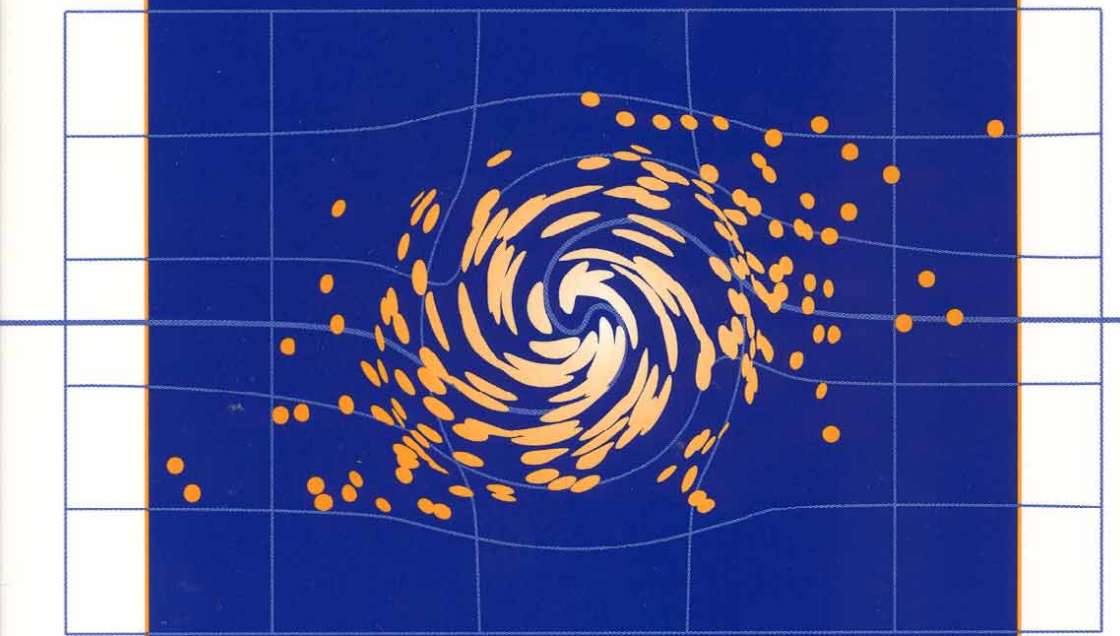
PREDICTION

Science, Decision Making,



and the Future of Nature

Edited by
Daniel Sarewitz, Roger A. Pielke, Jr.,
and Radford Byerly, Jr.

Library of Congress Cataloging-in-Publication Data

Prediction : science, decision making, and the future of nature / Daniel Sarewitz, Roger A. Pielke, Jr., Radford Byerly, Jr., editors.

p. cm.

Includes bibliographical references and index.

ISBN 1-55963-775-7 (cloth : alk. paper) — ISBN 1-55963-776-5 (pbk. : alk. paper)

Science and state. 2. Decision making. 3. Forecasting.

I. Sarewitz, Daniel R. II. Pielke, Roger A., 1968– III. Byerly, Radford.

Q125.P928 2000

363.1′07—dc21 00-008179

Printed on recycled, acid-free paper

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Decision Making and the Future of Nature: Understanding and Using Predictions

Roger A. Pielke, Jr., Daniel Sarewitz, and Radford Byerly, Jr.

A Prediction Enterprise

The story, by now, is familiar. A danger or opportunity is lurking out there, perhaps ill defined, imminent or in the more distant future, and decision makers must take action. While our ten cases are diverse, each is rooted in an effort to mobilize predictive science to pursue desired outcomes on behalf of society.

