United States
Department of
Agriculture

Forest Service

**Southern Forest
Experiment Station**

New Orleans,
Louisiana

General Technical Report
SO-110
July 1994

Assessing Uncertainty in Expert Judgments About Natural Resources

David A. Cleaves

SUMMARY

Judgments are necessary in natural resources management, but uncertainty about these judgments should be assessed. When all judgments are rejected in the absence of hard data, valuable professional experience and knowledge are not utilized fully. The objective of assessing uncertainty is to get the best representation of knowledge and its bounds. Uncertainty assessments can be used to compare alternative projects and risks, help prevent use of extreme decision strategies, help communicate and justify decisions, guide research, and establish monitoring programs to improve learning.

Uncertainty assessment is the art and skill of judging the level of knowledge that backs up estimates. Use of structured processes, use of an objective analyst or facilitator, and development of rewards for honesty and introspection in professional judgment can increase the accuracy of assessments. Techniques and performance aids are available for structuring decisions and uncertain elements, guarding against motivational and cognitive biases, and dealing with rare but consequential events. Experience and increased understanding of decision analysis, artificial intelligence, and behavioral decision theory will also enable decisionmakers to make increasingly accurate assessments of uncertainty in professional judgment.

ACKNOWLEDGMENTS

The author thanks the following people who contributed to this publication:

Bill Bradshaw
Steve Derby
James Granskog
Minnie Haten
Denise Hutchinson
Dale McDaniel
Joan Norman
Alan Salmon
James Saveland
Paul Slovic

Contents

INTRODUCTION 1
A CASE FOR ASSESSING UNCERTAINTY 1
Nature of Uncertainty and Human Judgment 1
Uncertainty and Risk 2
Uncertainty Assessment 2
ASSESSMENT PROCESS AND GUIDELINES 4
Forms of Expression 4
Verbal Expressions. 4
Numerical Expressions 4
Assessment Process 5
Assessment Strategies 5
Role of the Analyst/Facilitator 7
Assessment Process Outlined 7
Structuring the Decision and the Uncertain Element 8
Decision Structures 8
Element Structures 8
Response Frames for Assessments 10
Guidelines for Conducting Assessments 11
Judgmental Biases 12
Modes of Judgment—Sources of Bias 13
Anchoring 13
Reliance on Availability 13
Reliance on Representativeness 14
Reliance on Internal Coherence 14
Reliance on Hindsight 14
Group Assessments 14
Assessing Rare Events 15
Assessment Quality 15
LITERATURE CITED 16

Assessing Uncertainty in Expert Judgments About Natural Resources

David A. Cleaves

INTRODUCTION

Those who make decisions about natural resource policy usually rely on experts to interpret biological, social, and economic systems and to forecast the effects of proposed policy alternatives. Similarly, natural resource managers rely on professional specialists to predict the future and the consequences of decisions. Estimates by these experts are sometimes regarded as accurate and certain representations of the real world. However, as professional judgments and extrapolations from existing data, such estimates are laced with uncertainty. The degree and nature of this uncertainty is often not revealed or is described only in vague, qualitative terms. Research and experience in decision analysis have shown that uncertainty, although it may not be capable of being measured, can be expressed in quantitative terms that can be understood and used in decisions.

All natural resource management decisions involve some uncertainty. Decisionmakers are likely to develop elaborate and costly alternatives or forgo important opportunities in order to relieve or avoid vague uncertainties that they can feel but cannot articulate. Quantitative analysis can indicate possible consequences of some--but not all--proposed actions. Moreover, quantitative analyses can be unacceptably expensive, or of poor quality, or not in forms usable in managerial and political decision processes. Even if available, quantitative analysis may not be any more accurate or more useful than direct human judgment. The appearance of precision given by numerical outputs of such analyses can mask the fact that the analyses and their interpretations reflect human judgment.

Given the limitations of analytical models and published studies, it may be necessary or desirable to directly elicit judgments from scientists and specialists. Where this is done, the scientist becomes an interpretive filter between the body of scientific knowledge and the policy or decisionmaking process. Many sources of uncertainty can influence these judgments. Decisionmakers may be forced to choose among alternatives that embody unexpressed judgmental uncertainties. When uncertainties are not expressed properly, those who develop or choose among alternatives can be misled. A working definition of uncertainty and a rationale for quantifying uncertainty are presented, and general strategies and guidelines for estimating uncertain quantities, values, relationships, and events are offered here.

A CASE FOR ASSESSING UNCERTAINTY

Nature of Uncertainty and Human Judgment

Judgment is a process of estimating, valuing, or choosing--in essence it is a thinking process. Judgments can be expressed in various ways. They can be viewed as data collected directly from human cognitive processes. The quality of judgments reflects the level of rationality and scientific rigor in the judgment process. An expert may have technical understanding of processes and relationships, but there is no assurance that the expert's judgment process will follow the rules of rational thought or scientific review or that the judgment will constitute useful information. Rationality in judgment means that the person who estimates, values, or chooses thoroughly uses available information and is aware of alternatives and their implications and that the judgment is coherent, consistent with similar judgments, in agreement with general laws of probability, and understood by users. Scientific judgment is rational judgment that has been challenged, has stood up to or been revised as a result of experience, and could be repeated if necessary.

"Soft" and "subjective" are terms often applied to judgments that have not passed scientific muster. However, there are no absolutely "hard" or "objective" judgments. All judgments--even those we consider to be facts--have a subjective component. Theoretically, the more rational and scientific the judgment process, the better are the policy and operational decisions made.

Uncertainty is a condition of not knowing. All judgments are made with imperfect knowledge, and thus no estimate is completely accurate or perfectly predictive. Uncertainty can be about present conditions or about future events. The judgment process can hide or ignore uncertainty or it can seek to explicate it. When structured procedures for explicating uncertainty in judgment are followed, understanding of the judgment and its implications is increased by opening the judgment up to productive scrutiny and critique.

David A. Cleaves is principal economist in Evaluation of Legal, Tax, and Economic Influences on Forest Resource Management, U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, New Orleans, LA 70113.

Finkel (1990) lists four sources of uncertainty:

(1) Parameter uncertainty emanates from random and systematic errors in measurement or estimation. Discussions of forest amenities, long-term projections of timber prices, estimates of effects of forest policies on employment, and estimates of project costs and completion times provide many examples of parameter uncertainty.

(2) Model uncertainty results from imperfections in representations of the real world--that is, in estimates of relationship among model parameters. For example, it is difficult to model the causal links among various parts of the economy accurately; macroeconomists and regional scientists struggle constantly to estimate changes in these relationships and their influences on human welfare. Model uncertainty is present when biological responses of forest ecosystems to new silvicultural approaches and responses of private landowners to changes in incentives or regulatory policies are estimated.

(3) Decision-rule uncertainty surrounds the quantification or comparison of social values and preferences. These values can affect decisions profoundly but are difficult to quantify and are subject to the vagaries of public opinion. Participants in policy processes often understand what they don't want more clearly than they understand what they want, and their notions change rapidly. The setting of forest practice standards involves selecting a level of risk that balances protection of the environment with maintenance of landowner rights and initiative and feasibility of implementation. Comparisons of such disparate elements can be so uncertain that policymakers opt for simplistic regulations that may turn out to be ineffective, unduly restrictive, or difficult to administer.

(4) Variability in the natural occurrence of model parameters or decision variables across time, space, and human interactions is the classical notion of frequency and is amenable to statistical modeling and testing. Forest plantation failures, prescribed fire escapes, historic stumpage prices, and many other parameters are often analyzed for their variability. Even with complete certainty in parameter measurement, model representation, and decision-rule specification, variability in parameters can make outcomes uncertain.

Uncertainty and Risk.--Risk is exposure to a chance of loss. A full description of any risk includes clear articulation and quantification of the exposure, chance, and loss dimensions. Such a description would be "a 5-percent chance of fusiform rust hitting a loblolly pine plantation within the next 10 years and causing a loss of 10 percent in growth and yield." Here, the chance is 5 percent, the exposure is to our loblolly plantation over 10 years, and the loss is 10 percent of growth and yield.

Evaluation of risk is becoming increasingly important in natural resource and environmental protection decisions (Talcott 1992). Risk assessments are usually made up of point estimates or averages. As mixtures of objective and subjective estimates of quantities, frequencies, and other components, these assessments are subject to their own second order uncertainties. Quantitative risk assessment should include techniques for incorporating uncertainty into regulatory decisionmaking.

Uncertainty Assessment

Assessment of uncertainty in professional judgment is necessary in a world that demands quick answers to complex scientific questions. Gregory and others (1992) identified inadequate treatment of uncertainty in environmental assessments as one of five main problems reducing the usefulness and applicability of environmental impact statements required by the National Environmental Policy Act. Informal and formal assessments of uncertainty abound in the decisions about silvicultural strategies and forest protection, especially protection from fire. A recent analysis of alternatives for management of late successional forests in the Pacific Northwest (Johnson and others 1991) incorporated biologists, and managers, subjective estimates of risks of extinction for groups of fish and wildlife species. The most notable and recent use of uncertainty assessment is the report of the Forest Ecosystem Management Assessment Team (FEMAT), a group of scientists commissioned by President Clinton to develop a plan for embattled Federal forests in the Pacific Northwest (USDA FS 1993). Uncertainty assessment procedures were used by FEMAT to elicit judgments about habitat viability for hundreds of old-growth and late successional plant and animal species (Cleaves 1993).

In these and other examples, it is often the elicitation process more than the assessment itself that gives rise to contention or threatens the credibility of the organization or implementation of the decision. Problems in the process of eliciting judgments can produce inaccurate or incomprehensible results. Armstrong (1978) provides a thorough and provocative discussion of subjective judgment in the making of long-range forecasts. Cleaves (1987), Cleaves and others (1987), and Saveland and others (1988) outline implications for eliciting subjective judgments in the construction of expert systems and forest protection models.

