure." And, finally, the experimenter must be *willing* to start again.

There is now increasing difficulty in meeting these minimum conditions. Our trials are capable of producing errors larger and more costly than society can afford. While the individual parts of a nuclear plant, for example, can be tested to the point of failure, the full integrated system cannot. Moreover, when this integrated system is viewed as not just an engineering system, but one that links ecological and social aspects as well, then the variety of unexpected events – from coolant failure to sabotage – and the scale of the consequences make trial-and-error truly a way to live dangerously.

Moreover, even when errors are not, in principle, irreversible, the size of the original investment of capital and of prestige often makes them effectively so. This behavior has its roots in a very human characteristic of industrial man: we do not like to admit and pay for our past mistakes; we prefer to correct them. And the consequences of correcting an inflexible plan is often increasing investment, increasing costs for maintaining and controlling the system, and progressive foreclosure of future decision options. Retreat from error is difficult for three reasons: because of the scale and consequence of possible "irreversible" physical changes; because changes in expectations for future returns make traditional goals politically or economically unacceptable; because reserves of capital and faith are lost, and the governed rise up against the governors, forcing them to invest in order to satisfy basic constraints newly perceived.

But the search for a solution should not replace trial-and-error with some attempt to eliminate the uncertain and the unknown. That could only result in tighter monitoring, regulation, and control based upon an illusory assumption of sufficient knowledge. Rather, the proper direction lies in the design of policies and economic developments that can allow trial-and-error to work again. Efforts to reduce uncertainty are admirable. Much of this book concerns just that. But if not accompanied by an equal effort to design for uncertainty and to obtain benefits from the unexpected, the best of predictive methods will only lead to larger problems arising more quickly and more often. This view is the heart of adaptive environmental management – an interactive process using techniques that not only reduce uncertainty but also benefit from it. The goal is to develop more resilient policies.

STABILITY AND RESILIENCE OF SYSTEMS

Our concept of resilience emerges from a very specific understanding of the structure and behavior of ecological systems (Chapter 2). It seems to have a counterpart in the behavior of institutional and other systems. The way a system responds to a planned or unexpected disturbance depends on its stability properties. One view, implicit in many of man's past efforts to manage, assumes that there is global stability. That is, no matter how large the disturbance, the system will recover to its original stable condition, once the disturbance is removed. This is a view of a Benign Nature that can comfortably accommodate trial-and-error on any scale. In this view, "big," which is necessary for economies of scale to be achieved, is always allowable.

A contrasting view infers a high degree of instability of ecological systems. They are fragile and caught in a natural rhythm of small-scale extinctions. They persist because of diversity in structure and over space. Outside sources provide the source of recovery. This view of Ephemeral Nature naturally leads to "small-is-beautiful" and to concentration on the need for spatial variety, diversity of opportunity, and fine-scaled, local autonomy.

But the burden of examples and of analysis leads to a combination of these extremes. Natural systems often have more than one stable mode of behavior. As long as variables like population density, amount of nutrients or even level of unemployment stay within a certain range, small disturbances can be absorbed. Quantities may change, but qualitative behavior does not. Small disturbances can be introduced incrementally, particularly if no apparent danger signal appears in the system. Then one additional increment can "flip" the system across the boundary into a totally different mode of behavior. A river can become an open sewer, or the economy of a nation can suddenly begin to prosper. In this world, the prudent manager would be wise to view nature less as benignly forgiving than as a Practical Joker.

The "small-is-beautiful" theme can still operate much as before with a more focused sense of optimal spatial scale and a recognition of the need for a balanced dependency on outside forces. But "big-is-necessary" can also be accommodated; one need only be more cautious. Thus if boundaries exist separating "desirable" from "undesirable," then the task is to control the variables carefully to keep them well away from the dangerous boundary. In addition, the boundary itself may be made less permeable; the strength of the guardrail can sometimes be more critical than the characteristics of the highway. To achieve less permeability effectively, big might well be necessary as the only way to gain sufficient knowledge of the boundary, to monitor the distance to it, and to institute control procedures to maximize that distance.

Maximizing the distance from an undesirable region is within the highly responsible tradition of safety engineering, of nuclear safeguards, of environmental and health standards. It works effectively if the system is simple and known – say, the design of a bolt for an aircraft. Then the stress limits can be clearly defined, and the bolt can be crafted so that normal or even abnormal stresses can be absorbed. The goal is to minimize the probability of failure. For bolts, this approach has succeeded. The probability of failure of bolts in aircraft, for example, is extremely small. But in parallel with that achievement is a high cost of failure – the very issue that makes trial-and-error as now practiced so dangerous.

One additional view of stability is needed. The three views – of Nature Benign, of Natural Ephemeral, and of Nature the Practical Joker – have been described thus far in three steps of increasing reality and comprehensiveness. In each case, however, it was implicitly assumed that the rules of the game were fixed. But ecological – and for that matter, economic, institutional, and social – systems are not static or completely determined. Variability and change are the rule and provide the next step toward reality.

Chance events dominate some ecosystems. Fire, rather than being a disaster, is the source of maintenance of some grassland ecosystems. Shifting patterns of drought determine the structure of some savannah systems in Africa. In addition, the variables themselves can move, through internal forces, from one region of stability to another. That is one of the lessons derived in the case study of forest pest management discussed in Chapter 11. There, we see that periodic insect outbreaks can be triggered by chance patterns of weather, by dispersal of moths from other areas, or by the natural growth of the forest. Populations increase explosively from low stable numbers to high. While the high numbers are stable for the insect, the forest cannot absorb the level of defoliation. The forest dies back, regeneration occurs, and the clock is started again. Such large swings and movements between stability regions contribute directly to forest renewal and to the maintenance of diversity.

Hence the variables of natural ecosystems do not reside in one stability region far from boundaries. Locally, species may even become extinct, to be reinstated through contributions from other localities. The variables are moving continually and the stability boundaries are being tested periodically. There is an internal monitoring of boundaries.

And now the central issue: not only do the variables shift and move, but so do the boundaries between stability regions. In ecosystems, this "stability landscape" owes its features to natural selection, which responds to the variability that occurs naturally. The reason boundaries exist where they do is that they are tested periodically.

This dynamic pattern of the variables and of the basic structure lies at the heart of coping with the unknown. However much we may be sure of the stability landscape of a physical system, we will rarely know the societal or ecological stability landscape in any detail. Policies often attempt to reduce variability within these partially known systems, either as a goal in itself or as an effort to meet standards of safety, health, or environmental quality. That constricted variability in turn may itself shift the balance of natural, cultural, or psychological selection so that stability regions will contract. Paradoxically, success in maximizing the distance from a dangerous stability boundary may cause collapse, because the boundary may implode to meet the variables. If surprise, change, and the unexpected are reduced, systems of organisms, of people, and of institutions can "forget" the existence of limits until it is too late.

This final view is of Resilient Nature, where resilience is a property that allows a system to absorb and *utilize* (or even benefit from) change.

But, of course, a different perspective on the generic issue, even with a concept to give it form, is not enough. Flowing from it must be some effort to design and test specific procedures and techniques that allow at least one step to be taken in harmony with this perception.

PROCEDURES AND TECHNIQUES

Our recommendations for a specific procedure and a range of techniques come from our particular experience with a number of studies of renewable resource problems in different national settings: renewable resource management and disease control in Venezuela and Argentina; range and wildlife management in the United States; developmental and oceanographic problems in Europe; ecological process studies in the Soviet Union; renewable resource and pest management systems in Canada.

We provide five specific case studies (Part II) so that examples of the results of applying these methods can be exposed. The first is a detailed example of the lessons learned in developing and evaluating policies for a problem of forest pest management. This one has gone farthest in coping with existing management questions, validating alternative modeling techniques, and generating management alternatives and evaluating their consequences. It has resulted in the adoption by agencies of two Canadian provinces of the approach for setting research priorities and developing and evaluating management options. The second case study is an example of an analysis of new procedures to enhance and manage fish stocks in North America, in which adaptive management approaches are proposed that provide, as an integral part of the policy design, a way of reducing uncertainty. It has gone farthest in affecting and modifying a proposed new development to enhance fisheries populations. The third is an example of the results of one of the intensive 5-day workshops (whose details will be described shortly) that resulted in a preliminary but broad assessment of the consequences of development in a high alpine region of Europe. The fourth is a modeling and policy analysis of a major regional development in a sparsely populated region of Venezuela involving hydroelectric, forestry, and agricultural development. The fifth and final example deals with the impacts on wildlife populations of oil-shale development in the western United States.

In each case, the purpose was to develop a set of alternative policies or plans and assess their environmental, economic, and, in some cases, social consequences. At first thought, therefore, the process we recommend would seem more appropriate for environmental management than for assessment. Before addressing that question in the next section, however, we shall compare our recommendations with two procedures that are in common use.

At one extreme is the approach of having a small core planning staff contract out parts of the study – the hydrological analysis, vegetational or wildlife survey, and so on. Integration occurs on receipt of the contracted reports. But two difficulties emerge. First, the contracted pieces typically drift farther and farther from the question posed, and, since the parts are not linked with each other throughout, useful integration of the pieces becomes very difficult. Second, it is unlikely that a small core planning team will have sufficient breadth and depth of knowledge to identify those key elements or processes that deserve analysis. To protect themselves, there is a natural tendency for them to wish to measure everything they can think of. Typically, these are static quantities, both environmental and economic, or the more obvious physical processes. But the problems are not static; they are not simply physical; their behavior comes from the integration of the parts and not just from the parts themselves. As a result, much of the information gathered is unnecessary, and key items are ignored entirely. The cost is unnecessarily large, and the product incomplete.

At the other extreme is the large interdisciplinary team that attempts to develop the integration missed in the above approach by mobilizing most of the expertise within one organization. In order to avoid bureaucratic growth, a task force is sometimes established only for the duration of the study, with staff provided from a number of existing institutions. Such large teams, however, have a high financial, organizational, and emotional overhead attached to them. We suppose this could be overcome by appropriate organizational techniques, but the common experience is that it is not. Anarchy and fragmentation often develop. Separation of the team from the policymaker is common, and internal goals evolve that have more to do with survival of the team than with the original purposes.

In contrast, the process we have evolved depends on a small core group of two or three analysts and a support staff of one or two. The core group should have experience in two or three of the disciplines involved – for example, forestry, fisheries, economics, or ecology. At the same time, their prime experience should be in integrating information and coordinating people. In our case the integration comes from application of systems techniques – e.g., computer modeling of dynamic systems, mathematical analysis, optimization, utility analysis, and communication. The coordination comes from the development of a series of steps, each of which is initiated by a workshop that brings together key cooperators for short periods of intense interaction. The time between the workshops is spent in consolidation: the core group refines the model(s), develops initial alternative policies, analyzes data; the collaborators collect and integrate data and information both on behavior of the system and on goals of the project. The workshops that define the sequence of steps are the heart of the approach (Chapters 3 and 4). They provide a series of sequential targets, maintain integration while minimizing organizational and emotional overhead, and allow involvement of a wider spectrum of key actors than is normally possible. The policymaker, busy as he is, is involved at key points for short periods.

Each workshop draws upon up to twenty specialists, the choice depending on the particular stage of the process. The first workshop is critical, for it is then that the problem is defined and focused. It is essential to have all prime "actors" present at that time – scientists, managers, and policy people. The policy people and the managers provide a balance to the scientist's penchant for exquisite detail and excessive resolution. The scientists provide the rigor and understanding of fundamental physical, ecological, and economic forces. During such a workshop, impact categories are classified, key information needs defined, alternative actions described, and the framework and crude working version of a computer model developed. Even if, through lack of expertise, facilities, or time, a model is not developed, the techniques of organizing elements in preparation for a formal modeling effort are themselves of fundamental value. The point is that, at the very beginning of the study, all elements – variables, management acts, objectives, indicators, time horizon, and spatial extent – are jointly considered and integrated. Even the crude model that is developed at this stage can be a powerful device to explore the significance of unknown relationships. By testing alternative extremes, priorities can be established for data and for scientific and policy analysis.

That first workshop is followed by a period of consolidation. The model is further refined and tested by the core group. Some of the attending specialists assume responsibility for collecting detailed information on both scientific and policy questions. Subsequent workshops further define management objectives, construct alternative policies, and explore uncertainties. Some workshops involve only scientists when the goal is critical scrutiny of underlying assumptions. Some involve largely managers, when the issue concerns operational feasibility. Some involve only decision makers, when the purpose is to ensure relevance and understanding. In every instance, a period of consolidation follows the workshop.

One key technique makes it possible to set this process in motion. That is the ability to abstract the essential properties of at least some ecological and environmental systems and to represent them in a model that mimics behavior over time for a variety of conditions. By essential, we mean those properties that generate the minimal natural behaviors that must be retained in order for the model to be responsive to the management questions. The models, therefore, are not designed for general scientific purposes but for very specific management ones. Hence, they attempt to be both parsimonious (and hence tractable) and realistic (and hence useful).

Our professional experience is ecological and environmental. But it is obvious that at least regional economic systems can be treated in the same way and integrated with the ecological and environmental systems. Because this integration occurs in the very first step in the analysis, it is possible to achieve designs that work with rather than against natural forces. In so doing, more opportunity is provided for less costly and intrusive economic developments and even for the enhancement of natural systems rather than simply for their protection. We provide examples of this integration, as well as examples in which simple social phenomena, such as demographic and market processes, are included. More complex social behaviors are well beyond the state of the art and are better dealt with as they are ideally treated now – through experience, sensitive perception, and public dialogue.

The models conceived in the workshop process focus on one or more of the ecological, environmental, economic, or simple social forces underlying many developmental problems. They provide a credible "laboratory world," which makes it possible to mobilize a set of techniques for prescription and evaluation – techniques to allow the following:

- Generation of a range of alternative objectives
- Design of effective policies to achieve alternative objectives
- Generation of indicators (social, economic, resource, and environmental) of relevance for decision
- Evaluation of each policy in terms of the behavior of the indicators over space and time
- Partial compression of indicator information to facilitate screening of the most appropriate policies
- Communication and interaction between and among those who design, choose, and endure policies (staff, decision maker, and citizen)

The particular techniques chosen to represent or model the dynamics of a system need not be numerical simulation models. Beyond the constraints set by expertise, the characteristics of the problem in part suggest the technique chosen. There are three key characteristics; (a) the number of variables, management actions, and spatial elements; (b) the level and breadth of understanding of underlying physical, ecological, and economic processes; and (c) the number and quality of data. No matter what combination of these any specific problem has, there is a technique available.

Our exploration of techniques covered a range from nonquantitative cross-impact matrices, to "qualitative" modeling techniques that generate dynamic changes over time without data on magnitude, to simple simulation methods, and finally to fully detailed, large simulation modeling techniques.

If the level of understanding of processes is low, and data are scarce, all of these techniques seem to perform equally well or poorly. But even when data are scarce, there is usually more understanding of processes available than is generally recognized. And there are techniques available that organize and focus understanding of processes even in the face of scattered data. If these techniques are used, then we have found simple or complex simulation models to be clearly superior in predictive capacity, responsiveness to questions, and relevance of results (Chapter 5).

Even if we have a satisfactory dynamic model, however, one further step is helpful. Such models are complex. They are so difficult to understand that many are tempted to play computer games with them in a blind, undirected exploration. But there are ways to simplify these models so that we can understand the essential behavior. The structure of such a model can be analyzed in order to reduce the number of variables and interrelations to those that are key determinants of the qualitative behavior. Often a simplified set of equations can then be devised that is used to provide a depth of understanding that is enormously useful as a guide to intuition and judgment. Alternatively, topological or graphical representations can sometimes be designed to achieve the same purpose in a form more readily understood by nonmathematicians (Chapter 6). All these techniques provide a clear direction to this search for policies and impacts, and allow us to convey our understanding to the decision maker more effectively.

Before a model can be used as a laboratory world to test the consequences of alternative policies, its degree of credibility must be explored. Note that no model – mental or mathematical – is "true." But degrees of credibility and usefulness can be defined, not, as is often done, by attempting to tune parameters to fit a given set of historical data; rather, the effort should be directed to *invalidate*, and not to validate, the model (Chapter 7). That is in harmony with the scientific method, where only disproof, not proof, is possible. Invalidation requires information from extremes of behavior that can then be compared with model predictions for similar extremes. The data on extremes come from natural experiments that have been historically recorded – for example, the extreme weather that occurs in some particular geographical region or that has occurred at some past time. Further information on extremes comes from the behavior of the target system or similar systems that have been subjected to management by man. The more robust the model at these extremes, the more confidence can be placed in its behavior under newly designed policies.

That leads to the final set of methods, which use this laboratory world to develop, explore, and evaluate alternative policies (Chapter 8). These methods include the formulation of objectives, the definition of indicators, and the touchy job of evaluation.

There may be many ways of attempting to achieve a given objective. For example, maximum sustained yields from a fishery can be reached by controlling fishing effort through manipulation of open fishing days or by setting catch quotas. The role of the model at this point is to generate those indicators that best match the objective. Because costs and benefits arise in many forms, the manager usually needs a large number of indicators. One necessary step then becomes the compression of this mass information into a comprehensible form. There are several ways this can be done. Because of the breadth of their comprehensibility, we prefer indicator compressions that are graphical. The relative merits of alternative management actions can be evaluated using the indicator output from the model. Both formal and informal evaluation techniques are useful here, but in either case, they are used only to point out policies that should be more thoroughly explored. The object is not to derive some mythical "optimal" policy, but rather to compare and then combine alternative policies in order to illuminate the range and nature of available choices.

Methodologies are only parts of the process, however. Communication holds these parts together. The thick volumes that characterize the products of many impact assessment programs are an inefficient and ineffective way to communicate results. There are other ways to present the information, in which the level of detail and attention required are determined by the particular user (Chapter 9). The resulting reports, graphical summaries, and audiovisual materials become, with the workshops, an integral part of the procedure, allowing interaction and adaptive modification throughout.

That, then, completes our summary of the issues, concepts, procedures, and techniques. In closing this chapter, we discuss not their merit or lack of it, but whether it is at all practical to implement them in the face of present institutional realities.

THE PROBLEM OF IMPLEMENTATION

DEVELOPED COUNTRIES

To those conditioned by North American approaches to environmental assessment, where assessment is viewed as a passive reaction to an independently developed proposal, the process described above would seem too inclusive. Proposals are generated according to guidelines and are then reviewed by an informal panel. Certainly, the modeling techniques, at least, would be useful in forming a judgment. But, we would argue, the other techniques and the procedures themselves would make a qualitative improvement in even this reaction mode of assessment.

In order to assess something properly, there has to be a yardstick against which performance is measured. And that yardstick is some alternative policy or development. One is clearly the "no-policy" world. If that is the only alternative, then confrontation between no development and development is encouraged. But other explicit policies would provide a richer menu of alternatives that would sharpen and focus response to a proposal and suggest specific modifications. Once there is a requirement for such alternatives, if only for internal comparison by the assessment panel, then assessment is in the game of policy design. At that point all the procedures and techniques described above can apply.

Despite the breadth and depth of such an adaptive assessment approach, the cost is small. An experienced core group of two or three analysts and a support staff of two could comfortably undertake one major assessment a year, together with perhaps four to six preliminary "rough cut" assessments. Each, of course, would draw heavily on available expertise within the agencies concerned with the problem. Hence the benefit is not only the assessment itself but a growing body of experience within agencies. In nearly every instance, there are enough *existing* data, however, scarce, to begin, since we argue that the design of a program to collect data for establishing baselines or for monitoring must follow and be integrated with the approaches described here, rather than precede them. A modest budget is necessary to mobilize and organize existing data, but this can typically be managed within the cooperating agency. At most, it is a one-man-year effort. Similarly, computing budgets can be as small or large as facilities and expertise warrant. The resource in scarce supply is rarely money; it is expertise and experience in the techniques and procedures described here. If this expertise and experience are available or can be developed, the costs are an order of magnitude less than those typical in North American impact assessment efforts.

DEVELOPING COUNTRIES

When a new approach, such as the one put forward in this book, appears, it is useful to examine it from different viewpoints. Here an attempt is made to focus on some aspects that seem particularly relevant from the point of view of developing countries. Moreover, by adopting that perspective the lessons for industrialized countries might, paradoxically, emerge more vividly.

Problems are perceived very differently in developing countries, and, in addition, there is often a high within-country cultural diversity. Because of these differences, developing countries can sometimes more easily explore new ways of looking at problems and new solutions. An example is the perception of eutrophication in Southeast Asia. There, high nutrient loads, abundance of algae, and aquatic weeds like water hyacinth are considered desirable in rice fields, fishponds, and even in some natural water bodies. They are viewed as a resource and as enhancing the production process, rather than as a nuisance. Also, it was not coincidental that a totally different way of measuring socioeconomic growth in global models was originated in developing countries. In the Latin American World Model this was life expectancy at birth. Differences in the perceptions of problems made it necessary to look for alternative solutions.

In this book, a nontraditional perception of the behavior of ecological systems is presented. We link this with the potential richness of perceptions emerging from the present cultural diversity on our planet. This variety is prized because it is not yet possible to decide whether one, some, or many perceptual frameworks, or paradigms, are necessary in order to cope with different problems in different regions of the world. It is likely that new and evolving paradigms will be needed. And it is also likely that some of these will originate in developing countries, and will modify and enrich the views presented here.

Too often in the past, socioeconomic development and environmental quality have been perceived, or construed, as if they were quite opposite, antagonistic concepts. The conceptual framework proposed here is not only absolutely compatible with the dynamic concepts of development and the rational use of natural resources, but it also tends to promote the generation of self-reliant and endogenous approaches to the environmental problems – approaches appropriate to local conditions, needs, and socioeconomic structures.

For any one management or developmental objective, there are usually many alternative ways of implementation. We emphasize that it is essential to generate and consider a wide range of alternatives, especially in developing countries. Inadequate search for alternatives can make plans and projects fail utterly because they are not adapted to the local realities. This is evident in tropical agriculture, where there are many examples of attempts to introduce temperate-zone, capital-intensive technologies. And more important than alternatives for implementation, the generation of alternative objectives, or goals, is viewed as a fundamental process.

The emphasis throughout the book upon a permanent and inherent state of change in ecological systems suggests a richness of qualitatively different behavior modes that might be an appealing concept for the developing countries. It is often shown that attempts to force classical stability or constancy may lead to a shift of behavior into undesirable modes. But changes need not be catastrophic. By the same token, an explicit search could be made to discover desirable stability regions. Strategies might then be devised to move the environmental or socioeconomic system from an undesirable condition to more desirable ones.

Developing countries, perhaps more than others, are in a permanent state of change. Although it is an open question whether the perceived goal is always the desirable one, in most cases in the developing countries it is good to move away from the present state. Thus, developing countries, having no vested interest in constancy, might find the concepts of resilience, of managing with uncertainty or even managing uncertainty itself, appropriate and suggestive. The concepts might also have an influence upon the socioeconomic theories, approaches, and strategies of national, regional, and global development. For instance, the concepts emphasized in the book might help one to understand how some decisions and strategies reduce the stability region of a system, showing how some policies lead to a narrowing of the set of future options. So, even though the set of case studies utilized as examples in the book pertain to a small class, the implications of the approach impinge upon a much wider set of problems.

It seems clear that any approach that attempts to deal explicitly with uncertainty would be of particular relevance for the developing countries. Considering the needs for rapid socioeconomic development, the existence of unexploited natural resources, and the availability of technology for wide-scale projects, the uncertainties involved are not only great but are often of a qualitatively different nature than in developed countries. This, coupled with the great vulnerability of major segments of the population, suggests that the explicit consideration of uncertainties is a fundamental concern of developing countries.

While it might be argued that some of the techniques presented here are not universally adaptable, the main emphasis throughout is on an overall approach to the problems. This is why a range of techniques, from the simple and naive to the more sophisticated, has been explored. The choice and usefulness of a particular technique depend very much upon the particular situations and available resources. As an example, for a group of experts engaged in a regional planning project, even the simplest approach to the first steps in the workshop procedure has proved to be very valuable in reidentifying the relevant issues, promoting integration among disciplines, and producing a more global and coherent view of the problem and its solutions. This happened with a 2-day workshop in the Bermejo River basin in Argentina. Thus, the relevant question is not whether the approach presented here is the best possible one, but whether it is better than the traditional ones.

The adaptive approach is particularly useful in helping to make fast decisions where data are incomplete and uncertainty is great. All of the techniques discussed are accessible at a moderate cost, and some are very cheap. For a fixed budget, whatever its size, the approach can allow a substantial saving in terms of data collection, in the sense that the emphasis is put upon collecting only the relevant data, without following the traditional massive data collection procedure.

Finally, it is important to emphasize the value of the workshop procedure (one of the cores of the approach) in terms of its efficiency for mobilizing and organizing scarce critical resources (expertise, funds, time). It also has a high demonstration potential, thus encouraging institutional flexibility and the dissemination of integrated views about the relevant issues.

CONCLUSIONS

We have attempted, in this overview, to present the issue of uncertainty that underlies the major resource and environmental problems facing mankind. The concept of resilience, in which the different distinct modes of behavior are maintained because of, rather than despite, variability, is suggested as an overall criterion for policy design. The more that variability in partially known systems is retained, the more likely it is that both the natural and management parts of the system will be responsive to the unexpected. The very process and techniques we recommend, while aimed in part at reducing uncertainty, are designed as a changing adaptive process of policy design. It is the combination of the issue, the concept, and the process and techniques that makes for adaptive environmental assessment and management.

Although we see assessment as an integral part of management, in some countries these are viewed as separate activities. Because of this, we will separate our detailed conclusions into those most relevant for management and those most appropriate for assessment. First, the recommendations for adaptive management:

1. Environmental dimensions should be introduced at the very beginning of the development, or policy design process, and should be integrated as equal partners with economic and social considerations, so that the design can benefit from, and even enhance, natural forces.

2. Thereafter, during the design phase, there should be periods of intense focused innovation involving significant outside constituencies, followed by periods of stable consolidation.

3. Part of the design should incorporate benefits derived from increasing information on unknown or partially known social, economic, and environmental effects. Information can be given a value just as jobs, income, and profit can.

4. Some of the experiments designed to produce information can be part of an integrated research plan, but part should be designed into the actual management activities. Managers as well as scientists learn from change.

5. An equally integral part of the design is the monitoring and remedial mechanisms. They should not simply be *post hoc* additions after implementation.

6. In the design of those mechanisms there should be a careful analysis of the economic trade-offs between structures and policies that presume that the unexpected can be designed out, and less capital-expensive mechanisms that monitor and ameliorate the unexpected.

There are also specific conclusions relevant to the techniques of environmental assessment, some of which are summarized here:

1. Structural features (size distribution, age, who connects with whom) are more important to measure than values of individual variables.

2. Events at one place can re-emerge as impacts at distant places.

3. Monitoring of the wrong variable can seem to indicate no change even when drastic change is imminent.

4. Impacts are not necessarily immediate and gradual; they can appear abruptly some time after the event.

5. Variability of ecological systems, including occasional major disruptions, provides a kind of self-monitoring system that maintains resilience. Policies that reduce variability in space or time, even in an effort to improve environmental "quality," should always be questioned.

6. Many of the existing assessment methods (e.g., cost–benefit analysis, input–output, cross–impact matrices, linear models, discounting) assume none of the above occurs, or at least, that none is important. All such methods should be used with caution.

Part One

The Approach

2 The Nature and Behavior of Ecological Systems

Our perceptions determine the methods we use and the solutions we see. That is why puzzles fascinate and challenge, for their solution requires a shift of perception. Without that shift, the method for solution eludes us. Puzzles of ecological evaluation are the same. If present methods seem to be inadequate or even to magnify problems, perhaps the perceptions of the way ecological systems behave or are structured are partly at fault. Certainly, the different methods and approaches described in the following sections emerge directly from a very specific view of how such systems behave. It is important to make that view clear. At the least, by making our biases visible we make them testable.

Long before man appeared on the scene, natural systems were subjected to traumas and shocks imposed by drought, by flood, by geological changes. The systems that emerged are the ones that were able to absorb and adapt to these traumas and to their continual occurrence. Such systems hence are not fragile but are the creation of change. They are not, however, infinitely resilient. A forest can be turned into a desert, or a river into an open sewer. But to do so, man must often try very hard.

The evaluation of ecological policies is an attempt to assess how an ecological system will be affected by disturbances, both man-made and natural. Those disturbances may threaten survival, but they can, with care in design, enhance benefits. Examples of how ecological systems respond to shock and disturbance provide the core of our understanding of their structure and behavior.

Four properties determine how ecological systems respond to change and, as a consequence, how policies should be designed and how impacts should be assessed:

- The parts of an ecological system are connected to each other in a selective way that has implications for what should be measured.

- Events are not uniform over space, which has implications for how intense impacts will be and where they will occur.
- Sharp shifts in behavior are natural for many ecosystems. Traditional methods of monitoring or assessment can misinterpret these and make them seem unexpected or perverse.
- Variability, not constancy, is a feature of ecological systems that contributes to their persistence and to their self-monitoring and self-correcting capacities.

These will be discussed by example in the following sections.

THE ORGANIZATION OF ECOLOGICAL SYSTEMS

Everything is not strongly connected to everything else.

Smith and van den Bosch (1967) have prepared a particularly well-documented example of the response of a cotton ecosystem to disturbance. There is a series of valleys on the coast of Peru formed by streams running from the high Andes to the Pacific Ocean. Many of these valleys are under intensive agriculture and, because of the low rainfall, are irrigated. As a result, each valley is essentially a self-contained ecosystem isolated from the others by barren ridges. In one of these valleys, the Cañete, the crop was shifted from sugar cane to cotton during the 1920s. Over the years a group of seven native insects became significant cotton pests. The pest problem was essentially modest and the farmers of the region lived with the resulting economic damage. In 1949 chlorinated hydrocarbons like DDT, benzene hexachloride, and toxaphene became widely available, and the opportunity arose to dramatically decrease pest damage and increase crop yields.

The initial response to the insecticide treatment was a pronounced decline in pests and a 50 percent increase in cotton production. After two or three years, however, six new species of insects became as serious a problem as the original seven had been. The reason for the appearance of these new pests was the elimination of parasites and predators that were killed by the insecticides. Within six years the original seven insect pests began to develop resistance to the insecticide, and crop damage increased. In order to control this resurgence, the concentration of the insecticide had to be increased and the spraying interval reduced from two weeks to three days. As these control measures began to fail, the chlorinated hydrocarbons were replaced by organophosphates. But even with this change, the cotton yield plummeted to well below those realized before synthetic insecticides.

The average yield in 1956 was the lowest in more than a decade, and the costs of control were the highest: the agricultural economy was close to bankruptcy. This forced the development of a very sophisticated ecological control program that combined changed agricultural practices with the introduction and fostering of beneficial insects. Chemical control was minimized. These new practices allowed the re-establishment of the complexity of the food web, with the result that the number of species of pests was again reduced to a manageable level. Yields reached the highest level in the history of cotton production in the valley.

This example emphasizes the point already mentioned: many ecosystems are remarkably forgiving. The surprise is that such frequent application of a blanket of insecticide over an entire isolated valley did not have a more dramatic and destructive effect. But the effect that was triggered suggests the importance of the linkages within ecosystems. The complex of the original seven pests, the six induced pests, the cotton and other food sources, and the natural enemies represents a subassembly of the larger ecosystem of the valley. The insects are linked by various intensities of competition for different species of their food resource, part of which is shared, part specific. The parasites and predators establish further links and connections -- some connecting a single parasite with a single host, some connecting with several hosts. This ecosystem provides an example of a food web, through which energy flows and material is cycled.

But note that the connections are organized in a special way. Each species has a limited number of connections with others that give a distinct organization to the ecological system. This organization results in a unique capacity to absorb or funnel impacts.

Before we explore these capacities, however, we shall cite one additional example from our own experience that emphasizes the importance of simply knowing who is connected to whom. The large open-sea fishes of the North Sea, like herring and mackerel, have been nearly eliminated by fishing pressure. At the same time there has been an increase in the number of bottom fishes. At first thought the spatial separation of these two groups – one living in the upper waters, one living in the lower waters – would make such a response unexpected. But removal of herring and mackerel relaxed the competition with smaller open-sea fishes such as sand eels, Norway pout, and the young stages of the bottom-feeding fishes. Since these species, unlike herring, migrate between upper and lower regions, they provide a conduit that carries energy and material to fishes living near or on the bottom. With their herring competitors and predators removed, this conduit could carry more resources downward so that bottom-dwelling populations increased. Thus, it is the number and kinds of these links that can induce unexpected consequences.

But note that the simple thought (often expressed in species lists or the popular rhetoric of ecology) that everything is intimately connected to everything else is simply not true. One might have expected the removal of large pelagic fishes to have had effects on many other groups, especially their ecological neighbors, the pelagic invertebrates. To the contrary, the available energy appears to have been diverted through one specific channel to a relatively distant part of the food web.

The persistence of a species would be precarious indeed if its fate depended on every other species in the system. Analyses of studies such as those reported above suggest that ecosystems exhibit patterns of connections resulting in subassemblies that are tightly connected within themselves, but loosely connected to others.

Simon (1962) has shown that such structures have remarkable survival properties. First, removal of one subassembly does not necessarily destroy the whole. Because of the minimal connection between subassemblies, the others can persist, often long enough for self-recovery. Second, for the same reason, these structures rapidly adapt to change. As long as the same connections are maintained to other subassemblies, major changes and substitutions can take place within the subassembly. Species can substitute for other species as long as the same function or role is performed.

The conclusion for environmental assessment is that even qualitative measurements of structure are more important than measurements of numbers of every organism possible. The structure depends on who is connected to whom and how.

SPATIAL BEHAVIOR

Impacts are not gradually diluted over space.