We know what undesired outcomes look like: more than ten thousand deaths from Hurricane Mitch in Central America in 1998; losses of $20 billion as a consequence of the 1993 Midwest floods; serpentine lines of cars waiting for gasoline during times of shortage; acidified lakes in temperate forests. Other outcomes are yet left to the imagination: toxic effluent leaching from pit mines and nuclear waste repositories into groundwater supplies; huge conflagrations ignited by giant asteroid impacts; fragile ecosystems collapsing under the pressure of rapid climate change. Of course, the future of human interaction with nature does not offer only disaster. Changing weather patterns might allow for more efficient agricultural harvests; the discovery of new hydrocarbon reserves (or new energy technologies) might enhance economic well-being and lessen the incentive to drill for oil in ecologically sensitive areas. But underlying every such scenario, whether pessimistic or hopeful, is the assumption by those demanding action that knowledge of the future is necessary to prevent negative outcomes and to capitalize on opportunities for gain.

It is not surprising, then, that each year policy makers invest tens of billions of dollars of public funds into technologies ranging from satellite-based observational platforms in the sky, to stream gauges on the ground, to seismometers in the deep ocean in an effort to monitor the environment and provide an ever expanding database for scientific prediction of the future of nature. Prediction has been central to such organized efforts as the U.S. Global Change Research Program, the U.S. Weather Research Program, the National Earthquake Hazards Reduction Program, the Advanced Hydrological Prediction System, the National Acid Precipitation Assessment Program, the Yucca Mountain nuclear waste repository site assessment process, and the Near-Earth Asteroid Tracking Program. Each of these science programs has been justified in terms of the need to support decisions in the present through better scientific understanding of the future.

The quest for prediction of earth systems exists in a dynamic social and political milieu that we call the "prediction enterprise." The public demands action or useful information that can facilitate action. But because the public comprises a great diversity of interests and values, it rarely, if ever, speaks with one voice about what that action or information ought to be. Other participants in the prediction enterprise include policy makers looking to satisfy (or at least address) conflicting demands made by their constituents and a scientific community looking to help define and resolve problems while at the same time satisfying its own desire to expand the frontiers of knowledge.

The prediction enterprise also involves institutions. At the international level, the United Nations coordinates activities to address climate change and natural disasters. Within the United States, the Federal Emergency Management Agency and its state and local counterparts together help citizens prepare for and respond to disasters; the National Weather Service disseminates the latest meteorological information; the Bureau of Land Management seeks to manage public lands according to its legal mandate. Universities and federally funded laboratories are integral parts of the prediction infrastructure. Private-sector institutions are also involved: the insurance industry seeks profit from investments based on a balancing of risks, airlines depend on weather forecasts to maintain safety and schedules, and the construction industry implements building standards aimed at preventing damage from a variety of natural hazards.

How effectively does this prediction enterprise serve the common interest? Its sheer complexity—diverse participants, conflicting perspectives and values, numerous institutions representing different sectors of society, and significant resources at stake—makes evaluation a daunting task. In fact, the existence of a prediction enterprise has not been recognized as such, in part, perhaps, because prediction seems like such a "natural" part of science (chapter 2), society (chapter 15), and

policy (chapter 1). Yet the prediction enterprise is as real and pervasive as "the economy" or "the medical system." As with the medical system, for example, one can look in many directions for accountability: to scientists, the media, government regulators, politicians, special interests, the nonexpert public. But unlike the economy or the medical system, little attention has been focused on the prediction enterprise. We therefore lack insight that can be applied to decision making at the intersection of predictive science and environmental policy. The cases in this volume begin the task of developing such insight.

Predictions are commonly viewed simply as pieces of information, as quantitative products of scientific research. From that perspective a prediction is understood as a "set of probabilities associated with a set of future events" (Fischoff 1994). To understand a prediction, one must understand the specific definition of the predicted event (or events), as well as the expected likelihood of the event's (or events') occurrence. When predictions are seen in this light, then the goal of the prediction enterprise is simply to develop *good predictions*, as evaluated by objective criteria such as accuracy and skill.

Yet once we have recognized the existence of a prediction enterprise, it becomes clear that prediction is more than just a *product* of science. Rather, it is a complex *process*. This process includes all of the interactions and feedbacks among participants, perspectives, institutions, values, interests, resources, decisions, and other factors that constitute the prediction enterprise. From this perspective, the goal of the prediction enterprise is *good decisions*, as evaluated by criteria of common interests.[1] The common interest is often invoked in areas such as social security and health care policy, but it should also be a rationale for the prediction enterprise.