Uncertainty assessment is the process of assessing an individual's or a group's state of information about a given outcome or event. It involves telling not only what but also how much one knows. For example, one might assess the uncertainty in a fire expert's judgment that a prescribed fire will escape under given conditions. Formal assessments of uncertainty can be useful to those who make major decisions about policies and rules, capital investment, and strategic planning. They can be useful when there is disagreement about the problem definition or the outcomes of proposed alternatives, where there is insufficient data for statistical analysis, or both. A good assessment of uncertainty about an estimate accurately represents the estimator's degree of confidence in the estimate and fully taps the estimator's most rational and consistent judgment in a form that can be used in the decision process. Assessments are most accurate when rigorous and systematic processes for eliciting judgments are employed. These processes are designed to control and compensate for inconsistencies and eccentricities inherent in human judgment processes. The effort expended in improving these judgments should be matched to the importance of the information. The uncertainty of estimates that are critical in influencing choices that are economically or socially important must be assessed with some rigor.

Decisionmakers too often evaluate the quality of a judgment solely on the basis of the estimator's self-confidence or credibility as an expert. Some decisionmakers accept expert opinion without questioning whether they have captured the true extent of the expert's knowledge. Many of these decisionmakers have been disappointed when actual outcomes have differed from those predicted or when the public strongly disagrees with the expert. At the other extreme are decisionmakers who reject all expert judgments as arbitrary and baseless guesses--and who substitute their own judgments for those of the better-informed experts.

Is it important to express the uncertainty in expert judgments? A "yes" answer assumes that confronting uncertainty leads to better decisions. Good decisions are based on the full range of information available and are consistent with the beliefs and preferences of the decisionmaker. Decisionmakers must deal with uncertainty even if it is not well described or measured. Many decisionmakers try to resolve uncertainty by relying on highly confident and specialized experts. Others ignore uncertainties, treat uncertain events as if they were certain, hide uncertainties, gather information compulsively, or build in safeguards to ensure against negative outcomes. These strategies can misallocate intellectual, financial, and physical resources if minor uncertainties get more attention than major uncertainties.

Formal uncertainty assessment can give decisionmakers a positive attitude about confronting uncertainty. This benefit is not obtained where the decision has already been made and where the uncertainty assessment becomes a justification exercise. The benefit is not obtained when decisionmakers want to build in elaborate safeguards and contingency plans against all types of uncertainty, and these decisionmakers will likely ignore or circumvent attempts to quantify and compare uncertain elements. Reward systems in organizations encourage confidence and control; they do not benefit those who discuss what is not known or controllable.

Another important benefit of assessing uncertainty is in relieving the "tension between analysis and action" (Finkel 1990). Having to identify, explain, and quantify different sources of uncertainty can help decisionmakers steer a rational path between "analysis paralysis" and impulsive reaction. One strategy for responding to uncertainty is to defer action "until more information is available" or "until the facts have been established." The opposite strategy is to take strong and immediate action because "the future is too uncertain to permit inaction." In their extremes, these approaches can lead to indecision or to frenzied overprotection, and they have too often become enduring strategies that polarize interest groups and block compromise.

Expressing uncertainty can also increase public involvement and communication. In natural resource issues, many organizations participate in decisionmaking, and opportunities to miscommunication are many. If the public is to participate effectively, the public must be able to distinguish facts from myths, values, and unknowns. Honest expressions of how little or how much is know about important points can help the public understand the issues. Moreover, the public may place greater trust in organizations that express uncertainty candidly than in organizations that claim perfect knowledge.

Quantification and comparison of uncertainty can help guide information gathering and research. Collection and analysis of information is costly and can represent a misallocation of resources. Without a guiding assessment of knowledge gaps, data collection can evolve into a smokescreen to disguise decisions already made or sleight of hand that creates the illusion that uncertainties have been dealt with.

Uncertainty assessment as an ongoing process can also encourage organizations to monitor decision outcomes as a way of helping specialists and decisionmakers improve their judgments. Evaluation of the quality of judgment requires a refined ability to distinguish between choice and chance. Actual events provide feedback to specialists and suggest ways in which specialists might adjust their judgments. However, without some rigor and documentation in initial assessments of uncertainty, feedback from actual events provides little basis for comparison. As experience accumulates, estimators can be rewarded for being candid about states of knowledge, and decisionmakers can be rewarded for following good decisionmaking processes. According to Fischhoff (1990), organizations "must create (and demonstrate) an incentive structure that rewards experts for saying what they really believe, rather than for exuding confidence, avoiding responsibility, generating alarm, or allaying fears."

Finally, uncertainty assessments can be used as a basis for comparing proposed projects, risks, or policies. When described with point estimates or even averages, alternatives may appear equally uncertain or may not appear to be subject to any uncertainty. These conditions often lead decisionmakers to make poor choices. Incorporating uncertainty assessment into ranking of options can provide a more complete perspective.

ASSESSMENT PROCESS AND GUIDELINES

Forms of Expression

Verbal Expressions.—Individuals usually treat uncertainty as a feeling or belief but are usually not forced to describe or quantify it. Common verbal expressions indicating levels of uncertainty include:

- adapted from beyond reasonable doubt
- commonly
- conceivable
- fair, even, or poor chance
- frequently
- highly likely (or unlikely)
- high, medium, or low probability
- incredible
- possibly
- probably
- rarely
- remotely possible
- seldom
- usually.

Verbal expressions are most useful as indicators of levels of uncertainty where differences between levels of uncertainty are obvious. Words are easier to use than numbers; they do not demand data or disciplined precision. According to Behn and Vaupel (1982), verbal expressions "reflect more than imprecise communication. They reflect imprecise thought. . .and mask. . .unwillingness to think carefully about uncertainty."

As indicators of uncertainty, verbal expressions have several major disadvantages (Beyth-Marom and others 1985). First is their imprecision. "Likely" can mean different things to different people and even different things to the same person at different times (Capen 1984). For a particular event, "likely" may mean a probability of occurrence of 0.4 to 0.6 to one person and 0.9 to 0.95 to another. "Beyond reasonable doubt" means from less than 70-percent certain to some judges to 100-percent certain to other judges. Jurors range from less than 50-percent to 100-percent certainty in their interpretation of "guilty by preponderance of evidence."

Second, the meanings of verbal expressions are very specific to the context. The numerical probability of a "likely" event in one situation may be quite different from that of a "likely" event in another situation. Third, verbal expressions are not always used with sensitivity to small but potentially important changes in beliefs about events. An event may be called likely both before and after its probability doubles. Finally, words may confuse belief about the occurrence of an outcome with belief about the value of the outcome. For example, the statement that an ecosystem has a "good" or "promising" chance of recovering from a natural disaster may be presented as a description of relative likelihood, but it also reveals that the assessor considers recovery a good thing. The assessment could be more the result of wishful thinking than of a dispassionate evaluation of uncertainty. Words are vague enough to conceal human biases and evoke emotions that can persuade people to choose alternatives they would not choose on a more rational basis.

Verbal expressions can be misleading unless the rationale or the information base for the assessment is understood fully. Verbal expressions of uncertainty can be useful but should be presented in a comparable scale. Repeated experience with verbal assessments can be used to construct relationships between a probability scale and commonly used words, as in figure 1 (Kent 1964).

Numerical Expressions.—Numerical expressions can be understood by all users, allow comparison across different elements of uncertainty, can convey small differences in uncertainty, and separate a quantitative judgment of an event's probability from a qualitative judgment of an event's consequences. However, numerical expression of the uncertainty requires more discipline, analysis, training, and expense. Because the vast majority of uncertainties

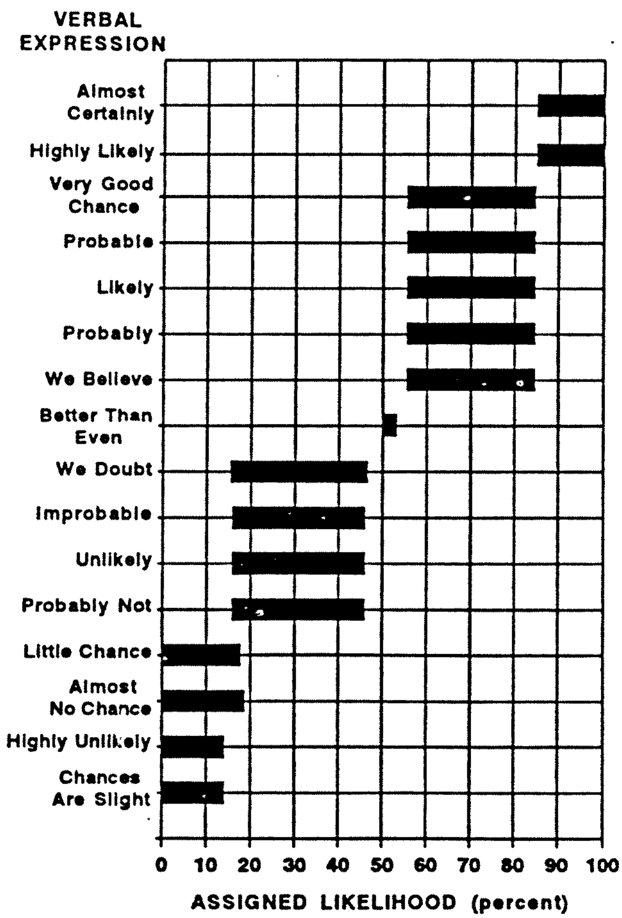


Figure 1.—*Proposed relation between verbal expression and assigned likelihood adapted from (Kent 1964).*

are not critical, numerical assessments should not be made unless precision is needed.

Figure 2 portrays several options for graphically displaying numerical assessments of uncertainty. Not all of these forms are equally comprehensible, so users should test them thoroughly and adopt them with caution.

Forms of numerical expression differ greatly in their information content and difficulty of application. It is often simplest and most useful to assign events or values to "high," "medium," and "low" levels of uncertainty. Ranges can be specified as intervals that capture the "true" value of the estimate with some level of confidence, say 90 percent. A wide interval indicates a high level of uncertainty. For example in figure 2a, the range-and-box shows the extremes, and the central portion (the box) contains a fixed percentage of the possible values.