Both the cotton and North Sea fisheries examples also demonstrate an important spatial property of ecological systems. One of the reasons the cotton mangement system eroded so rapidly was the application of insecticide over the whole of a self-contained, isolated ecosystem. Hence, no recovery from outside the system was possible to either slow or reverse the disruptions. The North Sea example emphasizes that events can be very different in different parts of space. The fishes and associated organisms in the upper waters are different from those in the lower. And yet they are uniquely coupled to each other. Moreover, if we were to look in greater detail, we would see a mosaic of spatial elements – of patches – that differ in their biological and physical characteristics. The parts of this mosaic are not totally isolated from each other but are linked by movement of material, energy, and some of the organisms; movement dictated by winds, by currents, or by active dispersal of organisms.

The consequence of this spatial mosaic and the linkages within it have been well demonstrated in a study by Huffaker (1958) in which he examined the interaction between populations of a plant-eating mite and a mite-eating predator. When there was unimpeded movement throughout the experimental universe (a homogeneous world, therefore), the system was unstable and the populations became extinct. When barriers were introduced to impede dispersal between parts of the universe, small-scale heterogeneity was introduced and the populations persisted. Thus, populations that began to collapse in one small area could be reinforced by invasion from other populations that happened to be at the peak of their numbers.

This view of spatial behavior is different from that implied in many ecological evaluations. The more usual assumptions concerning spatial effects are shown in Figure 2.1*a*: the greatest impacts are expected to be nearby, with decreasing effects as we move away from the location of the change. We call this assumption the "dilution of impacts" paradigm. Harmful physical effects (pollutants) are assumed

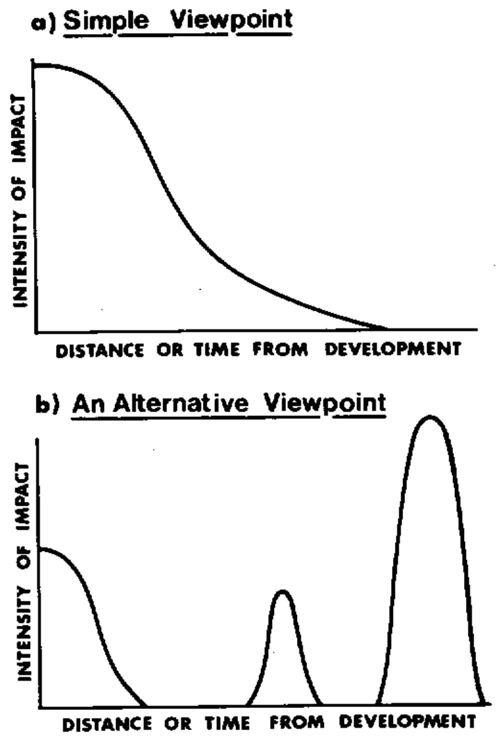


FIGURE 2.1 Alternative paradigms for the distribution of development impacts.

to diffuse in space, damages are assumed to repair themselves over distance, economic perturbations are assumed to be damped in a complex network of economic transactions, and so forth.

An alternative view is shown in Figure 2.1*b*. In this view impacts and problems are not related in any simple way to the location of the development. We would obviously not take this view seriously in dealing with many physical problems (though some pollutants can be concentrated to dangerous levels far from their source by biological and physical mechanisms). But it is not clear that the physical analogy holds in dealing with other subsystems. In particular, we argue that, within broad geographical and temporal limits, impacts mediated by social and economic processes need bear no obvious relation to the initial investment. For example, the local environmental impacts of a pipeline project in a developing region can usually be identified and ameliorated. But the induced effect of the invasion of capital and of construction workers on settlements remote from the pipeline can have dramatic social consequences that cause more significant environmental impacts than the pipeline itself.

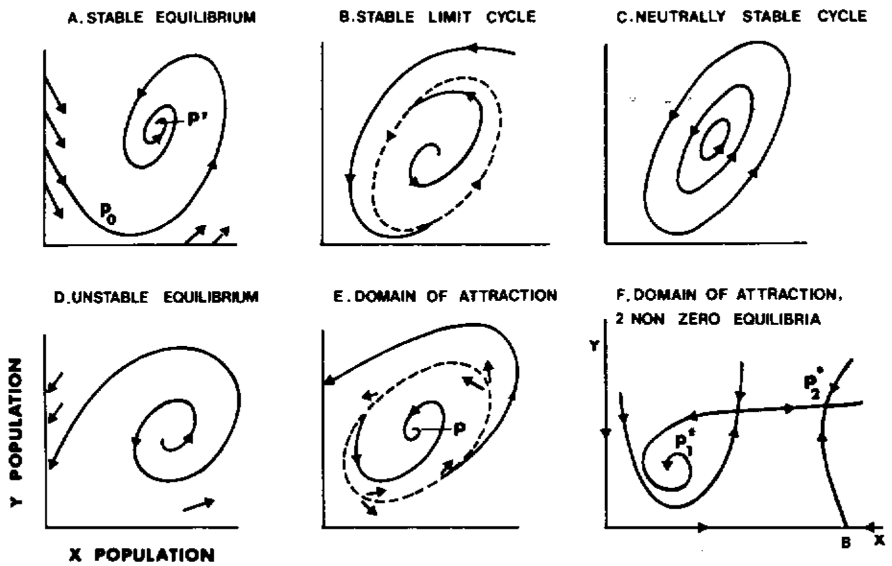


FIGURE 2.2 Ecosystem stability portraits. *A* to *E* are stylized and *F* is a specific example from Bazykin (1974).

STABILITY AND RESILIENCE

The unexpected can be expected.

Much of traditional ecological evaluation, policy design, and even ecological science itself implicitly or explicitly presumes that if a disturbance is removed, the system will ultimately return to its original condition. That is a view of an infinitely forgiving Mother Nature. But responses to disturbance can in fact take a number of different forms that can conveniently be represented by stylized portraits of stability (Figure 2.2). These representations are technically called phase portraits. The trajectories simply represent the moment-by-moment change in the value of two variables, given one starting point. The variables may be predator and prey, two competitors, or a herbivore and its food.

Consider the consequences in Figure 2.2*A*, Stable Equilibrium, for which the initial condition is at some point P_0 on the spiral. Given no intervention by man and no stochastic effects, the tendency of the system is to move inward along the spiral-like trajectory, taking steps of varying size in each successive time interval and in the limit approaching the equilibrium position (P^*). Stochastic influences derail the process, the size and direction of the random component usually being a function of location in the phase plane. But apart from these details, it is clear that systems characterized by case *A* will always migrate toward equilibrium. While long recovery times may be associated with larger displacements from an interior or quasi-equilibrium position in the phase plane, the fact of recovery itself is a certainty.

Case B, Stable Limit Cycle, demonstrates similar convergence. Any point in the plane converges to a closed loop that shows dynamic rather than static equilibrium. Any disturbance of the stable limit cycle produces ecological pressures that ultimately drive the system back to the cycle. If X and Y alone are plotted against time, the time series would show patterns characteristic of sustained oscillation; this is a typical consequence of simple predator–prey behavior.

The special property of case C, Neutrally Stable Cycle, is that any displacement results in a new, sustained oscillatory time series. This phenomenon has not yet been identified in real biological systems; indeed, there may not exist any such systems, but the case is included for completeness.

In Case D, Instability, every point leads to ultimate extinction. Recovery is possible only through reinvasions from other areas. This viewpoint reinforces the notion of the need for spatial heterogeneity as the only way to maintain persistence.

Cases E and F are of great interest for environmental management; case E is a general stylization and case F a specific example that will be discussed later. In case E, there is a closed region from which inward displacements converge on an equilibrium position or from which outward displacements diverge to some new domain of stability or to extinction of one (or more) species. Of course, small displacements will not necessarily result in these terminal positions, because movements in the phase plane contain random components that might push the trajectories across the boundary in either direction. It is useful to think of a domain of stability as a mesa. A particle moving on the mesa has a nonzero probability of falling off in one step, and the probability varies according to the location of that particle on the plane and according to the size of the step at any time. Once fallen, a particle can climb back onto the mesa and re-enter the domain of stability; the likelihood of such re-entry is smaller than that of falling off.

Another possibility is that the particle, having fallen, comes to reside on a new mesa. In biological terms, the system flips from one domain of stability to another.

If cases A and B can be viewed as Beneficent Nature, then case D is Ephemeral Nature and cases E and F are Mischievous Nature. In the last, the system will seem to be absorbing incremental disturbances but will then suddenly jump to another, unexpected mode of behavior. Such portraits are not simply mathematical curiosities. They find their counterparts in the behavior of the real world.

One of the more dramatic and extensively documented examples is the fisheries of the Great Lakes in North America. Data on catches exist from as far back as 1880, and a remarkably similar pattern has occurred in each of the seven most important commercial species in each of the five Great Lakes (Beeton, 1969; Christie, 1974). There was first an extensive period of sustained and modestly fluctuating catch. In a number of examples the catch suddenly increased briefly, but whether that happened or not, there was then a precipitous decline in catch over 2 to 3 years. In some instances the populations became extinct. In others the populations were driven to very low numbers. The populations were not held there by continued fishing pressure or the additional mortality from an introduced predator. Even when both fishing pressure and predators were reduced, populations did not return to their original levels; they persisted in this new configuration, this new equilibrium.

This is an example of a system that, in all likelihood, has at least two equilibria – one high and one low. If populations are displaced a small amount from either one of the equilibria, they will tend to return to it. But there is a limit to how great the displacement can be before the populations unexpectedly flip into the other equilibrium region. There are distinct stability regions and separations between them.

Even this picture of two separated regions of stability is oversimplified. The borders between equilibria of high and low densities are not simple "straight lines" determined only by the particular nature of the species in question. The unique relationships of a food web may allow a population to reach its high equilibrium by first being pushed to densities *below* its low equilibrium (Bazykin, 1974). For example, the phase portrait produced by one version of Bazykin's general predator–prey model is shown in Figure 2.2*F*. If Y is a predatory fish of commercial value and the system is at the equilibrium $P_1{}^*$, it might be desirable to shift the system to $P_2{}^*$, where there is a higher equilibrium. But note that addition of this fish would still keep the system in the stability region associated with the lower equilibrium. A modest reduction, however, can cause the variables to cross the stability boundary, and the system would naturally evolve to the higher equilibrium $P_2{}^*$.

The Great Lakes case is not an isolated example. Similar behaviors have been shown for a variety of fish populations in North America and Europe (summarized in Holling, 1973); grazing systems in North America, Africa, and Australia (Glendening, 1952; Noy-Meir, 1975); and insect pest populations in Asia, North America, and Europe (Sasaba and Kiritani, 1975; Jones, 1975; Southwood and Comins, 1976; Isaev and Khlebopros, 1977).

Larger assemblages of organisms demonstrate similar multi-equilibrium behavior. A history of herbicide spraying in a forested region in the United States (Niering and Goodwin, 1974) has succeeded in suppressing tree regeneration and growth to the point where shrubs so dominate the system that even after cessation of spraying the system remains a persistent and distinctive shrub community. Clearing large areas of tropical forests can similarly lead to an irreversible treeless condition because of exhaustion of the soil and leaching of nutrients coupled with the very low dispersal properties of tropical tree seeds (Gomez-Pompa *et al.*, 1972).

As a final example, Hutchinson (1970) has reconstructed the series of events occurring in a small crater lake in Italy from the last glacial period in the Alps (2000 to 1800 BC) to the present. Between the beginning of the record and Roman times the lake had established an equilibrium with a low level of productivity that persisted in spite of dramatic changes in surroundings from *Artemisia* steppe, through grassland, to fir and mixed oak forest. Then suddenly the whole aquatic system changed. This change towards eutrophication, or high productivity, seems to have been initiated by the construction of the Via Cassia about 171 BC, which caused a subtle change in the hydrographic regime.

We have dealt with this multiequilibrium behavior in so much detail because it lies at the heart of the uncertainty of ecological evaluation and design. A system can seem to be behaving according to one set of rules, until it suddenly flips into a radically different state. Incremental nutrient input to a lake may for a long time cause no noticeable change in water quality. But at some point, one additional increment may trigger the sudden appearance of eutrophic conditions. A fisheries system in the Great Lakes can seem to be yielding a constant and stable catch and yet be on the verge of precipitous collapse. A productive flood plain in the Orinoco delta can turn into an acid desert rather than intended agricultural land after draining exposes sulfur compounds in the soil to oxidation.

Just as there has traditionally been a "dilution of impacts" paradigm for impacts over space, so has there been a similar presumption for impacts over time. Impacts have often been presumed to be immediate and to be gradual. That implies that even if unpredicted, these changes can be monitored and detected in sufficient time to be remedied. It assumes that incremental approaches to planning and design, or marginal assumptions in cost–benefit analyses, or smooth discounting functions are all appropriate techniques of ecological policy design and evaluation. None of those assumptions holds in a world that has more than one equilibrium or stability region, where sharp rather than gradual changes can occur.

If we think of one variable affecting another only as an entry in an input–output table or a cross-impact matrix, we are implying a straight-line relationship, or at most, a smooth one. But many relationships have a form in which thresholds separate regions of no-effect from regions of effect, or where effects increase in one region and decrease in another. These nonlinear relationships contribute to the existence of multiple stability regions. They can turn the traditional tools of policy design and evaluation into the source of the problem, not the source of the solution.

DYNAMIC VARIABILITY

Environmental quality is not achieved by eliminating change.

One additional property remains. Ecological systems are not static but are in continual change – change in numbers, change in equilibrium conditions, change in species composition – and this dynamic change determines part of the structure, diversity, and viability of ecological systems. From a long-term perspective, the frequent droughts in the plains of East Africa are probably an integral feature that establishes the remarkable diversity of animals and plants. An argument can be made that the periodic destruction of trees by fire or elephants involves two dynamic forces that maintain a savannah rather than a forest. Certainly, a combination of fire and grazing can maintain grassland systems in temperate regions of the world. Similarly, many forest insect outbreaks, like those of the North American budworm (Chapter 11), are part of the natural cycle of renewal that maintains the resilience and diversity of forest systems.

Some of the changes are induced by internal mechanisms that actually force change independent of outside intrusions. For a time one group of species might gain ascendancy through competitive advantages, and their very abundance can release or trigger counteracting forces which reverse that process. Again, the budworm–forest ecosystem provides an example. Hence, for impact assessment as a review process, the impact of insecticides should be assessed not only in terms of direct ecosystem contamination. In addition, the reduction of the pest itself might significantly alter the renewal mechanisms of the forest, unless they are replaced by harvesting and silvicultural practices. And, for policy design, a forest mangement policy can be designed so that the pest itself becomes the forest manager at places and times where it is not economically feasible for man to do so. Ecological policy design can, by working with natural forces, turn them to economic benefit.

Other dynamic changes are caused by outside events – the erratic or periodic occurrences of flood, drought, cold, heat, fire, and storm. Natural systems are hence continually being "tested," and their adaptation to that experience affects their response to new intrusions. Some paleoecologists (e.g., Bretsky and Lorenz, 1969) have suggested that the species complex within intertidal communities has changed less than that in deeper water communities. The former are exposed to continual extremes through tidal movement; the latter experience a much less variable world because of stabilizing properties of water. Hence, when the inevitable unexpected occurs, the intertidal species can adapt while the deepwater species cannot. Watt (1968*a*) provides more rigorous support for this contention in his detailed statistical analysis of indices of abundance of 988 forest insect species throughout Canada from 1945 to 1965. Populations from regions with less variable maritime temperatures were affected more by a unit change in temperature than those from regions with highly variable conditions. And it is obvious that the impact of a rare frost on tropical vegetation or crops is a consequence of their evolution within a stable temperature region.

In a sense, therefore, the continual "testing" of these systems gives them the resilience they have. Their self-correcting responses to the unexpected exist because they are used occasionally. Hence, for impact assessment as a review process, the intensity of a disturbance by man cannot be assessed simply by its absolute magnitude. It must, at the least, be measured in terms of the degree of variability that has been historically experienced. And the corollary to that for policy design is that reduction of variability could lead to the gradual loss of resilience through relaxation of selective pressures. Placing a system in a straitjacket of constancy can cause fragility to evolve.

The traditional paradigm of ecological evaluation often is that the world is or should be designed to be static or constant. The developed countries in particular have recently experienced a growing emphasis on ecological and environmental concerns, in part as a reaction against past emphasis on growth and social and economic issues. But when that leads to a goal of ecological or environmental "purity" and constancy, it can no longer be labeled ecologically sound. Ecological systems are dirty, changing, growing, and declining. That is the source of their resilience and diversity. And, paradoxically, the developing world might be more capable of responding to the need for constructive variability because they themselves have been so subject to change and rapid adaptation.

These four properties – organized connection between parts, spatial heterogeneity, resilience, and dynamic variability – underlie all our attempts to develop and test the techniques described in the following sections. Several broad lessons emerge from these four properties:

1. Since everything is not intimately connected to everything else, there is no need to measure everything. There is a need, however, to determine the significant connections.
2. Structural features (size distribution, age, who connects to whom) are more important to measure than numbers.
3. Changes in one variable (e.g., a population) can have unexpected impacts on variables at the same place but several connections away.
4. Events at one place can re-emerge as impacts at distant places.
5. Monitoring of the wrong variable can seem to indicate no change even when drastic change is imminent.
6. Impacts are not necessarily immediate and gradual; they can appear abruptly some time after the event.
7. Variability of ecological systems, including occasional major disruptions, provides a kind of self-monitoring system that maintains resilience. Policies that reduce variability in space or time, even in an effort to improve environmental "quality," should always be questioned.
8. Many existing impact assessment methods (e.g., cost–benefit analysis, input–output, cross-impact matrices, linear models, discounting) assume none of the above occurs or, at least, that none is important.

The above lessons relate to the methods and data required for assessment and policy design. But there are, as well, lessons for the way environmental issues are incorporated within an institutional process.

THE BEHAVIOR OF INSTITUTIONS

The behavior of ecological systems is only one side of the equation. The other is the social and institutional environment. We have been careful, thus far, to concentrate only on the behavior of the ecological and environmental component. That is where our professional experience lies. But our recommendations are so contingent on the way people, as well as biological and physical systems, behave that our perceptions of the behavior of man and his institutions need to be at least briefly highlighted.

The key point is that our experience suggests that human systems have the same four properties that ecological systems have. These four properties lead to the same conclusion. First, agencies are strongly connected to a limited set of other constituencies (however bewildering the variety of signals). Second, some are near at hand, but some are distant (the centralization-versus-decentralization issue). Third, individuals, institutions, and societies have multiple stability regions (so that sudden shifts of behavior can become crises). Finally, dynamic variability is a benefit in maintaining an adaptive response to the expected (unless reminded by occasional change, people and institutions develop rigidity).

The last two properties are the ones that particularly color our recommendations for incorporating environmental analysis within the policy process at the very beginning and our recommendations for an adaptive process.

Earlier in this chapter we illustrated the alternative modes of stability by using stylized phase portraits (Figure 2.2). These can be generated by fairly simple coupled differential equations, which in no sense represent reality but rather are highly simplified caricatures of the essence of behavior. Bazykin (1974 and Figure 2.2) has done just that for ecological systems. The same approach has also been applied to institutional systems (Holling *et al.*, 1976) and societal ones (Häfele and Bürk, 1976). Just as in the ecological equations, the assumptions built into these simple caricatures generate separate regions of stability, regions, moreover, that can shift and change if parameters evolve through the action of cultural selective forces. In an early version of the Häfele and Bürk societal equations, for example, one stability region leads to high energy consumption per capita and low population, and the other to the reverse. A flip across the line of separation would seem for a time to be little different from the past, but the ultimate consequences would be radically different. Such equations should never be used as reasonable laboratory worlds for the development of explicit policies, but they are useful as perspectives, or metaphors of reality whose relevance depends on whether they match common sense and practical experience.

Certainly our experience with a number of institutions – management agencies, research laboratories, "think tanks," businesses, and universities – reinforces these metaphors (Holling and Goldberg, 1971; Walters, 1975*a*; Holling, 1976). Our conclusions are supported by more formal analyses as well (Cyêrt and March, 1963; Crozier, 1964; Etzioni, 1968). Those institutions that have developed policies that induced a rhythm of change, with periods of innovation followed by consolidation and back again, maintain a flexible and adaptive response. Expected problems and opportunities are detected and can be turned to benefits. Those institutions that have evolved toward stability, toward minimizing disturbance, toward being risk-averse, tend to react to problems and opportunities as crises. The adaptive response withers, and instead there is, paradoxically, an attempt to further reduce these uncomfortable intrusions. Options are rapidly foreclosed.

This matching of the metaphors and behavior of ecological systems with those of institutional systems leads us to specific recommendations for using environmental assessment as well for doing it. With this perspective, even the word assessment, implying passive reaction, is inappropriate. Adaptive environmental management is a more relevant term. Several broad recommendations emerge that will be reinforced in later chapters:

1. Environmental dimensions should be introduced at the very beginning of the development or policy design process and should be integrated as equal partners with the economic and social dimensions.

2. Thereafter, in the design phase, there should be periods of intense, focused innovation involving significant outside constituencies, followed by periods of stable consolidation.

3. Part of the design should include benefits attached to increasing information on unknown or partially known social, economic, and environmental effects. Information can be given a value just as jobs, income, and profit can.

4. Some of the experiments designed to produce information can be part of an integrated research plan, but others should be designed into the actual management activities. Managers as well as scientists learn from change.

5. An equally integral part of the design are the monitoring and remedial mechanisms. They should not simply be *post hoc* additions after implementation.

6. In the design of those mechanisms there should be a careful analysis of the economic trade-offs between structures and policies that presume that the unexpected can be designed into insignificance and less capital intensive mechanisms that monitor and ameliorate the unexpected (Holling and Clark, 1975). (That issue is explicitly addressed for the design of pollution control standards in Fiering and Holling, 1974).

7. The above points imply changes in institutions and legislation. We find, unexpectedly, that such changes seem more feasible in "less efficient" developing and developed countries. Whether intended or not, the unexpected has been part of their history, and adaptive change can be perceived as a modest shift from past experience.

3 Steps in the Process

As a relief from the philosophical, conceptual, and abstract discourse of the first two chapters, we turn now to a more concrete and pragmatic discussion of the steps that are involved in the process of adaptive assessment and management. Although we would ideally like to integrate assessment into management, we realize that this is not yet institutionally possible in many cases. Therefore, we treat them separately, first describing the major events and aims that are critical ingredients in environmental assessment. Even here we see two types, each with its own tactics: a long-term (1-year) assessment project and a quick (2-month) project. Many of the steps described also apply to an environmental management program. In addition, there are also steps needed for effective communication to, and implementation in, the responsible management agencies.

We must emphasize that this chapter is *not* intended to provide a "cookbook"; such a prescriptive device is the antithesis of the proposed adaptive management process. Rather, we hope this chapter provides readers with enough of a sense of the order of events that they can begin such an adaptive process on their own. Each situation will be different, however, and the steps described here should be molded to meet specific requirements in each case.

ENVIRONMENTAL ASSESSMENT

A 1-YEAR ASSESSMENT PROJECT

This section is written for the person charged with preparing an assessment of the environmental consequences of some proposed action. He is responsible for gathering together and coordinating a team to examine the problem, analyze the possible consequences, and prepare a report that will be used as an aid for decision. While we suggest a hypothetical timetable (Figure 3.1) for the tasks and events that

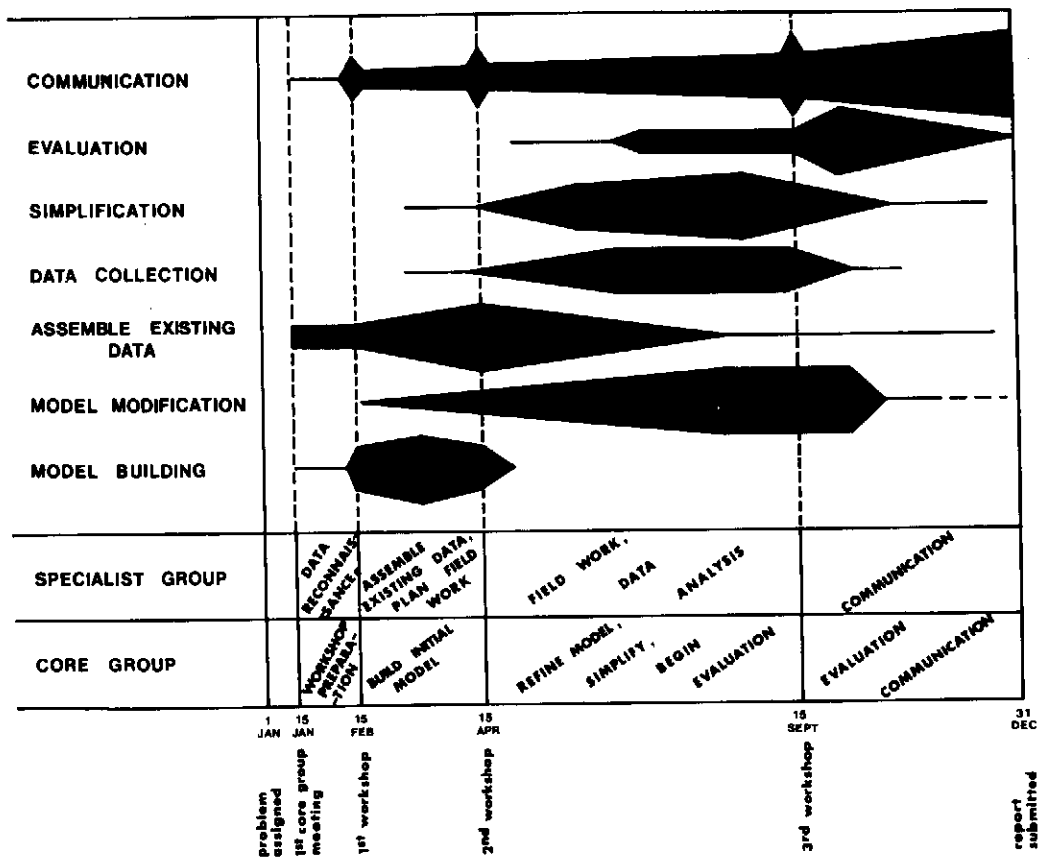


FIGURE 3.1 Activities and timetable for a 1-year assessment.

constitute the assessment, no two assessment problems are the same and they cannot be successfully treated with a fixed agenda. Therefore we have synthesized our experience into a "typical" scenario – flexibility and adaptability remain paramount. We have tested these procedures and are confident that they work. Specific procedures for operating the scheduled workshops are detailed in the next chapter.

January 1: The Assessment Begins

On January 1 the program manager is charged with preparing a report on the likely consequences of a major development. The report is to be completed within 1 year, and he may draw upon scientists and advisors both from his organization and from collaborating ones.

The program manager's first task is to identify the central members of his team. These fall into two groups, those who possess analytic skills (e.g., computer programming, data analysis, statistics) and the subject matter specialists, who might be biologists, geologists, economists, or engineers. The analytic group and one or two of the subject matter specialists will form what we call the *core group*. This group will run the workshops, do the computer modeling, and analyze alternative policies. The subject matter specialists outside of the core group will be called upon as their expertise is required. Workshops coordinate the activities of the core group with those of the specialists and methodologists.

January 15: First Meeting of Core Group

Before the entire team is assembled, the core group meets *in camera*, to outline the nature of the problem. This includes defining a range of management options, interest groups, and objectives. Additionally, and importantly, the core group should define the set of variables relevant to the decisions that must be made. At this meeting a first attempt is made to determine the physical boundaries of the problem, the temporal and spatial resolution required, and the level of detail the model should take. Other participants needed for the assessment groups are identified.

The products of this meeting are a list of participants for the first workshop, an understanding of the general form the model will take, and an assignment of responsibilities. The core group then begins to assemble the computer software and hardware for their modeling activities, and the specialists review the available data relevant to the problem.

The stage is now set for the first workshop. Although the core group has a preliminary definition of the problem, it is tactically important that these preliminary decisions remain invisible during the first workshop and that they be readily abandoned if it seems appropriate. In the workshop related decisions will be made again by all the workshop participants and will be modified as a consequence of the broader experience of the participants. It is important for these decisions to be made extemporaneously – and more important that they appear to be made so. The commitment of participants to the project in future workshops depends on their self-identification as creators of the model. However, it is also important that the first workshop establish momentum and that it does not become stalled over technical indecision. It is for this reason that the core group must have a set of "shadow decisions" in their back pocket to draw upon if the workshop falters.

February 15: First Workshop (2–3 Days)

This workshop is attended by the core group and all the specialists. In addition, it is critically important that the higher level decision makers and managers be involved as much as possible. Frequently, they will be able to attend only the first day, or even only the first hour, but it is of the utmost importance that they be there even for that hour, and at least two or three should attend the whole workshop. If the person who requested the report participates in the opening of the first workshop, he knows what is happening and feels a part of it. The ultimate decision makers can so guide the initial discussions as to ensure that the exercise remains relevant to their needs. A group of biologists left alone might produce a very interesting model of a game population, but one irrelevant to the management of that species. The presence of decision makers thus provides needed guidance in the early stages of the program.

This workshop follows the general rules described in the orchestration chapter (Chapter 4). The first days are concerned primarily with defining and bounding the problem, selecting the variables, and designing the framework of the model. Unless the core group is especially experienced, it is unlikely that they can have a rough model operating by the end of this workshop. The important point is that they have all the information and materials they will need to write the computer program before the participants leave. The core group must have the model structure defined for programming and must also have the estimates, however rough, of the parameter values for this model. The subject matter specialists must leave the meeting with a firm understanding of the data that are needed for further modification and refinement of a model that can be responsive to the management questions.

Three critical steps must be completed by the end of the workshop. First, the problem must be clearly defined – management actions, key variables, spatial extent and resolution, and time horizon and resolution. This definition should have led to at least a crude outline of a model. The core group will then use this information to develop, modify, and refine the model. Second, the key data needs must be defined, and preliminary research plans outlined by the specialists for the coming field season. Finally, the person requesting the assessment must have been so involved that he and the group are assured that the relevant information will be obtained. The more he is involved interactively in this critical 2 to 3 days, the more likely that this condition will be satisfied.

April 15: Second Workshop (2–3 Days)

By this time, two months later, the core group has a version of the model running on the computer. They have developed, as well, some alternative policies to the one proposed so that comparisons can be made. The specialists have obtained as much information as possible from the literature and have formulated their final research plans for the collection of the remaining data that are needed.

On the first day of this second workshop, the core group incorporates the specialists' data in the model and makes any necessary changes in the programming. Much of the technical work is done before the workshop, the actual meeting time being used to focus the activity and provide opportunity for communication. Once the changes are made and the data are incorporated, the model is ready to run. The workshop uses this running model to explore and test the suggested alternative policies and scenarios. Again, it is most useful to have the policymaker or manager present when policy options are being considered.

The last task of this workshop is to review each specialist's plans for data collection, thoroughly analyzing them to assure that the data are truly needed. Emerging from this meeting is a set of research plans for the specialists and a set of management options to be considered and tested rigorously by the core group.

The core group then begins the tasks of simplification, invalidation, and evaluation (see Chapters 6, 7 and 8). The model as it now stands is incomplete, since some major changes can be expected as a result of the specialists' field research, but the core group should start the analysis now. New data can be added when available, and in the meantime the analysis will help shape a better study.

September 15: Third Workshop (5 Days)

The first 2 days of this workshop are devoted to incorporating the revisions in data and model structure from the past 5 months of research. Again, this need not all be done within this workshop, as the core group will have begun this effort as data became available from the specialists. The final 3 days of the workshop are set aside for gaming with the model and evaluating alternative policies. A top policy person should be involved during these sessions. He can see the types of results generated and the direction that the final report will take.

The job of everyone involved for the remaining months of the year is communication. The core group must complete evaluation runs, produce informaton packages and graphs, and describe the likely outcome of options. Numerous demonstrations of the model should be made for the higher level administrators, as the final report constitutes only a part of the assessment output. The purpose of the entire program is to affect decision making, and all of the creativity of the team should be employed to that end.

December 31: Final Report Handed In

With the report finished, the 1-year task is now complete. The above schedule is fairly ambitious. As described, it involves 4 core group members and perhaps 15 specialists for 1 year. Frequently, these people would not work full time on this one project: the core group might have 3 or 4 similar simultaneous projects, and the specialists might devote half of their time or less to this project. Full-time commitments might, however, be appropriate for the analysis of a very large power generating station or transmission corridor, for example. For such projects the specialists might have several assistants who do much of the field work.

Lessons from the Guri Study

Of the five case studies reported in Part II, that of the Guri hydroelectric development (Chapter 14) comes closest to the intensive assessment scenario described above. The purpose of the study was to compare alternative forestry and agriculture practices in a $3 billion hydroelectric development, proposed for an undeveloped region of Venezuela. It was not, however, meant to be a comprehensive environmental study. The entire process of model building, evaluation, nomogram construction, and report writing required one coordinator for a year and twelve other participants for three months, full-time. This is considerably less than the 10- to 20-man-year program described above. No data collection was done in the field; all data were available from government maps, the scientific literature, and other commonly accessible forms of information. All computations were performed on a Hewlett-Packard 2000 (32,000 words); computers of this capability are commonly available in most cities around the world.

A SHORT-DURATION ASSESSMENT PROJECT

How can this workshop procedure be used if there are only 2 months instead of 12 to prepare the report? The first two workshops will have to be very close together, and there will be no chance for serious data collection or extensive evaluation. We have frequently been called upon to do a full assessment in 5 days, including model construction, alternatives definition, and policy evaluation.

The Obergurgl study (Chapter 13) serves as a prototype for such a short-term study. Its purpose was to examine the likely consequences of several options available for this high alpine region of Austria: zoning changes, building subsidy or taxation, ski-lift construction. In a 5-day workshop a model was built, and the alternative futures under the different options were examined. The results of this exercise became a topic of major consideration in the region, and we believe they made a significant impact on decision making. After a 1-day planning meeting, a core group of 5 methodologists and 15 participants met for a 1-week workshop. Some of these participants were specialists from the University of Innsbruck, some were regional government planners, and some were residents of the village itself. After the workshop, one person spent 2 weeks writing a report on the results. A PDP-11 computer (28,000-word memory) was used – again a computer of a size commonly available throughout the world. The investment in time and money was small, and the payoffs were great. This type of workshop could probably be used in many short-term evaluation programs; some parallel examples are outlined in Walters (1974).