Prediction as a Product

A central irony of this book is that the quest for prediction products can in some cases undermine the societal goals that originally motivate the quest. In the cases of earthquakes, global climate change, beach nourishment, nuclear waste, and mine impacts, for example, decision making might be improved through *less* reliance on predictions. Effective decisions are not necessarily promoted by "good" prediction products and not necessarily prevented by "bad" ones. Even so, there will be cases in which reliance on prediction is unavoidable. Knowing when to depend on predictions is itself a challenge of the prediction process and one taken up in greater detail below. (See box 18.1.)

BOX 18.1

Design of Critical Facilities without Time-Specific Predictions

Thomas L. Anderson, Construction Engineer

The engineering community is moving away from traditional prescriptive building codes toward performance-based design criteria. Experience shows that prescriptive codes do not ensure that a critical facility will continue to function in case of a natural hazard event, while performance-based codes are designed to ensure a specified level of performance in the face of specified hazards. Natural hazards of traditional concern include earthquakes and high winds.

Performance codes do not depend on prediction of specific events at specific times or places; rather, they use information based on past events and general understanding of hazard phenomena to determine the maximum expected level of stress placed on a building by a potential event.

Two examples illustrate the effective use of earthquake-related information in the design and construction of critical facilities. The examples involve the Fire Command and Control Facility (FCCF) in Los Angeles County, California, and the proposed trans-Alaska Natural Gas Pipeline (ANGP). The FCCF receives all fire and medical 911 calls for Los Angeles County and is responsible to over three million county residents. The proposed ANGP was to carry natural gas from the Prudhoe Bay fields to a terminal in southern Alaska along a route that generally paralleled the trans-Alaska oil pipeline.

The information needed for the FCCF design was the ground motions for the largest earthquake that could be expected from the several nearby faults in the region. The design requirement was for the FCCF to remain functional during and following such an event. Design alternatives were subjected to detailed calculations of their response to expected earthquake motions. All uncertainty factors were provided to the design team so that they could compare building designs based on equal values of assumed parameters and other forms of uncertainty. There was a very close and collaborative relationship between the seismologists and the design engineers at every step in the process. The final decision was based on lowest life-cycle cost and lower first cost to achieve the performance level demanded by the county.

The information needed for the ANGP was the expected ground movement where the pipeline crossed active faults. The performance requirement was to lower the risk of pipeline rupture to less than 1/2,500 a year. The project geotechnical team provided the required fault motion descriptions in great detail based on extensive research and fieldwork, and uncertainties were fully disclosed and expressed in terms understandable to the design engineers. Design engineers have a "tool kit" of strategies that allow a buried pipeline to withstand a wide range of abrupt fault movements, i.e., without having to resort to placing the pipeline above ground on sliding supports, where it is exposed to many other hazards. But use of those tools requires the predicted fault motions to be fully defined, including the nature of uncertainties in those motions. Armed with those data, lowest life-cycle cost designs were readily developed for each of the fault crossings to keep the risk of rupture below the acceptable level.

These examples show how in many situations experiential information coupled with understanding can be more useful than uncertain predictions. The construction engineer wants to build safely, i.e., to an acceptable level of risk, no matter when an earthquake occurs.

Accuracy

Given that one has decided to rely on predictions for decision making, how does one know whether a particular prediction product is a good one? A critical assessment criterion is accuracy—a measure of how closely a specific prediction product conforms to the actual event (Ascher 1981). The value of accuracy may seem too obvious to merit discussion, but sometimes accuracy is impossible to evaluate; and other times, when evaluation is possible, decision makers fail to do it. The case studies of beach erosion and mining showed that once a forecast is produced and used in decision making, there may even be disincentives to looking back and assessing predictive accuracy.