Other numerical expressions are used to describe the likelihoods of particular values or events. These are called subjective, personal, or judgmental probabilities or likelihoods. These probabilities can be treated like statistical probabilities in calculations in order to combine or isolate elements or to show how one event is conditional on another. A judgmental probability is not a frequency per se, but rather represents the assessor's degree of belief in a possible outcome. Assessors can assign judgmental probabilities to events that have true probabilities. Judgmental probabilities treat uncertainty as a property of knowledge about events or outcomes rather than as a property of events themselves.

Judgmental probability uses the assumptions and rules of probability theory to model uncertainty. Under this framework, individual probabilities are greater than 0 and less than 1, and all probabilities in a set (the probabilities of all possible events) sum to 1. The probabilities of independent events can be multiplied together to give an estimate of the probability that all events will occur together. These joint probabilities are always less than their individual components. Events that are interrelated can be modeled as conditional probabilities.

Where there are multiple events or values, uncertainty assessments can be expressed as distributions. The simplest distributions are histograms (fig. 2b) and pie charts (fig. 2c). Distributions can also be expressed as probability density functions (the familiar bell-shaped curve) (fig. 2d), or as cumulative density functions (fig. 2e) that show ranges of likelihoods that specific values will not be exceeded. Sometimes it is useful to estimate fixed positions in an uncertain distribution. Quartile points are the most commonly used fixed positions, which require an estimate of the median value, the extreme highs and lows, and the midpoints between the extremes and the median. A simplified form is the single percentile value, the 10th or the 90th percentile, that focuses on particular values that are critical in the decision process.

Another form of numerical expression of uncertainty is the index that describes in a single measurement the overall variation in the distribution. Such indices include the variance, the standard deviation, and the coefficient of variation (standard deviation as a percentage of the mean).

Numerical expressions of uncertainty can present difficulties, however. The more sophisticated the measurement of uncertainty, the more difficult it is for assessors to estimate it and for decisionmakers to comprehend it and incorporate it into their decisions. Percentile values provide only partial information about the range of possible outcomes and can be misleading if they are not fully understood by both assessors and decisionmakers. Some factors in decisions are so vague that they are difficult to think about, much less break down into discrete events or values. However, in attempting to define and determine the uncertainty of such factors, both assessor and decisionmaker may learn more about the element and clarify what they do and do not know.

Assessment Process

Judgments of uncertainty are predisposed to many human biases and inconsistencies. However, a structured process for eliciting these assessments can help guard against biases, create more consistency, and make assessments easier to communicate, understand, and repeat. Decisionmakers can have confidence in assessments that are demonstrably unbiased, consistent, and comprehensible.

There are no right or wrong uncertainty assessments; the assessment process is designed to elicit information about the state of knowledge, not to elicit a prediction or a commitment. Over time, comparing assessments with actual occurrences can improve a person's ability to model uncertainty. However, the major purposes of uncertainty assessments are to codify judgments and to help decisionmakers make choices that are consistent with their own feelings and knowledge.

Assessment Strategies.--Finkel (1990) describes two approaches to the assessment task. The "bottom-up" approach decomposes the uncertain element into subelements. The uncertainty of each subelement is assessed numerically. The component assessments are then combined--usually mathematically--into an overall assessment of uncertainty for the element. This approach offers natural checkpoints for testing overconfidence and other forms of bias. Wrestling with the subelements and the way they fit together can show where information and research activities are needed and encourage a more complete understanding of the system being modeled. Uncertainties of subelements can be assessed by appropriate specialists, strengthening the technical foundation and credibility of the overall assessment. The bottom-up approach also has disadvantages. It requires heavy commitments of time and can magnify systematic biases

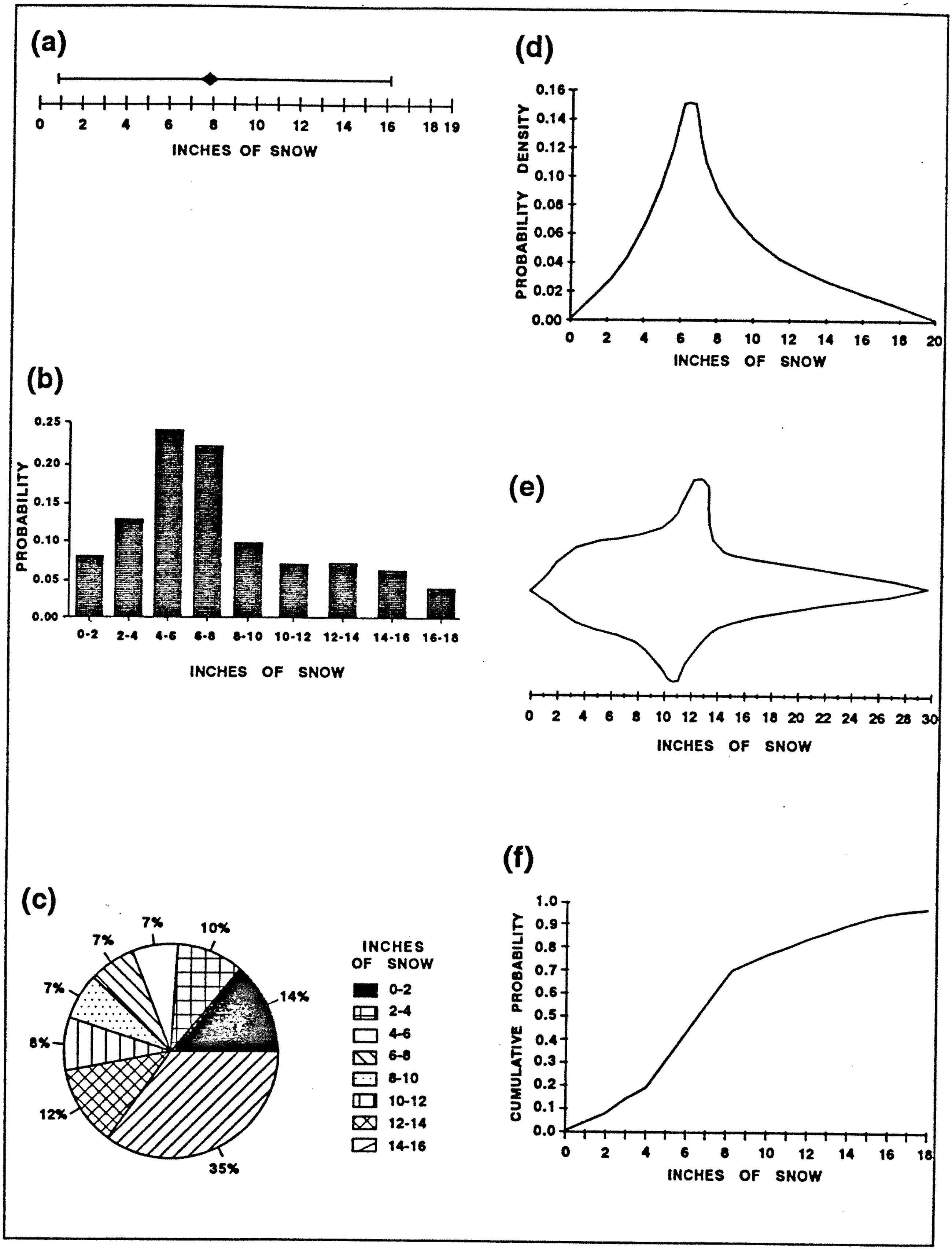


Figure 2.--*Options for displaying numerical expressions of uncertainty adapted from (Ibrekk and Morgan 1987). (a) box and range plot, (b) histogram, (c) pie chart, (d) probability density function, (e) cumulative density function.*

in the subelement assessments. Assumptions about relationships among the components are very critical; errors in mathematical equations can negate rigorous efforts in assessment of the individual parts. Finally, intense concentration on smaller elements can lead to unprofitable debates about unimportant questions.

The "top-down" approach seeks an overall assessment for the main element without breaking it down. This approach is easier, less time consuming, and avoids many unresolved scientific controversies. However, it is a daunting task to deal "all at once" with a complex phenomenon for which there is little direct experience or data. Experts are often so specialized that they have no feel for broader questions. Furthermore, some elements are just too complex to assess directly.

The bottom-up and top-down approaches can be hybridized by cross-checking bottom-up assessments provided by scientists and technical experts with top-down assessments by managers, planners, or others. The first group contributes scientific and technical perspectives, while the second group contributes their unique observational experience and managerial perceptions. The sensitivity of the overall assessment can be checked against the rigor and precision being required in the bottom-up tasks.

Role of the Analyst/Facilitator.--Uncertainty assessments should be led by one or more skilled, analyst/facilitators (A/F's). Assessment is very difficult without this assistance because assessors can become overwhelmed with the task and can be unaware of biases that creep into their judgments. The A/F leads the assessors (scientists, specialists, and/or decisionmakers) through the process, motivating them to recognize and deliver useful and bias-free judgments. The A/F knows the larger decision context and how the assessment results will be used, and can match assessment techniques to the personalities of the assessors and the characteristics of the uncertain element.

The A/F should challenge the assessors to uncover new sources of information and to confront inconsistencies. The A/F can provide continuity across assessments by different people and disciplines and can mediate disagreements during group assessments. The A/F can also encourage assessors to persevere through ambiguities and tedium, and can provide feedback for improving assessment skills.

The A/F can organize the process to make efficient use of the assessors' time and efforts, and can document the process for justification to decisionmakers and the public. An A/F can also help translate assessors' responses to decisionmakers and the public and vouch for the assessment's procedural rigor.

Assessment Process Outlined.--Figure 3 sketches an assessment process. This process was designed to produce numerical assessments of uncertainty (McNamee and Celona 1989, Merkhofer 1987, Spetzler and Stael von Holstein 1975), but it could be a framework for producing assessments in any form.