Several important problems were defined and clarified by the Obergurgl model. The initial concerns about environmental quality receded to minor significance. Of more concern was the obvious inability of the village to maintain its current style of life, which is associated with continued growth of the hotel industry. The land will run out; subsidization, taxation, and zoning changes can only alter the date. When the Obergurglers returned to their village after the workshop, they initiated a series of public discussions about the future of the village. This period of discussion reached a peak during a 1-day presentation in the village of the results of the model by the modeling group. The need for a change in life style and expectations became obvious to many of the villagers; the search for a solution began. The model could not provide a solution, but the people can. They are now actively exploring means of expanding the economic base to provide nonhotel employment, and more important, the children who are now growing up are doing so with a better understanding of their future.

ENVIRONMENTAL MANAGEMENT

It is more difficult to prescribe a generalized sequence of steps for the process of designing policies for management. In many assessment situations the institutional authority, however narrow, is at least clear and undivided, and a useful sequence can therefore be generalized. Most environmental management situations, however, are much more complex. There is often a division of responsibilities for research from those for policy design and management. In such instances, as a consequence, the research often drifts from a focus on management and policy questions to a focus on general scientific questions. And those developing policies find themselves isolated from appropriate research information either because it was never obtained or because it is hidden behind institutional barriers. Moreover, in many problems of development or resource policy design a bewildering number of agencies seem to have, or desire, some voice. Finally, policy design, more than environmental assessment, must face the conflicting objectives of different governmental, industrial, and public interest groups.

Because these problems and the cast of actors concerned will be different in different situations, the best we can do now is attempt to identify the lessons we have learned from our various case studies. All our studies have contributed insights, but the budworm (Chapter 11) and salmon (Chapter 12) work, having gone farther toward introducing concrete change within agencies, have been the major learning experience. Both these case studies give the flavor of the institutional complexity that faced us.

In the broadest sense, the steps described above for the assessment process still apply. There is, however, greater explicit emphasis on designing a range of alternative policies and on involving a larger variety of institutions, role players, and constituencies in the actual design and evaluation. As a result it takes more time, more flexibility, and more adaptive response to opportunities as they emerge.

The major conclusions drawn from our efforts to implement the process and techniques within operating agencies follow:

1. Transfer of analysis, of the process, and of techniques means more than mailing the computer codes and writing a report. It also requires a program of workshops and intense "user" involvement so that the local scientists and managers end up as the real and acknowledged experts. A measure of success is the extent to which the original analysis group becomes less and less visible and the local groups more and more visible as the program moves into implementation. The initiators' very strong and markedly parental inclinations to keep control too long must be resisted, or transfer will fail.

2. Vigorous institutional support and protection is necessary but not sufficient; the policy design approach can be transferred only to people, not to departments. Respected local leadership of the program is essential.

3. The analysis must be made fully transparent and interactive. Hence extensive use of graphic presentations (Chapter 9) and an interactive computer environment are important to allow easy examination and modification of model assumptions. Cooperating scientists and managers can therefore explore their own experience and assumptions in the context of the models and so develop a critical understanding of the strengths, weaknesses, and limitations of the analysis.

4. Communication of the results must go beyond the traditional written forms. Modular slide–tape presentations describing the approach, the problem, and the model can communicate the essential features vividly and rapidly without compromising content (Chapter 9). In the budworm study, for example, a 4-minute motion picture of space–time dynamics under various management regimes better revealed that behavior than any amount of static discussion and analysis.

5. A sequence of participatory workshops beginning with scientists, proceeding to managers, and finally involving policymakers builds a foundation of confidence and understanding. A "top-down" sequence would, by contrast, force the technical analysis group into a premature position of prominence, alienating local experts and promoting little but suspicion.

6. The final – and perhaps the most restrictive – requirement of effective transfer is time. The budworm policy analysis *per se* took less than 6 months; the full program to implementation more than 3 years. Some of this time was spent in the workshops described above and in Chapter 4, but much was an incubation

period. A prerequisite for effective implementation seems to be time for the analysis group to appreciate the real options and constraints, time for the local managers and scientists to become truly conversant with new concepts, and time for the policy people to credit the analysis group with relevant intent. In retrospect, we doubt that the process could be rushed without fatally prejudicing the results in one way or another. Successful implementation requires patience.

Responsible policy choices by the decision maker are based on understanding and control of, not necessarily belief in, the technical analysis. If such understanding is not clearly communicated, if such control is not effectively transferred, then mere technique surreptitiously replaces political judgment as a basis for public policy decisions, with no accountability for the results. That would simply be the promulgation of another undesirable myth – the one Lewis Mumford has called the Myth of the Machine – in systems analytic disguise.

4 Orchestrating the Assessment

In Chapter 2 we discussed many characteristics of ecological systems that make them particularly difficult to understand and manage. In addition, it has become obvious in recent years that environmental management problems encompass biological, economic, and sociological factors, and that these must all be considered when evaluating development plans or when assessing alternative resource management options. The complex nature of environmental problems raises three questions of special concern to the resource manager or impact assessment team:

- How can the problem be bounded or delimited so that it is tractable and manageable?
- How can information and expertise that is scarce or widely dispersed best be applied to the problem?
- Finally, once the analysis is done, how can the complex results or recommendations be most effectively transferred to the decision makers and to the public?

CURRENT PRACTICE

Two major responses to the complex characteristics of environmental problems have emerged recently: the formalization of environmental impact assessment procedures and the creation of large interdisciplinary teams to tackle resource management problems. There is little argument about the need in assessment studies to call upon expertise from a number of disciplines. In most cases it has been deemed sufficient to establish a series of study tasks, or consulting contracts, with only minor provision for coordination in administrative matters, data gathering, and preparation of the final report. Statements are elicited from different specialists about the probable impact of a given development or management decision on their particular area of concern. Thus, a wildlife biologist might be consulted about the effects of a dam on big game animals, an economist about effects on recreation, a hydrologist about water flows, and a fisheries biologist about effects on fish. However, this approach often omits consideration of cross-disciplinary interactions, such as the effect of changing recreational demand on big game and fish populations (Walters, 1974).

In contrast, the interdisciplinary team approach exemplified by many recent research programs has attempted to promote communication among disciplines, which was lacking in the first alternative. Computer models are usually the focus of these team efforts, and because these teams involved many disciplines, the models are usually large and complex. However, it is now believed that the original goals of many of these team efforts were not met (Holcomb Research Institute, 1976; Mar, 1974; Mitchel *et al.*, 1976; Watt, 1977). The research was not significantly more integrated than in nonteam programs (Mitchell *et al.*, 1976), and models originally developed for research purposes were not necessarily appropriate for decision making (Holcomb Research Institute, 1976; Peterman, 1977a). In addition, the large number of people, large budgets (\$1–2 million/year) and long time frame for project completion (~ 5 years) created an environment where studies within disciplines became bogged down in details irrelevant to the management questions, where cross-disciplinary interactions were ignored, and where group activities drifted off in different directions (Ford Foundation, 1974; Holcomb Research Institute, 1976; and Mar, 1974). Moreover, the highly complex models that resulted from these large team efforts often defied understanding by either the modelers or the client decision makers (Lee, 1973; Holcomb Research Institute, 1976).

Both the interdisciplinary team approach and the formalization of the environmental assessment process were nobly motivated efforts, often expensive and experimental because they were so new. It is the history of that experience, of successes and of failures, that has led to a thread of tested concepts and techniques that deserve broader application. The failures were both expected and necessary; that is how we learn. Since the approaches have been admirably reviewed elsewhere (Ackerman *et al.*, 1974; Council on Environmental Quality, 1976; Dasmann *et al.*, 1973; Ford Foundation, 1974; Holcomb Research Institute, 1976; Lee, 1973; Mar, 1974; Mitchell *et al.*, 1976; O'Neill, 1975; Peterson, 1976; Schindler, 1976; Watt, 1977), we will only comment that these failures appear to have been consequences of inexperience in bridging the gaps between disciplines, data, techniques, knowledge, institutions, and people.

WORKSHOPS, THE CORE OF ADAPTIVE ASSESSMENT

In contrast to the individual-discipline or large-team approaches to environmental impact assessment and resource management, we have used an approach to bridging some of the above gaps that depends upon a small group of people that interacts with a wider set of experts during a series of short-term, intensive workshops. Most of our workshops have used the construction of a quantitative model as a focus for discussion, but as we will demonstrate later, many benefits will arise from workshops even if other predictive methods are substituted. Both the process and the product of these workshops are directly applicable to assessment and management problems.

Involvement of small teams and short time spans in these workshops circumvents the scientist's natural tendency to break problems down into components, and those components down into subcomponents, and so on. This tendency is a natural response to complexity and is deliberately encouraged in disciplinary training, especially in biology. But it is often not suitable for dealing with management concerns that are at a different level from those of the scientist (Mar, 1974) and that are likely to lie between usual areas of disciplinary interest and training. Instead, a small group of people working with a specific goal (model) in a well-structured atmosphere over a short period of time has advantages. Participants are forced to recognize that not all the components of biological or economic systems are of equal importance and that judgments will have to be made about the relative importance of the various pieces of the problem. Some details of workshops, such as size of group and budget, have already been discussed in Chapter 3.

From experience in more than two dozen cases (e.g., Himamowa, 1975; Clark *et al.*, 1977; Walters, 1974; Walters and Peterman, 1974; Walters *et al.*, 1974; Part II of this volume), we have found that small teams interacting through modeling workshops over a relatively short time can successfully carry out an assessment while addressing the three issues raised at the beginning of this section. Watt (1977) and Mitchell *et al.* (1976) have also concluded that small teams are most productive. However, success can be achieved only if appropriate people are involved at the various stages of analysis. The main participants are disciplinary specialists; methodologists who are familiar with techniques of analysis such as modeling; and decision makers who will ultimately use the information that results from the analysis.

There are obviously many environmental problems that cannot be solved without long-term studies by large research teams. But it is pointless and wasteful to initiate such studies without a clear and reliable strategy for insuring continued coordination and cooperation, particularly on issues that the individual specialists will tend to avoid. We suggest that modeling workshops can help to provide a brain for the body of the research team – they provide periodic reassessment and redirection.

We have used workshops in three ways during our studies of environmental problems. First, workshops are an effective way to begin a problem analysis, that is, to bring people together, to define the problem clearly, to examine existing data, to formulate some initial predictive scheme, and to identify future steps in the analysis. Second, workshops can form the backbone of a longer term, in-depth analysis in which alternative models or predictions are made and alternative management or development schemes are evaluated. Finally, workshops are a useful mode for transferring and implementing the results of the problem analysis to individual clients or agencies that did not participate in the assessment. While we will discuss the characteristics of all three types of workshops, we will concentrate on the most critical of these, the workshop that begins the problem analysis.

THE INITIAL WORKSHOP

THE WORKSHOP MODEL

We have found that it is critical to have the development of some sort of model predictions as an enunciated workshop objective. At this stage the model is not viewed as an end in itself; indeed, its predictions are usually not very precise. Rather, the model provides a focus for communication and a point of departure, allowing objective discussions of the importance of various components. The model is a device to promote objectivity and honesty. In interdisciplinary discussions that do not have such a focus, much time is wasted in general discussions of what is "important." When factors are brought into the open and quantified as part of a larger model, their importance can be judged by all the workshop participants. It should not come as a great surprise that many specialists find modeling workshops exceedingly painful: many of the "important" factors always turn out to be irrelevant for prediction.

Before describing the steps involved in a workshop, we must emphasize an important idea about simulation models: they should never be more detailed than is necessary to capture the essential behavior of the system being studied (see, for example, the spruce budworm case study described in Chapter 11). There are two reasons for this, one pragmatic and one technical. First, we wish the model to be as understandable as possible; a complex model may end up being as unfathomable as the real world and therefore unlikely to be understood by decision makers (Ackerman *et al.*, 1974; Holcomb Research Institute, 1976). Second, more detailed models do not necessarily result in greater predictive power. In fact, more complex models *may be* less reliable than simple ones (Lee, 1973; O'Neill, 1973): as one includes more detail (variables) in a model, the number of explicit assumptions made about interaction between those variables rises exponentially (imagine the implied interaction matrix). Therefore, the probability of making a *wrong and critical* assumption increases rapidly, and it has been found that the predictive power of a model usually declines after some level of detail has been exceeded. Unfortunately, there are no specific rules for how detailed a model should be; this judgment usually is a result of experience and intuition. Finally, we have found that breadth rather than depth is usually more appropriate for answering complex management questions of the sort that concern us here. Rather than concentrating on a few disciplines in great detail, models should include many disciplines (see also Watt, 1977).

From our experiences with models at many levels of detail, it is easy to look back at the field of ecological modeling as it was in the early 1970s and point out the difficulties inherent in the approach of building very large, detailed models of complex ecosystems. But at the time this approach seemed the obvious path to follow; computers were getting much bigger, faster, cheaper, and more accessible, and more data were becoming available. We have now gone through that unfortunate yet necessary phase in the development of ecological modeling that exactly parallels the trials with large models in atmospheric, water and urban modeling (Holcomb Research Institute, 1976; Lee, 1973). The approach we are proposing in this book incorporates many of the lessons learned from that experience.

PROBLEM ANALYSIS

Let us review the general steps of problem analysis to illustrate what is done and what the benefits are. First, an environmental problem arises, such as a proposed dam in a valley rich in wildlife or the extension of territorial claims on the ocean to 200 miles. One of the first steps in problem analysis is to recognize the institutional situation that governs the way decisions are made in the problem area at hand. It is best to choose that level of analysis that most closely fits the needs of an easily identifiable client (Mar, 1974). For example, it may make more sense to work on problems on an entire watershed than on those of subsections within the watershed if the planning commission or other decision-making body acts at the watershed level. Generally, it is possible to identify several levels of decision making within the client's responsibility, from broad and long term (investment strategies, facilities siting, and so on) to narrow and short term (construction tactics, remedial regulations, and the like), corresponding to levels in the organizational hierarchy. The problem analysis should state clearly which levels are to be addressed, and which are to be taken as given constraints or minor issues to be resolved as they arise in the field. However, as noted in the discussion of the myths of environmental management and assessment in Chapter 1, one should be very careful to look for impacts that may occur beyond jurisdictional boundaries.

Soon after the client and the problem have been defined, problem analysis should start by involving a small group of people in an early workshop to build an initial model. These people should include the required disciplinary specialists and a few of the decision makers and methodologists. It is best to involve decision makers at this point to ensure that management objectives are made clear and that appropriate management variables are considered. Early involvement of a few decision makers or administrators will also smooth the path for the specialists and methodologists. An assessment program is doomed to failure if administrators are not willing to invest sufficient people, facilities, money, and time in the project. To increase the chances that such an investment and commitment will be made, the decision makers should be given and should accept a role in shaping the course of the analysis through participation in one or several early workshops. Moreover, higher level administrators, along with other participants, should be provided with a series of payoffs during the course of evaluation (Holling and Chambers, 1973). The problem analysis can often result in substantial reordering of research priorities and identification of new data requirements, a benefit to researcher and administrator alike.

The first workshop for the specialists, administrators, and methodologists can take the form of one or two 3–5-day sessions whose goal is to produce a working first-approximation model that can be used for testing alternative management or development schemes. A common reaction to an early attempt to build a model is the feeling that not enough data are available. However, we have found that if useful data are ever going to be collected in a research program, some conceptual models must exist to guide the collection. In an attempt to quantify those conceptual models, the assumptions underlying them are brought out into the open and appropriate test data are more clearly defined. Thus, with a modest amount of basic survey information and knowledge of similar systems, the first workshop can begin.

The key element of this first workshop, as well as of subsequent ones, is the small core team, in our cases made up largely by people with some background in both the methodology (simulation modeling) and some resource discipline. This group integrates the information provided by specialists and managers. If and when subsequent workshops are conducted to deepen and broaden the analysis, this core group provides the continuity of experience needed to carry on the problem analysis. For those readers that have little experience with workshops of this type, we must emphasize that most of the art of conducting them is in dealing with people, not in facility with techniques. Holling and Chambers (1973) and Walters (1974) discuss some of the "people" lessons revealed through our own experiences, but the best and quickest way to learn modes of successful operation of workshops is to build a body of experience by conducting some. A full description of the steps we have taken in first workshops, those devoted to initial problem analysis, follows.

THE WORKSHOP PROCESS

First, some management goals need to be defined; even for a development scheme there must be some overall objective. Even if the decision makers present agree on an objective, a wide range of alternative objectives should still be considered so that the model can be responsive to possible future changes in objectives (Holling and Clark, 1975). By a *range of objectives*, we mean goals as extreme and as simple as maximizing economic return from a renewable resource versus preserving the natural state of that resource. While no one of these goals would be realistic, together they would cover a wide enough range that any real objective would fall somewhere within it (Clark *et al.*, 1977). The importance of an early statement of questions to be answered by the exercise cannot be overemphasized. As Brewer (1975) points out, too many models have been built with unclear program goals, resulting in too many inappropriate models.

Next, it is necessary to identify the variables, or indicators, that the client decision makers can use to judge how well alternative management actions meet given objectives. These indicators are really performance measures, such as level of employment, number of animals harvested, or kilowatts of electricity produced. As a consequence of the identification of objectives and indicators, the problem to be analyzed begins to be bounded. Further decisions have to be made concerning the range of management actions to consider, the temporal horizon and resolution, the spatial extent and resolution, and the ecosystem variables to be included. For example, should a salmon fisheries model consider a set of management actions ranging from building of enhancement (artificial propagation) facilities down to specific controls on insurance against bad times? Should the model consider only one small fishing area and the boat movements within it, or should it consider the whole coast and movement of boats between areas? Should the model explicitly consider all species of fish that potentially interact with salmon, or should only the major salmon species be accounted for? These questions are of the type that define the problem, and their answers are, in large part, determined by the management needs established earlier. A detailed example of problem definition in the spruce-budworm/forest-management case study can be found in Chapter 11. This first step of defining or bounding the problem through indicator identification is very critical; the rest of the analysis will in large part reflect decisions made at this early stage. Too narrow a conceptualization of the problem can eliminate from consideration a perfectly viable set of management options, or lead to predictions that overlook some key management concern.

One of the main purposes of the workshop is to promote interdisciplinary communication and to focus the scientist's expertise on the real management questions that the assessment is to address. To initiate communication, we have found it effective to use a process we call "looking outward." In the usual kind of impact assessment or management design program, each specialist is asked to predict how his own subsystem, such as the fish population or the vegetation, will behave. His natural tendency is to devise a detailed conceptual or numerical model consisting of many variables and relationships that reflect current scientific knowledge within his discipline. However, this conceptual model is usually more complex than is necessary to predict the behavior of a subsystem at the level of management indicators. Worse, each narrow conceptual model usually does not consider important links with other subsystems. In the "looking outward" approach we simply reverse the standard question asked of the specialist. Instead of asking "what is important to describe your subsystem *X*?" we ask "what do you need to know about all the other subsystems in order to predict how your subsystem *X* will behave?" Thus, the specialist is asked to look outward at the kinds of inputs that affect his subsystem.

After each subsystem has been subjected to this questioning process, each specialist possesses a list of "output" variables whose dynamics he has to describe so that these variables can serve as inputs to other disciplines. These cross-transfer variables that link the subsystems are essential in describing a picture of the overall system dynamics, and the modeling of each subsystem can be greatly simplified when the desired outputs from the subsystems are known precisely. For example, it may not be necessary to calculate changes in ten different classes of vegetation if the animals that utilize the habitat only distinguish between two classes of vegetation. Only after cross-transfer variables and variables needed to calculate management indicators are established should the specialist be permitted to add other variables that are of interest only to him.

The "looking outward" process, which is a modification of interaction matrix methods such as the Leopold matrix, is normally done by setting up an interaction table in which the system variables (deer population size, vegetation type and abundance, water level, and so on) are listed both down one side of the table and across the top. Then one asks for each element in the table, "Does the variable on the left in this row affect the variable in this column? If so, how?" In this way, cross-disciplinary information flows are identified. Systematic use of such an interaction table reduces the probability of leaving out some important interaction. During the "looking outward" process, there may be some disagreement about what variables or interactions should be omitted. Often, a bit of simple calculation can determine whether some detail is important to the final management indicators. If a decision cannot be made, then the disputed variable or relation can be held for later testing in the model as an alternative hypothesis to see if it makes any difference to predicted impacts (see Chapter 7).

Finally, some quantitative description needs to be made for each possible interaction identified in the "looking outward" table. Small subgroups of specialists can do this in a relatively short time by drawing upon existing information. Compared to the initial bounding and conceptualization steps, this step is generally surprisingly easy.

Finally, at the end of the first workshop, as submodels are quantified and interfaced, some validation and evaluation of management alternatives can be begun. This evaluation is the workshop product that is of most relevance to assessment (see Chapter 7 and 8).

BENEFITS

A number of benefits usually are realized from the first few steps of the workshop. Gaps in existing information are exposed, so future data collection programs, which are a major part of any assessment, can be more efficiently designed. The specialists get a better feeling for how their subsystem fits into the total system, and they gain an appreciation of the management questions. Similarly, managers learn of the importance of the various subsystems within the total management system. The need to clarify management goals and performance criteria is also established. Note that these benefits emerge even before a working model is produced and persist even if no credible model is built. Thus, this initial workshop can be valuable almost regardless of which predictive method is being used, and even if the time constraints on problem analysis are such that the first workshop is the only workshop. In such a case, which unfortunately occurs too often, the resulting model is probably the best synthesis of data and knowledge that can be produced over a short period. We therefore see a role for this first, intensive workshop both as a mechanism for making first-cut predictions that will then point the way for future study and as a means of making "best guess" predictions under severe time constraints. In addition, because of its nature and form, the workshop is an effective way to use scarce resources efficiently, be they data or people.

Because the process of putting together almost any kind of model, but particularly a quantitative one, results in recognition of new data needs, an assessment program or problem analysis can benefit significantly from a data-gathering program that is intimately tied to the modeling program. Often masses of data gathered before the synthesis begins turn out to be superfluous or irrelevant. It is for this reason that we suggest that modeling is more useful when it is done early in a program instead of as a final synthesis.

STEPS IN THE FIRST WORKSHOP

After holding several of these workshops, we have been able to compress all of the above steps into an intensive 5-day session. In this section we describe the sequence of steps by assuming that they will occur over 5 days, but we fully expect that initial workshop attempts by readers may stretch over two weeks or more. Nevertheless, the order and relative length of the steps should still be the same.

The first day is devoted to clarification of the problem, conceptualization, and definition of indicators and state variables. During the second day, interactions between variables are generally listed, and responsibilities of subgroups (those dealing with particular sections of the overall system) are laid out. Then four or five subgroups begin to define the interactions that need to be considered and data (which participants have brought with them) are applied in these submodels. On the third day, subgroup meetings continue, and subgroup coordinators begin to program and test submodels. Late on the fourth day the submodels, with luck, can be integrated. Serious debugging, validation, and policy evaluation can begin on the last day. Clearly, a special kind of leader is needed for such workshops. He must be someone with broad perspective on the problem, who is willing to make bold assumptions and move onward when proceedings bog down and who can channel trivial arguments into useful directions. Except for this individual, requirements for expertise and facilities for such an undertaking are not great, as was discussed in Chapter 3.

Two logistical details help to make workshops successful. First, they should be held at a neutral location where everyone is removed from his normal responsibilities and other distractions. Second, it is important that participants have the opportunity to run through some of the analyses themselves. For example, computer terminals that permit individuals to ask "what happens if . . ." questions of the model can be extremely beneficial in making model assumptions and limitations clear, in suggesting further refinements, and in revising performance criteria. Only modest investment in computer software and hardware is needed to create this important "hands-on" gaming capability (see Chapter 3 again).

SECOND-PHASE WORKSHOPS

The kind of workshop just described serves to start a problem analysis. The resulting model is clearly incomplete, and further efforts may be required to clarify data needs. The next phase of analysis can involve additional workshops, the number depending on the problem being studied. These workshops aim to revise the model and define new information needs, particularly as new data become available. In some cases a credible process of evaluation can be completed with only two workshops, held several months apart; other cases may require a series of workshops that are held over a year or two. The same mix of people, though not necessarily the same individuals, should participate in these later workshops: methodologists, specialists, and decision makers. The time between workshops is spent in data collection, model testing, and evaluation of management policies (Chapters 7 and 8), the last two activities largely being carried out by the small core team.

Again, the second phase of workshops can be equally valuable, whether participants are operating in an active, integrated policy design mode or making a relatively independent assessment of proposed policies. The value derives from the more careful focusing on critical issues, data needs, and questions. Some of these second-phase workshops were illustrated in Chapter 3.

TRANSFER WORKSHOPS

Finally, as the analysis or assessment nears completion, the phase of transfer to the contracting agency or other clients who were not involved during problem analysis begins. Here again workshops have proved valuable (Gross *et al.*, 1973; Clark *et al.*, 1977; Peterman, 1977a) in both an impact assessment setting and a resource management program. When the model is used as a focus for discussion, the assumptions underlying the analysis are clarified and the "client" decision makers can ask various questions of the model through interactive gaming. This so-called "implementation" phase is quite critical; without a smooth transition, even the best analyses are incomplete. Thus, attention must be given to the best ways of communicating the information. Chapter 9, on communication, illustrates some of the most effective ways we have found to transfer information.

5 Choosing a Technique

There are a great many analytic techniques and modeling styles, and the environmental assessment team must choose among them. The choice is important: the factors considered, the scope of the evaluation, and the eventual credibility and usefulness of the effort are tied closely to the techniques chosen. However, the choice is not immutable. Adaptive modeling contributes to adaptive assessment and management, and therefore we expect that the number and nature of techniques employed and of models constructed will grow, evolve, and shift as the analysis progresses and as understanding emerges.

Many of the chapters in this book call for the comparison of alternatives: alternative objectives, alternative developments, alternative models. Equally, alternative analytical and predictive techniques should be mobilized – each chosen for its usefulness and appropriateness for some particular aspect of the study. In this chapter we shall offer our views of the strengths and weaknesses of several of the techniques that we have utilized in our own environmental assessment and resource management problems.

The choice of technique follows from the nature of the problem at hand. The scope of that problem demands a complementary capacity in the tools used to address it. At the same time, however, the limitations of available data and information constrain and modify the selection of techniques and the means by which the assessment proceeds. All too often, it is the technique that grabs the lead, and the problem is then bent and redefined to suit. Every analyst or consultant has his favorite methods for solving problems, and it is only natural for him to advocate their use. The authors of this book lean heavily toward simulation modeling, but we feel it very important to maintain as much breadth and flexibility in our methods as possible in order to be responsive to a wide range of environmental and management problems.

To emphasize the importance of putting the nature of the problem ahead of technique, we first compare and classify nine of the major case study problems with which one or more of us has been involved. Some of these are described in detail as the supporting case studies of this book (Part II). Other problems are introduced here to enlarge the present discussion.

These nine problems cover three broad types of environmental concern. The first type of problem concentrates on the social and economic system and focuses on the dynamics of human behavior and associated economic causes and effects. For the most part ecological phenomena are not treated explicitly but are handled by transforming the socioeconomic variables into indicators of environmental effects. The problems of this type that we consider here are

Obergurgl. A study of land use development in a high-alpine Austrian village. The conflict between resort development and farming in the face of an expanding population is a central issue (see Chapter 13).

GIRLS (Gulf Islands Recreational Land Simulator). A study of land use and development in the Gulf Islands of western Canada. A strong emphasis is placed on the effects of speculation and perceived quality on the real estate market (Chambers, 1971; Holling, 1969).

Georgia Strait. A study of the interaction and conflicts between recreational sport fishing and the commercial harvest of salmon in British Columbia's Strait of Georgia.

The second type of problem concerns large-scale resource development projects. These problems call for an exploration of the dynamics of the environmental changes that will result from extensive interventions. Typically, many biological species and habitats are considered, but the socioeconomic system is not treated in depth. Problems of this type include

James Bay. A study of a large (440,000 km^2) hydroelectric development in the Canadian subarctic. Wildlife preservation and native Indian welfare are two major facets considered (Walters, 1974; Munn, 1975).

Guri. A study of an extensive regional development program in connection with a hydroelectric project in the Orinoco River basin in Venezuela (see Chapter 14).

Oil Shale. A study of the impact of oil-shale mining and exploitation on wildlife communities in the western United States (see Chapter 15).

The third type of environmental management problem concerns the population dynamics of a few species. Typically, only the dominant species of interest and its immediate prey and predators are considered. This is true whether the central population is a harvestable resource, a pest, or an endangered species. The dynamics of the socioeconomic system in which the biology is embedded are not treated explicitly: rather, the ecological variables are translated into the appropriate social and economic indicators for management decisions. We consider in this chapter the following three studies as problems of this third type:

Budworm. A study of forest management in the face of a major insect pest, the spruce budworm. This study focuses on the design of ecological policies for the Canadian province of New Brunswick (see Chapter 11).

Caribou. A study of the population dynamics of caribou herds in northern Canada (Walters *et al.*, 1975).

Capybara. A study of the capybara, a large and commercially important rodent, in Venezuela.

These nine sample problems of resource management and environmental assessment are also useful because they represent a broad range of variation in many characteristics besides the three problem types under which they were presented. In the next section we develop a classification scheme to organize our perceptions of the important aspects of any problem. We propose three broad measures that, for all our case studies, characterize the challenges to, and opportunities for, creative and adaptive management. If we think of these as three axes of a graph, it is possible to locate the nine case studies, and others, on the graph (see Figure 5.1). The three axes of this problem classification scheme are

- The common, though usually subjective, measure of problem *complexity*. This complexity comes from several sources, which we describe in the next section.
- The amount and quality of *data* available. Of course, the amount of relevant and usable data may be a small fraction of the total.
- The degree of conceptual *understanding* we have of the inner workings of the system in question. This understanding reflects our ability to identify and analyze the causal relationships of the principal ecological and social processes involved.

When we organize our perceptions of a problem's characteristics along the three axes of this classification scheme, we are in fact characterizing the model that will be used to analyze the problem. The way that the model is conceived and constructed depends on whether the problem is complex or simple, has many or few data, or involves processes of which there is considerable or little background understanding. How the model, or other analytic technique, relates to the problem will be clearer after we locate the nine sample case studies according to the classification criteria and then consider what modeling technique was used in each of these cases.

In the third section of this chapter we move from a general classification of the whole problem along the three axes – complexity, data, and understanding – and begin to consider how the problem analysis can be addressed with the analytic techniques available. Operationally, of course, headway can best be made by dealing with submodels of individual ecological or social processes, rather than by treating the entire problem in one lump. Each of these constituent processes will have its own location along the complexity, data, and understanding axes and thus will have its own requirements for analytic technique.

The various mathematical assessment and analysis techniques can be thought of as sitting on a continuum that stretches from highly qualitative to highly quantitative. On the qualitative end would be such non-numeric procedures as species checklists and cross-impact matrices, while on the quantitative end we place detailed simulation models and other more analytic procedures, such as formal optimization methods.

When we examined the mathematical techniques we have used, we found we had no modeling techniques that could address incompletely specified problems – systems that had few available data and that were poorly understood. One candidate technique for filling this gap we call "qualitative simulation." In the fourth section of this chapter we describe a modest effort to explore the effectiveness of such qualitative simulations when applied to problems with various amounts of data. This exploration served primarily as self-education, and we present as its principal product a list of the major lessons learned.

COMPLEXITY, DATA, AND UNDERSTANDING

The classification presented in this section highlights some of the sources of complexity in a problem analysis and points to ways to minimize and organize that complexity. Additionally, much attention is given to the distinction between quantities of data and extent of understanding. These two are often confused and interchanged. However, the type of analysis employed is very much affected by the mix of these aspects. Specifically, we show that one can proceed farther than is normally thought possible in the face of meager data by mobilizing available insight into the system's constituent processes. As an illustration we shall take one of the case studies and examine some of its processes and how they are analyzed from the viewpoint of this classification.

COMPLEXITY

Complexity is a relative concept at best, and in the world of modeling it has been used to mean so many different things that it no longer conveys much information. We can explicitly list some of the attributes contributing to complexity, but whether the whole model is called simple or complex remains a matter of opinion.

A quantitative measure of complexity has several parts. Perhaps the most obvious is the number of variables required to describe adequately the dynamic conditions of our system at any moment. Typical variables used in our models include the number of spawning salmon, the flow rate of a river, or the fraction of available capital that Obergurglers hold in their savings accounts. In the budworm case study one variable is the number of insects, two other variables keep track of the amount and condition of the foliage, another represents the weather, and seventy-five variables account for the number of trees in seventy-five single-year age classes. We view a model with 79 variables as modestly large, but, in this case, the fact that 75 of these variables have nearly equivalent functions somewhat reduces the effective complexity.

Most environmental and ecological problems are not contained in a single location, and it is often necessary to disaggregate a model into several spatial areas. In hydroelectric developments, large areas are involved, and separate impoundments must often be treated as explicit units; the Obergurgl village/farm/ski-resort region is subdivided into ten spatial units. In the budworm study the tremendous dispersal capabilities of the moth and the operational needs of the forest managers require modeling 265 separate land areas. When the 79 variables from one area are replicated 265 times, we suddenly have 20,935 state variables! Spatial disaggregation results in an explosive increase in the state variable count.

A third component of model complexity is the number of different management acts being considered. These acts represent the interface between man's intended activities and the subsequent alterations in the environment. Again in a hydroelectric development, the construction of a dam of a certain size at a certain place in the watershed is an act. Complexity arises when the variety of ways to design a network of dams and the variety of possible construction sequences are considered. In the budworm study the available acts are "cut trees, plant trees, or kill insects." Even here, however, one must ask: Cut trees of what age? Kill budworm at what life stage and at what time in their outbreak cycle?