Attempts to "retrodict" or "hindcast" past events can give a measure of the accuracy of predictive methods and has been central to assessment of global climate models (see chapter 13). Comparing different prediction methodologies can also give some indication of accuracy because if the results of independent predictions diverge, they cannot all be right (although they could all be wrong). However, the case of nuclear waste disposal showed that convergence of different predictions on a similar result is not necessarily a sign of accuracy, either. Shared scientific assumptions and political incentives may cause "independent" predictions to converge on a result that is palatable, even if incorrect (see chapter 10). Predictions of beach erosion and oil and gas reserves show similar evidence of such "convergence of convenience."

The ultimate test of a prediction, of course, is to evaluate its accuracy against actual events as they unfold, which is not as straightforward as it might seem. Consider the case of early tornado forecasts (Murphy 1996). In the 1880s a weather forecaster began issuing daily tornado forecasts in which he would predict "tornado" or "no tornado." After a period of issuing forecasts, the forecaster found his predictions to be 96.6 percent correct—a performance that would merit a solid A in school. But others discovered that simply issuing a standing forecast of no tornadoes would result in an accuracy of 98.2 percent. This finding suggested that in spite of the high accuracy, action based on the forecaster's predictions could result in costs rather than benefits. In other words, *simply comparing a prediction with actual events does not provide sufficient information to evaluate its performance.* A more sophisticated approach is needed.

Scientists use a range of techniques to assess the skill of a prediction—skill being defined as the improvement of a prediction over some standard (Murphy 1997). One way to evaluate skill is to compare the accuracy of a prediction with the accuracy of some naive baseline. For example, historical weather information provides such a baseline because it yields the best estimate of the future occurrence of weather events, absent any other information. Thus, a forecast is considered skillful if it improves upon a prediction based on such climatological data. For instance, the average high temperature over the past one hundred years in London on September 6 (the climatological mean for that date) might be, say, 10 degrees Celsius. Absent any other information, the best prediction of the temperature on the next September 6 is thus 10 degrees. Any forecast for that particular day would be considered skillful if it were closer than the climatological mean to the actual temperature recorded on that date.

Such considerations suggest that our capacity to evaluate prediction as a technical product depends strongly on what is predicted. The accuracy of some types of predictions is clearly amenable to evaluation. Weather is the best example, because of the huge number of forecasts, their wide use by decision makers, and the ease of comparing forecasts with actual events, which reflects what Byerly (chapter 16) has termed the short "characteristic time" of weather events. In contrast, if an event

has a long characteristic time, predictive accuracy often cannot be evaluated. This situation applies to cases such as global climate change, long-term mining impacts, and nuclear waste disposal. Decisions on such issues will have to be made long before the skill of the prediction can be assessed. The case of floods (see chapter 5) represents an intermediate case, amenable to some evaluation of skill, yet considerably less than weather.

Uncertainty

The science fiction writer Isaac Asimov introduced, in his *Foundation* series, the notion of "psychohistorians," who could predict the future with scientific certainty based on complex mathematical models. We know that Asimov's characters lie squarely in the realm of science fiction—there can be no psychohistory. Yet the quest for a scientifically legitimated view of the future is no recent phenomenon; it dates back at least to the efforts of ancient Egyptian hydroengineers and astronomers to predict the stages of the Nile. Fifty centuries later, the future, as a weather predictor might say, still looks partly cloudy. Given today's circumstances, there are many possible ways that tomorrow might unfold (and even more possibilities for tomorrow's tomorrow). Prediction promises to narrow the range of possible futures so that decision making can be more successful. Occasional clearing can occur—we *can* predict some events with skill—but uncertainty can never be eliminated.

Stewart (chapter 3) distinguished between *aleatory* and *epistemic* uncertainty. Aleatory uncertainty is irreducible, because it is introduced by random processes in a closed system—for example, a deck of cards or a pair of dice. Epistemic uncertainty, on the other hand, derives from incomplete knowledge of a system—perhaps the dealer is a cheat, or the dice are loaded. Epistemic uncertainty can sometimes be reduced through more and better knowledge.