During the motivation phase, the A/F establishes rapport with the assessor and describes how the assessment fits into the overall structure of the decision(s). The A/F also orients the assessor to the basic goals and philosophy of uncertainty assessment, emphasizing that there are no right or wrong answers and assuring confidentiality of responses. This may also be a time for orientation to fundamentals of probability.

During the structuring phase, the A/F and the assessor define clearly the uncertain elements to be assessed, choose an overall strategy for approaching the task, and organize the work. They also select a form of expression and an appropriate scale. The A/F helps the assessor explore the implications of the decisions that are to be made and uncover hidden assumptions and feelings that could bias the assessment of uncertainty.

During conditioning the A/F and assessor list sources of data, research reports, or other information that could be used as background for the assessment. They also list key assumptions that will serve as guideposts if the A/F and the assessor become distracted or gridlocked. The A/F describes possible biases that may affect the assessment and explains techniques that will be used to guard against them. The A/F asks the assessors to recall significant events or organizational

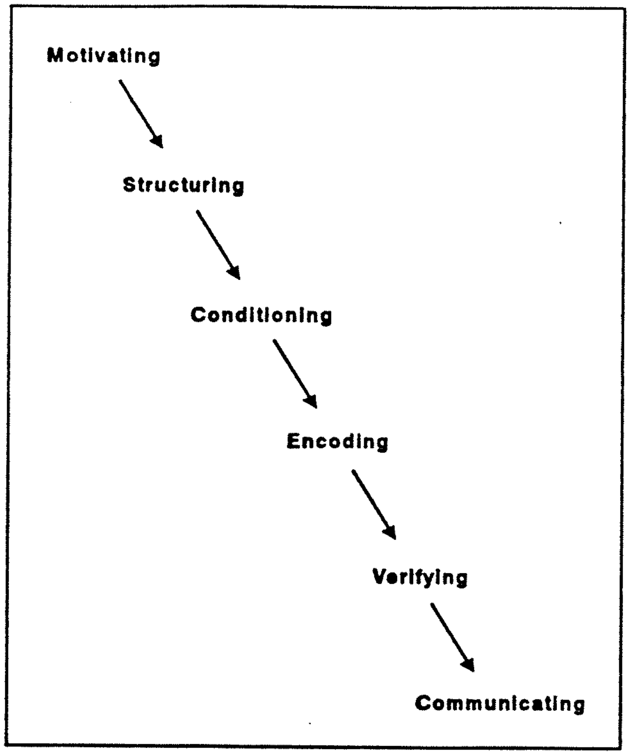


Figure 3.--*Process for assessing uncertainty.*

standards or rules that might unduly influence the assessment results. The A/F also encourages the assessor to talk about or write out descriptions of extreme values and scenarios about the uncertain elements to widen the assessor's thinking before actual encoding begins.

Conditioning is also the time for practice in assessing the uncertainty of common elements (unrelated to the decision). The assessors estimate subjective probabilities and ranges of probabilities, which are compared to actual data. These tests provide feedback and instill confidence in assessment skills.

During encoding, the A/F leads the assessors through each of the uncertain elements, using any of a number of techniques to elicit their responses. The A/F reviews the results as they are produced and looks for inconsistencies, biases, and disagreements between assessors, often repeating the assessments using a different technique to check for further inconsistencies.

During verification, the A/F shows the assessment results to the assessors and gets agreement that the results represent the assessors' best judgments. Assessors are asked to provide additional justification for extreme probabilities or likelihood distributions with irregular shapes. The assessments can sometimes be compared with existing data for similar types of elements or with assessments of the same element by other assessors. Sometimes assessors will adjust their judgments to reflect this "outside" information. In most cases, however, outside data is not available.

During the communication step in the process, the results are delivered and explained to the decisionmakers. The A/F and the assessor help the decisionmaker understand the assessment's implications. The A/F may have to further analyze the results, or restructure the decision based on the range of outcomes implied by the assessment. According the Finkel (1990), the decisionmaker should understand four things: (1) implications of replacing single estimates with uncertainty assessments, (2) the costs of over-estimating or under-estimating uncertainties, (3) sensitivity of the decision to unresolved controversies among specialists or scientific explanations, and (4) implications for further data gathering or research about the elements.

Structuring the Decision and the Uncertain Element

Understanding the decision is a prerequisite to providing useful assessments of uncertainty. Also, the uncertain elements must be defined fully and organized so that people can judge them. Even complex and apparently chaotic problems can be taken apart and arranged so that there is a common understanding of the assessment task among the participants. Several techniques, or decisionmaking aids, can assist this structuring process.

Decision Structures.--Decision trees (McNamee and Celona 1989, Merkhofer 1987, Morgan and Henrion 1989) are tools for displaying alternatives and the results of each alternative's interaction with important sources of uncertainty (fig. 4). A decision tree is composed of decision nodes, which represent sets of decision alternatives, and event nodes, which represent sets of outcomes of the sources of uncertainty. The probability of events are described at the event nodes. The end points of the decision tree are the outcomes of various interplays of actions and events. These consequences are either completely or partially out of the decisionmaker's control.

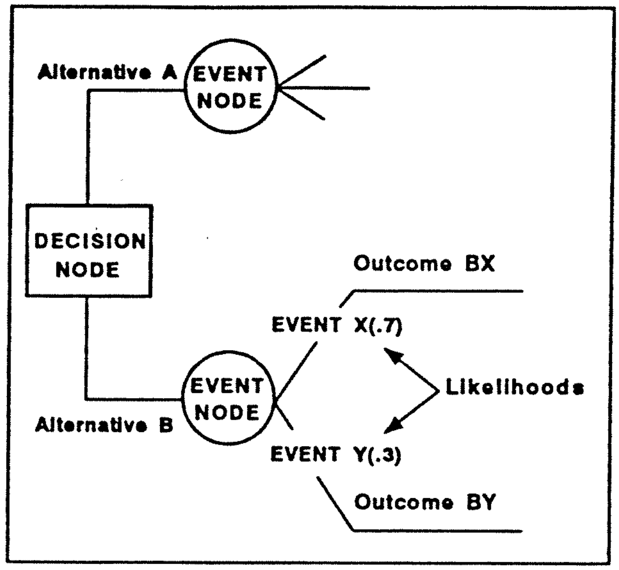


Figure 4.--*Decision tree.*

Decision trees can be analyzed for optimal decision strategies, the implications of mistakes in judgment, and for monetary values of information that might support revision of the uncertainty assessments (probabilities). They can also help identify elements that are most critical to decisions.

Influence diagrams (Shachter 1987) present flowchartlike depictions of interactions of decisions and the uncertain outcomes of elements without imposing decision tree hierarchies or sequences (fig. 5). Influence diagramming can show how probabilities can depend on other probabilities, other decisions, and revealed outcomes. Influence diagrams can then be turned into decision trees and analyzed.

Element Structures.--Assessment trees are devices for visualizing complex events as hierarchical branching patterns of component events or values (von Winterfeldt and Edwards 1986). They can be helpful in bottom-up uncertainty assessments. Unlike a decision tree, an assessment tree contains no decision points, but it suggests possible interventions that might change the overall outcome. The probabilities assigned to branches can be combined mathematically to give an overall rating or probability. One special approach to assessment trees, the analytic hierarchy process (Saaty 1988) uses paired comparisons to integrate