Acts are man's inputs to the system, and various social, economic, and environmental indicators are the outputs. These output indicators are a fourth component contributing to model complexity. The natural system may operate according to state variables, but the people who are concerned with, or who manage, resource and environmental problems respond to other measures of performance. Winter tourists in Obergurgl may respond to crowded ski slopes, while those who come in summer may object to roads, clearings, and pylons obscuring the alpine vistas. A small sample of the indicators generated for the budworm study is given in Table 8.1 of Chapter 8. These include the costs and profits to the logging industry, the volume of wood "in reserve" as young trees, and the number of high-quality recreational areas.

A final component of complexity concerns the way time is handled in the model. Often a simple, uniform time step is adequate. During one time period (a year, say) all current variable values interact to create new values for the next time period. In the budworm study we had the happy congruence of a once-a-year insect generation and a yearly management operating period. In other cases processes operate on different time scales, time lags between events occur, or the dynamics of some variable depend conditionally on variable values from previous time periods. Such mixed-time-period dynamics contribute to a model's complexity.

TABLE 5.1. Components of Complexity for Nine Sample Environmental Case Studies

Case Study	Number of State Variables	Number of Spatial Units	Number of Management Acts	Extent of Socioeconomic Impacts Considered	Time Resolution
Obergurgl	Many	Few	Moderate	High	Simple
GIRLS	Many	Few	Many	Moderate	Simple
Georgia Strait	Moderate	Very few	Few	Moderate	Simple
James Bay	Many	Moderate	Many	High	Simple
Guri	Few	Moderate	Few	High	Complex
Oil shale	Very many	Very many	Many	Moderate	Simple
Budworm	Many	Many	Few	Moderate	Simple
Caribou	Few	Very few	Few	Low	Simple
Capybara	Few	Very few	Few	Low	Moderate

These five components start to describe complexity, even if they do not define it. The important point to remember is that the total complexity is not the sum of these components, but rather the product. The benefits of parsimony at any stage are multiplied in the final product. Even so, the final working management model may still be too complex to allow useful interpretation. If the model appears to be nearly as complex as the real world, it will be difficult to achieve creative assessment and management. In the next chapter we describe some steps to cut through the remaining complexity of the working model and to reach a level of simplification for improved understanding and interpretation.

To make this discussion of complexity more concrete, in Table 5.1 we subjectively score our nine sample problems for each of the five components. These nine particular case problems were selected to illustrate a wide range of variation among these components of complexity. The Obergurgl, Guri, oil shale, and budworm studies are documented in Part II; the others can be visualized in relation to these. The numbers of state variables and spatial units are not given precisely because the model may exist in several adaptive versions of different size, the number of state variables may differ between spatial units, or the spatial disaggregation can be changed by the model user. From this table we see that Capybara and Georgia Strait are the least complex while Oil Shale, Budworm, and James Bay are the most complex.

DATA

The second axis of our problem classification scheme represents the amount of data that can be brought to bear on the problem. Some data are required for the calculation of the parameters in the descriptive functions of the model. Assignment of numbers to these parameters is what actually makes a model quantitative. Some data are needed for invalidation – the process of establishing a "degree of belief" in a model. This is done through an active search for comparisons of model and real-world behavior that show where the model is wrong, not where it is right (Chapter 7). Ordinarily, the time behavior of only a few of the state variables is known. Because the duration of a dynamic system depends on its starting conditions – different starting conditions lead to different outcomes – we need data that give a complete description of all variables at some specific moment. Without this, any direct comparison between real and simulated history is hampered by an extra burden of ambiguity.

The data need not all have been procured as part of the resource development program. Many usable data, for example, may have been gathered incidentally or may concern similar situations.

Sheer volume of data is not necessarily helpful in and of itself. Too many of the data normally collected prove to be utterly useless for constructing a management model, even when the data are scientifically sound. What science and scientists emphasize often bears little relation to what is needed for establishing environmental policy. And even research that is undertaken for management will surely end up with information missing if the research is not organized with at least a hypothetical management model in mind. It is for this reason that we advocate model-building workshops at the very early stages of a project. The benefits in organizing the research and identifying problems that would have been overlooked make the effort worthwhile.

The models associated with the nine case examples in Table 5.1 were built from a wide range of data bases. One reason that the budworm was selected as a case study for the development of ecological policy design techniques was its rich research foundation – both intensive and extensive. Few ecological systems have been studied as much. Detailed life history studies of budworm had been made; significant information was available about such biological processes as parasitism, reproduction, the effects of foliage condition on survival of trees and budworm, and the effects of insecticides on the target species. Additionally, population estimates had been made for over 25 years at many locations in a 50,000 km^2 area.

For the oil shale problem, a broad range of data was available, most of which were not as statistically sound as those available for the budworm study. There was some information on many species but very little information on the relationship between species and between other ecological factors. In Obergurgl a surprisingly large amount of data could be extracted from the village records: birth and death records were used to build a very reliable demographic model; other records established patterns between economic profiles of groups and investments in savings accounts and hotel construction. For Guri, on the other hand, there were virtually no data other than those pertaining to the strict engineering specifications and basic hydrology.

UNDERSTANDING

On the final axis of our classification scheme is the extent of basic understanding we have of the processes that underlie the behavior of the systems. This information can be derived from a growing literature of laboratory and field experimental research: with it, we can know in advance the necessary and sufficient attributes that characterize a particular process. Without this prior knowledge of form, we would require a great many observations, over a range of variation, to establish a functional representation. However, as soon as we know that a particular mathematical function will describe a process, the information requirements are suddenly reduced greatly. Now we need only estimate values for the few parameters of that function. In some cases parameters will have a strict physical or biological interpretation that makes their evaluation direct.

When faced with the problem of sending a spacecraft from the earth to the moon, the "managers" know and use the equation describing gravitation and other well-developed laws of physics. Parameters must still be set, such as the mass and location of the moon and the configuration of the craft, but these are specific parameters for known functional relationships. Here, the known and understood processes of gravitation and thrust reaction are the core of the controlling "management model."

Many ecological problems can be treated in an analogous fashion. Rather than using arbitrary relationships between variables – such as those provided by statistical regressions – we can mobilize a substantial body of theoretical and experimental work and place the representations of relationships on a firmer foundation. Predation is one ecological process that is particularly well documented. It is now possible to take a predation equation "off the shelf" and use it in a model. An example of this is discussed later in this chapter and in Chapter 11, on the budworm case study.

Of the nine case examples, Budworm and Caribou had the most supporting knowledge of the constituent processes. Human social phenomena as found in the Obergurgl and GIRLS studies were not so well understood, and in the oil shale problem there was insufficient knowledge, even of which variables were connected to which, so that the potential of using process understanding could not be realized.

CLASSIFYING OUR EXAMPLES

We can make a loose, subjective placement of our nine examples within the dimensions of complexity, data, and understanding (Figure 5.1). The variation among these nine studies is evident in the figure. The models and other analytic procedures applied to each of these studies can in some measure be determined by the location of the study in this figure. The nature of the problem – whether it is a socio-economic question, a resource development project, or a population dynamics problem – does not influence the style of analysis nearly as much as does its location in this classification.

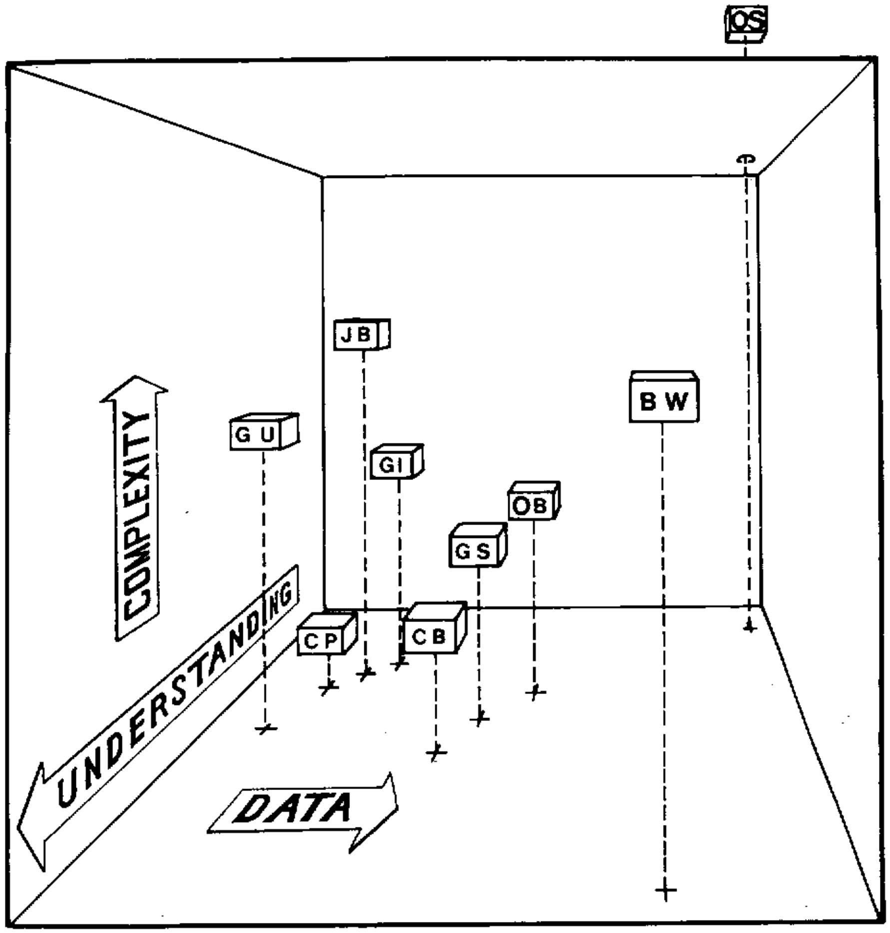


FIGURE 5.1 The location of nine sample case studies of environmental assessment and management in a problem classification scheme measuring degree of complexity, amount of available data, and degree of background conceptual understanding. The case studies are OB: Obergurgl; GI: GIRLS; GS: Georgia Strait; JB: James Bay; GU: Guri; OS: Oil Shale; BW: Budworm; CB: Caribou; CP: Capybara.

For example, the oil shale problem is isolated in the high-complexity/high-data/low-understanding corner of Figure 5.1. This problem was also treated very differently from the others (as can be seen in Chapter 15). The budworm study is also at the periphery of this constellation, being rated high for each of the three measures; this accounts to some extent for the relatively advanced development of the budworm case study. It also accounts for the ubiquitous appearance of the budworm to illustrate points in this book. A larger number of lessons have been learned through the challenges and opportunities afforded by the available data, the prior understanding, and the inherent complexity of this system.

Clearly, any management problem has some parts for which there are sufficient data and others for which there are not; some parts whose processes we know from other sources and some not; and some parts that can adequately be described by a simple function and others that require more elaborate mathematics. It is precisely this that had led us to utilize simulation modeling as a technique for assessment and analysis. With simulation models we have the flexibility to program a wide variety of functions and relationships and thus make full use of the knowledge we do have. Simulation model construction also helps us identify those areas where information is scarce and needed.

Placing an entire model on a chart such as Figure 5.1 requires subjective aggregation of all the parts – the strong with the weak. In the following section we shall look in more detail at the parts of one of these studies – the budworm study – to see how their location in this classification affects the way they were treated.

MODELING THE PROCESSES

An effective management model requires an explicit causal structure in its formulation. The quest for realism, however, should not lead to the inclusion of excessive detail. The challenge is to restrict what is included to the minimum, while still retaining an accurate and "workable" representation of the key phenomena.

A model that accurately describes the ultimate behavior of the variables is not enough. Almost any arbitrary model, given enough parameters to tune, can be made to match a set of historical observations. This is, of course, the essence of regression-type models and other forms of analysis whose structure is determined not by the problem but by extrinsic motivations – such as the desire for mathematical tractability. Any useful environmental or resource management model must be able to respond to unique changes and unprecedented perturbations that alter the system's conditions. New management acts will cause the system to move into new regimes of behavior; the model, to be useful, should be responsive to these same shifts. If the model has an appropriate causal structure, it will respond to these new conditions more faithfully.

There will always be some uncertainty about a model's flexibility in responding to novel conditions. Whole new mechanisms may enter the picture, or elements that were excluded from the original model may become important in unexpected ways. But this will always be the case, no matter what form the analysis takes. If the model has a logical causal structure, new items can be easily incorporated as they are discovered. This is an important aspect that makes the modeling procedure part of the entire adaptive process.

The most direct way we have found for ensuring a causal model structure is to focus on the level of the constituent processes. These are the operating subdivisions that link the variables of a system. Looking at processes also has the advantage of capitalizing on generality – because processes extend across many situations, we can draw upon the knowledge and understanding gained from other cases and other research.

The examples of processes in this chapter are primarily ecological and are illustrated by the budworm. However, the other cases are formulated in a similar manner. In Obergurgl there is a market process relating tourist demand to hotel and ski-lift construction, as well as an inverse process relating existing facilities to demand. Hydrological developments such as James Bay and Guri involve such processes as stream scouring and erosion. The Georgia Strait study must consider how the commercial catch affects fleet investment, as well as the effect of angling success on sport fishing activity.

Ecological processes include such things as growth, reproduction, competition, predation, and natural selection. Such "natural" processes exist across a very wide variety of situations. The ecological processes are very like those that a meteorologist would list: advection, convection, evaporation, and the like. The analogy is worth pursuing, for the meteorologist seeking to explain or predict a given pattern of weather does not start each study *de novo*. Rather, he makes extensive use of the discipline's existing stock of well-tested process theories, parameterizing and combining them in modular fashion as each specific situation demands. The individual modules provide an *a priori* structure for interpretation of the data, can often be individually tested, and inevitably highlight the weak or missing aspects of the analysis.

THE BUDWORM PROCESSES

The major processes in the biological phase of the budworm study are shown in Figure 5.2. These processes represent the important phenomena that affect budworm population growth, forest development, and the interaction of the two. Details of these processes can be found in Chapter 11 and in Yorque *et al.*, 1978).

In Figure 5.2, we locate the individual processes of this study upon the axes of data availability and conceptual understanding. We do this to emphasize the range of variation that is inherent in any environmental study. To develop a management model, all the parts necessary for a holistic picture must be included. The scattering of the parts on the data/understanding plane to a variety of challenges and approaches for any study. The axes of Figure 5.2 are in many ways complementary – a low value on one can be compensated for by a high value on the other. Too often, however, amount of data is assumed to equal amount of understanding. In traditional environmental assessment work and in some large ecological modeling projects, data acquisition becomes an end in itself, and there is too little creative exploitation of the existing background understanding.

The budworm processes Figure 5.2 span four distinct areas of the plane (I–IV in the figure); and each area requires its own type of analysis. To address the corner

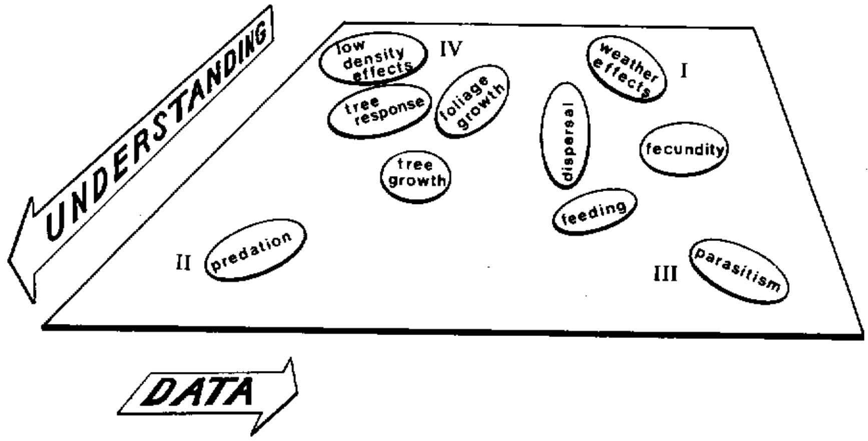


FIGURE 5.2 The location of individual ecological processes from the budworm study on the axes of amount of available data and degree of background conceptual understanding.

with much data but little supporting understanding (I) we have available a whole battery of techniques from statistical analysis. Though statistically fitted curves do not "explain" (despite the misuse of that word in the context of statistical tests), they can describe a relationship in a mathematical form that will at least allow the analysis to continue. The more data and the broader the range of observations, the more comprehensive will be the resulting submodel. But without a foundation based on theoretical understanding, any extrapolation of this submodel to new situations will be dangerous.

Masses of data have been collected on the relationship between a variety of weather parameters and budworm survival rates. Missing from these data is information concerning weather-induced shifts in insect "quality," shifts that could lead to selection of different "types" of individuals that would alter future generations and the dynamics of the outbreak cycle. Also, without knowledge of the mechanisms that actually link weather with survival, we have little guidance for suggesting policies of forest management that could alter the microclimate of this pest species.

In the opposite corner of Figure 5.2 (II) we have few data but considerable conceptual understanding. We have come to the conviction as a result of our case study experience that much can be done when data are scarce but good backup knowledge of process exists. Predation in the budworm system provides an example. In this particular case we are fortunate because predation has been well analyzed at the level of process needed for the model. On the other hand, data are scarce because predation has it major impact when budworm are scarce; at low densities it is very difficult to obtain meaningful samples.

From our knowledge of predation we can specify in advance the mathematical characteristics of the governing functions. Once we know this and have picked a candidate function that meets the requirements, then even scattered data can be applied to establish parameter values. In the budworm case we were able first to classify the various bird predators into distinct parameter classes and then to establish for each class feasible maximum and minimum parameter values. The sensitivity of the simulation model to this range of parameter values can easily be tested through simulation runs, the emerging behavior being used as one criterion for judging the importance of predation.

Along the diagonal region of Figure 5.2 (III and IV) specific data and understanding are more in balance. When both components are large, modeling and analysis are straightforward. The difficulty comes when a process is modeled with glorious sophistication simply because the information is available. In the budworm example, enough was known to construct an elegant and detailed submodel of parasitism. However, such an effort would have been out of keeping with the rest of the model and in violation of our rule of parsimony. The result, in this case, was to use only a single, simple equation that expressed rate of parasitism as a function of budworm density.

Where there are fewer data and where the functional form is not known, it is best to set up alternative testable hypotheses. In the case of dispersal, for example, two extreme alternatives were taken. The first was that dispersal was a random "diffusion" process dictated by weather. At the other extreme was the hypothesis that insect movement was highly clumped and directed. Again, sensitivity tests were made with the model using the recorded spatial dynamics for comparison. In this case the choice between alternatives depended on field data on the overall system behavior. If such field data are not available, then we have identified a research priority. (But note that in highly periodic systems there are often qualitative data available on such things as frequency and amplitude under various conditions.)

When there are few data and little understanding, the requirement for alternative hypotheses becomes even more critical. Sensitivity tests must always be made to check for important shifts in management effectiveness. Technically, there are no "tricks" for modeling such processes other than ensuring logical soundness and checking to be sure that the functions adopted have not introduced unwanted mathematical artifacts into the computations.

CONCLUSIONS

From the predation, dispersal, and other examples we have concluded that one can indeed go farther than usually thought with qualitative analyses of processes. When we are able to use such analyses to complete a causally structured process model, the results are superior to those obtained with any prepackaged modeling "language" or externally imposed mathematical framework.

A SPECTRUM OF TECHNIQUES

Our emphasis on causal relationships based on processes comes from our experience with numerical simulation models. This same process orientation is also appropriate for other "problem-solving" methodologies, such as dynamic programming and other techniques of optimization. Unfortunately, the mathematical structure of these techniques very often places severe constraints upon the way a model can be expressed. At least with a simulation model you are free to "say it the way you want to." However, simulation is such an open forum that it is easy to say too much – this is why we put such strong emphasis on parsimony. One successful technique for reducing the problem to the bare-bones essentials is "looking outward," as practiced in the workshop setting (Chapter 4). These efforts help to keep the resulting management model itself manageable. The next chapter discusses additional steps that can be taken to further simplify and gain understanding.

Simulation modeling, however, covers only a part of the spectrum of mathematical techniques available for environmental assessment and management. As suggested earlier, we think of this spectrum as spanning a range from qualitative to quantitative.

Techniques on the qualitative end, such as interaction matrices, rely on intuition and deep understanding for useful projections of the environmental effects of man's proposed interventions. However, these techniques founder where there are too many variables and relationships linking them, too many nonlinear processes, or too many available actions and potential consequences. Basically, difficulties arise when the problem becomes too big and complex or when its internal interrelationships differ radically from the rather simple form implicit in matrices.

On the other hand, numerical techniques, such as simulation models and optimization procedures, rely on accurate identification of relevant variables and the form of their interrelations, on data for parameterizing those relationships, and on accurate descriptions of the available actions that can be taken. Unfortunately, these models can fail through the mind-numbing barrage of complexity that sometimes appears to exceed that of the real world. Additionally, simulation models built from a base of too few data and, more important, with too little understanding, can lead one quickly and easily to false conclusions.

Within these two extremes, how do we steer a course toward a model that will adequately address any particular management or assessment problem? Our bias toward simulation is stressed throughout this book. Many other workers performing environmental assessments have extensively used and described a variety of cross-impact techniques such as the Leopold matrix and its improved descendants. Our own preconception was that such matrices were probably the best techniques available when very little was known about a situation. Nevertheless, it seemed unlikely that there would be much gain in understanding for improved management – and that any gain might be deceptive because these methods were formalized procedures.

A question arose as we examined a growing range of environmental studies: were there techniques available that would be appropriate for only partially defined systems? We were thinking of situations where more was known than an impact matrix could utilize but perhaps not enough to embark on a normal simulation modeling effort. We thought that if such techniques did exist and were useful, they would have particular merit in developing countries, where the call for development and action is strong but the background of research is limited. We describe in the next section some explorations we and our colleagues have made in response to this question.

EXPLORATION OF QUALITATIVE TECHNIQUES

We were sure that there were ways to effectively analyze systems that possess insufficient information to allow construction of a normal simulation model. Often, all that is known is the major variables and how they interact qualitatively – when *A* is large, *B* will decline. We realized that most environmental studies do not rely on simulation models, but the techniques that are employed in these studies often fail to utilize the information that is available.

In response to this perceived need, and to satisfy our own curiosity, we set out to explore the possibilities offered by assessment techniques thay lay between static impact matrices and more complete dynamic simulation models. We call these intermediate methods "qualitative simulations" because they are formulated on a qualitative rather than numeric base, yet they dynamically project the implications of their interactions into the future. We focused our explorations on the performance of qualitative modeling across a range of data quality and quantity in order to determine if there was a useful matching of these methods to a certain level of information.

These explorations took the form of a gaming exercise. We enlisted ourselves and several of our colleagues in a series of mock environmental assessments. Preliminary sets of data from a few of our well-developed case studies were given to "assessment teams" who attacked them with one or more analytic methods. Others who were very familiar with the case studies were the "judges," comparing the mock assessments with their own hindsight. The real evaluations, however, came from the users' own experiences of the advantages and disadvantages of each technique. An ideal experimental design would use a number of test projects and have several teams of experts analyze each one using a different assessment methodology. We would then wait 10 to 50 years and see how well each methodology predicted the impacts and why some techniques performed better than others. In lieu of this ideal, we approached these explorations as a learning experience for ourselves; consequently, the major product was a set of lessons and observations. These are reported below.

One auxiliary feature of this exercise was its cross-cultural character. In all, about 50 people participated in this series of assessments over a 2 year period. They came from groups in Venezuela, Argentina, and Canada, and they had varying amounts of background skills, though most had been schooled in ecology. We were surprised to find that there were no apparent differences in the groups' ability to utilize various assessment techniques; there was also an unexpected uniformity in their judgment of the relative strengths and weaknesses of the methods.

Because this exploration was a mock exercise, and therefore somewhat artificial, we decided to "anchor" it to our previous experience and to the experience of others who have undertaken environmental assessment. To accomplish this, we subjected the data packages from the sample problems to assessment by simulation modeling and by the Leopold matrix (Leopold *et al.,* 1971) as well as by "qualitative simulations." Although the Leopold matrix is no longer widely used in its original form, it is the precursor of many currently advocated techniques and so was taken for the present purpose as representative of that class of methods.

The product of these explorations was a scorecard like that shown in Figure 5.3. A rating was placed in each box indicating how well each technique did at each level of data quantity. The success of a technique consists of how well it does at, among other things, accurately predicting impacts, adding to our understanding of and insight into the problem, and providing a means for guiding policy. In keeping with our noncookbook style, we will not fill in Figure 5.3, but will let the reader draw his own conclusions from the participants' comments given below and from his own experience.

We next briefly describe the techniques that were used in this exploration, expand on the description of the assessment protocol, and present the lessons and conclusions that we drew from this activity.

THE TECHNIQUES USED

The techniques used were qualitative modeling, the Leopold matrix, and simulation modeling. Since this gaming exercise was primarily a reconnaissance into qualitative modeling, we examined two different qualitative modeling techniques – GSIM and KSIM. We describe both of these, plus the Leopold matrix, below; simulation modeling has already been discussed thoroughly throughout the text. More detailed descriptions of all four techniques can be found in Appendix A.

GSIM

GSIM is a qualitative modeling approach requiring the least information of the four techniques evaluated in this exercise. The user need only specify the relevant system variables and then decide whether the relationship between each pair of variables is positive (an increase in A leads to an increase in B), negative (an increase in A leads to a decrease in B), or zero (an increase in A does not directly

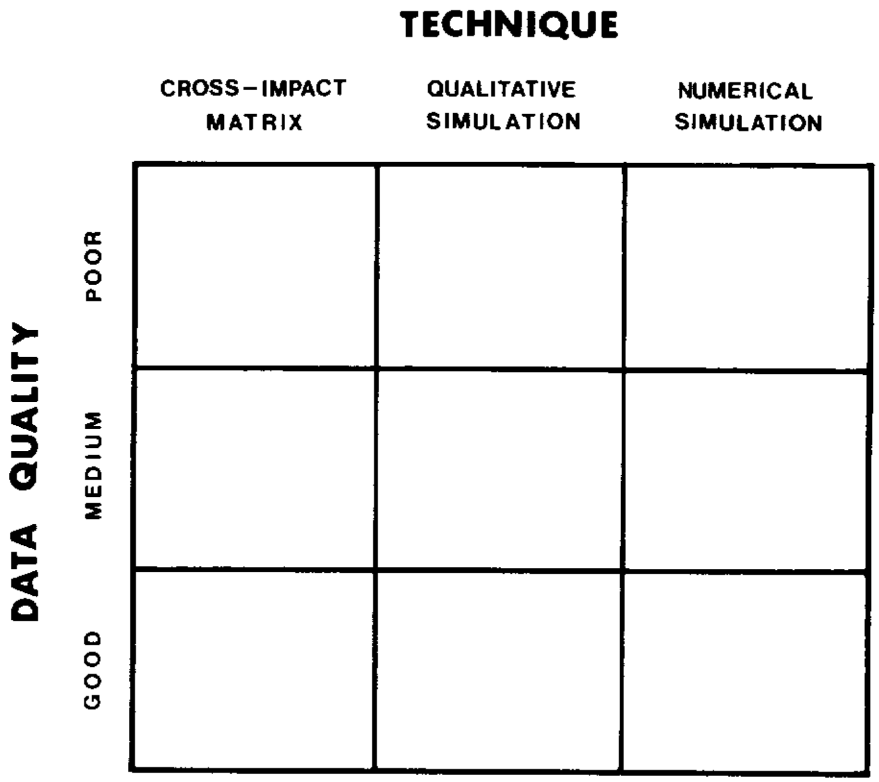


FIGURE 5.3 A hypothetical scorecard for ranking three types of techniques given three levels of available data. This matrix of combinations guided the exploration of techniques described in the text.

affect *B*). The GSIM technique, readily implemented on a computer, evaluates the dynamic implications of these specified relationships. If additional information is available on the relative "importance" of the variables, this is easily incorporated into the evaluation. The principal advantage of this approach is that it allows one to consider the dynamics of the systems and the interactions among variables at an information level too sparse to allow the construction of a standard simulation model. Other advantages are the speed with which the user can structure the model and the very low hardware requirements (a desk computer or even desk calculator is sufficient). This kind of model can provide only rough qualitative trends of the variables and cannot reliably handle situations sensitive to precise numerical balances of the variables.

KSIM

KSIM is a qualitative simulation technique that begins with the same information used by GSIM but also incorporates data on the relative magnitude of interaction effects (a doubling of *A* leads to a halving of *B* and so on). The two basic assumptions behind KSIM are that everything has a potential maximum and minimum and that if among factors of equal importance there are many that cause some variable to increase but few that cause it to decrease, it will increase. KSIM allows some factors to be more important than others and also allows factors to act, for example, more strongly when they are near their maxima than when they are near their minima. The technical details of KSIM are moderately complex, and readers desiring an in-depth understanding should consult the technical description in Appendix A. KSIM may be adapted to accommodate a great deal of quantitative detail, but it then becomes more of a direct simulation than a qualitative technique. For this reason, our tests of KSIM were restricted to a version that did not require quantitative information.

Leopold Matrix

The Leopold matrix and its many variants utilize an impact table that lists a set of possible actions (water diversions, road construction, and so on) down the side of the table, and a set of potentially impacted indicators (water quality, wildlife populations, and so on) across the top. The impact assessment team fills in the appropriate boxes with its impression of the strength of each action's impact on each indicator as well as the importance of the impact, using a subjective scale of 1–10. The result of the Leopold matrix is a very large table describing the effect of each action on each impact indicator. Matrices of this form are a common predictive technique used in environmental impact assessment in North America.

We use the original Leopold matrix here. Some of its defects have been eliminated through various modifications, but the general structure remains substantially the same.

WHAT WE DID

Our initial belief was that the properties and capabilities of a technique should be matched to the characteristics of a particular problem. In the present context, we felt that the extent and detail of the data associated with a problem were the most critical characteristics. We have stated above that background conceptual understanding of the processes can compensate for missing data. Although we knew how this compensation is made in a simulation model, it was not clear if either the Leopold matrix or the qualitative models would have this flexibility. Hence no effort was made to draw benefits from this conceptual understanding. It can be accommodated easily only in a quantitative simulation environment and would unfairly bias the results toward numerical simulation. Therefore in these explorations the only characteristic that was varied from trial to trial was the amount and quality of the data available to the analysis and assessment team.

A group very familiar with one of the case studies was the "expert" during this exploration of techniques. That group took all the material from the problem and assembled three packages of data in a form that might be available to an assessment team charged with analyzing such a problem and predicting the effects of alternative management options. The lowest level data package consisted of only a general description of the system and a minimum of quantitative information. The highest level package was very detailed and included most of the relevant data at the expert's disposal. The third package was intermediate.

The experts also drew up a set of specific questions about the nature and behavior of possible impacts of developments specific to their particular case. The experts, having been intimately involved with the study, knew the answers from hindsight, and in retrospect felt that an environmental assessment team should have been able to predict them.

These data packages and questions were given to other groups – the "assessment teams" – who knew little or nothing about the particular case study. Each team applied one or more of the four techniques, using one of the data packages, and attempted to answer the management questions. As participants we found the project exceptionally useful. As we explored the possibilities of these techniques in various situations, we were frustrated, we were excited, we were angry, but above all we learned a great deal. We attempt to convey the flavor of that experience in the next section.

WHAT WE LEARNED

One lesson of this experience confirmed our original bias: as we moved from poor to good data, only numerical simulation models were able to use the additional data effectively. The qualitative models did not have the capability in their intrinsic structure to utilize numerical data. Indeed, when a group using such a technique was given a set of good data, they often abandoned the qualitative techniques and started doing numerical calculations with pencil and paper.

This exercise also crystallized our feelings about the Leopold matrix. Despite its ubiquitous use, it is in no way a predictive technique. However, it was often a great help in guiding intuition and as a check for overlooked relationships.

In the course of these explorations we were surprised to find that simulation models often fared poorly, failing to answer some of the critical questions about impacts properly. This failure of the assessment teams' models was underscored by the fact that a simulation model built for the original case studies had performed so much better. We attribute this failure of the simulations to two factors.

First, there was a lack of time. This led to misinterpretation of data, logical mistakes, and computer programming errors. But this can happen in any real

TABLE 5.2. Advantages and Disadvantages of the Leopold Matrix

Disadvantages	Advantages
The 88 × 100 matrix is oriented toward construction projects so, categories of actions and characteristics incomplete and not general	Easy to use, no computer facilities needed
	Promotes communication between disciplines
	Relatively little hard data required
Categories too broad, cannot look at specific interactions for which information is available	Useful as a check against other methods to see if particular categories of actions or system characteristics have been omitted
Gives false sense that all possible interactions have been considered once the matrix has been filled in	
Not really a predictive technique – predictions based only on the user's intuition and experience	
Time and effort required large relative to the technique's value	
User not forced to articulate assumptions	
Cannot distinguish between rare and common interactions	
Hard to separate "importance" from "magnitude"	
Rankings of interactions from 1 to 10 highly subjective	
User not forced to define mechanisms of the interactions	
Cannot handle nonlinear impacts	
Relations or interactions assumed constant through time	
Results cannot be summarized in a form easily communicated to the decision maker	
No distinction between processes at different levels in the hierarchy of natural processes	
Uncertainties cannot be included	
Many actions and characteristics have different levels of resolution: some very specific and others very general	

environmental study where deadlines loom and bugets are tight. Errors of these types are always waiting in the wings. Practice, learning, and interactive model construction help reduce these problems, but they never eliminate them. The solution, to the extent that there is one, is to acknowledge the possibility of errors, establish a "degree of belief" through invalidation, and design policies that are robust to these technological difficulties.