Even though epistemic uncertainty can be reduced, if one is dealing with open systems (as is generally the case for environmental predictions), the level of uncertainty itself can never be known with absolute certainty. Seismologists assigned a probability of 95 percent to their prediction of the Parkfield earthquake, but the earthquake never occurred (chapter 7). Were the scientists confounded by the unlikely but statistically explicable one-out-of-twenty chance of no earthquake? Or was their probability calculation simply wrong—i.e., was the uncertainty associated with the prediction in fact larger than initially thought? We would need many more Parkfield-like predictions to begin to answer such questions. Similarly, regardless of the sophistication of global climate models, many types of unpredictable events (changes in solar output, volcanic eruptions that cool the atmosphere, new energy technologies that reduce carbon emissions) can render climate predictions invalid and associated uncertainties meaningless (e.g., Keepin 1986). One way scientists deal with such "unknowable unknowns" is by introducing fudge factors into their predictions, as we saw with beach models, asteroid impact predictions, and global climate models.

Moreover, many of our cases show that efforts to reduce uncertainty reveal vast, previously unrecognized complexities. In such cases, decision-relevant uncertainties can actually *increase* with more knowledge. This dynamic of spiraling uncertainty can have the perverse effect of increasing political controversy rather than reducing it, leading to calls for even more research to reduce uncertainties, while the problem that motivated the research goes unaddressed. As Robinson (1982, p. 249) observes, "By basing present decisions on the apparent uncovering of future events, an appearance of inevitability is created that de-emphasizes the importance of present choice and further lessens the probability of developing creative policy in response to present problems." The counterintuitive lesson for decision makers is that *uncertainties about the future can often be reduced more successfully through decision making than through prediction.* (See box 18.2.)

Predictability

Asteroid orbits can be calculated from observations of the asteroid's positions combined with well-understood physical laws (chapter 6). But in the realm of earth science prediction, asteroid impacts are atypical. In most other cases, predictability is limited because knowledge of the future depends on knowing the present, which itself can never be completely or accurately characterized. For example, weather forecasts depend on knowing the present state of the atmosphere and then projecting the future behavior of the atmosphere, based on computer models. Because the future is dependent on initial conditions, small changes in these conditions can add up to large differences in outcomes. That is why maximum weather predictability is about two weeks: even though the system is well understood, measurement of initial conditions is invariably subject to error and omission.

The complexity of earth sciences phenomena of interest to policy makers increases when human and earth processes interact. Consider nuclear waste disposal. Predicting the performance of a waste facility ten thousand or more years into the future depends on knowing, among a multitude of other potentially relevant factors, how much precipitation might be expected at the site. Precipitation is a function of global climate patterns. And global climate patterns might be sensitive to human processes such as energy and land use. Energy and land use are functions of politics, policy, social changes, and so on. What at first seems a narrow scientific question rapidly spirals into unbounded complexity.

Finally, decision makers sometimes are led to believe that the sophistication of a prediction methodology contributes to greater predictive skill, i.e., that in a complex world, a complex methodology will enhance predictability. In reality, the situation is not so clear-cut. An evaluation of the performance of complex models in energy, economics, population, and other areas has shown that "methodological sophistication contributes very little to the accuracy of [predictions]" (Ascher 1981, p. 258; see also Keepin 1986). Yet energy, economics, and population are integral to any long-term, policy-relevant predictive capability in the areas of global climate change, nuclear waste disposal, and oil and gas reserve assessment. Overall, more sophistication can introduce more uncertainty and more sources of error into a prediction. Our case studies suggest that better prediction products arise more from the feedback between predictions and experience than from the introduction of more sophisticated predictive methodologies. The lesson for decision makers is that they should not be overly impressed by claims of sophistication, unless those claims are backed up by demonstrable increases in accuracy. (See box 18.3.)