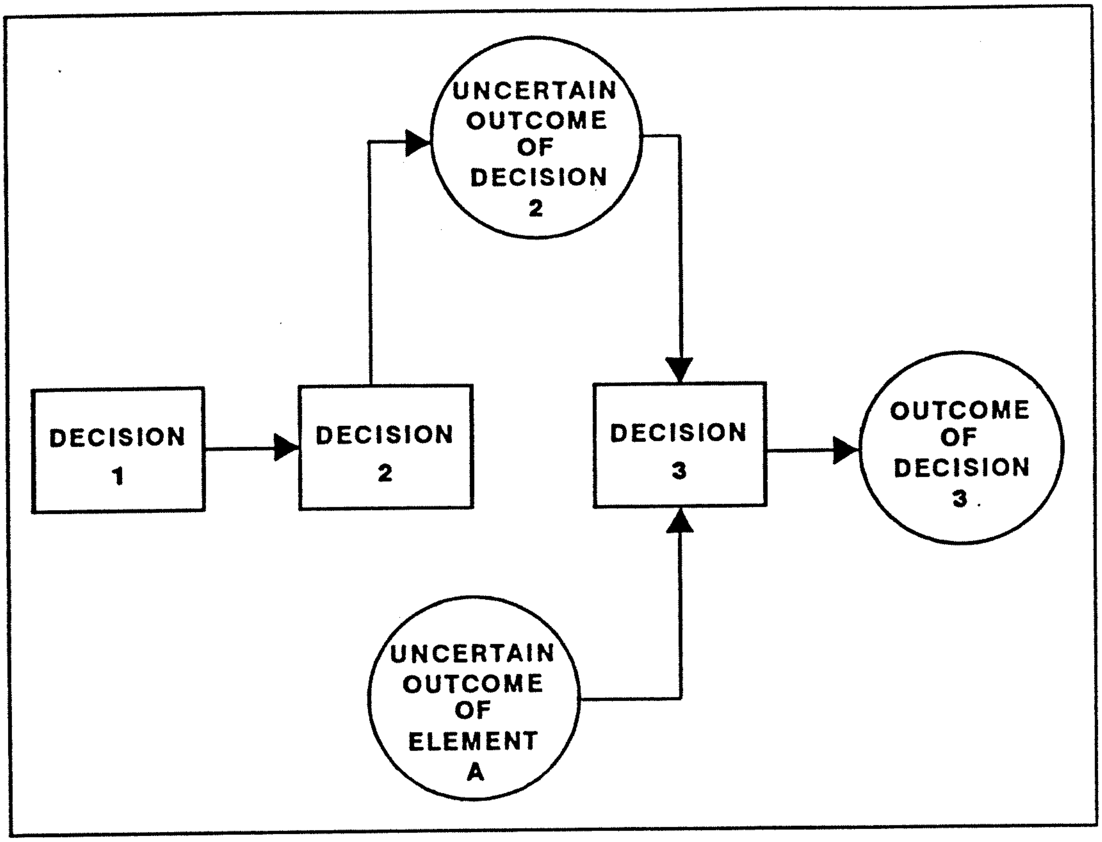


Figure 5.--*Influence diagram.*

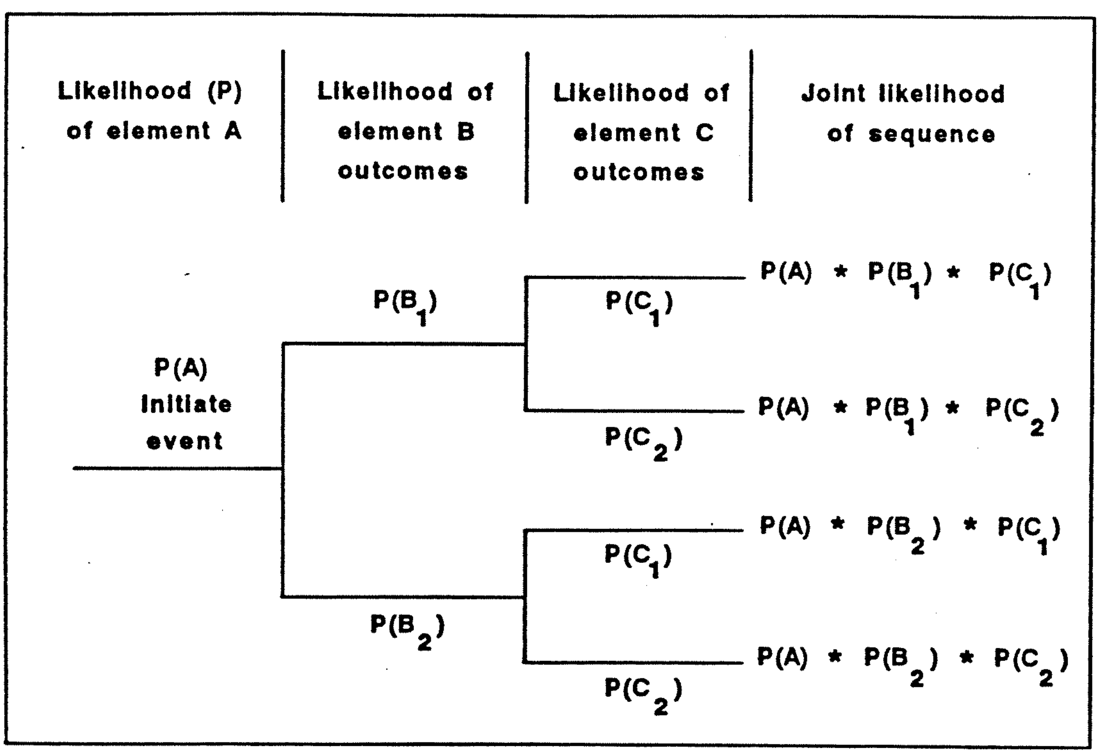


Figure 6.--*Event tree.*

tangible and intangible elements, to analyze their relative importance mathematically, and to check for logical consistency.

The premise of assessment trees is that uncertainties of component events or values are easier to comprehend and assess than uncertainties of larger elements. Sensitivity testing of assessment trees can indicate which component assessments are more critical and should get more attention in the encoding process. There are four major types of assessment trees (figs. 6, 7, 8).

Event trees (fig. 6) start with an initial event and show what events or states of the system it might lead to. Each subsequent event leads to its own set of possible final events or conditions. The uncertainties of the various branches are stated as probabilities, and the probabilities of the branches leading to the individual end conditions are multiplied to give probabilities for the end conditions. Event trees can be used to develop logical scenarios of easily imagined events and can reveal inconsistencies in thinking during the assessment process.

Fault trees (fig. 7) start with events or conditions and trace their possible causes or precipitating events. "And" nodes indicate that all lower level events must occur before the higher level event can occur. "Or" nodes indicate that any

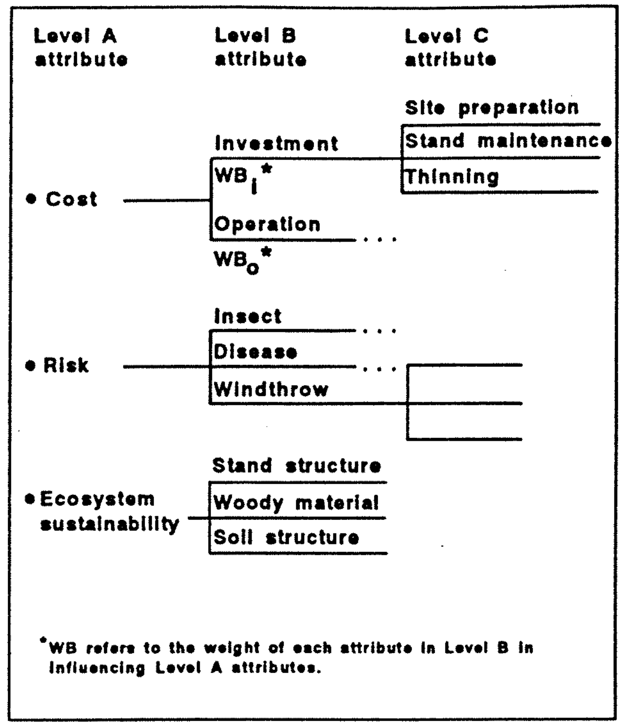
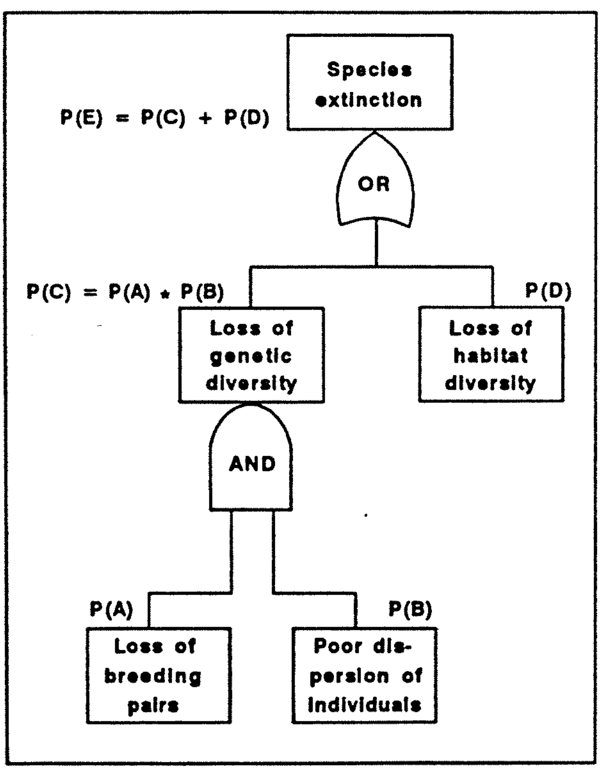


Figure 8.–*Value tree using an example of silvicultural strategy evaluation.*



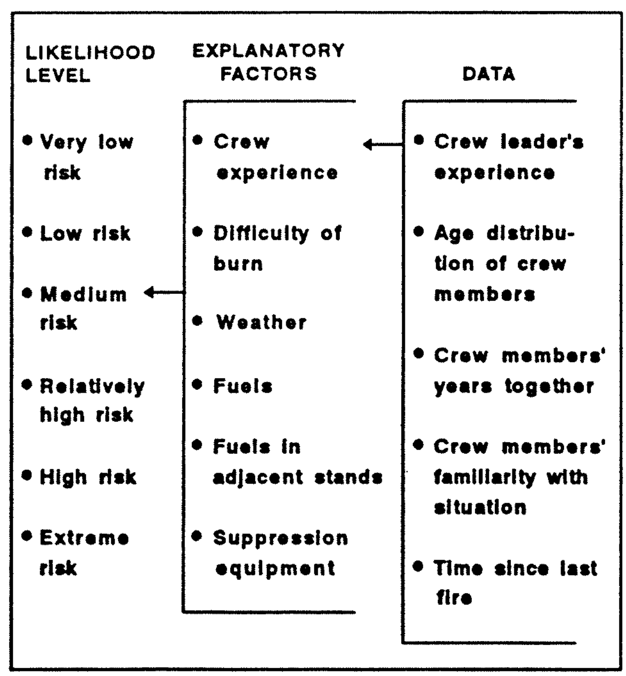


Figure 7.–*Fault tree using an example of species extinction.*

Figure 9.–*Inference tree using an example of likelihood a prescribed fire will escape.*

of the lower level events is sufficient to cause the higher level event. Probabilities of higher level events can be calculated by multiplying the probabilities at the "and" nodes and adding the probabilities at the "or" nodes.

Value trees (fig. 8) organize attributes, goals, or values into intermediate attributes and finally into tangible measurements. Each of the branches in a set is weighted according to its importance in fulfilling the higher level attribute. The attribute "risk," for example, could be broken into subattributes such as dread, voluntariness, surprise, and manageability. Value trees can help identify pivotal factors in the decisions and help assessors understand their own preferences and biases.

Inference trees (fig. 9) start with hypotheses or scenarios that are as yet unobserved. Preceding these hypotheses are layers of events that explain the hypotheses or the scenario. At the lowest levels of the tree are sources of observable data. Probabilities of the hypotheses are assessed after data and observations are fitted to the structure of the tree. Inference trees can help resolve disagreements between assessors and encourage use of the existing knowledge and data.

Response Frames for Assessments

Techniques for eliciting assessments of uncertainty differ in their response frame--the form of numerical expression and the approach to the estimate. Some techniques ask for probabilities for a given outcome; some ask for outcomes at given probability levels; others ask for both. Some techniques ask that a probability be estimated directly, while others compare the likelihood of the outcome to the likelihoods of others. Some techniques are better for assessing uncertainties of discrete events and others for estimating distributions of continuous variables. Various techniques are covered well in Finkel (1990), Kahneman and others (1982), Merkhofer (1987), Morgan and Henrion (1989), Spetzler and Stael Von Holstein (1975), and Von Winterfeldt and Edwards (1986). It is common to assign risk levels (sometimes in probability format) to future scenarios. Here, the assignment of probabilities to discrete outcomes or scenarios is discussed. Uncertainty assessments for continuous distribution are more elaborate and are covered in each of the references cited.