TABLE 5.3 Advantages and Disadvantages of GSIM

Disadvantages	Advantages
Cannot handle numerical effects or behavior modes directly dependent on precise numerical balances	Handles very imprecise or qualitative data without introducing too many unwarranted assumptions
Time units arbitrary	Only small computer facilities required
Because of sequential discrete structure, only rough approximation to contiuous processes	Easy to conceptualize, program, and understand the causal determinants of the response
Care necessary about the order of the variables in a causal chain, taking into account whether the impact of some variables upon others should be in phase or out of phase	Handles a large number of causal chains
Changes in variables assumed to be unitary, so GSIM does not differentiate among variables that change at numerically different rates	Handles multiple relations, feedback relations, logical decisions ("IF" statements), time-lags, simple nonlinearities, threshold effects, discontinuities, etc.
Results sensitive to assignment of possible ranges of values of the variables	Forces the user to think about very basic forms of causal connections in terms of the user's own conceptual background, thereby reducing the probability of being caught in the details of the system
	Handles short-term, transient behavior as well as long-term outcomes

The second factor that led to poor model performance was the modelers' unfamiliarity with the underlying processes of the system being modeled. The modelers depended completely upon the data packages and did not have access to the breadth of knowledge needed to supplement the always incomplete supply of data. The mock assessments failed in this regard because we did not follow our own recommended procedures – the models were built by modelers and not by a workshop. A major reason for beginning with workshops is to bring together those people who do have the breadth of familiarity to address the problem adequately.

What was learned by the participants while exploring these techniques is much more important than any scoring and rating of them. We have collected their specific comments in Tables 5.2 through 5.5. Some comments could reasonably be applied to other techniques; some reported advantages and disadvantages are mildly contradictory. We make no attempt to resolve these contradictions but retain them as part of the record to illustrate the need for flexible and adaptive attitudes toward technique selection.

All these classes of technique have a role in environmental assessment and management. The Leopold matrix, or its descendants, are useful for screening but are not intended to be predictive tools. Qualitative simulation models like GSIM and KSIM provide an easy way to formulate a trial dynamic model and to experiment with alternative policies but are of little help for detailed predictions. Numerical simulation models provide the best prediction when the data are good and are still

TABLE 5.4 Advantages and Disadvantages of KSIM

Disadvantages	Advantages
Behavior essentially logistic	Relatively little knowledge about the mechanisms of interactions between variables needed
Built-in assumptions not necessarily made clear to the user	Good at promoting interdisciplinary communication and getting decision-makers involved
Arbitrary time scaling possibly confusing	Helps to identify some variables and interactions that should be investigated or used later in a more detailed simulation
Relations between variables assumed constant through time	Helps to bound the problem, that is, limit the variables to be considered
Difficult to assign values to relations in the input interaction matrix, particularly if observations on the real system are of a "process" type instead of time series	Good for a "quick and dirty" simulation
All variables bounded between 0 and 1, making it difficult to compare the relative impact of each variable	Graphic output a good way of communicating impacts
Difficult to guess what initial conditions should be assigned to variables (e.g., are 60,000 trout equal to 0.2 or 0.8 of the maximum number possible?)	Alternative management schemes can be compared relatively easily by changing values in the input matrix and rerunning model
Detailed information on processes often cannot be used in the KSIM framework	Handles large numbers of different *kinds* of variables (physical, sociological, biological, etc.)
Graphic output can delude; gives false sense of security in precision of predictions	
Fails to allow measures of degree of belief in data or assumptions to be reflected in final results	
Users often adjust values in input interaction matrix in order to give "reasonable" output: i.e., data are adjusted to fit preconceived notions of what should happen – obviously not useful in the context of environmental impact assessment	
Users cannot distinguish between processes at different levels in the hierarchy of natural processes	
Computer facilities needed	
Cannot include uncertainties	

useful for guiding research when the data are poor. There is no reason why all these techniques could not be used if the assessment process is to be adaptive. The judgment of proper timing and mixing of techniques comes best from experience.

TABLE 5.5 Advantages and Disadvantages of Simulation Modeling

Disadvantages	Advantages
Requires computer facilities	Promotes communication between disciplines
Requires expertise and a fair amount of time	User forced to clarify assumptions and causal mechanisms
Results may be too easily believed by decision makers	Any form of relationships can be handled – linear or nonlinear
Results are usually complex (if there are many variables) and are therefore difficult to communicate to decision makers	Helps to identify key variables or relationships that need to be investigated or are sensitive
Fails to allow measures of degree of belief in data or in the assumptions to be reflected in final results	Can include uncertainties of various types
Relations between variables usually assumed constant through time	Can easily compare alternative management schemes
	Can use detailed information concerning processes in the natural system
	Graphics output a good way of communicating impacts
	Can utilize information about known processes that have not been investigated for the particular system of study but that have some generality (e.g., predation, population growth).

We mentioned above that the simulation models built during this exercise differed from those originally constructed for the case studies. Although extenuating circumstances rooted in the nature of these explorations contributed to these differences, it still remains true that models of the same situation built by different groups will not be the same. If they are not the same, then which is the right one? Our answer, which should be easily anticipated by now, is that there is no "right" one. A model is only one piece of evidence that contributes to creative design of environmental policy and assessment. An adaptive approach to technique selection relies on alternative models emerging from alternative forms of analysis. The broader the range of evidence, the better, it is to be hoped, will be the conclusions.

Many environmental decisions must be made now, and we hope they will be made well. The developing countries should not be asked to stop resource development simply because our predictive tools are not perfect and therefore we cannot foresee and avoid all the unwanted consequences. The shortage of food and material for the people of these countries is real, and doing nothing solves nothing. Actions will not, and cannot, wait in the developed world either, where the pressures to develop are also real. On the other hand, the pendulum can swing too far the other way. All development should not go blindly ahead simply because we lack the tools to confidently predict the bad effects.

We need to learn how to gain information as we proceed with management. We need to choose an adaptive analysis that utilizes a variety of techniques so that insight from one will help foster understanding of another. We need to learn how to avoid irreversible decisions at the beginning, when data are being acquired. Above all, we need creative methods for acknowledging uncertainty and progressing in the face of it.

6 Simplification for Understanding

Complexity and simplicity each have a place in the adaptive analysis of environmental problems. A model that adequately represents the real world will necessarily contain some of the world's complexity. Although we strongly advocate parsimony, there is always a limit to the number of complications that can be removed from a management model if reliability is to be maintained. Ecological behavior stems directly from nonlinear dynamic linkages, time lags, and spatially heterogeneous distributions – each of which promotes model complexity. A model that is too simple will lack credibility, and one that fails to address a level of detail coincident with management operations will not be usable.

Simplicity, on the other hand, permits comprehension – a prerequisite for developing understanding and gaining insight. Simplified versions of the "working" management model provide alternative perspectives and avenues of analysis that foster innovative policy design. These same simplified versions are also useful for making trial assessments of candidate environmental policies and for identifying and investigating the system components that are sensitive to perturbations. Additionally, effective communication between analysts, managers, and the public depends on concise, unencumbered, but accurate formats that are easily developed from a formal process of simplification.

An adaptive approach to environmental problems avoids choosing a single level of complexity. Rather, it deliberately seeks to meet the requirements of reliable representation and credibility by using an adequate degree of realistic complexity. The adaptive approach also addresses the requirements of understanding, critical evaluation, and communication by using creative simplification. Failure to address both sides of this dichotomy will jeopardize important elements of assessment and management.

We propose an active and deliberate blending of the simple with the complex. We accomplish this by creating a collection of simpler, but complementary, representations of the management model. The simplifications are caricatures that help describe the properties, behavior, and possibilities of the environmental situation that confronts us. Because these simplified versions are unified by the detailed management model from which they were derived, exchange of ideas between them is facilitated. Interpretations from one version provide a backdrop for others.

These various representations form a hierarchy of alternative models, each providing a different perspective or a different level of detail. In no case are these simpler versions substitutes for the complete, "official" model.

Although this is a "technical" chapter, simplification is not a technique, but rather an attitude based on curiosity and a desire to get the most out of an analysis. This attitude is made operational by iteratively transferring ideas developed at one level into another level for testing and evaluation. Thus we take a policy suggested by one of the graphical techniques described below and implement it in the complete management model, where a fuller range of constraints and interactions is brought into play. The performance of the model under this new policy is one piece of evidence used to corroborate or reject the potential of this proposed policy. Similarly, ideas generated by the management model are tested at a higher level of complexity – a carefully designed and monitored field trial. Eventually, the ideas and analyses that have performed successfully at all levels available are applied to the real world.

There are no fixed procedures to follow in these modeling extensions, but we shall indicate through some detailed examples the range of things that can be done and the benefits both to us as analysts and to the case study clients – the people in the various management positions to whom the case study materials will be ultimately transferred.

We shall discuss three types of simplification:

- Smaller models created by extracting submodels that are explored independently of other submodels.
- Sets of differential equations incorporating fewer variables and parameters than the complete simulation model.
- Pictorial diagrams that display the underlying structure of the model. These serve as powerful analytical tools for penetrating to the heart of the model, and they require no "mathematics" to use.

SUBMODEL ANALYSIS

As mentioned, the complexity of the budworm model reflects both the large number of state variables and its variety of behaviors. We can eliminate much of the numerical complexity by extracting the biological submodel from one forest area so that its behavior may be explored separately from the other 264 areas. While this circumscription reduces the direct forest management relevance, the simpler

79-variable biological model still contains much of the dynamic character of the complete spatial model. By treating this biological model (which we call a *site model*) as a stand-alone entity, we can cheaply, easily, and more thoroughly explore the causes, range, and significance of those dynamics.

Operationally, we found it very helpful to embed this site model in a computer software environment that allowed quick graphical interaction between the model and any user. This interactive program package (Hilborn, 1973) was the first computer software item that was installed during the process of transfer to management personnel in New Brunswick, Quebec, Maine, and elsewhere. With this simulation system it was possible for a person, upon his first exposure to the model, to ask questions, make changes, propose alternative hypotheses, and receive an immediate graphical response. When changes produced significant results worthy of further investigation, those changes were made in the complete spatial model and examined in detail. Thus the simplified site model served both as a convenient experimental tool for the analysts and as a convenient "doodling-pad" for the potential policy maker.

There are a great number of submodels and combinations of submodels that can be isolated in a similar way. When a part is examined separately, it is necessary to set the conditions explicitly for all the excluded variables. We are completely free to set them at realistic or at interesting values. Thus in the case of the isolated site model mentioned above, the effect of dispersal from other forest areas was partially mimicked by establishing a particular fixed background of immigrating insects. Behavior of the site model with and without management controls under various levels of constant immigration was a stepping-off point for examining the more complex space–time patterns that the complete spatial model exhibited. (The complex behavior of the spatial model is shown in Figures 11.8 and 11.9 in Chapter 11.) This leap from one site to many was bewildering enough that an intermediate model with only a few sites and simple geometry and meteorology was useful (Stedinger, 1977).

At the other extreme it is often necessary to add a more complex level to the hierarchy. Baskerville (1976) chose to expand the model from 265 to 450 spatial units and to record 120 tree ages explicitly rather than 75. This expansion was operationally necessary because of the questions and concerns of a particular set of administrators.

SIMPLE ANALYTIC MODELS

The second class of simplification steps back from the complete model and seeks a smaller, less complex alternative using only a subset of the variables and functions. This subset aims at retaining the major causes of the system's behavior but, being more amenable to analysis, helps to crystallize our understanding of the important interactions and the possible effects changes will have.

In the case of the budworm, this simplified model took the form not of a simulation but of a set of three coupled differential equations (Ludwig *et al.*, 1977). One variable was the budworm density, the second was the developmental stage of the forest, and the third was the physiological state of the trees. These equations were constructed from our assessment of the important components and adapted by continually comparing their mathematical behavior with the complete model.

To give some indication of the economy achieved in this way, all of the elaborate programming of budworm biology and survival reduced to the following equation:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = rB\left(1 - \frac{B}{k}\right) - \beta\,\frac{B^2}{\alpha^2 + B^2},$$

where B is the budworm density and r, k, α, and β are parameters that depend to some extent on forest conditions. We will not go into detail here, as a complete description is available in the paper cited above. We wish only to highlight the possibility that simple alternative models can pinpoint important relationships and provide the raw material for rigorous penetrating mathematical analyses.

The interaction among a collapsed set of variables was also formulated as a set of differential equations in the study of recreational development in the high alpine valley of Obergurgl, Austria, described in Chapter 13. In this case a few differential equations were able to replace the complete simulation model without a significant loss of capacity to mimic the full behavior of the larger model. While the Obergurgl model was not complex by "modern" standards, it still contained sufficient detail to prohibit adequate analysis of its internal workings. The full model contained more than 100 variables, each with a value representing the condition of some piece of the system, such as the number of villagers in various age groups. The major variables were collapsed into a set of five coupled differential equations. Each equation was much simpler than its analogous submodel but faithful to the main interrelationships. These equations produced behavior qualitatively equivalent to the behavior of the full model. The payoff was an increased ability to explore the model's calculations, and to discern why the output changed when alternative starting conditions and hypotheses were used.

These differential equations alone are inadequate for the design of economic policies for Obergurgl. For one thing, the ten spatial areas were lumped into one. To the villagers, each of the subareas has special meaning in terms of things that affect their lives. Even so, by using only five variables, we obtain important clues about "how the system really works." The awareness that these five variables could account for a large fraction of Obergurgl's socioeconomic structure was a conceptual advance over what was believed before the first workshop. Actual policy and social decisions must, however, address the more complex features reflected in the complete model.

MANIFOLD ANALYSIS

The third and final class of simplification requires more detailed description – not because of any inherent difficulty, but because of its novelty. The product is a set of pictures or graphs that can be easily comprehended and require for their understanding no mathematical skills (although the graphs themselves are founded upon mathematical principles). These diagrams are not models in the sense of a simulation, but rather are alternative representations of the internal structure of the model. They are analogous to medical x-rays that reveal the structure of the skeleton without removing the surrounding flesh (this was done in the simplified models described above). And as with x-rays, our perception of structure improves with several complementary views taken from different orientations or perspectives.

These pictures are useful and usable because they make strong use of qualitative information rather than opting for the quantification espoused by most scientific disciplines. The qualitative property of interest in the budworm example is the classification of forest conditions into those that cause budworm numbers to increase and those that cause them to decrease. At first this may appear to be a minimal criterion, but in many management situations knowledge of gain or loss would be prized information, if available. (Imagine the profit to be made with the same information on the stock market.)

The powerful aspect of this qualitative division is its inclusion in a topological view of the system. The interface between regions of increase and decrease defines conditions for no change – that is, equilibria of the system. Our topological view links the basic dynamic behavior to the number and interrelation of equilibrium states and focuses as well on our central concern for ecological resilience and policy robustness. Just as the skeleton determines much of an organism's appearance, the structure of the equilibrium states determines the system's dynamic behavior.

Our first step is to use the complete simulation model to generate a population growth rate, or "recruitment rate," curve of the sort introduced by Ricker (1954) for the analysis of fish populations. The recruitment rate is

$$R = \frac{N_{t+1}}{N_t},$$

that is, the ratio of the population in the next generation $(t + 1)$ to the population in the present generation (t). This is the number of times bigger, or smaller, next year's population will be than this year's. In Figure 6.1 R is plotted against the present density of budworm for particular forest conditions. The recruitment rate curves condense all the reproduction and survival functions within the model, and a unique curve can be calculated for each state of the forest. Three selected curves are shown for three levels of forest development – immature, intermediate, and mature. In reality there is a continuum of curves, each representing a particular forest state. Each point is computed simply by starting the simulation model at

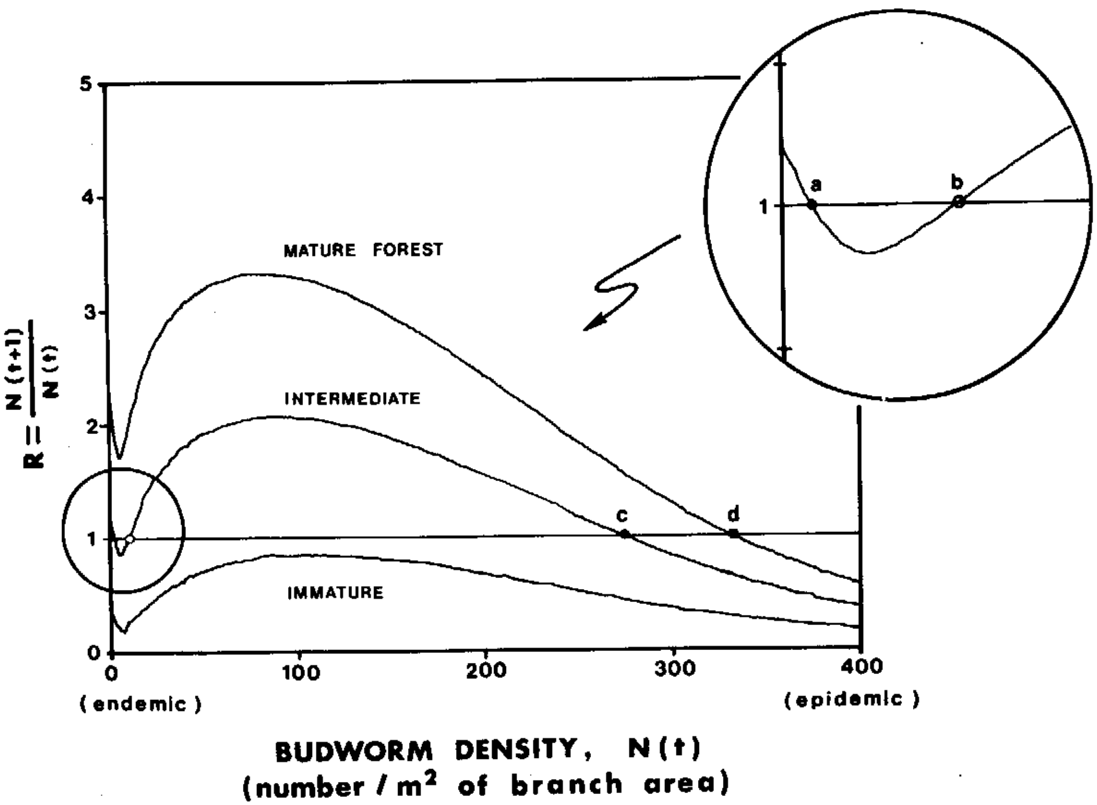


FIGURE 6.1 Recruitment rate curves for budworm. R is the rate of population growth from one generation to the next as a function of current population density. Each of the three curves represents a particular state of forest maturity; all other variables are assumed fixed at their nominal values. See text for a discussion of the significance of points a, b, c, and d. The insert expands the circled part of the intermediate forest curve.

the specified values [here, $N(t)$ and forest state], running it for one time interval, and noting the resulting R.

Interpretation of the curves is straightforward. We start by focusing on the location and properties of the equilibrium points – the points where the recruitment rate takes a value of 1.0. These equilibria may be stable or unstable, depending upon the slope of the curve as it passes through the $R = 1$ line. Briefly, if a slight increase in density from the equilibrium point results in further increases in the next generation (i.e., if $R > 1$), or if a slight decrease results in further decrease ($R < 1$), then the equilibrium is unstable (represented as an open circle in Figure 6.1). In contrast, where a slight increase in density from the equilibrium point is offset by a decrease in the next generation ($R < 1$), and a slight decrease is offset by a subsequent increase ($R > 1$), then the equilibrium is stable (shown as solid dots in Figure 6.1).

Subsequent discussions draw heavily on these recruitment curves, so it is useful to consider their structure in some detail. The high-density equilibrium points (*c, d* in Figure 6.1) are established largely through competition among budworm for the available foliage. Although these points are stable equilibria for budworm, they are unstable for trees. At such high budworm densities, defoliation is so heavy that older trees die and are replaced by seedlings and understory growth. This shifts the system onto the immature forest curve with a lower budworm growth rate. Since $R < 1$ for the immature forest at budworm density d, the insect population declines. In summary, when the forest is immature, R is less than 1 for all budworm densities and no outbreak is possible. With a very mature forest, however, budworm will increase from all densities less than d, rising until they reach this upper equilibrium. The ensuing defoliation and tree death bring the population back to low numbers.

There is almost no information available about the fate of budworm at very low densities (lower than can be shown on the arithmetic scale of Figure 6.1). Either the local populations become extinct in immature areas of the forest ($R < 1$ for all densities) and dispersers must re-establish populations at the site, or the local populations can be maintained at some very low level ($R > 1$ at densities less than this low level). In either case there is a lower equilibrium, which is zero or some low density. The remaining curves are appropriate for either situation.

The dip in the recruitment rate curves at low budworm densities reflects the activity of avian predators, augmented to a degree by parasitism. When the forest is of intermediate age, this dip introduces two low-density equilibria – one stable at a and one unstable at b (see insert, Figure 6.1). The population may persist at density a until improved forest conditions raise the bottom of the dip above the $R = 1$ line. When this happens, only the high equilibrium remains and an outbreak occurs. But an outbreak can occur even in an intermediate-aged forest if a sufficient number of budworm are imported by dispersal from outside areas. Thus, in Figure 6.1, a small number of budworm added to the population that is at equilibrium a will result in an increase in density above the unstable equilibrium density b. As R is greater than 1, an outbreak starts.

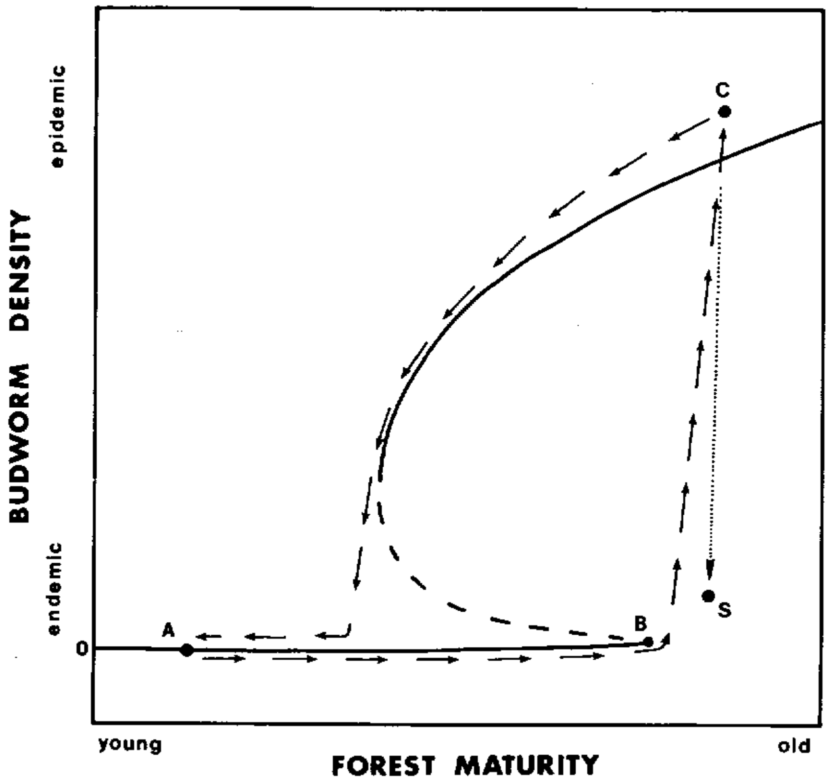


FIGURE 6.2 The equilibrium manifold of budworm densities for different forest conditions. The solid line represents the location of equilibria; the dashed line separates the high and low budworm densities. A normal cycle begins at *A* (young forest, few budworm) and progresses to *B*, where the low equilibria are lost and the system can no longer maintain a low budworm population. An outbreak is triggered. The budworm density is drawn toward the upper curve and arrives at point *C*. The feeding stress at this magnitude of budworm density causes tree mortality, and the forest is forced back to a younger condition, taking the budworm population down with it. The cycle returns to point *A* and begins anew. If 80% of the population at *C* were killed by insecticides, the system would move to point *S*, where there is little loss to the forest but high vulnerability to any suspension of spraying.

The recruitment curves as described do not yet include the stochastic elements of weather that affect both survival and dispersal. When these effects are included, there is a third trigger for outbreak – a sequence of warm, dry summers, which can raise normally low recruitment rates above the replacement line.

A more complete and succinct summary of these multiple equilibria can be obtained by plotting the location of only the equilibrium budworm densities (the dots from Figure 6.1) for all levels of forest maturity. The heavy curve in Figure 6.2 shows just such a relationship. The lower, solid segment corresponds to endemic densities such as *a* in Figure 6.1; the middle, broken segment corresponds to the unstable points such as *b*; and the upper, solid segment traces the epidemic densities such as *c* or *d*. Note that, just as in Figure 6.1, when the forest is immature there is only one low equilibrium, and when the forest is mature there is only an epidemic equilibrium, but when the forest is of intermediate maturity, there are two stable equilibria separated by an unstable equilibrium.

We call the collection of equilibrium points such as drawn in Figure 6.2 an *equilibrium manifold*. In the remainder of this section we shall examine some of the useful properties of this manifold and explore the ways that its shape changes under the influence of changing conditions. The shape of the manifold governs much of the dynamic richness of this system.

With these manifolds we can follow the shifts in the number and position of equilibria. The same is true with simple two- or three-variable models where the equilibria are easily determined analytically. As was indicated in Chapter 2, the organization of the equilibria of a system has a fundamental effect on its dynamic behavior. The equilibria are easy to find in a simplified model, and, having found them, we know where to look in the complex model. It is also important and useful to study the positions of the boundary lines separating different areas of stability. Some configurations of these boundaries can lead to unexpected outcomes. For instance, in some situations a decline in the population of a pest species can lead directly to an "explosion" to high densities (Bazykin, 1974; and Figure 2.2*F*, Chapter 2).

The focus and use of equilibrium manifolds are suggested by that part of the field of mathematical topology evocatively called "catastrophe theory" (Thom, 1975; Zeeman, 1976). An expanded exposition of this theory in terms of budworm outbreak dynamics is given in Jones (1975), and Jones and Walters (1976) and Peterman (1977b) have related it to fisheries management.

Returning to Figure 6.2, we show how the particular configuration of this manifold dictates the essential features of the classic outbreak cycle. A normal sequence begins with a young forest (at point *A*). Such forest conditions will support very few budworm, as reflected by the single low equilibrium. The ruling property of these manifolds is that the budworm densities will either increase or decrease as governed by the population growth curves illustrated in Figure 6.1 until they reach a point of equilibrium – a point on the solid branch of the manifold. If the budworm densities are on the manifold, then they will try to remain there even as the level of forest maturity changes.

Thus, as our typical forest grows older, the budworm densities follow smoothly and evenly along the lower branch from point *A* to point *B*, showing very little change in density. However, the moment the forest grows beyond point *B*, the lower equilibrium is lost, and the only one available to the system is the upper, epidemic level. An outbreak is triggered. As the budworm population begins its rapid increase, the forest continues its growth, and the system trajectory moves upward toward point *C*.

The manifold we are following portrays the movement of budworm numbers in response to forest conditions. There is also a manifold that portrays changes in forest conditions as the forest is affected by budworm densities. Rather than show this second manifold graphically, we shall rely upon a verbal description of how it comes into play and influences the trajectory that has reached point *C*. The manifold at *C* is an equilibrium for budworm only if forest conditions remain unchanged. However, the feeding stress imposed by this density of insects causes severe tree mortality, and the forest reverts from a mature one to one that is young. As the forest condition collapses, the budworm population falls along with it. The cycle returns to point *A* and begins anew.

We can immediately draw several very broad and important conclusions from Figure 6.2. First, it is clear that if the forest has the capacity to reach a condition beyond point *B*, then an outbreak is inevitable. Much of the mystery about the "cause" of outbreaks disappears when we view them as a simple playing out of the mechanism inherent in this manifold configuration. We also see that once an outbreak is triggered, it is destined to continue its course even if we could restore the forest to a pre-outbreak condition slightly below point *B*.

The second conclusion is that if we were to prevent the forest from ever reaching point *B* (by logging or thinning, say), we could happily maintain the budworm at an endemic level. However, it is clear that such a system is extremely vulnerable to invasions of budworm from outside areas. This is the same conclusion we drew earlier: even though an intermediate forest would not suffer outbreak spontaneously, outbreaks could be triggered by a pulse of immigrating insects. Through this mechanism a central mature stand can initiate an epidemic that spreads throughout surrounding less mature areas. We will return to this point later and develop a manifold that expresses these conditions directly.

A third obvious conclusion from Figure 6.2 has important policy relevance for budworm control. If during an outbreak (point *C*) insecticide spraying is initiated, the system would be displaced to a state such as point *S*. Because this point is far from an equilibrium, it is being held "unnaturally" in an unstable condition. The longer this policy is followed, the larger the area that requires spraying – both because more areas are maturing and because surrounding less mature areas are being invaded by insects leaving the sprayed areas. The maintenance of desired system behavior is therefore extremely sensitive to any intervening failure in implementing the policy, be it through evolved genetic resistance, errors in spray formulation and delivery, or legal restrictions on spray dosages, targets, and frequency. The entire system would collapse. This is the predicament in which eastern Canada now finds itself.

For the purposes of easy understanding of the nature of the manifolds, we have defined "forest condition" in a causal and intuitive manner. The measure of "maturity" of relevance to the budworm is the surface area of branches, which is the area available for habitat. As a forest stand ages, the total area of branches increases monotonically. However, there is an additional component of forest condition that affects a budworm's life. That is the foliage quantity – the amount of food available

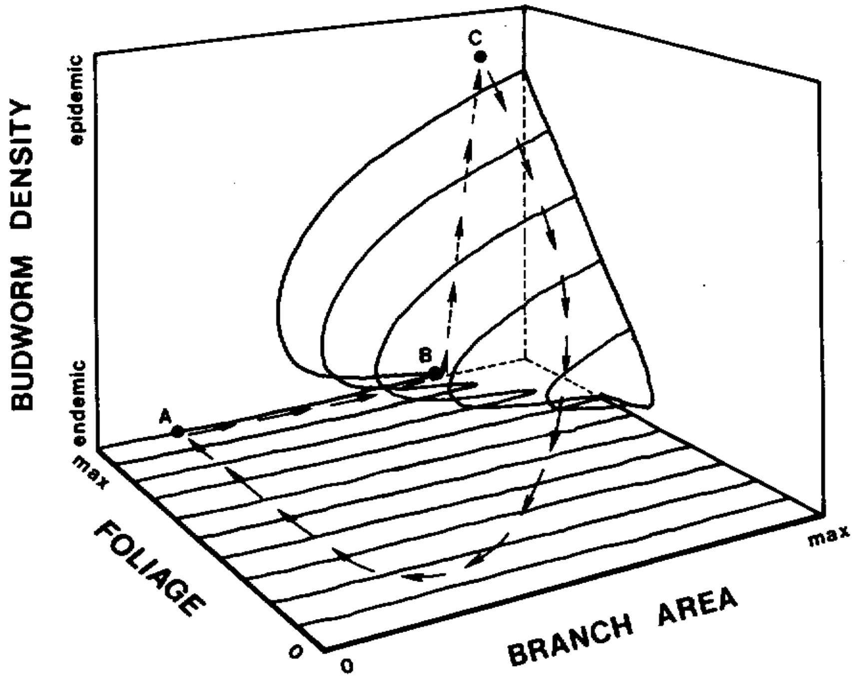


FIGURE 6.3 The equilibrium manifold of budworm densities as a function of the two measures of forest state: foliage condition and branch area. Branch area was called 'forest maturity' in Figure 6.2; the curve at the back of the box (foliage = max) is the same as the manifold in Figure 6.2. The typical budworm outbreak cycle is repeated here (points *A*, *B*, and *C* are the same) to show how foliage and branch area interact during an outbreak collapse.

per individual. When we include foliage as a second measure of forest condition, the budworm manifold becomes a surface in a 3-dimensional box, the axes now being foliage, branch area (what we earlier called "forest maturity"), and budworm density. The manifold surface for these variables is shown in Figure 6.3. Note that the curve at the back of the box (where foliage is maximum) is exactly the same as that of Figure 6.2. The same budworm cycle trajectory is repeated in Figure 6.3, with points *A*, *B*, and *C* as before. Now we see that, starting at point *C*, the foliage goes first, and its loss leads to the death of trees and a reduction in branch area.

The equilibrium manifold representations also prove to be a powerful device for exploring the consequences of changes in ecological processes or management approaches. In progressing from Figure 6.2 to Figure 6.3 we saw how the manifold changed shape as foliage quantity varied from its maximum down to zero. In any ecological model there will be a great many significant factors whose variation would also change the manifold. The number of predators, the number of parasites, the weather condition, the intensity of immigration, and the intensity of insecticide

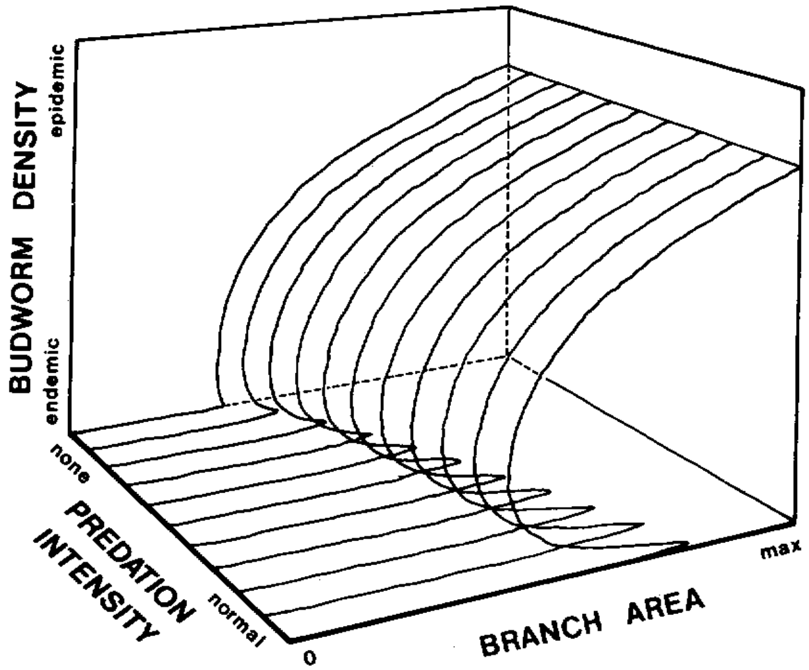


FIGURE 6.4 The predation manifold. This shows the changes in the budworm equilibrium manifold for different intensities of predation by insect-eating birds. The curve at the front with normal predation is the same as that shown in Figure 6.2.

spraying have all been mentioned as important components of the budworm/forest system. On any one three-dimensional figure, such as Figure 6.3, we can only look at the effects that two factors have on the budworm equilibra; all other factors are fixed at their nominal values. To look at a new factor graphically we must sacrifice explicit portrayal of one of the variables in Figure 6.3. In the present case, it is most useful to return to Figure 6.2 (with foliage fixed at its maximum value) and implicity retain our understanding of how the foliage dynamics produce the cyclic trajectory shown initially on Figure 6.2. We now can start with this simpler manifold as a base and investigate how it changes under the influence of other factors, one by one. We know that, in the background, the foliage will continue to operate according to the scheme shown in Figure 6.3.