The Interface of Product and Process: Understanding Predictions

Given the many factors that influence the generation of a scientific prediction, one can see why accurate, useful predictions are so hard to make. Those same factors also ensure that the predictions are hard to understand. How should the numerical or statistical output of a given predictive effort be interpreted? That is a problem that plagues scientists (with their "unknowable unknowns") as well as the decision makers who try to use predictions.

The challenge of understanding predictions was aptly illustrated in the case of the 1997 flooding of the Red River of the North (see chapter 4; also Pielke 1999). In February 1997, forecasters predicted that the river would see flooding larger than at any time in modern history. At Grand Forks, North Dakota, forecasters expected the spring flood to exceed the 1979 flood crest of 48.8 feet sometime in April. Forecasters issued a prediction that the flood would crest at a record 49 feet, hoping

BOX 18.2

Ranching and Prediction

Rob Ravenscroft, Rancher

As a rancher, I deal with physical and biological systems, as well as economic and social systems. Since these are dynamic and interconnected, attempts to predict their behavior are probably wrong more often than right. If honestly done, there are no "bad" predictions. It's my responsibility as a manager to use them properly.

"Proper" use can be measured only by progress toward a goal. Technology (including predictions), financial and biological capital, labor, and (most important) creativity pretty well sum up the tools a manager can use to devise and implement a plan of operation.

The dynamic nature of the systems involved, and the inherent possibility of errors in predictions and assumptions made, means that a plan must have two essential characteristics to be effective in achieving personal, family, and business goals. It must be monitored, and it must be flexible.

In ranching, as in most other businesses, the obvious monitoring areas are financially oriented. But cattle and beef production are just part of the entire biological system. Healthy plants, animals, and soils are critical to long-term sustainability of families, ranches, and communities. Biological alarms we watch for on our ranch are decreases in plant and animal diversity, which usually indicate some flaw in our plan that could hinder our ability to deal with future adversity.

Any plan that aims to achieve quality-of-life goals (which is the real need of the individuals and families implementing the plan) must be monitored for social impacts, too. If our ranching and business practices endanger our neighbors and community, our long-term goals can't be achieved. This is more difficult to measure but must be kept in mind.

Early warning is the most effective first step in reversing a planning mistake or reacting to a change in conditions. Flexibility built into the plan and the business is the next. For us, weather and prices are the major risk factors. Those are also factors that are regularly predicted. Experience shows that neither can be forecast with great reliability. This means that we can't afford to direct all our assets and efforts to best capitalize on any one set of predicted conditions. Here again, diversity enhances flexibility. Diverse plant communities support animals through a wider range of weather conditions. Diversity in the cattle enterprise can supply staying power as prices cycle from low to high. Monitoring lets us know when we're not achieving our goals; flexibility gives us the chance to replan and get back on track.

Science-based predictions can be enormously helpful. People are responsible for using such predictions appropriately. In most cases, that means recognizing that prediction is just one of the tools that can be used to help achieve goals. Quality-of-life-based, goal-driven plans that are economically, ecologically, and socially sound should be applied with flexibility and with a monitoring system that provides early warning when straying from the goal occurs. There are no bad predictions, only inappropriate uses of predictions.

to convey the message that the flood would be the worst ever experienced. But the message sent by the forecasters was not the message received by decision makers in the community.

Decision makers in the community *interpreted the event* being predicted and the *probabilities* associated with the prediction within the context of their own experience. First, the prediction of 49 feet, rather than conveying serious concern to the public as the forecasters hoped, instead resulted in reduced concern. Local residents and officials interpreted the forecast in the context of the record 1979 flood, which caused damages but was not catastrophic. With the 1997 crest expected to be only a few inches higher than the record set in 1979, many expressed relief rather than concern, perhaps thinking: "We survived that one. How much worse can a few inches be?" Second, decision makers did not understand the uncertainty associated with the prediction. All flood forecasts are uncertain, but predictions of record floods, i.e., floods for which there is no experience, are especially uncertain. Yet forecasters issued a quantitative prediction with a simple qualitative warning about uncertainty. Hence, many decision makers could interpret the forecast