The indirect response frame asks for outcomes for a fixed probability and is generally preferred over direct probability assessment. Many assessors are not comfortable in directly expressing probabilities, odds, or chances.

The usual procedure for indirect responses is to ask for 10th and 90th percentile and median outcomes. The analyst often uses props to help the assessor visualize the probability relationships. The assessor compares probabilities with proportional slices of a pie or with segments of a horizontal bar. The entire pie circle or length of a bar represents a probability of 1. The slice or segment is adjusted until the assessor feels it represents the correct likelihood. Both indirect and direct assessments are very susceptible to overconfidence bias, so the analyst should word questions carefully and encourage assessors to imagine extreme circumstances.

The direct response frame can be effective after the assessor gains experience. Probabilities of less than 0.10 are usually handled better with direct estimates of probabilities or odds. A response can be expressed as an absolute number (0.20), a percentage (20 percent), a fraction (1 in 5), or even as the logarithm of a probability or odds ratio. Probabilities of rare outcomes are more easily expressed in fractions such as "1 in 100" or "1 in 1,000."

Analysts/facilitators can check assessments by shifting from one response frame to the other. If the initial assessment was in terms of direct probabilities, the A/F can ask for descriptions of events at key probability levels like the 50th (median) or the 10th and 90th percentiles. Another check is to change the visual prop or use a different analyst. These procedures help identify inconsistencies and build confidence in the assessment results.

Guidelines for Conducting Assessments

1. Use an analyst/facilitator. Select an A/F who is patient, impartial, and experienced in helping people through judgment tasks. The A/F should not be a decisionmaker, staff specialist, or one of the assessors. Beware of A/F is who try to fit answers to particular analytical forms, even when the assessors do not seem to think in those terms. Use more than one A/F's to mitigate fatigue and provide constant vigilance against judgmental inconsistencies. The A/F should understand the decision thoroughly but should not advocate any alternatives.
2. Select assessors carefully. Expertise in the subject may ensure that the uncertain element is understood but does not guarantee high-quality assessments. Look for people who have had experience making judgments and receiving feedback in managerial and political decisions. Check their understanding of basic rules of probability and their willingness to express their judgments in numerical forms. Beware of "experts" who may have ulterior agendas or preconceived preferences for particular outcomes. Assemble a portfolio of assessors who can cover the technical territory and counterbalance each other's biases and weaknesses in assessment skill.
3. Create an environment where the assessor(s) can focus on the task. Remove distractions and any references to the decision that could bias the assessors' thinking.
4. Relieve assessors' fears. Emphasize that judgments of probability are no more difficult or dangerous than any of the judgments of condition or quantity that the assessors make every day. Assure them that there are no wrong answers. Judgments of complete uncertainty are perfectly acceptable and are valuable in the decision process. Assessors should try to ensure that they have given their best judgments. They should be rewarded

for their honesty and effort and not for displaying high confidence in particular outcomes. Promise and deliver anonymity and confidentiality.

5. Make efficient use of assessors' time and efforts. Only the uncertainty of the elements most important to the decision should be assessed. Isolate these critical elements before beginning the assessment process.
6. Define the uncertain element as specifically as possible. The definition should include all outcomes that could actually occur and should exclude all others. Involve assessors in the definition process. Brainstorm "possibility lists" before the assessment begins to avoid discovering "new" outcomes while the numerical assignments are under way. Check element definitions for consistency with the decision structure. Ask assessors to describe the element in their own words. All members in a group assessment should agree on the definition and be able to provide examples to one another. Challenge assessors to think about the complements of single outcomes and to probe for possible outcomes that might be omitted. The clairvoyant test (Spetzler and Stael Von Holstein 1975) is a good check for completeness and clarity in outcome definition. In terms of this test, the clairvoyant knows all but only in the form of facts and cannot make interpretations or inferences. If a clairvoyant could identify the actual outcome with a true or false reply or without requesting clarification, it is well defined.
7. Be prepared for problem elements. Growth rates and scenarios present special pitfalls and difficulties. For example, people find it difficult to assess the uncertainty of estimates of compounded quantities. It is better to ask for the uncertainty surrounding an actual value at the end of a specified growth period and calculate the implied rate of growth (McNamee and Celona 1989). The probabilities of conjunctions of events are usually overestimated because descriptions of such conjunctions make them seem more plausible. Likewise, probabilities of single events or scenarios with less detail are often underestimated. Test scenario-type outcomes with simple inference trees or questioning that encourages the assessors to imagine alternatives. It is easy to assume that assessors understand what information is contained in a probability distribution or in parameters such as the mean or the variance. Assessors may not know how the mean, median, and mode differ or that distributions can have different shapes. People tend to shy away from asymmetric distributions because of long conditioning with the normal bell-shaped curve, but many uncertain elements have their most consequential events in the tails (low probabilities). Leave assessments of continuous variables to experts who know both the subject matter and the mechanics of probability density functions. Always cross-check assessments of distribution parameters against estimates for specific outcomes. People tend to underestimate the uncertainty of outcomes that are observed frequently and often assess the likelihood of unique events poorly.
8. Describe outcomes in terms that are meaningful to the assessor and to the decisionmaker.
9. Separate the task of describing the outcomes or events from that of encoding the probabilities. Asking assessors to do both things at the same time makes their work harder and can cause assessors to overlook possible outcomes.
10. Use a mixture of assessment formats, scales, and questioning protocols with the same assessors rather than relying on only one. Cross-check results and average them if differences cannot be resolved.
11. Challenge assessors to explain their rationale and to consider extremely likely and unlikely outcomes. Ask for estimates for extremely unlikely outcomes first and then obtain estimates for more likely outcomes. People tend to "anchor" on initial estimates unless they are assisted in recognizing the true variability in outcomes.
12. Don't discourage low levels of confidence or wide ranges in credible intervals. A high level of uncertainty is a legitimate response, reflecting a true lack of predictability that should be considered in the decision process. High levels of confidence (narrow ranges) are not necessarily good, because confidence can emanate just as easily from illusions as from accurate representations of the knowledge base or future events.
13. Check estimates against any records of similar elements. Don't overlook similar types of events or outcomes and any recorded experiences. Ask assessors to reconcile differences between their assessments and historical records.

Judgmental Biases

Assessment bias is a discrepancy between the assessor's estimate and an accurate numerical description of the assessor's knowledge. Calibration is the degree of agreement between the assessment and the actual statistical probability of an outcome. Both assessment bias and genuine lack of knowledge contribute to poor calibration.

In many instances, calibration cannot be determined because the relevant probabilities are unknown. However, the A/F should strive to obtain the best and most justifiable assessments possible. Assessment processes should be adjusted to correct for modes of judgment that give rise to biases. Bias can result from errors in defining the assessment task or from the thinking processes of the people involved. Although the following discussion concentrates on assessor biases, much of it could apply to equally serious biases of A/F's or decisionmakers.

Task bias occurs when the assessment task is not well defined, and when assessors do not understand what they are being asked to evaluate, but are reluctant to admit it. They provide answers anyway, often out of a sense of duty or pride. Task bias can be prevented when A/F's and assessors know exactly what is going to be done and exactly what information is needed. The assessors should understand the task so well that they can explain it clearly to others. Defining the assessment task may take 60 to 80 percent of the assessment time, but it is worth the effort.

Conceptual bias can be motivational or cognitive. Motivational bias results from the assessor's values or perceptions of rewards and punishments associated with the assessment results. Assessors may be predisposed in favor of one decision alternative and may unconsciously orient their assessments to support it. Personal feelings about people or organizations may also cloud assessors' judgments. Some specialists may be accustomed to being rewarded for being confident about proposals and inspiring this confidence in others; they may find the process of expressing uncertainty difficult.

Disciplines often imply pledges to anticipate certain events; to the specialist, these events may loom larger because they are a central focus. Experts often overestimate the likelihood of dire events, especially when they perceive that their predictions are bases for action. They would rather be wrong by predicting events that do not happen than by failing to predict events that occur. For example, ecologists who focus their work on ecosystem fragility may overestimate the likelihood of an ecosystem disaster. Engineers, focused on controlling natural systems, might underestimate the probability of the same event.

Managers, specialists, and planners usually give assessments that are too narrow, but for different reasons. Managers tend to concentrate on results as "done" (even before they are actually done). Specialists try to fulfill expectations that they are knowledgeable. Planners try not to disagree with in-house predictions or assessments that have already been made. Some assessors may rely on information that looks technical because they think it is objective, but it is actually poor in quality or irrelevant to their tasks.

There is no cure for motivational bias, but it can be combatted in several ways. One approach is to divide the uncertain element into subelements, creating enough detail so that assessors do not know how their answers will serve their self-interests. Assessors can also be reminded directly about motivational biases and asked to recognize and list symptoms in their own thinking. The A/F should listen for signs of wishful thinking or inappropriate pessimism or optimism and should challenge assessors to explain their assessments. Be alert for words such as "should" and "ought" that may indicate strong preferences for certain outcomes, and ask about personal experiences that could influence the assessor. Assure assessors that their work is confidential or that they are otherwise protected from reprisal or reward. Assembling a mixture of assessors who may hold offsetting motivational biases may help.

Cognitive bias is introduced by the way in which the assessor psychologically processes information. The most common cognitive bias is to be overconfident, that is to estimate intervals or probability distributions that are too narrow. This is also called conservatism. Other biases include predicting events that happen at the same time (conjunctive), insensitivity to available data about frequencies of similar events (base rates), expecting sequential patterns of high and low results to respect themselves, overestimating human control, and mistakenly perceiving of patterns in random phenomena. Bazerman (1986), Dawes (1988), Kahneman and others (1982), and von Winterfeldt and Edwards (1986) provide good explanations of these and other biases and corrective procedures.