As an example, Figure 6.4 shows an equilibrium manifold that looks at the effect of different intensities of predation. When predation is at the level occurring in nature ("normal" on the scale), the "pit" responsible for the lower equilibrium is pronounced (again the same curve as in Figure 6.2). But as predation is relaxed, the pit gradually disappears, along with the folded character of the manifold.

Under such conditions, the behavior of the system is radically and predictably

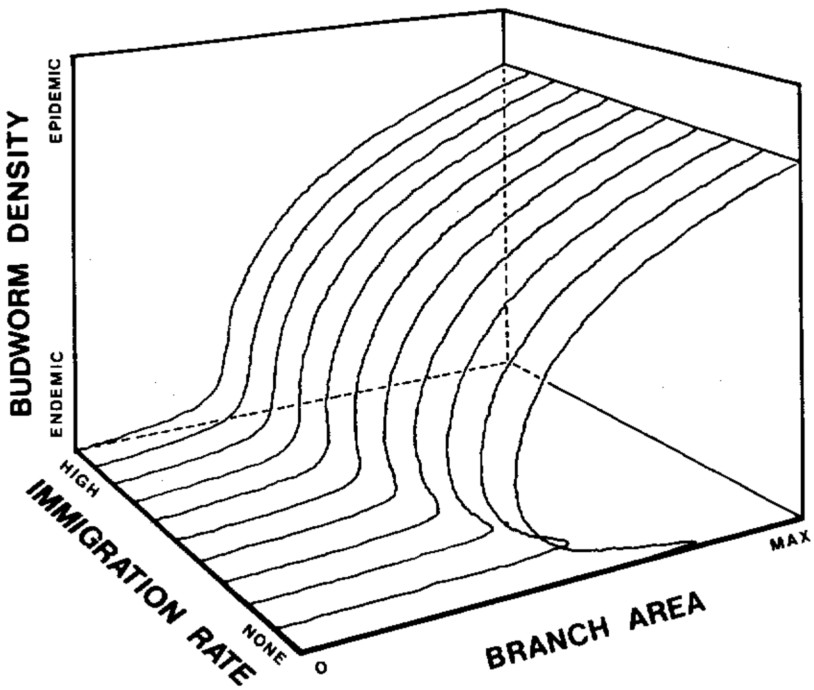


FIGURE 6.5 The dispersal manifold. This shows the changes in the budworm equilibrium manifold for different intensities of immigration by budworm from other forest areas. The curve at the front with no immigration is the same as that shown in Figure 6.2.

altered, since the natural "boom-and-bust" pattern is intimately associated with the reflexive form of the manifold. Simulation runs conducted to check this topological implication of reduced predation show a world with a perpetually immature forest, where moderate budworm densities oscillate with a 12–16-year period. This residual oscillation is a typical "predator–prey" cycle between budworm and foliage. Since insecticides have exhibited a potential for reducing vertebrate predation directly through mortality or indirectly by affecting food availability, the significance of this finding for management is obvious.

Another example is shown in Figure 6.5, where the manifold is used to explore the qualitative implications of dispersal. The immigration-rate axis reflects the intensity of budworm moths immigrating from outside areas. The similarity of this dispersal manifold to that for predation is striking and significant. An increased rate of immigration clearly has qualitative properties much like those of a decrease in predation. This is in keeping with the earlier analysis of recruitment rate curves (Figure 6.1) where the quantity of immigrants necessary to release a budworm population from its low density equilibrium was directly related to the size of the predator–induced pit. As would be expected from the comparison of manifolds, a systematic increase in immigration rate affects the dynamic behavior in very much the same way as a systematic decrease in predation, flipping the budworm–forest system into its alternative mode of a sustained outbreak with a 12–16-year insect–foliage cycle.

The greatest payoff from the topological simplifications comes in their implications for policy. In discussing the recruitment rate curves of Figure 6.1, we noted that a forest could be so immature that no outbreak was possible under any conditions ($R < 1$ for all budworm densities), or so overmature that an outbreak would ensue if any budworm were present ($R > 1$ for all subepidemic budworm densities). This phenomenon is reflected more clearly as the budworm–foliage–branch manifold in Figure 6.3.

We have shown the policy consequences of spraying outbreak populations – the system is perched precariously at point S in Figure 6.2. In our discussion of policy evaluation procedures (Chapter 8) we describe two new policies for budworm management that explicity recognize the form and flexibility of the budworm manifold. We briefly outline one of these policies here.

We saw previously that an outbreak occurs whenever the forest matures beyond the end of the low-density pit (point B). This suggests a policy of "pit enhancement," emphasizing management at *low* densities. A specific agent or management act is not stipulated, only a broader description of a reshaped manifold with a deeper pit. There are many possible management acts that would accomplish this; for instance, any mortality agent applied only at low insect densities. To have a significant effect, the added mortality need not be anywhere as high as the 80 per cent common to epidemic spraying. We could combine this new act with a supplementary insecticide capability to push outbreak populations back into the newly deepened pit whenever unexpected events occur. Because predation by birds is primarily responsible for the basic pit, we know we must also include efforts to maintain them as an important budworm control resource. When this policy was introduced into the complete simulation model, it proved very effective, with radically reduced spraying requirements.

In summary, a compressed and simplified version of a dynamic model can be captured in topological manifolds that focus upon its multiple equilibrium properties. These manifolds are then exploited to improve understanding of the system behavior and structure and to qualitatively diagnose regions of policy sensitivity and potential.

Clearly, if the descriptive part of the analysis stops at the development of a complex simulation model, the clarity of understanding needed for creative environmental management and assessment is seriously compromised. Creative simplification is necessary for understanding.

7 Model Invalidation and Belief

Once we have formulated a model and have subjected it to analysis through simplification, the natural question is whether the resulting products should be believed. Are they valid representations of reality?

The so-called validation process is really nothing but hypothesis testing because models are merely statements of hypotheses. We have little new to say on this subject, and our treatment here largely reviews some of the more fundamental guidelines and dangerous pitfalls involved.

The majority of environmental modeling efforts are silent on the model testing issue, apparently assuming high-quality predictions once all known relations between variables are included (Mar, 1974). Most studies that do address the validation problem seem intent upon proving models to be correct (see Ackerman *et al.*, 1974; Ross, 1972). They tend to emphasize "tuning" to historical data and elaborate statistical testing against replicate study areas or against independent data withheld from the model development exercise. None of these approaches is worth much for assessing the value of management model predictions, simply because management actions often move the system toward conditions that have not been historically encountered.

In fact, it is the central tenet of modern scientific method that hypotheses, including models, can never be proved right; they can only be proved wrong (Popper, 1959). This is why the frequent claims of – and demands for – "valid" models in ecological management, impact assessment, and policy design are so unsound. Provisional acceptance of any model implies not certainty, but rather a sufficient *degree of belief* to justify further action. In practice, the problem is one of model invalidation – of setting the model at risk so as to suggest the limits of its credibility. The model is subjected to a range of tests and comparisons designed to reveal where it fails.

There is no checklist approach to intelligent invalidation, just as there was none for model formulation. But our experiences have suggested three major considerations relevant to the critical assessment of model credibility:

Data, model structure, and invalidation
Evidence for invalidation
The analysis of alternative models

DATA, MODEL STRUCTURE, AND INVALIDATION

THE MODEL AS CARICATURE

Any model is a caricature of reality. A caricature achieves its effectiveness by leaving out all but the essential; the model achieves its utility by ignoring irrelevant detail. There is always some level of detail that an effective model will not seek to predict, just as there are aspects of realism that no forceful caricature would attempt to depict. Selective focus on the essentials is the key to good modeling, and invalidation tests must recognize this as a strength and not a weakness.

WHAT WE PREDICT

There is no sure way to decide what to predict and what level of detail to include in order to produce a believable model. This depends in large part on the bounding decisions made earlier and the sorts of predictions needed for the assessment. At a minimum, however, a believable model should accurately predict qualitative properties of the temporal and spatial patterns characteristic of the historical system.

An extreme example of the distinction between predicting exact numerical detail and predicting qualitative behavioral properties is provided by the budworm–forest analysis presented in Chapter 11. The model of this system predicted insect numbers and tree condition for each of 265 geographical cells, representing a continuous area of about 50,000 km^2. Historical data were available for the same variables at each location over a 25-year period.

No model, however detailed and accurate, could be expected to reproduce the historical detail exactly. The bounding decisions leading to parsimony described in Chapter 4 make this impossible. Random effects and unique but unrecorded events in the historical record also prevent an exact mimic. But independent of this fine detail, historical data showed general, stable patterns in space and time: they revealed a characteristic 30–45 years between insect outbreaks, a local outbreak duration of 3–6 years, and an outbreak spread rate of about 50 km per year. Model predictions corresponded very closely with each of these qualitative characteristics of the historical record, although there were quantitative discrepancies when predictions and history were compared at individual points in space and time. This qualitative comparison of time–space predictions and behavior served to substantially strengthen our belief in the model, though it did not, of course, "validate" it. Further invalidation tests, which we describe below, strengthened our belief in other ways – no one test was sufficient or even dominant.

The opposite effect, that of definitive invalidation, can be demonstrated with a study of an oceanographic model. Marine plankton data required for fisheries studies are usually highly variable, making most space–time models effectively untestable. However, by looking at the data in a different way, one finds that this variance from place to place consistently *increases* when larger and larger areas are compared. With a focus on this pattern, it becomes possible to use the variation as an aid to invalidation rather than treating it as a hindrance. It is often assumed that this pattern in the variance results from the interaction of growth rates of the organisms with the effects of horizontal mixing. A model incorporating simple prey–predator interactions and lateral diffusion was developed (Steele and Henderson, 1977). The output was expressed explicitly in terms of variance as a function of horizontal scale so that it could be compared with a set of data from the North Sea. In this case, predicted variance *decreased* with increasing scale, thereby invalidating this simple picture of reality and requiring the development of alternative models (Evans *et al.*, 1976). These models in turn will require further testing before they can be used in a fisheries management context.

While this example illustrates that a single critical test can invalidate a model, there is no predetermined number of tests that will establish a sufficient degree of belief in it. This depends on the use to which the model will be put.

SOME CAVEATS

Two caveats must be mentioned with respect to treatment of historical observations. The first is that comparison must be carried out with verified observations, not with second-hand interpretations or impressions. It is appalling how often in ecology we find that supposedly well-established past observations or case examples turn out to have been badly distorted by well-intentioned researchers wishing to support some hypothesis or to report something interesting. One example of this is the Kaibab Plateau deer irruption reported in most ecology texts. There is now good evidence that it never occurred at all (Caughley, 1970). Another example occurred in our own budworm work (Chapter 11), where the model predicted that forest volume would decline independently of insect damage, while it was "common knowledge" that volume was high and would remain so if insects were controlled. We spent 2 months checking the model for errors when we should have been spending 2 days looking at the available raw data on forest volume. When we belatedly took this obvious step, the model was vindicated and "common knowledge" was shown to be at variance with the data on which it should have been based. We suspect that this is not a rare occurrence.

The second caveat is the obvious one that correlation does not imply causation. Lack of reasonable model correspondence with the historical picture speaks strongly for invalidation. But the achievement of such correspondence, while gratifying, really only lets us move on to the next step in the process. It does not "validate" anything, and it tells the manager precious little about how much he should believe in his model as a predictor of future impacts. This is true because practically any complex model can be "tuned" to fit practically any given pattern of historical data. Since the causal structure of such a "tuned" model need have nothing in common with that of the real world, its predictions under the new conditions of development or management are highly unlikely to correspond to reality. This situation is similar to the well-recognized danger of extrapolating (or, for that matter, interpolating) from general polynomial regressions to situations outside the range of observations.

MODEL STRUCTURE

A few additional points regarding the relationship of model structure to the invalidation process should be mentioned here.

Our view of model building emphasizes the advantages of modeling in terms of causal or "functional" components. To the extent that such causal modeling is possible, one's ability to assess the resulting model's credibility will be greatly enhanced. Although belief must certainly relate to the total model's prediction, it is also a function of the logical consistency and clarity of the model's structure. Relationships involved in the prediction should agree at least qualitatively with experimental experience. Biological relationships should make sense when interpreted in terms of lower levels of organization (physiology, behavior); economic relationships involving market situations should be consistent with known behavioral characteristics of firms; and so forth. In short, it should be possible to see how the predictive model could arise by aggregation of more detailed components than those actually employed. If the model is not cast in the form of functional components, then the path to establishing credibility is obscured – we lose the benefits of analogy in understanding the model. We will show in the next two sections that when the model has been causally structured, its comparison with historical evidence and alternative models is also greatly facilitated.

Finally, we have one observation regarding model structure that is very much at odds with conventional wisdom. A great deal of present practice in environmental management and impact assessment modeling implies that the more detailed the model structure, the more boxes and arrows and variables considered, the better will be the model's predictions (e.g., Goodall, 1972). Our own experience and other explicit tests of this notion (Lee, 1973; O'Neill, 1973) suggest that it is often, perhaps systematically, false. Those scientists, managers, and administrators who call automatically for more detail often produce giant reports rather than useful predictions. As emphasized in Chapter 6, it is not detailed complexity but rather comprehensible simplification that gives rise to understanding. And it is on understanding alone that a critical assessment of model credibility must ultimately be based.

EVIDENCE FOR INVALIDATION

TRIAL-AND-ERROR EVIDENCE

Historical data reflect behavior of the system only within the narrow range of circumstances encountered in the past. New programs or developments will change those conditions, and our principal concern is in the believability of the model's predictions for the new situations. We are, after all, interested in a management model. In order to assess the model's credibility as a predictor of new management impacts and future uncertainties, we need to assess the range of possible behaviors over which the model is applicable.

The usual but often impractical approach to this problem is explicit trial-and-error. For example, our model might predict that if a proposed equipment restriction is implemented in a particular fishery, then fish harvest will decrease by 20 per cent. If we adopt the new equipment restriction policy in an actual fishery and the predicted harvest decrease occurs, then our belief in the model's predictive ability is appreciably enhanced.

The problem with trial-and-error evaluation of predictive limits is that it always takes time, is frequently expensive, is limited to the particular trial undertaken, and often risks disaster if the predictions prove wrong. Nonetheless, the potential benefits of combining operational activities with experimental goals may be great enough to justify or even demand trials. The rationale for considering such experiments as an integral part of the management program is discussed in Chapter 10 and is treated at length by Walters and Hilborn (1976) and Peterman (1977b). When opportunities for trial-and-error invalidation of the model are limited, however, we must look for natural trials as well.

NATURAL TRIALS AND EXTREMES OF SYSTEM BEHAVIOR

Useful natural trials exist wherever there are examples of ecological or environmental systems that are similar to the one we have modeled but that exhibit qualitatively distinct behaviors. In reference to three of the case studies in Part II, we might look for comparable situations where an alpine village still farms its potential hotel land; where a salmon stream provides unusually high yields; or where a previously mined area supports a particularly low diversity of wildlife. If minor, plausible changes in the parameter values or structure of the model replicate these extreme forms of actual behavior, then the range and degree of belief in the model as a predictive tool under future extremes of management and uncertainty are enhanced accordingly. We at least gain confidence that no significant component of the system has been left out.

The procedure for comparing the model with the results of natural experiments is best conveyed by example; we draw again upon the budworm–forest management study. As noted above, the original budworm model predictions corresponded well with the historical patterns of insect outbreak in the Canadian province of

New Brunswick. But an explicit search for atypical behaviors uncovered some patterns that did not match the New Brunswick norm (Holling *et al.*, 1975). In northwestern Ontario, for instance, outbreaks are more intense and tend to occur at intervals of 60 or more years rather than the 30–45-year period observed in New Brunswick and predicted by the model. The principal differences between the regions are a lower proportion of susceptible trees and better weather for budworm in northwestern Ontario. When these differences were introduced into the New Brunswick model, the Ontario behavioral pattern was reproduced.

A similar opportunity for invalidation was presented by consideration of outbreak histories in Newfoundland, an island more than 200 km off the New Brunswick coast. Historically, outbreaks there have been extremely rare and short-lived. This pattern changed only recently, coinciding with management activities in New Brunswick that produced an increased outbreak frequency there and consequently a source of emigrating budworm. In Newfoundland, the proportion of susceptible trees is greater than in New Brunswick, but the weather is worse for budworm. Again, these parameter changes were introduced into the New Brunswick budworm model, which then predicted the very rare, very brief outbreaks typical of Newfoundland. When pulses of immigrating budworm from New Brunswick to Newfoundland were also introduced into the model, the predicted outbreak frequency, though not the duration, increased, again matching actual behavior in the real world.

A final invalidation test consisted of adding to the basic New Brunswick budworm model a management submodel mimicking insecticide application and harvesting activities introduced there in 1950. This test, described in detail in Chapter 11, showed that the unprecedented outbreak pattern actually experienced in the 1950s and 1960s could in fact be reproduced by the basic biological model linked with the management rules.

The set of extreme behaviors tested during the invalidation studies directly increased our belief in the model's predictive abilities under a range of weather conditions, susceptible tree densities, and insecticide-induced mortalities. Indirectly, these tests supported a provisional belief that the model's credibility was not limited to the narrow range of circumstances defined by local history.

The sort of highly qualitative natural experiment or "extreme behavior" data necessary for invalidation studies almost always exists. The manager's challenge is to find the data and mobilize them in spite of the invariable insistence of the scientists and specialists that they do not know enough to say what the effects of extremes will be. The result is usually worth the battle.

THE ANALYSIS OF ALTERNATIVE MODELS

THE NEED FOR ALTERNATIVE MODELS

A model could make all the testable predictions referred to above and still be the wrong representation of reality. The chance always exists that other models will meet these historical tests equally well but give very different predictions of future impacts or management success. For example, budworm outbreaks could be largely caused by changes in the nutritional quality of the foliage or by changes in the genetic structure of the insect population instead of by the interaction among predators, parasites, and budworm as presently formulated in the model. We can never eliminate the possibility that these other models could adequately represent historical observations, but we can take further steps to refine our degree of belief in the impact predictions of the model(s) upon which decisions must finally be based. The basic approach is to design alternative models of the system under study.

The critical need to seek alternative interpretations (or models, or explanations) rather than try to seek validation of any single one is most obvious in the statistical concept of "the power of tests." We can establish belief or disbelief in any hypothesis only by reference to some alternative. The closer the alternative is to the original hypothesis, the more difficult it becomes to tell which one is more likely to be correct with a given set of data. When we make only a vague assertion like "this model must be wrong because it is too simple-minded" (or too complex, or whatever), we must have at least some criteria by which to judge "rightness" or "wrongness"; that is, an alternative model that predicts better or worse than the model being examined.

The greatest hope of any search for alternative models is always to find one that passes a greater number of significant invalidation tests than the original. Failures are almost as useful as successes, however. Each alternative considered and rejected on the basis of available evidence eliminates one way of modeling the impact problem that might well have been acceptable but is now known to be wrong. The general goal of the comparison exercise is to generate two lists from the alternative models considered: models rejected, and models possibly useful for prediction. The characteristics of these lists – specifically, the range of alternatives considered, the plausibility of the rejected models, and the variability in results of the remaining (unrejected) models – will strongly influence our degree of belief in the eventual impact predictions. This degree of belief is one of the most significant pieces of information communicated to the decision makers. We will first discuss these properties of alternative models and then outline some specific ways of generating candidate alternatives.

PROPERTIES OF ALTERNATIVE MODELS

Range

The greater the range of the models considered, the more confident we will be that the ones offering adequate explanations of historical data are in fact good models on which to base future predictions. By a wide range of models, we mean models that involve a variety of different assumptions about how the causal mechanisms are represented. For predicting effects of salmon enhancement, for example, one might consider a model that assumed that salmon populations were largely limited by mechanisms operating during their stay in fresh water, or an alternative one that emphasized mechanisms in the marine environment.

Clearly, one of the most valued and effective traits a manager can possess is his ability to see (and therefore to model) a problem from a wide range of perspectives. In practice, most interpretations (i.e., models) offered for a problem tend to be shaped by habitual ways of thinking, and effective "new looks" are most difficult to establish. Consensus-breeding techniques are your enemy in this situation, and imagination is your only sure friend. A few technical crutches for broadening the range considered are discussed below, in the section on generating alternative models.

Plausibility

Clearly, if we cannot (or cannot be bothered to) imagine any alternatives, then we might as well not have a model at all. This is just the same as saying "any model will do, none predicts better than others." Equally clearly, however, it is not the sheer volume of alternatives considered by the end of the study that counts. If we go out on the street and ask the first ten people (or ten consultants) for their opinions (i.e., models) on the relationships of age structure and land tenure to erosion in Obergurgl, their predictions should not affect our belief in the model one way or another. What counts is not the number of silly or trivial alternative models discarded, but rather the number of plausible ones. The real payoff comes when we can generate alternative models that give credible performance for all our historical tests. Critically designed experiments may allow rejection of some of these models, adding substantially to the credibility of those remaining.

Variability

When a broad range of models has been considered, a set of plausible alternatives identified, and a number of these rejected on the basis of available evidence, there will generally remain several different models. Any (or all, or none) of these might provide a realistic basis for predicting future impacts, but we have no way of choosing among them. To the extent that all the remaining alternatives give the same predictions, there is no problem. If the alternatives give different predictions, then there exists a problem of choice under uncertainty. You may elect to reduce the uncertainty through further data collection and experimentation or as part of your management program (Chapter 10), or to consciously gamble on the basis of other factors influencing your belief in one or another of the alternatives. Finally, you may seek to change the development or management program so as to minimize the variability and uncertainty of impact predictions. These are problems of evaluation and choice rather than invalidation *per se* and will be taken up again in the next chapter. One invalidation issue does remain, however.

Almost all parameters in almost all environmental or ecological models cannot be fixed exactly. It is often convenient, nonetheless, to treat them as though they were fixed throughout most of the analysis, using mean or, occasionally, extreme values for model predictions. Before these predictions can be "believed," however, it is necessary to examine their sensitivity to realistic variation in the parameter values. Such variability in parameter values is to be expected as a result of measurement errors or future variation, and if the predictions change radically as a result, then these predictions must be treated very cautiously during assessment.

Some authors (e.g., Miller, 1974) claim that the most "valid" ecological models are those with predictions that are least sensitive to changes in parameter values. But both ecological systems and the models that realistically reflect them may in fact be acutely sensitive to small differences in their structure or parameters (Gilbert *et al.*, 1976). In the budworm and many other insect–plant systems, for example, it is clear that differences of a few days in temperature-dependent development rates can determine whether a potential host plant species is fed upon at all by a particular defoliating insect. Thus, the question is, given a set of best estimates and measurements of parameter values, how sensitive the resulting model's predictions are to changes in those parameters.

The techniques of sensitivity analysis are well known and have been applied to a number of impact assessment models (Ackerman *et al.*, 1974; Hamilton *et al.*, 1969). It should be noted, however, that simultaneous variation of the parameters in question is necessary to give reliable results. A good example of this is given in a study by Scolnik (1973) on the Meadows world model. Conventional analysis had shown the model's predictions of population boom and collapse to be stable to small perturbations in many parameters. But when several parameters were simultaneously varied over ranges of less than 10 percent, the results changed dramatically, giving an increase of populations to a density that was maintained thereafter. Since simultaneous variation of the parameters is to be expected in the real world, the model's predictions of catastrophe are not necessarily credible.

An opposite result was reported by Herrera *et al.*, (1976), who examined the agricultural sector of the Latin American World Model for sensitivity to small simultaneous variation in the parameters. In this case, the model predictions were found to be stable and therefore comparatively believable, even in the face of a search for "worst case" combinations.

Where acute sensitivity to small changes appears to be a true property of the system under study and not simply an artifact of the model, the only recourse is to seek management policies and programs that can tolerate the range of possible variation.

GENERATING ALTERNATIVE MODELS

At one extreme, the notion of alternative models can be approached by conducting independent workshops from independent data bases, independent assumptions, and independent perspectives, each generating an independent set of hypotheses or models. However, a multiworkshop model approach is usually prohibitively inefficient and expensive, and a more practical view of the alternative model issue is necessary.

The most obvious set of alternative models to consider are those implied by the issues left unresolved or the components deliberately excluded during development of the process model (Chapter 4). Recall that during model development explicit lists were kept of (a) those things that were left out of the analysis because of bounding considerations and (b) the functional relationships and parameter values for which reliable data were least available or disagreements most acute. We now construct alternative models for comparison with our original by adding the suspect factors initially left out and exploring the most likely alternative functional forms and parameter values. This process creates a number of "plausible" alternative models, fairly similar in structure and predictions to the original. Some will be rejected on the basis of comparisons between their predictions and available data; others will be retained for use in the evaluation exercise.

For example, in a lake model we have worked with, it was thought necessary to calculate nutrients added to the water by zooplankton and fish excretion. However, when these calculations were added to the simpler model, virtually no difference was seen in the overall system behavior because the amounts of excreted nutrients were an insignificant fraction of the total nutrient inputs from the watershed. In another model, it was thought that caribou feeding on snow-covered lichens during winter did not cause intraspecific competition. However, when the effect of feeding behavior on the trampling and packing of snow in the surrounding area was added to the model, very different results were obtained. In fact, one of the most critical parameters in the model turned out to be how much food was made unavailable through compaction of snow per unit of food eaten (Walters *et al.*, 1975).

The models produced by examining the workshop bounding and choice decisions may well span a fairly narrow range of alternative structures. In order to expand that range so as to better assess the limits of credibility, it is necessary to develop more extreme alternatives of model structure and to explore their predictive consequences. Our experience suggests that if the initial model is in fact a very good representation of reality, then most of its extreme structural variants are likely to make very bad predictions. But only by actually verifying that this is the case can we develop a confident belief in a given model's credibility.

The method for generating these extreme structures is essentially that of systematically adding entire functional components or processes to a basic version of the model and removing others. In the Obergurgl study we examined the consequences of such functional components as the effect of ski-lift construction on farming or on the perception of erosion by summer tourists. In the budworm analysis very substantial insights were gained from the alternative models developed by adding vertebrate predation and removing dispersal processes. In fact, the addition of predation effects produced such markedly superior predictions that the "best-guess" model was revised accordingly. The detailed budworm case study (Chapter 11) further shows how the qualitative, simplified model forms discussed in Chapter 6 can be used to facilitate the generation of extreme types of model structure.

When you have finished the invalidation procedures, you will not have a valid model, you will not have eliminated all uncertainties, and you will not even know probabilities. However, you will have a critical understanding of the weaknesses and strengths of available models that is extremely valuable. You will be able to meet criticisms that "such-and-such was left out" by saying why and what difference including them would have made. Most important, by understanding both the extent and limits of your models' predictive capabilities, you will be able to proceed with the design and evaluation of development proposals in the most responsible manner possible.

8 Evaluation of Alternative Policies

The invalidation process generates one or several models that elicit the greatest degree of belief. These models can then be used to predict impacts and to compare different ways of management. Some traditional environmental assessments consider only a single proposed development or management scheme. We argue that alternative development programs should always be considered because there may be other ways to achieve the desired goals while avoiding some disadvantages of the original proposal. Thus, the process of choosing between alternative development schemes becomes analogous to choices faced in resource management problems in general, such as choosing between managing a population by setting kill quotas or by directly controlling hunting effort.

Before going further, we should clearly define our usage of some terms that have rather varied meanings in practice.

Actions Specific deeds available to the manager of some environmental system. For example:

- Harvest trees
- Release x cubic feet of water from a reservoir
- Spray insect pests
- Build a fish hatchery

Policies Rules by which these actions are initiated. They state at what time or under what conditions actions are taken. For example:

- Cut all trees above a given age
- Spray insects when populations surpass a certain density
- Release enough water from a reservoir to maintain a given minimum flow downstream

Indicators Measures of system behavior in terms of meaningful and perceptible attributes. For example:

The number of trees of harvestable size
The crop loss due to insects
The stored volume of a reservoir
The costs of a program

Preferences The trade-off rates between one indicator and another.

Objectives Desired goals in terms of indicators. For example:

The reservoir to remain at least 90% full
The catch to sport fishermen to stay above 1965 levels
The cost of management to grow at a rate less than the national budget

One should remember that decision structures are hierarchical, and what is a goal at one level in the structure may be a policy at the next higher level. For example, a manager of a fishery of a given species has a harvest goal that he attempts to achieve by regulating the number of days open for fishing, the allowed gear types, and so forth. But his harvest goal is only a part of the policy designed at a higher level to achieve a broader goal of maximum sustained yield over many stocks.

We view evaluation as the entire iterative process of combining actions into policies, using a model (or some other predictive device) to enact the policies and generate time streams of indicators, and using objectives to choose among the different time streams of indicators.

The traditional view of evaluation assumes that there is a given set of management objectives and decision preferences. It sets out to characterize these in a quantitative fashion, to reduce them to a single measure, such as a cost–benefit ratio, and then to rank several policies from "best" to "worst" according to this measure. The rankings are then presented as a list to the decision maker. However, this traditional outlook is static and fundamentally inadequate for adaptive environmental management and assessment.

The approach we have used treats evaluation as an essentially adaptive communication process. It assumes that neither policies nor objectives are immutable and that the critical assessment and modification of both is one goal of the analysis effort. It therefore concentrates on those aspects of evaluation that promote understanding rather than on the numerical products – products that all too easily become goals in themselves.

So defined, adaptive evaluation takes on a broad and varied character with which we shall not deal in any systematic fashion in this book. Rather than presenting a superficial overview, we have chosen to discuss in detail two fundamental aspects of adaptive evaluation – namely, indicator generation and an informal process of policy comparison. These we view as both essential and feasible steps for every assessment. In addition they constitute the foundation of attitudes and understanding upon which any critical application of more subtle evaluation concepts must be based.

Utility analysis and objective functions, discounting and intertemporal trade-offs, uncertainty, and conflict resolution are some of the many evaluation topics you will *not* find treated here in any depth. We *do* feel that they are important – often critically so – and we have therefore included a brief review of some of our own experience toward the end of this chapter. The case studies document this importance in more detail and illustrate some of the benefits and pitfalls inherent in the various techniques. This experience has left us with strong biases regarding the opportunities for use and abuse of commonly advocated numerical evaluation techniques. In the last section, the more obvious of these biases are explicitly stated along with a few key references to further reading on the subject. It is essential to emphasize, however, that we believe that no one, including ourselves, is yet equipped to write a general "how to" manual for applying the more complex techniques of evaluation to environmental assessment and management. The issues involved are subtle in the extreme. You will need expert help, and the experts will disagree profoundly on each subject. This is not necessarily a bad thing, provided that you can use the disagreement to stimulate dialogue and communication. Here, perhaps more than in any other aspect of environmental management and assessment, it is the adaptive process rather than the numerical product that should be your pre-eminent concern.

INDICATOR GENERATION

The first requirement of evaluation is a suitable language or vocabulary to describe objectives and the outcomes that result from applying given policies. Up to now we have dealt with this issue rather informally, usually describing the output of assessment and modeling activities in terms of fundamental "state variables" such as number of fish or proportion of trees over a given age. But socially relevant and responsible evaluations cannot be based on the behavior of these elements alone. State variables must be translated into a broader set of indicators relevant to those who make, and those who endure, the ultimate policy decisions. Indicators can usually be broken down into a few broad but overlapping classes – e.g., ecological, economic, recreational. Several examples are given in the case studies, and a typical list drawn from the budworm analysis is shown in Table 8.1.

Appropriate indicators for evaluation are readily generated in any assessment problem, provided that an essential constraint is understood: there is no "comprehensive" list of indicators, and there is no "right" set of indicators for any problem, ever. This is the same issue encountered earlier in our discussion of choosing variables to include in a model. There we stressed the importance of bounding many variables *out* of the dynamic model to make it parsimonious and more understandable.

Evaluation is also essentially a model formulation process in which we develop ways to prescribe "better" policies. Therefore, attempts to include everything as an indicator will likewise result in an incomprehensible and misleading monstrosity, rather than an aid to assessment. This attitude is implicit in the "looking outward"

TABLE 8.1 Examples of Indicators of Known Interest Taken from the Budworm Case Study

Socioeconomic Indicators
Profits to the logging industry
Profits as a proportion of total sales
Cost per unit volume of harvested wood
Cost of insecticide spraying
Unemployment rate reflected by the proportion of mill capacity utilized
Resource Indicators
Volume of wood in trees older than 20 years
Volume of wood in trees older than 50 years
Volume of wood harvested
Proportion of total volume harvested
Volume of wood killed by budworm
Mill capacity
Total forest volume
Environmental Indicators
Visible damage due to budworm defoliation
Damage due to logging operations
Age class diversity of the forest
Number of high quality recreational areas
Insecticide impact in terms of fraction of province sprayed

approach to modeling presented in the chapter on orchestration (Chapter 4). Indicators, like variables, are included in the analysis when knowledge of their behavior is essential if the model is to respond to somebody's major policy choice or design question. When there is no client or potential user demanding the indicator, it is usually best to omit it from consideration. Of course, this presents a danger of leaving out something important and perpetuating habitual viewpoints, just as it did in the modeling work. One must use judgment and occasionally err on the side of inclusion. But, as we will argue below, implicit or explicit simplification to a few indicators is ultimately necessary for comprehensible comparison of alternative policies and objectives. There is consequently little to be gained from amassing huge lists in order "to be safe."