BOX 18.3

Perspective on Prediction Use in Funding Science

Jack Fellows, Government Executive

As someone who worked in the White House's Office of Management and Budget for many years and dealt with national public policy issues related to science, space, and the environment, the factors I would consider important for using or avoiding misuse of predictions include:

- *Problem dynamics.* Is using or improving a model's prediction even germane to the problem? In the situations I most faced, I was being asked whether (1) it was worth the cost of improving a predictive model, or (2) the output from a model would be relevant to a public policy decision. With respect to the first point, it was difficult to tell many times whether a model improvement would contribute to policy making or was only a challenging scientific topic. What could be a significant scientific advance might have little impact on those who might use the prediction in the real world. Depending on the nature of the situation, either outcome could have value, but if a model improvement was being proposed to address policy issues, then the value-added of the improvement to society needed to be demonstrated. Indeed, some problems are so oriented toward mitigation or adaptation that improvement in prediction is of little consequence. For example, better warning and storm shelter improvements probably would yield significantly more return to society than small improvements in tornado model predictions.
- *Uncertainty and risk.* Can scientists adequately characterize the level of uncertainty associated with a prediction and how best to quantify the benefits or risks associated with those uncertainties?
- *Range of predictions.* This might be viewed as another form of uncertainty, but there are uncertainties associated with a specific model, and then there is uncertainty associated with a range of models addressing the same issue. If most models tend to have similar results (assuming the physics, etc., are believable), then I would be more likely to accept the results than if the models significantly differed.
- *Fidelity of the model.* How well does the model replicate the historical record? If not well, then I would discount the prediction or not use it at all. Also, does it fit the scale of the problem? Is the output global in nature when my problem is local or regional—can I scale down or up to my situation?
- *Affordability.* Can I afford the model, and do I have the tools and data to run it for my application?

ncertainty in their own terms: Some viewed the forecast as a ceiling: 'The flood will not exceed 49 feet." Others viewed the prediction as ncertain, with different individuals estimating uncertainty in the crest rediction to range from 1 to 6 feet. The historical record showed that verage error for flood crest forecasts was about 10 percent.

On April 22, 1997, the Red River crested at 54 feet, inundating he communities of Grand Forks, North Dakota, and East Grand 'orks, Minnesota, and causing $2 billion in damages. In the afternath of the flood, local, state, and national officials pointed to inacurate flood predictions as a cause of the disaster. In fact, the ccuracy of the predictions was not out of line with historical perormance by any objective measure. Instead, forecasters failed to xpress, and decision makers failed to understand, the *meaning* of ne prediction, in terms of what was being forecast and the uncerainty associated with it. The failure was one of process, not of product. (See box 18.4.)

Other cases presented in this volume further illustrate that decision makers' understanding of predictive products has a profound influence on how—and how well—the products are used. Consider the following three examples:

1. Debate has raged for more than a decade about the policy implications of possible future human-caused changes in climate. This debate has been about "global warming" expressed in terms of a single global average temperature. But no person and no ecosystem experiences global average temperature. Each policy advocate is thus free to interpret that prediction product in support of his or her particular interests, ranging from pending global catastrophe to benign (and perhaps beneficial) change. Uncertainty and the inability to compare predictions to experience allow even more interpretive freedom. Predictive science is thus used (and misused) to justify and advance the existing interests of contesting participants in the political process.

BOX 18.4

Accuracy of Flood Predictions

Dennis Walaker, Public Works

My area of emergency decision making relates to weather-related events regarding straight-line winds, tornadoes, blizzards, heavy rains, and river flooding. Accuracy is what I most expect of predictions. Timeliness, of course, is also an important element. More lead time provides the ability to react efficiently to reduce damage, save lives, and make our communities more disaster resistant.

A flood forecast is a guide, not an absolute. If you expect absolute accuracy, you will be disappointed because weather-related events have numerous variables that are all subject to change. The best flood forecast model can be wrong if conditions (e.g., rainfall, temperature, etc.) change dramatically.