Modes of Judgment--Sources of Bias

People judge uncertainty as they judge distance--by relying on cues and landmarks that simplify complex computational tasks. These intuitive judgments become second nature in various modes of judgment or subconscious heuristic processes (Kahneman and others 1982, Tversky and Kahneman 1974). Judgment is necessary for survival, but excessive or inappropriate reliance on specific modes of judgment can lead to serious cognitive biases. The A/F and the assessor must recognize such causes of bias. Important modes of judgment include anchoring, reliance an availability, reliance on representativeness, reliance on internal coherence, and reliance on hindsight.

Anchoring.--When estimating an uncertain outcome, the individual searches for an approximate starting point. If asked to provide a range, the individual adjusts from the initial estimate. Two problems can beset this process. First is the possibility of a misleading starting point. With especially difficult or vague uncertain elements, assessors are readily tempted to adopt any suggestion--an organizational forecast, a strategic plan goal, a performance standard, or other value. Second is the common tendency to make only small adjustments from the starting point, which acts as an anchor on the assessment. If information that is contrary to the initial estimate is provided, the individual usually ignores it. This problem is worse with an easy assessment task, because assessors are very familiar with the element. These problems can lead to inappropriately precise assessments of uncertainty.

The A/F should encourage assessors to fully consider outcomes and probabilities beyond the ranges given in initial estimates. An important approach is to ask for extreme estimates first, regardless of response mode or technique. Asking first for the "most likely" or "best" or "average" estimate promotes anchoring and insufficient adjustment, and results in overconfidence that is hard to combat later. When focusing on very rare or very frequent events, it may be necessary to ask first for outcomes at intermediate probabilities, say 0.25 and 0.75.

Reliance on Availability.--Events that are more easily remembered are usually judged to be more probable. People forsake quality of information for convenience. Familiarity, saliency, recency, and imaginability influence judgments more than statistical patterns do because they influence what is available from memory. Information that is ordered, redundant, imaginable, sanctioned, or consistent with a person's worldview is more likely to be stored in memory and recalled (Kahneman and others 1982). The A/F should probe for experiences that may have receded in memory and should present extreme outcomes as straw men, asking the assessor to imagine that they have already occurred and to explain their occurrence. The A/F should pay attention to how outcomes are sequenced in the assessment questioning. People often grossly overestimate or underestimate values following extreme outcomes because they ignore the normal regression toward the mean. For example, they may expect an extremely hot, dry year to be followed either by another dry year or by an extremely wet year when an "average" year is actually more likely.

Assessors should also be warned about conjunctive fallacy--the tendency to overestimate the likelihood of scenarios composed of several events occurring together. For example, the conjunction of a bad fire season, an insect outbreak, personnel shortages, and a budget shortfall may be a vivid and plausible scenario, but the likelihood of all these events occurring together is quite small. Scientists who have been trained to look for connections among elements in ecological or economic systems can sometimes falsely imagine patterns in what are largely random coincidences. "Worst-case" scenarios, formerly required by the Council on Environmental Quality (CEQ) in environmental impact analyses were notoriously prone to conjunction bias. Decisionmakers were thought to pay more attention to the vividness of the scenarios than to the extreme smallness of the probabilities that the scenarios would occur. The CEQ abandoned the requirements for worst-case analysis in 1986.

Reliance on Representativeness.--People tend to judge events by their degree of similarity to other familiar events or to some stereotypic image. They extrapolate occurrences from even a small sample of events to the event they are being asked to assess, becoming insensitive to sample size, the reliability of the evidence, or their own knowledge of what could cause the event. Representativeness results in stereotyping outcomes, as in imagining disastrous consequences where a few characteristics of a widely familiar disaster are present. Objective information may be forsaken for vivid, detailed descriptions of what are actually rare events. Media coverage of disasters, tragedies, and scandals or other extremes (good or bad) can easily influence assessments.

The A/F should encourage the assessor to consider published data for similar uncertain elements and outcomes, and ask assessors for descriptions of events that seem to drive their judgments in the opposite direction. Too much talk about the "driver" event may lead the assessor to make the estimate more conservative. The A/F should also listen for stereotyping --quickly putting outcomes or information about them into classes--which could cause assessors to ignore important information.

Reliance on Internal Coherence.--People often make judgments conform to beliefs and experiences built up through the years. They reject new information that appears to be inconsistent with established beliefs or that describes events they haven't experienced. The result is that they underestimate uncertainty for unfamiliar events.

The A/F should take care that one outcome does not sound more logical and is not defined more clearly or described in more detail than others. The assessor should examine the uncertainties of subevents to check the plausibility of the overall outcome. Checklists of assumptions can help keep assessors from being bound by their own beliefs.

Reliance on Hindsight.--People overestimate the probability of past events. In hindsight, people find plausible explanations for events they did not foresee. Hindsight bias makes people forget that important information about an event was not known before the event occurred (Fischoff and others 1982). Hindsight bias causes people to overpredict events they have experienced and underpredict events that have not yet occurred.

The A/F should listen for words such as "inevitable," "ultimately," "should have (known)," and "looking back." These are clues of hindsight judgment. The A/F should be ready to produce contrary evidence, scenarios, and data that disagree with past experiences and perhaps challenge preliminary judgments in the assessment process. It may be necessary to withhold information about past outcomes to enable assessors to concentrate on developing a foresightful perspective.

Group Assessments

It may be desirable or necessary to use a group of experts to develop a consensus assessment. Several assessors can have a wider range of knowledge and a greater range of experience concerning the uncertain element. Also, the assessment may call for a mixture of key specialties, as in an environmental impact statement. Biases of some assessors may be compensated for by contrasting biases of other assessors. If the group is skillfully led through the assessment process, group members may gain insights and thus provide better judgments. Also, assessors may return to their specialized fields with a fresh perspective on how their organization faces the future and how their work complements that of experts in other fields.

Group assessments also have disadvantages. They can be costly and divert attention from other functions of the organization. They may also stir up old controversies and arguments between people of different disciplines. Group assessments do not necessarily provide better estimates of uncertainty than those provided by a single assessor. One very important group bias is "groupthink" (Janis 1972)--the tendency for group members to agree because they want to be cooperative or because they fear that they will be subject to reprisals if they disagree. This tendency can result in serious overconfidence. Groups of assessors from similar disciplines or interest groups are very prone to groupthink. However, groupthink can be reduced by selecting assessors carefully and using an astute and experienced A/F.

Techniques for developing consensus assessments are of two types: Behavioral approaches ask for initial assessments and encourage development of consensus through discussion (Seaver 1978). They use group interaction and can be useful in complex assessments where the element and its outcomes have not been defined clearly. They also are useful when assessing unique or rare events where interaction may be crucial to forming models of causation. Experimental comparisons have shown that there is little difference between results obtained through the two approaches for well-defined outcomes and tasks.

There is some evidence that group interaction leads to groupthink bias when extreme outcomes are estimated. Mathematical approaches combine assessments of respective probabilities by giving more weight to some assessments. There are several options for combining assessments mathematically (Winkler 1968). However, the arithmetic mean of the individual assessments has been shown to be just as reliable as values obtained by complex systems of weighting. For difficult assessments involving combinations of effects, differential weighting schemes may be necessary. Weights can be based on differences in technical ability, the assessors' own ratings of confidence, the importance of the subject area, and the assessors' records of accuracy in estimating known uncertainties.

Options for group assessments differ in their degree of face-to-face interaction (Van de Ven and Delbeog 1971). One is simply free-form discussion led by the A/F. In the familiar Delphi technique, assessors remain anonymous to one another. Feedback is coordinated by the A/F, proceeding through several iterations until further improvement is not worthwhile. The nominal group technique consists of silent individual efforts in a group setting followed by face-to-face discussion and perhaps voting until a consensus is achieved. There is no best technique, but one highly recommended hybrid consists of a Delphi-type process for initial estimates and one round of revision, followed by structured discussion among group members, then silent voting on two or three assessments for each outcome. This estimate-talk-estimate sequence works well in many applications but requires much time.

Assessing Rare Events

Some highly unlikely outcomes must be considered because their consequences are so great. Many natural or human-caused disasters and human health hazards have this characteristic. Sometimes little can be done to influence the likelihood of a rare outcome, but resources can be allocated to protect resources, people, or programs at risk.

Probabilities of events or outcomes that might occur less than 1 percent of the time are difficult to assess because few people have had any experience with such events or outcomes. Also, humans naturally have difficulty distinguishing among small probability values. Some authors (Finkel 1990, von Winterfeldt and Edwards 1986) recommend logarithms of odds ratios as a useful numerical frame for rare events. The logarithmic scale allows more resolution and assessor discretion in the tails of distributions.

Selvidge 1980 describes the following procedure. First, have assessors define the outcome in their own words and identify causes--sets of mutually exclusive events and sequences of events. The sum of the probabilities of these events is one. The rare outcome can be modeled using an assessment tree. Second, list the subevents in order of probability and estimate the probability of the least likely event as a fraction of the probability of the most likely event. Third, compare the likelihood of the most likely event with that of an event having a probability supported by data or experience. Fourth, from this base, assign probabilities to other events. Finally, calculate the probability of the rare event by multiplying across event sequences to find the most likely path.

Assessment Quality

Assessment quality is judged on knowledgeability, usefulness, and accuracy. It is possible to evaluate the quality of the assessment process or the quality of assessment results. Well-conducted processes can generate poor assessments when knowledge is inadequate. Knowledgeability is an intangible criterion, but adherence to procedural rules can ensure that the best subjective judgments are obtained.

Usefulness depends both on the output and on the confidence that decisionmakers have in the assessors' judgment. Decisionmakers are the ultimate judges of assessment usefulness. Assessments are useful if they are simple, clear, and in terms that are fitted to the decisionmakers' deliberations. Most people are not used to confronting uncertainty in numerical or graphic form, so decisionmakers may have to be trained and gradually introduced to concepts. The A/F can help decisionmakers acquire these new skills.