The "looking outward" criterion for indicators cuts two ways, however. It is not uncommon to find that an indicator that is clearly relevant to policy choice simply cannot be predicted with available models (e.g., the types of gear that will be used on fishing boats or the world demands for wood pulp). Sometimes the models can be changed, but often this is not feasible. The only defensible response in this situation is to record the indicator explicitly in a list of "things left out" and to weigh its significance and bearing on the policy choice question independent of the model part of the analysis. This might be accomplished by mobilizing expert opinion, by interfacing with other models or experience, or by some other means of resolution. An excellent example of the second approach is provided by Baskerville (1976). He used the budworm–forest model presented in the case studies to describe the effects of various management policies on forest harvests and inventory. The significance of these predictions for employment and industrial profitability was then evaluated through an independent economic analysis, using the model's forest inventory data as inputs.

INITIAL COMPARISONS OF POLICIES

Once the basic indicator set has been defined for an assessment problem, each decision maker can select those indicators of personal interest and compare their performance under alternative policies. Although there are rigorous techniques for making such comparisons, we find that simple visual inspection of the projected time series of the indicators is often a powerful and unambiguous first step in the evaluation process. Sometimes it is clear that certain policies dominate – they are better in all respects. More commonly, some policies will exhibit obviously desirable outcomes for a few indicators and indifferent or undesirable outcomes in others. For example, certain reservoir discharge policies will keep downstream water flow rates high for trout, but will also create a large, recreationally undesirable band of muddy lake shore.

Traditional static evaluation procedures seek to provide a common denominator or metric for ranking such complex alternative outcomes (cost–benefit ratios, dollar values, utilities, and so on). But we have found it useful to highlight the *differences* among indicators, at least initially, and to use these differences as starting points for policy modification and improvement. If we use the "laboratory world" of the assessment model, policies with complementary strengths and different weaknesses can be combined in an iterative, experimental effort. In this manner it is often possible to achieve more uniformly desirable indicator performance through "hybrid" policy design. A great deal of exploration of alternative policies can be made in this manner without worries about formal schemes of indicator combination or the rendition of objectives into numerical form. Furthermore, the process of policy comparison through direct reference to the individual indicators is the least ambiguous evaluation technique available. What it lacks in refinement is more than compensated for by the clear communication of relevant information.

As an example of this approach, we return to the budworm management policy evaluations mentioned earlier. Extensive experimentation with the system model and interviews with relevant decision makers identified five of the indicators listed in Table 8.1. as primary. The values assumed by these indicators in a simulation of the management policy historically used in New Brunswick are given in Figure 8.1. In an attempt to improve this policy, new spray and harvest rules were developed

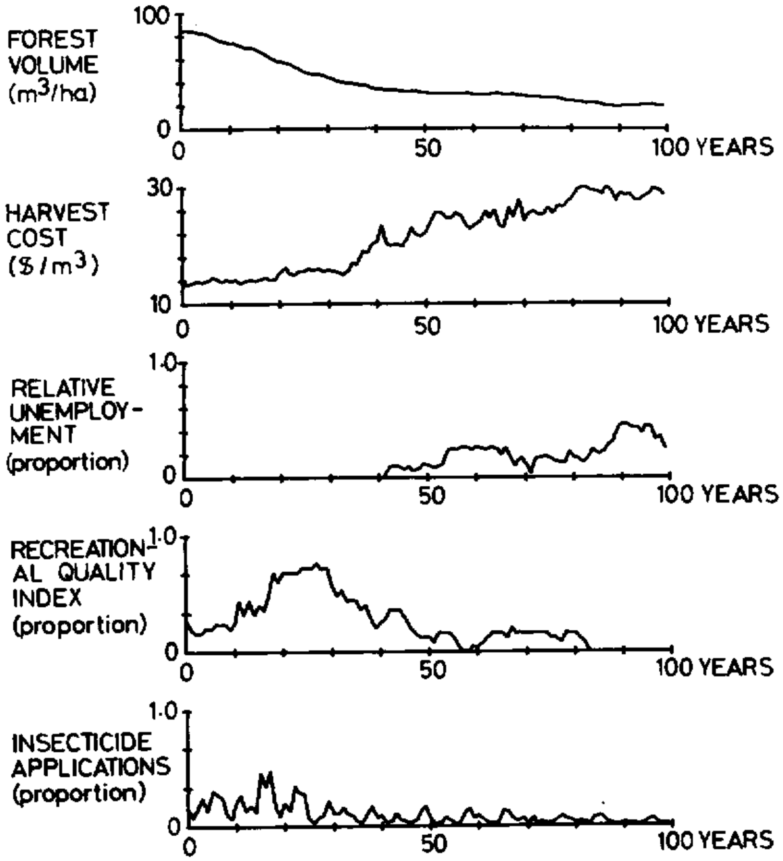


FIGURE 8.1 Value of indicators that resulted from the historical budworm management rules.

and then tested on the simulation (see the case study section for further details). The results, presented in Figure 8.2, show improvement in some indicators, notably total forest volume, profits to the logging industry, and recreation, but a somewhat worse situation with regard to employment and insecticide spraying. Without performing any but the most trivial analysis, we can say that it would be nice if a policy could be found that preserved the gains of this alternative policy, but repaired its failures.

A modification of the alternative policy was next designed, explicity tailored to decrease spraying by cutting down trees threatened by budworm. The results in Figure 8.3 show that spraying frequency is indeed reduced, but at a cost of even more irregular employment due to the sporadic antibudworm harvest. The "good" forest volume, harvest cost, and recreational performance have been reasonably

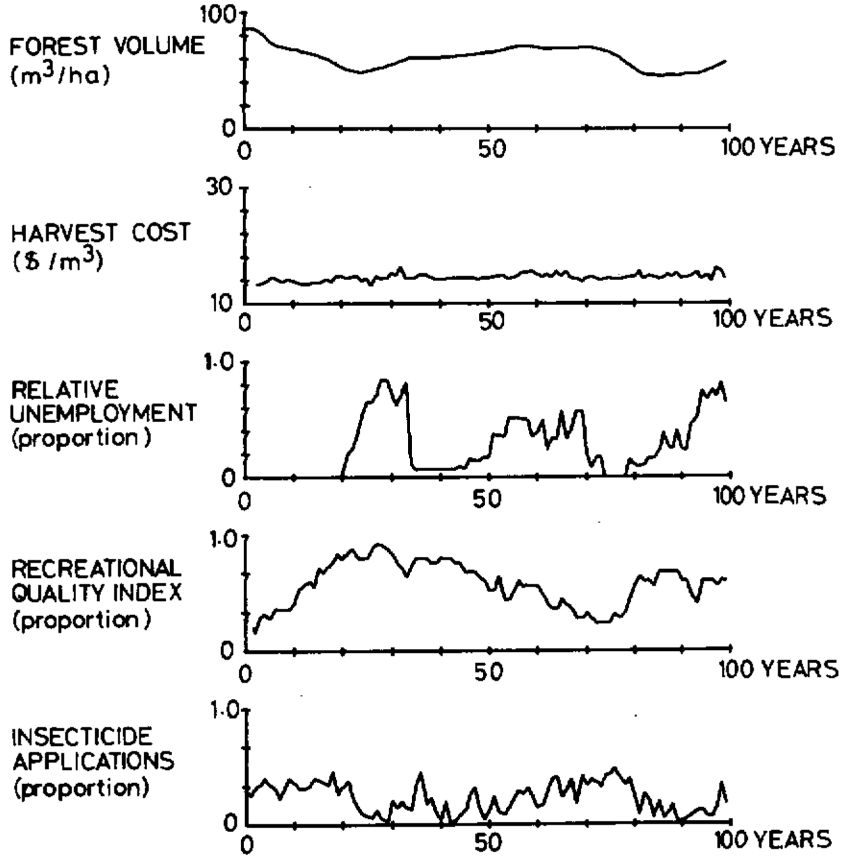


FIGURE 8.2 Value of indicators resulting from proposed management rules: first alternative.

maintained, however. Since any preventive harvest scheme seemed likely to incur this disadvantage, we searched elsewhere and attempted to reduce spraying by adding a hypothetical but realistic budworm virus to the model. As shown in Figure 8.4, this succeeded in reducing spraying substantially without radically increasing unemployment. Forest volume was better than with any other policy, and recreation was superior to any but the antibudworm harvesting policy.

At this point detailed utility analyses (quantitative statements of preference) could be made to identify the "best" of these four policies (see the next section). A good deal of careful study would have to be made of implementation costs and feasibility as well as of model reliability before such rankings would be meaningful. But to insist at this stage on a formal ranking would be to miss the whole point of adaptive evaluation. The benefits of the exercise just described are not

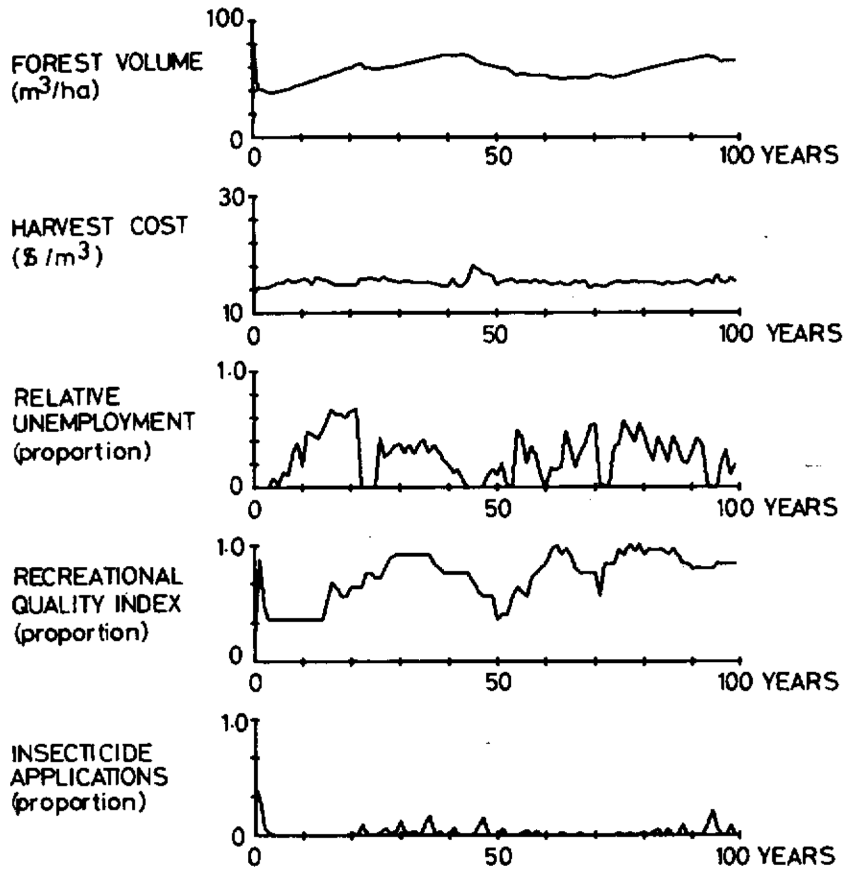


FIGURE 8.3 Value of indicators resulting from proposed management rules: antibudworm harvest.

found in the development of a ranking scheme, but rather in the design of policies for meeting specified objectives through creative exploration of policy alternatives.

FURTHER COMPARISONS

When the number of alternative policies becomes large, the problem of comparison and evaluation can hamper creative policy design. When the decision maker, or any interested party, embarks on a policy evaluation process, it is critically important that the trade-offs and compromises between competing policies (in terms of alternative indicator patterns) remain as visible as possible. If the evaluation process is too quickly given over to some numerical methodology, then important

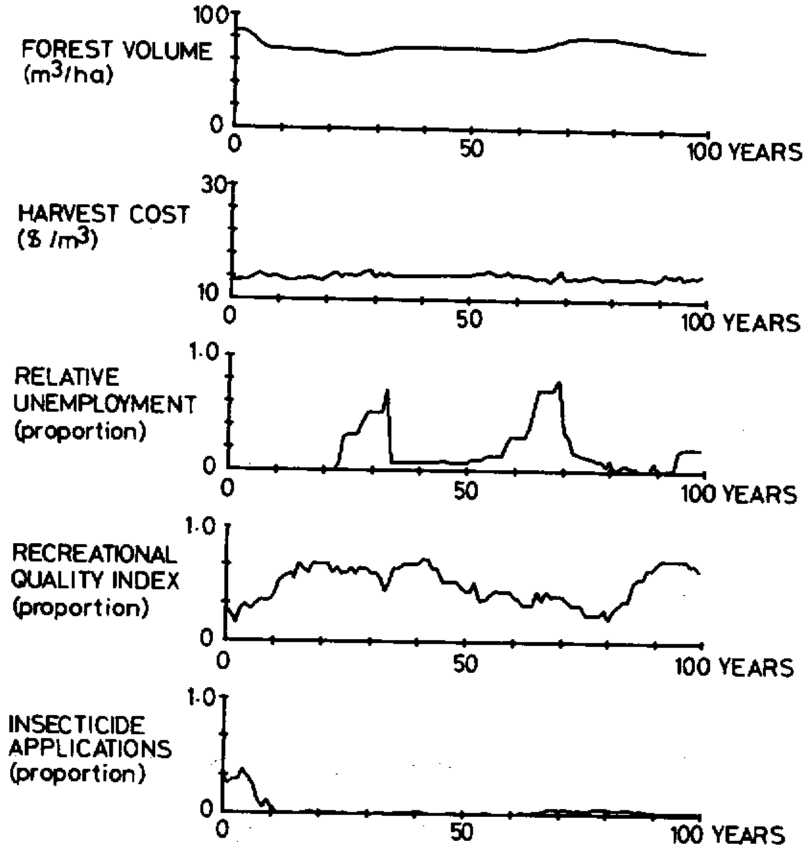


FIGURE 8.4 Value of indicators resulting from proposed management rules: virus addition.

opportunities for the exploration of preferences and discovery of objectives will be missed.

One approach to promote the dialogue between a manager, his problem, and groups that would influence policy involves a condensed graphical presentation of indicator values in a form that allows any user to have "hands on" access to the evaluation. The technique is, in effect, a "management slide rule" that can physically be pushed and pulled and moved about to reveal the consequences of different policies. This technique, sometimes called nomogram or isopleth diagram, is described in Gross *et al.*, (1973) and Peterman (1975). Examples of its application are given in the case studies and Appendix A.

Because the use of nomograms is integrally linked with the whole topic of communication, we defer explicit outline of their construction and use to the next chapter. It is sufficient to note here that they have proved extremely useful for policy evaluation. With these nomograms the manager can explore the consequences of different acts by manipulating the graphs himself. He can add political, economic, and other constraints, identify trade-offs, and begin to evolve realistic compromise policies. Done jointly with a number of interest groups, this becomes a powerful instrument for constructive dialogue and even conflict resolution. Because several indicators can be treated simultaneously using this technique, managers have found it an effective aid in learning to appreciate and creatively manipulate the intricate relationships among policies and indicators (Peterman, 1977a).

These graphical methods for preliminary comparison of policies and their resultant indicators quickly point out to the user (be he analyst or decision maker) the need to articulate goals and preferences clearly in order to make meaningful comparisons among the alternatives. More often than not, the methods provide all the technique necessary for evaluation in the adaptive management process. Under certain conditions, however, more quantitative considerations may be justified. We discuss some of the associated issues below.

As promised earlier, we now introduce some of the more subtle problems of evaluation by way of example. We emphasize that the techniques discussed in this section do more harm than good if employed superficially or uncritically. Expert advice from someone who appreciates or can be taught the needs of adaptive evaluation is mandatory. If this is not available (or believable), you will do best to stay with the solid and straightforward techniques already discussed. They are probably sufficient for most evaluation needs, anyway.

UTILITY ANALYSIS

When there are numerous indicators of interest, a quantitative method for defining preferences may be necessary. Utility analysis permits an individual (or an interest group) to define two things: first, the "satisfaction" or "utility" gained from different values of an indicator; and second, the trade-offs between indicators. In the case study of salmon fisheries, for example, a nonlinear saturating relationship was typically found to reflect the utility for different amounts of sport catch (Figure 8.5). (This is the case because adding 100,000 fish to the catch when the catch is small increases utility more than adding 100,000 fish when the catch is very large; demand becomes saturated.) There are formal questioning procedures to help a person define such utility functions (Keeney, 1977; Keeney and Raiffa, 1976) and these procedures can be repeated for all indicators of interest. It is also possible through another series of questions to determine trade-offs between indicators. For example, one could ascertain how much of a decrease in *utility* of commercial fish catch would be traded for a 20 percent increase in *utility* of native Indian fish catch.

In the salmon study the resulting quantitative description of objectives differed among interest groups such as commercial fishermen, sports fishermen, packing companies, and fisheries managers (Hilborn and Walters, 1977). These utility

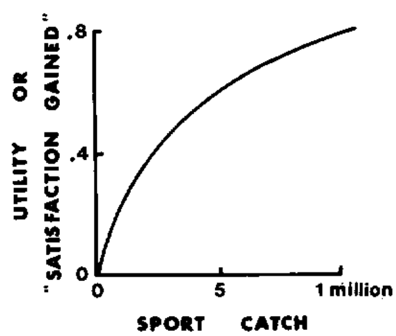


FIGURE 8.5 Example of a utility function for sport catch.

functions were then used in conjunction with appropriate indicators to determine which of the alternative management schemes would give the highest utility to each interest group. However, objectives are never fixed forever; new concerns may arise and interests may shift (witness the sudden importance of environmentalist viewpoints in recent years). While it is possible to deal with changing objectives (see the following section), the assessment team and the recipient of its report should recognize the dangers of ranking policies on the basis of fixed quantitative utility functions. By far the greatest benefit of the utility analysis process in our studies has been the triggering of dialogue about goals within and among interest groups (Hilborn and Walters, 1977; Hilborn and Peterman, 1977). Often, people are stimulated to articulate or at least think about their goals much more clearly than before, a useful result in itself.

More extensive utility analyses were conducted with affected decision makers in the budworm studies (Bell, 1975a, b). Again, the important result was not the production of some ultimate utility function but rather the creation of dialogues among the various decision makers, and especially between them and the analysts (Baskerville, 1976).

UNCERTAINTIES

Three major kinds of uncertainty are relevant to the evaluation process, and we will briefly discuss possible directions for each. Again, there are no simple solutions.

First, there is the uncertainty in objectives, which can change over time. A policy that was determined to be best for achieving one objective might be totally inappropriate for some new objective. Thus one must perform an analysis of the sensitivity of each policy to specified changes in objectives.

Just as we suggested a wide-ranging approach for formulating alternative models in earlier chapters, we also recommend consideration of extremely different future objectives here. Model results can be evaluated with these different objectives to identify preferred policies. A delicate issue may arise from such an exploration. A certain policy may be the least sensitive to a plausible range of objectives, but that policy may produce slightly worse indicator values than another policy. Which policy should be chosen? There is no fixed answer to this question; the decision maker must rely on his judgment of the likelihood that objectives might change.

The second sort of uncertainty arises from model assumptions. If several alternative models with different assumptions emerge from the invalidation process, then evaluation of alternative policies should be made with each of these different models in turn. If one policy comes out best under all assumptions, there is no additional problem. If, however, "best" policies are different under alternative assumptions, then the decision maker must again rely on that all-important measure of "degree of belief." When the degree of belief is not significantly different for the different assumptions, critical experimentation or data collection becomes necessary.

Third, there will always be some deviation from desired results. For example, a policy designed to result in a fish catch of 140,000 may actually result in 185,000 fish being caught, or a desire to maintain a minimum water flow from a dam of 400 ft^3/sec may actually result in a flow of only 300 ft^3/sec. The question is, "How significant are these deviations?" If the desired fish catch was near the overexploitation point, there might be serious consequences in terms of the indicators. The assessment model can be used to explore the effects of these "control errors," but only in a fairly haphazard way. Again, only a wide exploration of different possibilities can help minimize the likelihood of later surprises.

TIME HORIZONS AND DISCOUNTING

Finally, there is the problem of treating time in evaluations. Should the indicators produced by each policy be examined over a 10-year period or over 100 years? Should these yearly values simply be averaged, or should some years be discounted more heavily than others?

These issues are critically important in determining which policy alternatives will seem most appropriate for selection. We illustrate in Chapter 11 several policies that look good for controlling budworm in the short run but are clearly disastrous when their longer term consequences are included in the evaluation. In another study, Fox and Herfindahl (1964) re-evaluated 178 water resource development projects undertaken in 1962 by the U.S. Army Corps of Engineers. These projects represented a combined initial investment of over $3 billion and were all characterized by benefit–cost ratios of 1.0 or more when evaluated at the prevailing prescribed discount rate for federal project costs of 2.6 percent. Fox and Herfindahl re-evaluated the projects at discount rates of 4, 6, and 8 percent and found that the project adoption decision was reversed (i.e., the new benefit–cost ratio dropped below 1.0) for 9 percent, 64 percent, and 80 percent of the projects, respectively.

Similarly powerful cases for the dominating influence of time stream aggregation assumptions may be found in Baumol (1968), Krutilla (1969), and Koopmans (1974).

The theoretical literature on discounting and intertemporal evaluation in general is a perennial mess. Good examples of the prevailing arguments are assembled in Joint Economic Committee (1969), Layard (1972), and Lind and Greenberger (in press). We do not pretend to address the technical issues here, other than to note that there are excellent formal grounds for *not* applying the same discount rate ("market" or "social") to all evaluation problems (see, e.g., Feldstein, 1964). Our own biases and experience argue strongly that – subject to certain technical constraints of consistency – the choice of the "appropriate" time horizon or discount rate for evaluation is essentially a political or even ethical question. There simply is no extrinsically defined "technically correct" answer to questions like "How many fishermen should be put out of work today in order to increase the chances that their children will still have a healthy fishery available?"

Our own approach to this dilemma is to treat the discount problem as one of temporal preferences – i.e., of the trade-offs that a given decision maker is willing to make between future and present. Just as we earlier suggested discussions with decision makers of questions like "How much of a decrease in commercial catch are you willing to endure for an increase of 20 percent in native Indian catches?" so we now propose to ask those concerned, "How many fishermen would *you* be willing to put out of work today in order to increase the chances that their children will still have a healthy fishery available?" Answers to such questions (which should be more subtly posed – see Keeney and Raiffa, 1976) often suggest radically different time preference rates than those implied by standard discounting assumptions. For example, in the budworm work we found that managers using 5 percent or 10 percent "prescribed" discount rates in their formal economic analyses nonetheless exhibited 20 percent and higher rates when actually asked to choose freely among alternative time streams of indicators. And we doubt that this is an isolated example. The point is not that any of these particular discount rates is "right" or "wrong," but that the discussion of the contradiction forces all participants in the assessment exercise to explore the critical question of time preferences more deeply. Similar discussions are provoked by explicit comparison with the full indicator time streams, as recommended earlier in this chapter. As we have stressed repeatedly, it is only such a process of mutual exploration that can lead to understanding and meaning in the evaluation. To casually consign this fundamental question of values to the untender and unilluminating mercies of an extrinsically defined discount rate seems to us the epitome of unadaptive, irresponsible assessment.

SUMMARY

Every single exercise in adaptive evaluation can and should begin with the development of a set of specific indicators responsive to the concerns of those who will make, and those who will endure, the policy decisions. These should be followed by an explicit graphical comparison of indicator patterns. As we have stressed repeatedly, if you must address the more subtle issues of evaluation, you will require expert assistance. Obviously, you should have no patience at all with consultants hawking "answers" in such an uncertain field. But even the most well-meaning and self-critical experts tend to be bound to their own specialties and techniques.

A recent report by a U.S. study group critically reviews past efforts to apply decision-theoretical approaches to specific environmental problems and provides an excellent perspective for would-be evaluators (Holcomb Research Institute, 1976). There are several good texts on applied decision theory in which you can read about these formal approaches to evaluation. We have found those by Raiffa (Raiffa, 1968; Keeney and Raiffa, 1976) to be the most readable.

Since even good texts tend to concentrate more on strengths than on shortcomings of a field, however, we recommend several papers that provide effective self-defense against overenthusiastic technicians. Liska (1975) has edited a collection of essays on the so-called "consistency" issue. These show that preferences and utility functions of a given decision maker do change over time, and often as a consequence of previous interviews with decision analysts. Lipset (1976) presents strong empirical evidence that "objectives" dear to decision theorists simply do not exist on many issues except as they are elicited by the evaluation dialogue. This, of course, is just what adaptive evaluation hopes for.

The notion that each policy should be associated with a probability distribution of outcomes reflecting uncertainties in the analysis is attractive and probably formally correct. Decision theory is well adapted to coping with such probability distributions. Unfortunately, people are not. Slovic and Lichtenstein (1971) summarize a body of evidence that suggests that a probabilistic assessment of utilities is most unlikely to lead to meaningful evaluations, even in the simplest cases.

In retrospect, it should be clear that the real problem of evaluation is not one of technique, but of meaning. The ultimate goal is not to produce a set of numerical rankings, but to understand the strengths and weaknesses of alternative policies' performances. For it is on the basis of such understanding that meaningful, adaptive steps can be taken toward policy modification, improvement, and eventual implementation.

9 Communication

Effective communication is essential if environmental analysis is to have an impact on decision making. Our experience is that at least as much effort must go into communication as goes into the analysis. This has been confirmed by several other studies (Ackerman *et al.*, 1974; Ford Foundation, 1974; Holcomb Research Institute, 1976).

Individuals involved in doing an environmental assessment are generally not involved in the decision making. They are instead an advisory body that formulates and presents conclusions to the decision-making body. An analyst who wishes to convey the results of a detailed study faces a serious dilemma. The volume of information (data and future scenarios) is usually very large, too large to hope that decision makers will have the time to absorb and assess it. Yet findings that are condensed into an executive summary will carry little weight unless the reader has easy access to the supporting data and analysis.

In order to achieve successful communication, the assessors must clarify *what* information there is and to *whom* it should be transferred. The format or technique of communication depends on the answers to these questions. Several techniques are outlined below, but the general rule is that the sender must present the information in a language that is comprehensive and believable to the receiver.

WHAT INFORMATION?

Four types of environmental assessment information should be conveyed: first, the *data base*, both actual measurements and assumptions; second, the *technical method* used in the analysis and the assumptions of that method; third, the *results* of the analyses; and fourth, the *conclusions* derived from these. These last two have the highest priority. Each of these types of information has two facets:

the actual numbers or literal meaning, and the degree of belief. The believability of the information is by far the subtler and more difficult to convey – but it is certainly equally important.

METHODS OF COMMUNICATION

For illustration, let us assume that the assessment team reports to a single decision maker. Traditionally, a detailed report is produced that includes all the techniques, assumptions, and results, and an executive summary is prepared that is intended to be a set of recommendations for the decision maker. These detailed and lengthy reports are awkward documents that generally defeat their own purpose.

Instead of this traditional method, we have tried some alternative communication techniques, ranging from those requiring a high degree of involvement of decision makers in problem analysis to summarizing statements that only crudely represent the underlying complexity of the problem. From this spectrum we will discuss only four of the techniques we have used.

PARTICIPATION OR INTERACTION (THE WORKSHOP)

Of the communication techniques at the disposal of an impact assessment team, this one creates the most thorough understanding, and it is the most demanding of the recipient. As we discussed in Chapters 3 and 4, the managers should be involved in the original workshops when analysis begins. We have found, especially in the budworm and salmon case studies, that if managers can be involved from the start, they have at least a moderate understanding of the assessment techniques and assumptions. At the same time, they contribute insight and direction to the assessment and thereby develop a commitment.

In addition, at different stages of model construction, managers and policy people can be involved in short (2–4-hour) gaming sessions, where results of different policies are compared. The opportunity to sit in front of a graphics computer terminal and interactively try out alternative model assumptions or management options has several unique advantages. First, the decision maker gains an understanding of the underlying structure that generates particular predictions. When an unexpected result emerges from a run of the model, he can question the analysis team to discover what assumptions produced this result. This process makes the decision maker a member of the analysis team instead of an observer and gives him some understanding of the model itself. Secondly, by altering model assumptions, he can see how sensitive the predictions are to changes in these assumptions, to uncertainties in the data, and to uncertainties in the implementation of the policies. This leads to the third benefit, which provides a sense of the degree of belief that should be placed in the results. As more assumptions are explored, an appropriate level of confidence in the results is established. Finally, and perhaps most important, rapid interactive gaming with the model permits the decision maker to try out new alternative management schemes, which forces him to realize that the management alternatives are not necessarily limited to a few well-defined options. Thus he is encouraged to try new and unusual options and to begin to approach problems in an adaptive way.

It should be noted that such workshops will probably have value even if assessors use predictive methods other than simulation models. Any opportunity for the decision maker to analyze the predictive techniques' assumptions is beneficial. We emphasize that our experience in several case studies shows that the more decision makers can be involved *during* the analysis, the easier the transfer of information will be at the end. When the top-level manager is not available for workshop participation, we must turn to other communication methods.

NARRATED SLIDE PRESENTATIONS

At the other end of the spectrum there is an approach that requires little time or effort on the part of the receiver, but that does require considerable preparation by the sender of the information.

The basic premise of a narrated slide presentation (35-mm projection slides and an accompanying soundtrack on recording tape) is that technical language, mathematical formulations, computer programs, and even underlying theoretical concepts can be translated and condensed into a readily digestible form.

In the past it has been very difficult to communicate the technical methods and assumptions to a manager. Managers often are either not fluent in or comfortable with the "language" involved or else their time constraints are severe. Frequently, the evaluation technique has remained mysterious to them, and the credibility of recommendations resulting from the technique is low. In order to address these problems, we have prepared and used 10 different narrated slide shows (Bunnell and Tait, 1974; Bunnell, 1976) on subjects ranging from ecological and management history to actual models, techniques, and even ecological theory. Figure 9.1 shows a brief segment of the slide presentation of the spruce budworm simulation model.

The slide presentations usually last 10–25 minutes, yet they convey a great deal of information. They are short and to the point, they do not overwhelm with numbers or confuse with jargon, and they hold the attention of the audience. We have examined the usefulness of this approach by distributing evaluation questionnaires after slide shows. Audiences evaluated the usefulness to themselves (Figure 9.2*A*) and indicated for which other occupational groups they thought the slide shows were suitable (Figure 9.2*B*). Of all the types of viewers, our intended audience (decision makers and managers) was the group that found the material most useful. In addition, to our surprise, a much broader audience also found the shows informative. This suggests that narrated slide shows of this type may be useful in educating and involving the public.

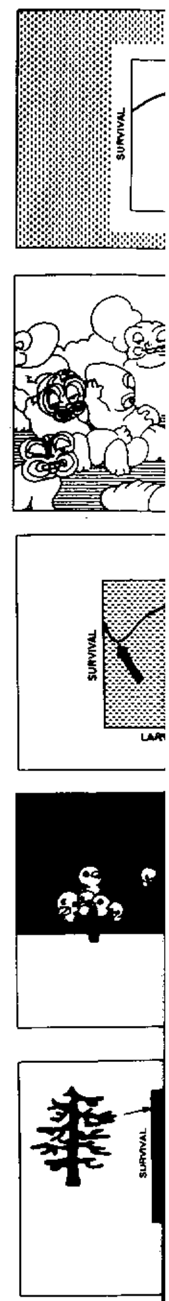
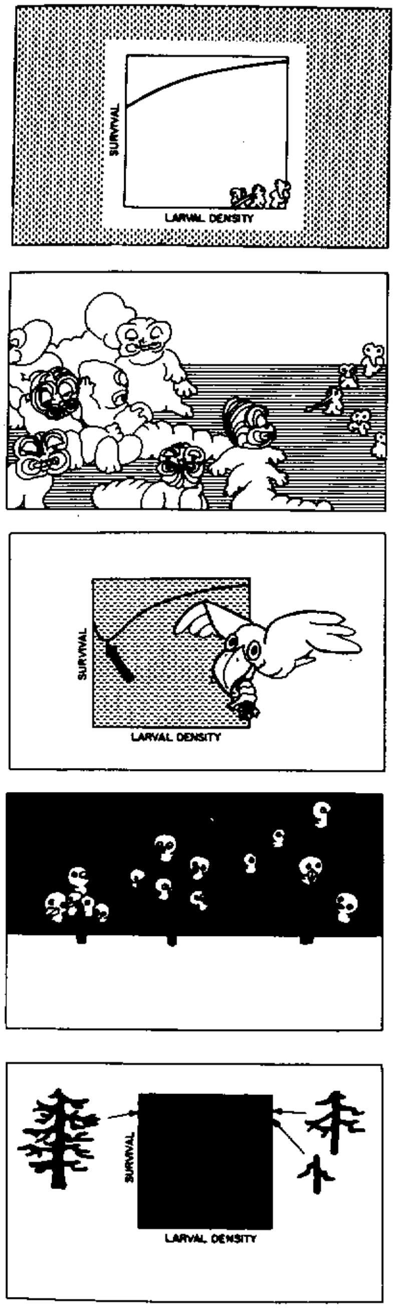


FIGURE 9.1
Model," a 10



25. Larvae may be killed by disease or by parasites.
As the larval density increases, a larger proportion manages to survive death due to parasites or disease. This is because the number of parasites is limited by other factors. Parasites can affect only a certain number of budworm.

26. When there are lots of budworm, the percent survival is high.

27. Birds eat larvae. There are several kinds of avian predators such as warblers, thrushes, ovenbirds, and finches. The model simulates the combined effect of all the birds. Like parasites, the birds cannot keep up to increases in the larval population. A higher percentage of larvae survives when there are lots of larvae.
An interesting feature of the bird predation pattern is the depression in the survival curve at low larval densities. The lowest survival occurs just above the lowest density – when larvae get extremely scarce, the birds cannot find them.

28. The ability of birds to find larvae is also influenced by the size of the tree on which the budworms and birds are living. When the trees are small, the birds are concentrated and their feeding impact is high. As the trees grow larger, the birds spend their time searching among more branches and the budworms have a better chance to escape.