Often, accuracy or calibration is not directly measurable because there is little information on natural occurrences of the outcomes. However, the concept of calibration can be used to familiarize assessors with their mission, improve assessors' methods of gathering information, and improve the communication of assessments. By assessing the uncertainty of elements for which there is data, the assessors can develop the skills and attitude required for assessing the uncertainty of elements that are more complex (Winkler and Murphy 1968). Few organizations collect data on outcomes after assessments have been performed, but feedback from experience can improve assessors' abilities and establish baseline levels of calibration.

Calibration is a comparison of observed outcomes to the probabilities assigned by the assessor. An assessor with high calibration will would assign a 0.20 probability to an outcome that data or later experience shows to occur about 20 percent of the time. Calibration is important when the

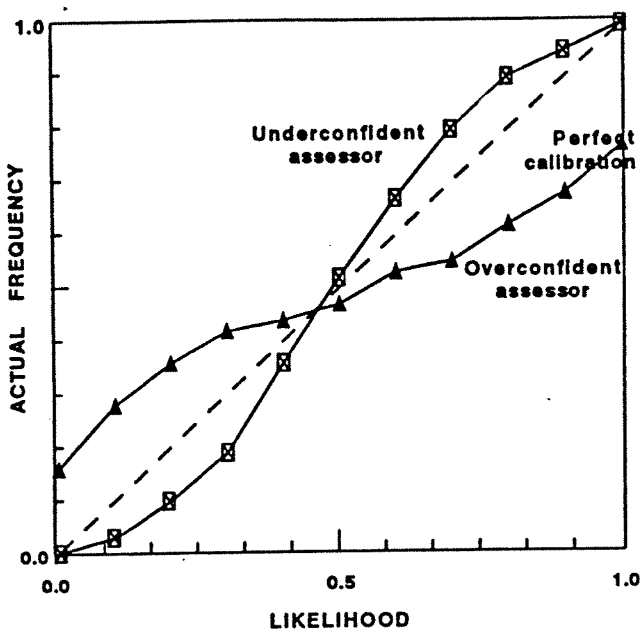


Figure 10.–*Calibration curves for assessments of likelihoods.*

number of alternatives in a decision is small, the payoff difference is quite large (e.g., in life-and-death emergencies), and when probabilities of several events are linked so that systematic errors can be compounded.

A calibration curve is a plot of actual or observed frequencies against the assessed probabilities for particular outcomes (fig. 10). The equation of a line representing perfect calibration would be $x = y$. Few people, even experts, provide perfectly calibrated assessments; most are overconfident. Overconfident assessors underestimate the probabilities of infrequent events and overestimate the probabilities of more common events. Ironically, the problem of overconfidence is worse in difficult assessments. Many people are actually underconfident in simpler assessments.

Calibration is determined by the assessor's knowledge of the subject matter and his or her ability to relate that knowledge to the A/F in probability statements. The assessor's recognition for scientific achievement or for his or her broad knowledge of a subject does not ensure reliable assessments. Training in probability assessment methods and rapid, frequent, and vivid feedback have been shown to improve the calibration of assessors.

LITERATURE CITED

Armstrong, Jon Scott. 1978. Long-range forecasting: from crystal ball to computer. New York: John Wiley & Sons. 612 p.

Bazerman, Max H. 1986. Judgment in managerial decision-making. 2d ed. New York: John Wiley & Sons. 195 p.

Behn, Robert D.; Vaupel, James W. 1982. Quick analysis for busy decisionmakers. New York: Basic Books, Inc. 415 p.

Beyth-Marom, Ruth; Shlomith, Dekel; Gombo, Routh; Shaked, Moshe. 1985. An elementary approach to thinking under uncertainty. Hillsdale, NJ: Lawrence Erlbaum Association. 154 p.

Capen, E.C. 1984. The difficulty of assessing uncertainty. In: Howard, R.A.; Matheson, J.E., eds. Readings on the principles and application of decision analysis. 2d ed. Menlo Park, CA: Strategic Decisions Group: 591-600.

Cleaves, David A. 1994. Out of uncertainty into forest policy: learning from the FEMAT risk assessments. In: Proceedings of the 1993 annual meeting of the Society of American Foresters; November 7-10; Indianapolis, IN. Bethesda, MD: Society of American Foresters: [page numbers unknown].

Cleaves, David A. 1987. Cognitive biases and corrective techniques: proposals for improving elicitation procedures for knowledge-based systems. International Journal of Man-Machine Studies. 27: 155-166.

Cleaves, David A.; Smith, Eric L.; Bradshaw, William G. 1987. Subjective assessments in the design and use of forest protection models. In: Proceedings of the Society of American Foresters symposium on system analysis in forestry; 1985 December 9-11; Athens, GA. Athens, GA: Georgia Center for Continuing Education, University of Georgia. 200 p.

Dawes, Robyn M. 1988. Rational choice in an uncertain world. Orlando, FL: Harcourt Brace-Jovanovich. 346 p.

Finkel, Adam M. 1990. Confronting uncertainty in risk management: a guide for decisionmakers. Washington, DC: Resources for the Future, Center for Risk Management. 68 p.

Fischhoff, Baruch. 1990. Understanding long-term environmental risk. Journal of Risk and Uncertainty. 3: 315-330.

Fischhoff, Baruch; Slovic, Paul; Lichtenstein, Sarah. 1982. Lay foibles and expert fables about risk--part 2. The American Statistician. 36(3): 240-255.

Gregory, Robin; Keeney, Ralph; von Winterfeldt, Detloff. 1992. Adapting the environmental impact statement process to inform decisionmakers. Journal of Policy Analysis and Management. 11(1): 58-75.

Ibrekk, H.; Morgan, M.G. 1987. Graphical communication of uncertain quantities to nontechnical people. Risk Analysis. 7(4): 519-529.

Janis, I.L. 1972. Victims of groupthink. Boston: Houghton Mifflin. 250 p.

Johnson, K.N.; Franklin, J.F.; Thomas, J.W.; Gordon, J. 1991. Alternatives for management of late-successional forests of the Pacific Northwest. A report to the Agriculture Committee and the Merchant Marine Committee of the U.S. House of Representatives. Washington, DC. 59 p.

Kahneman, Daniel; Slovic, Paul; Tversky, Amos, eds. 1982. Judgment under uncertainty: heuristics and biases. Cambridge, UK: Cambridge University Press. 555 p.

Kent, Sherman. 1964. Words of estimated probability. Studies in Intelligence. 8: 49-64.

McNamee, Peter; Celona, John. 1989. Decision analysis with Supertree. 2d ed. San Francisco: Scientific Press. 324 p.
Merkhofer, M.W. 1987. Quantifying judgmental uncertainty; methodologies, experiences, and insights. Institute of Electrical and Electronic Engineers (IEEE) Transactions on Systems, Man, and Cybernetics. 17(5): 741-752.
Morgan, M. Granger; Henrion, Max. 1989. Uncertainty: a guide to dealing with uncertainty in quantitative risk & policy analysis. Cambridge, UK: Cambridge University Press. 325 p.
Saaty, Thomas L. 1988. Decisionmaking for leaders: the analytical hierarchy process for decisions in a complex world. Pittsburgh, PA: University of Pittsburgh. 258 p.
Saveland, James M.; Stock, Molly; Cleaves, David A. 1988. Decisionmaking bias in wilderness fire management: implications for expert system development. Artificial Intelligence (AI) Applications in Natural Resource Management. 2(1): 17-29.
Seaver, David A. 1978. Assessing probability with multiple individuals: group interaction versus mathematical aggregation. Res. Rep. 78-3. Los Angles, CA: University of Southern California, Social Science Research Institute. 35 p.
Selvidge, Judith. 1980. Assessing the extremes of probability distributions by the fractile method. Decision Sciences. 11: 493-502.
Shachter, Robert. 1987. Probabilitistic inference and influence diagrams. Operations Research. 36: 589-604.
Spetzler, Carl S.; Stael Von Holstein, Carl-Axel S. 1975. Probability encoding in decision analysis. Management Science. 22(3): 340-358.
Talcott, Frederick W. 1992. How certain is that risk estimate? Resources for the Future. 107: 10-15.
Tversky, Amos; Kahneman, Daniel. 1974. Judgement under uncertainty: heuristics and biases. Science. 185: 1124-1131.
Winkler, Robert L. 1968. The consensus of subjective probability distributions. Management Science. 15(2): 61-75.
Winkler, Robert L.; Murphy, Allan L. 1968. Good probability assessors. Journal of Applied Meteorology. 7: 751-758.
U.S. Department of Agriculture, Forest Service, Forest Ecosystem Management Team [Interagency]. 1993. Forest ecosystem management: report of the forest ecosystem management assessment team. A report to the President of the United States. Washington, DC. 733 p.
Van de Ven, A.; Delbeog, A.L. 1971. Nominal versus interacting group processes for committee decisionmaking effectiveness. Academy of Management Journal. 14(2): 203-212.
von Winterfeldt, Detlof; Edwards, Ward. 1986. Decision analysis and behavioral research. Cambridge, UK: Cambridge University Press. 604 p.

*U.S.GPO:1994-565-016/00010

Cleaves, David A. 1994. Assessing uncertainty in expert judgments about natural resources. Gen. Tech. Rep. SO-110. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 17 p.

Judgmental assessments and predictions of biological, physical, social, and economic phenomena are necessary inputs for policy and managerial decisions about natural resource management. Uncertainty assessments can be used to compare alternative projects and risks, help prevent mistakes, communicate and justify decisions, guide research, and establish ecological monitoring programs. A working definition of uncertainty, a general rationale for quantifying and using uncertainty in decisionmaking, and a set of structured processes and guidelines for estimating uncertain quantities, values, relationships, and events are presented.

Keywords: Decision analysis, judgmental biases, modeling, risk, risk analysis, subjective probability assessment.