29. Thus, the size of the trees influences the ability of budworm to survive predation by birds.

FIGURE 9.1 Sketches and narration from a segment of the "Spruce Budworm Model," a 10-minute slide–tape presentation (Bunnell, 1976).

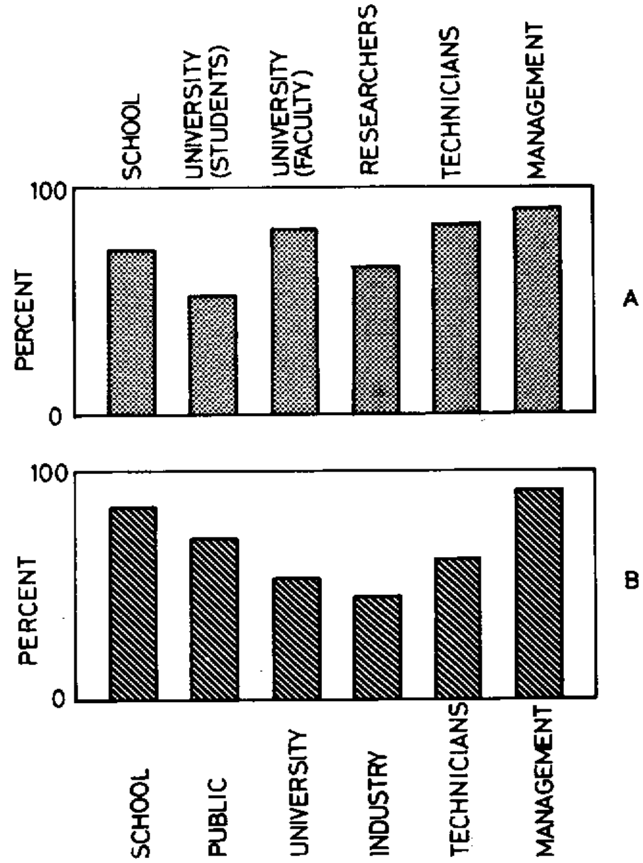


FIGURE 9.2 Audience evaluation of the usefulness to themselves (*A*) and to other interest groups (*B*) of the slide–tape presentations on the spruce budworm (Bunnell, 1976). Respondents (sample size = 139) indicated degree of usefulness in each of the potential categories; the histograms show average results. Zero percent indicates not useful, 100 percent extremely useful.

SUMMARIZING GRAPHICS

Between the two ends of the spectrum (lengthy participation in workshops and exposure to condensed slide presentations) are a variety of techniques that organize information. Two such techniques have proved particularly useful: manifolds, which reveal the essential inner workings of the model, and nomograms or isopleth diagrams, which condense simulation model outputs. Both allow conceptualization of complex phenomena. Nomograms furthermore permit gaming through manipulation of possible alternatives.

Manifolds

Equilibrium manifolds (described in Chapter 6) are extracted from a descriptive model. They represent the system's dynamics in a concise form and give an intuitive sense of how the model works. Manifolds are conceptually very simple, but because of their nontraditional nature, understanding them requires a modification of the viewer's perspective. People encountering a description of a system in the manifold format frequently go through a period of saying "So what?" followed by a feeling of revelation and understanding as a large number of apparently disparate observations fall into a logical structure. Because of this, it seems worthwhile to simplify the model into manifolds to communicate some of its characteristics.

Nomograms, or Management Slide Rules

In Chapter 8, the technique of nomograms, or isopleth diagrams, was mentioned as one way of permitting the decision maker himself to perform some evaluation of management alternatives. We re-emphasize the merits of this graphical technique in this chapter because of the method's proven value as an effective communication device. The communication of information takes place while the decision maker is using the nomograms. In order to illustrate this clearly, it is necessary to explain briefly how nomograms are created. (A more detailed discussion is presented in Appendix A.)

Nomograms are constructed from several runs of the same simulation model during which two management options are varied over some range. For example, in a deer management model the decision options might be percent of the population to be harvested and sex ratio of the harvest (Table 9.1). Each simulation run calculates the value of several variables or indicators that are relevant to decision makers – for instance, "annual harvest" or "long-term numbers harvested." Results of these several simulation runs are then plotted on graphs, one graph per indicator variable, whose axes are the two management options (Figure 9.3). Contours of values are then drawn through the values on the grid points (Figure 9.4). After this contouring, isopleth diagrams of several indicators are reduced in size and pasted onto a single page (Figure 9.5).

The nomograms, which now represent a considerable compression of numerous simulation results, are then ready to be used by the decision maker. Two benefits immediately emerge merely by inspection of the response surfaces (Gross *et al.*, 1973; Peterman, 1975). First, they provide a graphical information system that summarizes some of the data relevant for decision making. Second, limits of the system can easily be determined. For example, in step 3 (Figure 9.4) it can be seen that it is not possible, with the two management options shown, to achieve an annual deer harvest of more than about 325 animals for the herd modeled.

TABLE 9.1 First Step in Construction of a Nomogram

	Management Actions		Indicators		
Simulation Run No.	Proportion Males Harvested	Harvest Rate	Annual Harvest	Long-Term Harvest	...
1	0.0	0.0	0	0	
2	0.0	0.2	48	.	
3	0.0	0.4	35	.	
4	0.0	0.6	22	.	
5	0.0	0.8	18	.	
6	0.0	0.99	5	.	
7	0.2	0.0	0	.	
8	0.2	0.2	73	.	
9	0.2	0.4	40	.	
10	0.2	0.6	32	.	
11	0.2	0.8	25	.	
12	0.2	0.99	17	.	
13	0.4	0.0	0	.	
14	0.4	0.2	145	.	
⋮	⋮	⋮	⋮	⋮	
36	0.99	0.99	282		

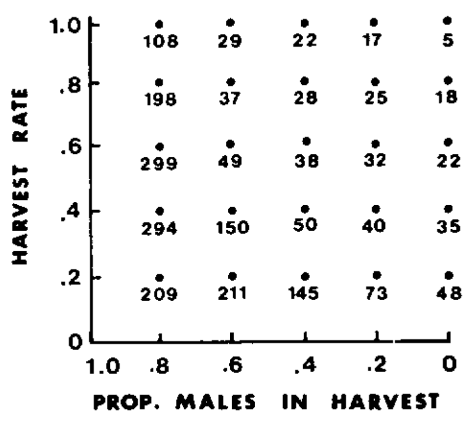


FIGURE 9.3 Second step in construction of a nomogram. See text and Table 9.1.

FIGURE 9.4 Third

The major benefi overlay is used with p One position of the o corresponds to a harv the harvest. It is then s this plastic overlay, th without touching the offs between indicat one indicator is at its The decision maker those indicators, until

This "experiment such as "managemen use of this method i in more detail in Ch to be an extremely assessment in a brief

GRADED SERIES OI

For any particular of reports or present messages or packages forms lie at one en other. With such a gr a starting point that ward the simpler en substantiates and mal

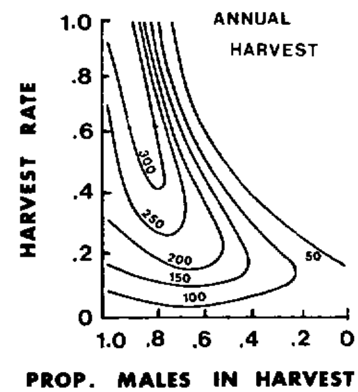


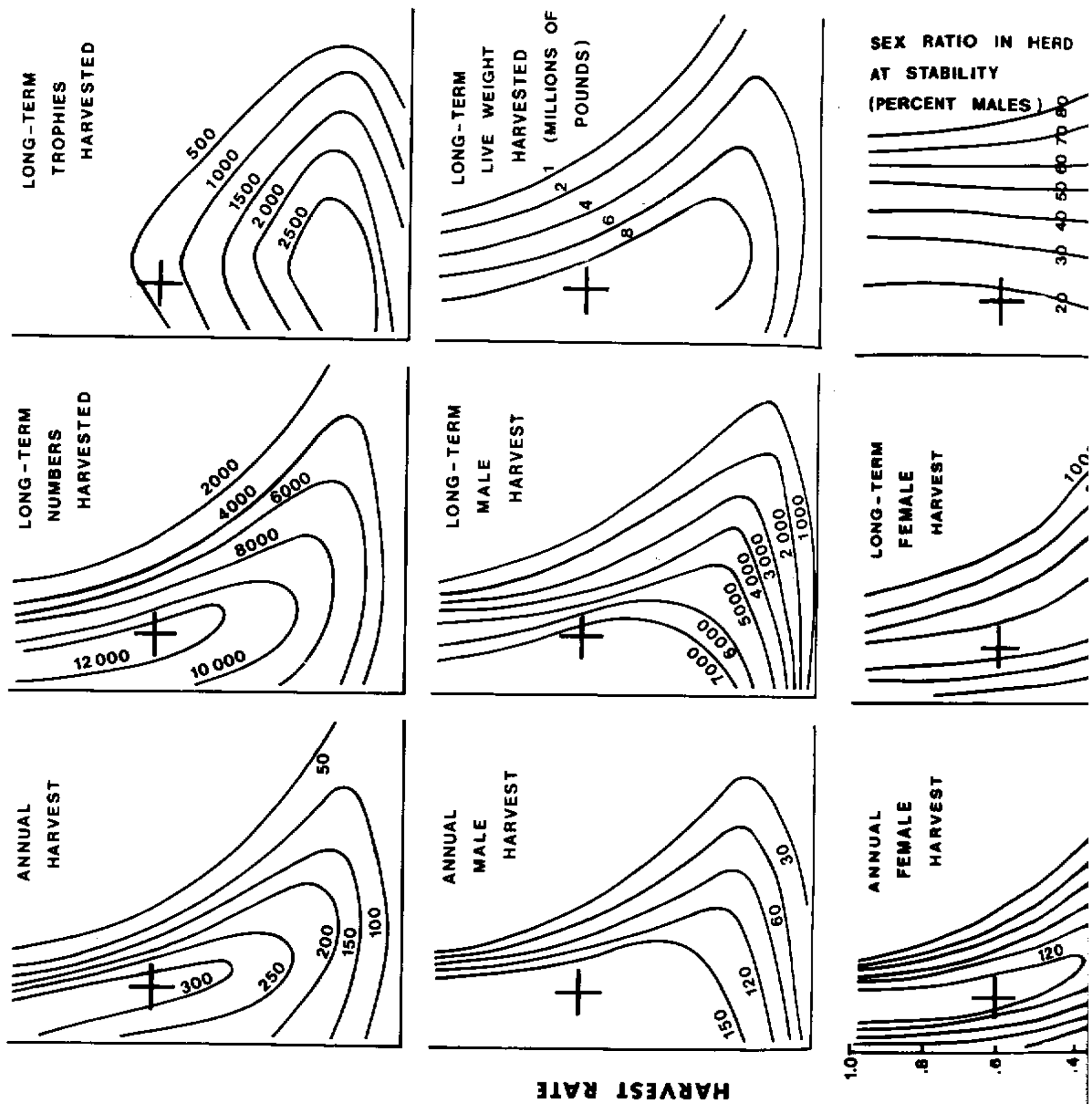
FIGURE 9.4 Third step in construction of a nomogram. See text.

The major benefits from the isopleth diagrams emerge when a clear plastic overlay is used with pointers indicating identical coordinate locations on all graphs. One position of the overlay pointers is shown by the + s in Figure 9.5. The position corresponds to a harvest rate of 60 percent and a proportion of 85 percent males in the harvest. It is then simple to read off the values of the various indicators. By moving this plastic overlay, the user can "experiment" with alternative management actions without touching the computer; the computer work has already been done. Trade-offs between indicators can easily be seen when, for a particular pointer position, one indicator is at its desired peak but another indicator is at an undesirable low. The decision maker then can "experiment" with alternative ways of trading off those indicators, until some satisfactory compromise is reached.

This "experimental" aspect of the nomograms has earned this method labels such as "management slide rule," "desk-top optimizer" or "ouija board." The use of this method in the budworm, salmon, and Guri case studies is described in more detail in Chapters 11, 12, and 14, but, in short, nomograms have proved to be an extremely effective way for decision makers to perform part of the assessment in a brief time and to understand some of the assessment's limitations.

GRADED SERIES OF COMMUNICATION DEVICES

For any particular assessment, the choice of components in a graded spectrum of reports or presentations is dependent on the methodologies used. A series of messages or packages is made available so that detailed and thoroughly explained forms lie at one end, and simply illustrated and briefly explained forms at the other. With such a graduated series of information packages the receiver can locate a starting point that suits his background and his time constraints. Anything toward the simpler end provides him with a summary, and anything more detailed substantiates and makes believable the simpler presentations.


LONG-TERM TROPHIES HARVESTED
500
1000
1500
2000
2500
LONG-TERM LIVE WEIGHT HARVESTED (MILLIONS OF POUNDS)
1
2
4
6
8
SEX RATIO IN HERD AT STABILITY (PERCENT MALES)
80
70
60
50
40
30
20
LONG-TERM NUMBERS HARVESTED
2000
4000
6000
8000
12 000
10 000
LONG-TERM MALE HARVEST
1000
2 000
3000
4000
5000
6000
7000
LONG-TERM FEMALE HARVEST
ANNUAL HARVEST
50
100
150
200
250
300
ANNUAL MALE HARVEST
30
60
120
150
ANNUAL FEMALE HARVEST
120
1.0
.8
.6
.4
HARVEST RATE

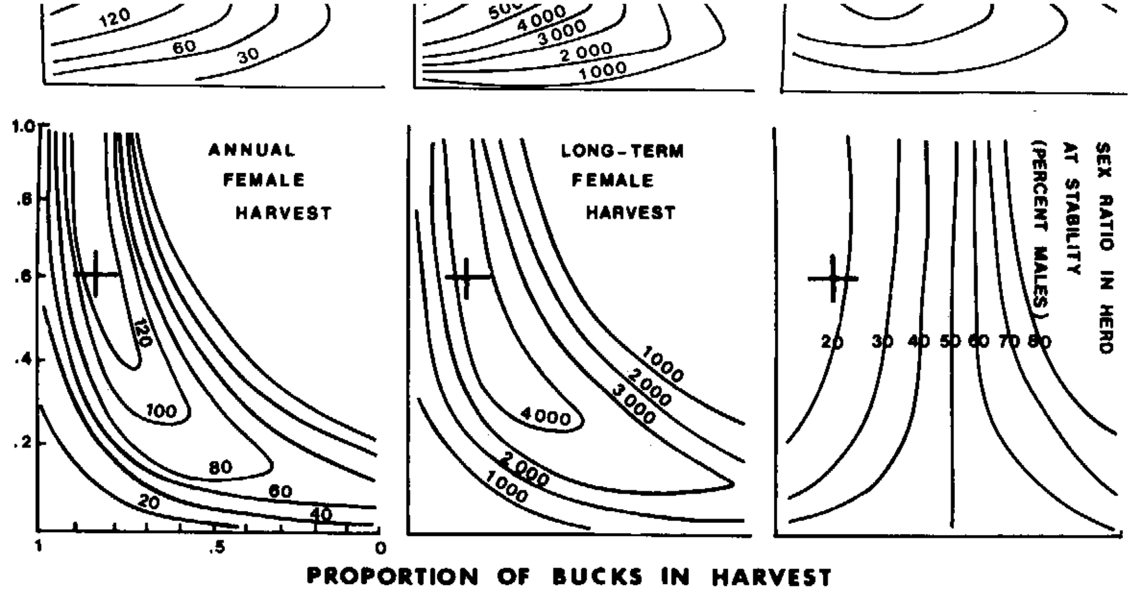


FIGURE 9.5 Fourth step in construction of a nomogram. A sample nomogram for a deer model demonstrating simulated responses of nine different indices, given varied harvest rates and proportion of bucks in the harvest. The X on each graph indicates the value of that particular indicator given the combination of the two management options (after Gross *et al.*, 1973).

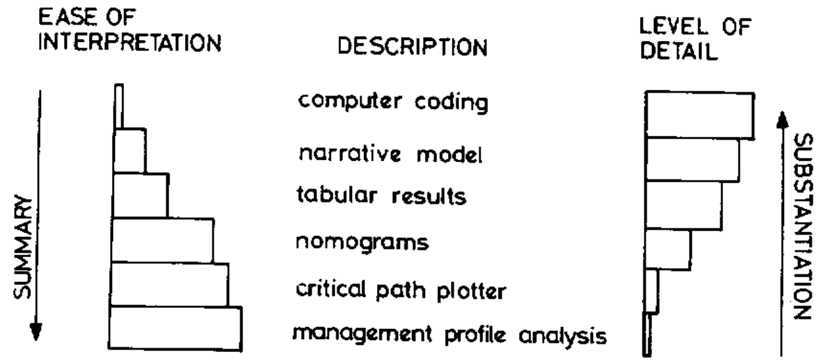


FIGURE 9.6 Graded series of descriptions used in explaining deer management models. The individual may enter at the level of detail suitable for him and proceed either to a summary or to a more detailed substantiation of the program. (Redrawn from Gross *et al.*, 1973.)

The Colorado Cooperative Wildlife Research Unit has used this approach successfully in explaining deer management models to administrators and decision maker (Gross *et al.*, 1973). The most detailed level provided is the actual computer program or coding. This is summarized by a narrative that follows the steps in the program and the results generated at each step. For example, the narrative may read, "In year 5, 50 male deer were harvested, 10 died of natural mortality, there were so many deer in each of the age classes, etc." Results of the model are summarized in tabular form, and from these tables (using several simulation runs), nomograms or isopleth diagrams are constructed. Next, feasible alternatives for moving from the present system state to some target several years away are summarized in yet a higher level, called a "critical path plotter." Finally, at the crudest level, there is a very brief summary of the alternative management strategies and their predicted consequences. The most important characteristic of this multilevel information system is that each level is visibly substantiated by the next most detailed one and summarized by the next less detailed level (Figure 9.6). Thus, the decision maker can easily consult any level of detail to answer questions or establish the validity of recommendations.

CONCLUSION

Communication is the bridge between environmental analysis and decision making. The strength of this bridge depends upon the methods of communication; our experience suggests two important criteria. First, communication should begin as soon as the analysis begins. Second, a variety of techniques of communication should be employed: the diversity of peoples' perceptions should be matched by a diversity of communication material. This material should span several levels of detail, so that decision makers can examine the analysis at any level appropriate to their interests and training. Most important, communication takes time.

Communication should be the sole responsibility of at least one member of a six-person staff. The others should spend up to one-third of their time in interactive communication. Note that this is not public relations. Rather, it is a vital aspect of environmental analysis and decision making. If the goal of the analysis is to produce better environmental decisions, then communication requires as much creative design as the analysis itself.

10 An Underview

Traditionally, a book that has spent nine chapters outlining and advocating a new perspective and a novel operational approach deserves a strong concluding chapter. So if you are expecting such a grand finale, you should be warned that there is none. To have provided such an ending would have been deceptive – creating the tidy but false impression that we have left no loose ends to our story. Unfortunately, our exposition of a new adaptive style of environmental assessment and management has left some important unresolved issues in its wake. These issues were raised in the first two chapters and remained an implicit backdrop thereafter. Therefore we end Part I by reviving these issues because we feel strongly that they must be woven into the mental framework of those who deal in policy, especially where environmental concerns are paramount. We also bring these issues back to the surface to emphasize the need for new conceptual and methodological tools to address them. In the meantime, they are reminders that our "solutions" are not ultimate and that we must operate without all the answers.

All of these unresolved issues relate in one way or another to the theme of *uncertainty*. We believe these issues to be philosophically important; our view of the world is inseparable from our view of uncertainty. We also believe these issues to be pragmatically important – first, because uncertainty is real and, second, because these issues need continual attention, creative conceptualization, and active research before useful procedures and techniques can emerge.

The phenomenon of uncertainty was raised in the opening chapters as a central theme. Although the word did not appear regularly in the middle chapters, it was implicit in our descriptions of an adaptive approach to environmental problems. As we return now to that theme, we emphasize that we have no theories or conclusions to report – only our concerns and speculations.

PREDICTION IS NEVER PERFECT

The future is uncertain. Few would disagree with this in principle; the debate, if any, would involve definitions and criteria. Moreover, environmental assessments are not, and cannot be, predictions in any real sense. First, we cannot measure everything, and, what is more, we should not try. The things left unmeasured will also be affected by man's interventions, and these effects will cause change in those things that are being studied. Initial bounding and selection of key variables aim to minimize this effect but cannot eliminate it.

Second, no amount of observation prior to a project will reveal what impacts the project will eventually have. Almost by definition, the impacts will be the consequence of disturbances that are unlike any the natural system has yet experienced. To some extent lessons can be learned from similar situations, and conclusions can be drawn from the general responses of disturbed ecological systems. But the post-project system is a new system, and its nature cannot be deduced simply by looking at the original one. If the project planning and development sequence fundamentally incorporates adaptive assessment throughout all of its stages, then the ecological response of both the new and old systems will be studied.

If assessment continues into the future, then prediction loses its status as a goal, and assessment merges into environmental management. Prediction and traditional "environmental impact assessments" suppose that there is a "before and after," whereas environmental management is an ongoing process.

If assessment techniques cannot make true predictions, then what are they for? Is assessment simply swallowed into the larger activity of environmental management? The activities described in the preceding chapters comprise a procedure of adaptive assessment, but their aim is no longer prediction of what will happen or even what will most likely happen. Environmental assessment should be an *ongoing investigation into*, not a *one-time prediction of*, impacts.

The people making environmental assessments often are the first to admit that their conclusions are not certain. But if they attribute their doubt to a lack of time, money, and manpower, then they have missed the point. Attempting to close the gap on imperfect predictions detracts from a proper focus on the consequences of the inherent uncertainties that will always remain. If prophecy is impossible, then go for understanding.

LIVING WITH UNCERTAINTY

As uncertainty is a very broad concept, it is useful to think of three classes of uncertainty. The analyses used for assessment and the strategies adopted for management will be different for each.

The first class involves those events that can be predefined, that have known direct effects, and that have known probabilities of occurrence. The coin toss is a didactic example, while varying weather patterns are an example with environmental significance. Statistical analyses, the study of stochastic processes, the subdiscipline of decision theory, and many other applied methodologies are founded on this class of uncertainty. It is natural that analytical advances should start here – when you know the probability distribution, a large proportion of the uncertainty is resolved.

The second class of uncertainty involves those events that are imaginable and at least partially describable, but for which neither the outcome nor the probability of occurrence are known. Nuclear reactor failure exemplifies this class of uncertainty, and the continuing scientific controversy highlights the absence of a conceptual or analytic framework for this class.

Many "natural" examples are either not entirely convincing or of minimal ecological importance. Being struck by lightning comes to mind. Earthquake and drought have large social importance, but with observation and experience, these events can move into the first class.

The situations that are rapidly dominating this class involve man-made interventions such as the development of a nuclear power economy, a possibility that currently enjoys a moderate degree of public attention. Climate modification, recombinant-DNA research, and heavy metal and synthetic chemical discharges are other relatively new items joining a rapidly expanding list. Two features make these "advances" potential horrors. First, they each introduce a perturbation into the environment that is unique in the evolutionary history of the biosphere. And second, modern technological and industrial capacity permits such perturbations to take place rapidly on a global scale.

The third class of uncertainty contains all those events for which we have no experience (or have forgotten) and events involving unknown processes of unknown functional form. Examples are to be found in the historical record; imagine, for instance, the character of a simulation model of disease had it been built before Pasteur.

Assignment of events to one or the other of these classes depends on what is "known," that is, on a changing constellation of ideas – new things being added while others are forgotten. A possible correlation appears between class and time scale. Things that occur on a "human time scale" (minutes to years) are more likely to fall into the first class.

The relationship of uncertainty and variability to the functioning of ecological systems received its prime emphasis in Chapters 1 and 2. The subsequent focus on the procedures of the adaptive approach that we recommend adds new elements for consideration. In an environmental analysis there will be things we know about but choose to exclude and things we do not know about and thus have no choice but to exclude. The distinction, though sounding simple, does have meaning. The former can be checked as discussed in Chapter 7. The latter should not be ignored simply because nothing can be done about them. Residual uncertainty should influence our decisions and policies now; it certainly will influence our world later. Watt (1974) accurately describes this "*Titanic* effect": when uncertainties are wished away and not planned for, the crises that follow are all the more intense.

Environmental decisions are made in a social setting. Of all the uncertainties of human and social behavior, the one of paramount significance to environmental policy is the shifting nature of individual and social preferences. The profound changes in policy stemming from the rise of environmental awareness in the late 1960s are more than obvious. The point to remember is that other equally radical shifts in social goals will occur in the future.

Even the relatively objective activity of environmental assessment is influenced by social preferences. The attributes selected as important for consideration, the time horizon chosen, and the treatment of alternatives are a few of the characteristics colored by public opinion. Emphasis can range from protection of unique scenic areas to smog to endangered species to the socioeconomic environment of the world's citizens.

In order to live successfully with uncertainty, our environmental management institutions must maintain their responsiveness to change. The ecological systems that have persisted have been those that were resilient enough to absorb the unexpected and learn from it. Our institutions, too, need a similar ability to cope. Institutions, like biological systems, learn to handle change by experiencing change. And as with other things learned, this ability will be forgotten if the experience is not occasionally reinforced. Insulation from small disasters leaves one ill-prepared and vulnerable to larger ones.

MONITORING AS "POSTDICTION"

The final draft of an environmental impact statement is stamped "approved," and then the bulldozers move in. Unfortunately, this has been the case too often. A major operational change required to shift assessment from its traditional role into meaningful environmental management is the continuation of assessment activities during and after the period of construction. Such an extension of activity requires the addition of a monitoring capability. At the very least, monitoring provides an opportunity to attempt an invalidation of the analysis that has already been done. Prediction may not be possible, but some postdiction is.

The choice of what to monitor presents many of the same problems that were faced in the choice of what to include in the original assessment. The easy solution would designate the key variables of the assessment as the quantities to monitor. However, such a choice would ignore some of the understanding gained in the analysis and would miss some opportunities to increase the scope of that understanding. Monitoring provides an opportunity to pursue model and assessment invalidation and to solidify our degree of belief in the investigation to date. This objective requires a testing and probing of our analysis that cannot be accomplished without stepping beyond the previously selected key variables and relationships.

Not all key variables are equally important. Some will have been found to be strongly implicated in possible future impacts. Others will be accompanied by larger uncertainties in the form or magnitude of their relationships with other variables. And some will be a combination – strongly implicated, but sensitive to the range of variation in our estimates. A monitoring plan should address these differences between variables.

A model or an analysis is characterized as much, if not more, by what is left out as by what is put in. Invalidation and monitoring have an obligation to "look outward" and include the excluded factors in some way.

Is this an open invitation to monitor everything possible? Clearly not. Limitations of time and money and the sheer incomprehensibility of masses of complex data call for restricted and focused monitoring. The question to ask is, "*What would I do with the information if I had it?*"

Some monitoring will also be needed to mitigate impacts. Almost all human activities have some impact on the environment. Some impacts will be acceptable prices to pay, and the project will proceed. However, we may wish to "fix up" a particular unwanted impact. For this, the monitoring and corrective actions are more focused than in the broader management problem, but the same general procedures and concerns apply. Mitigation often appears as a separate and distinct activity rather than as one component of good management. This distinction lies very close to the dichotomy between "externalities" and "internalities." Those undesirable effects that are mitigated are perceived as "side effects," as if they were somehow merely inconvenient intrusions from outside. But undesirable effects are an inherent part of the total problem, and management should treat them as such. Even the word "mitigation" reflects a perceived realm of responsibility: mitigation is left to other public agencies or is performed under legal obligation.

The above observation does not reject mitigation as an important activity. The development plan should include the mitigation of some impacts just as it should include steps to avoid others. There will always be impacts inseparable from the development itself that require remedial action – land recovery after strip mining and reforestation after logging, for example. Fiering and Holling (1974) discuss some of the properties and constraints of restoring a dynamic system to a desired condition.

Monitoring provides us with one other useful payoff – lessons for the next time. Future environmental investigations stand a better chance of improvement if monitoring and retrospective analysis contribute to the common experience. A catalogue of things that went wrong and impacts that were "surprises" could be a useful tool in future assessments (see, e.g., Dasmann *et al.*, 1973). Some of the same mistakes could be prevented. But the big take-home lesson is that the unexpected is to be expected.

ADAPTIVE MANAGEMENT

Adaptive management is not really much more than common sense. But common sense is not always in common use. Many industrial and engineering concerns routinely practice adaptive management. In developing a new product, not all the final details are planned and fixed before the first action is taken. Activities such as pilot projects, test modeling, and market surveys are all efforts to use information from the first stages to adapt the final outcome to greater advantage.

The extensions of relevance here are the inclusion of environmental considerations among the criteria for project adaptation and the integration of the assesment and planning processes. Such integration requires mechanisms that allow the assessment to continue along with the project evolution and mechanisms that allow the project to adapt in response to ecological considerations.

No particular set procedures will accomplish this task. But there are types of questions that can be asked: Are there times in the development plan when changes can be made and new directions followed? Will the analysis be able to respond at the right time with the information needed to influence the project development? Absolute replies to these questions are not possible, but the mere act of asking reorients the perspective from one of assumed certainty to one of prepared responsiveness.

Adaptive management can take a more active form by using the project itself as an experimental probe. In this context we place an explicit value on ecological information. A deliberate alteration in the project or the sequence of its stages may reveal detrimental ecological effects that can be avoided in the final form of the project. In many cases such alterations will be "inefficient" in a traditional sense, but a judgment must be made concerning the longer term value of the information to be gained. An explicit attempt to use the project itself can be used to address one element of the uncertainty surrounding environmental responses. Walters and Hilborn (1976) and Peterman (1977b) propose this strategy for the management of fishery stocks.

A note of caution should accompany these last proposals. There is small hope of gaining useful information by arbitrarily perturbing the environment or trying some action just to see what happens. Experimental probes of the type suggested here should be addressed to specific questions about environmental response. Experiments without clear questions are likely to give ambiguous answers.

Incrementalism is a very similar trap: Build a small dam and everything is fine; build a large dam and everything goes belly-up. The inherent nonlinearities, thresholds, time delays, and spatial redistributions of ecological systems may completely hide the potential effects that would result from a larger intervention. Small may be beautiful, but big is not simply several smalls (Holling, 1976).

FORECLOSURE OF OPTIONS

Without uncertainty in ecological behavior and without uncertainty in future societal preferences, finding the "right thing to do" would take on an entirely different character. However, along with these two very real sources of uncertainty comes the trap of irreversibility. Will the ecological system head off in an unanticipated and undesirable direction that is not amenable to recovery? In terms of the descriptions of Chapter 2, will the system be flipped into an entirely new equilibrium region? Or, on the other hand, will a project that is acceptable now be viewed as intolerable in the future?

Recovery and future flexibility present very real issues. We cannot always require a complete return to starting conditions or complete freedom to reach any other conceivable condition. But we can try to keep from getting locked into any one situation. No guarantees exist, but to ask honestly what options are being foreclosed reorients the planning and development process and makes dead ends less likely.

Besides the vagaries of nature and the swings of human preferences, decisions taken now have consequences for decisions to be taken in the future (Walters, 1975a). All decisions change the environment in which future decisions are made, but a pathological aspect arises when a particular decision sets up a sequence of following decisions from which there is no retreat. Developments involving large capital expenditures are especially apt to follow this one-way path.

Adaptive assessment should look ahead to identify at least some of the *de facto* future decisions that are being made by our present actions.

DESIGNING FOR UNCERTAINTY

Unless big disasters can be completely eliminated (which we take to be impossible), there remains the problem of designing our institutions and artifacts to cope with their occurrence. Occasional small disasters offer an important learning opportunity, but the choice between several small and one large calamity is intuitive at best. Nevertheless, we propose that some amount of change and uncertainty is necessary and healthy in order to maintain responsiveness and resilience.

Some systems are inherently more capable than others of absorbing insults and changes without losing their integrity. We would like to be able to conclude with a list of design principles that point the way, but, unfortunately, we do not know what those principles are. We do, however, believe there is one axiom that underlies any design for uncertainty. This axiom states: There exists a serious trade-off between designs aimed at preventing failure and designs that respond and survive when that failure does occur, (Holling and Clark, 1975).

We have no definitive picture of how this latter sort of system would look, but it probably would not be accomplished by the traditional means of maximizing engineering and economic efficiency. Our research into sources of persistence in ecological systems is beginning to point in some likely directions. Undoubtedly, some further lessons could be learned by examining the response and reactions of different societies to hazards and other disruptive forces. The anthropological literature should yield some clues how differing cultural structures react under stress.

Examples are few and a theory is lacking, but this will continue to be the case until we learn to see the world in a new perspective – a perspective that recognizes adaptability and responsiveness rather than prediction and tight control, and a perspective that actively views uncertainty as a fundamental facet of environmental life rather than as a distasteful transition to attainable certainty.

Part Two

Case Studies

In Part I we described the individual elements of the adaptive approach, drawing on appropriate case study material to illustrate our arguments. Part II treats the same issues, but shifts the perspective to that of the case problems *per se*. Each of the next five chapters documents one of the specific applied problems that figured in the development and testing of the general approach outlined so far. Because this approach evolved as a direct result of the studies, no one study represents what would be an "ideal" case of adaptive management or assessment. But together, these studies document the usefulness of the approach.

Each of these five case studies was developed by a different subset of the authors, together with their colleagues at each of their home institutions. The material presented was coordinated and prepared by the following individuals:

C.S. Holling, "The Spruce-Budworm/Forest-Management Problem," Chapter 11
R. Peterman, "Pacific Salmon Management," Chapter 12
C.J. Walters, "Obergurgl: Development in High Mountain Regions of Austria," Chapter 13
Jorge Rabinovich, "An Analysis of Regional Development in Venezuela," Chapter 14
J. Gross, "A Wildlife Impact Information System," Chapter 15.