In spring of 1997, when the flood of the Red River of the North occured, all flood forecasts were within acceptable ranges except in Grand Forks. We now focus on this one failure rather than the several other successes. In Fargo, we had severe weather, river gauge failures, and crest reversions, but we had time to adjust to those weather-related changes. Grand Forks had little time to react to a situation that overwhelmed them, as this was an unprecedented event.

We must rely on the National Weather Service and the Flood Forecast Center for reliable information. However, we must understand or at least better understand the variables, assumptions, and real value of their predictions. Society increasingly expects science to solve all problems. But not all problems are easily solved, and answers are not always absolute. Today when serious events are not accurately predicted, society seeks someone to blame.

Could the city of Grand Forks have been saved from the disaster if the forecasts had been absolutely correct? In my opinion, contingency dikes, earlier evacuation, and loss control were options, but they required difficult, unpopular decisions. Previous victories over floods gave a false sense of security. It would have been difficult to construct emergency measures against a 56-foot crest when dikes that were supposed to handle 52 feet failed at 51 feet.

Responsibility for protecting against disaster is a local one. Predictions are but one tool, albeit one of the most important, used by local officials along with political influence, historical reflection, acceptable loss, and other considerations.

In summary, we must support the scientists in gathering the information to achieve the best predictions. We then must question the predictions when life and property in our communities are at risk. We can't simply assume that the prediction is completely accurate without our own review. Predictions are based in part on historical data. If an event significantly exceeds all previous levels, accurate predictions may be difficult. Blaming others when failure occurs isn't enough. If we have done everything possible, we must accept the consequences—some events are beyond our control. An elderly woman victimized by a flood summarized this by saying, "Even if we lose the flood fight, you must feel that we have done everything possible to be successful."

2. In recent years scientists have increased their ability to observe asteroids and comets that potentially threaten the earth. In this case, the "event" is clear enough—possible extinction of life on earth if a large asteroid slams into the planet. But public reaction to the discovery of asteroid 1997 XF11, and the associated prediction that it could strike the earth on October 26, 2028, illustrates that scientists, as well as the public, can fall prey to misunderstanding. Blame for the misunderstanding can be apportioned among scientists who hastily issued an erroneous prediction; the media, which jumped on the prediction because it was spectacular; and the public, which responded to the magnitude of the potential event, rather than its uncertainties or probabilities.

3. Weather forecasts afford decision makers the best opportunity to understand prediction products. It is well worth repeating that in the United States the National Weather Service issues more than 10 million predictions every year to hundreds of millions of users. (In contrast, we have seen less than a dozen scientifically legitimate earthquake predictions.) This activity provides a basis of experience from which users can learn, through trial and error, to understand the meaning of the prediction products they receive. Of course, in the case of weather prediction there is still room for confusion. People may fail to understand predictions even for routine events. Murphy (1980) documents that when forecasters call for a 70 percent chance of rain, decision makers understand the probabilistic

element of the forecast but do not know whether the rain has a 70 percent chance of occurring at each point in the forecast area, or whether 70 percent of the area is expected to receive rain with a 100 percent probability, and so on. Even so, one of the important lessons of weather prediction is that decision makers, including the public, are in general able to use probabilistic information, and such products can have significant value.

These examples illustrate how viewing prediction solely as a product is inherently problematic, because doing so conceals the context that gives the product meaning. Thus, if one wants to improve the use of prediction, one must do more than simply develop "better" prediction products, whether more precise (e.g., a forecast of a 49.1652-foot flood crest at East Grand Forks), more accurate (e.g., a forecast of a 51-foot crest), or more robust (e.g., a probabilistic distribution of various forecast crest levels). While better prediction products are in many cases more desirable, *better decisions* in the common interest require attention to the broader prediction process. From this standpoint, better prediction products may be neither necessary nor sufficient for improved decision making, and hence desired outcomes. To effect better decisions, it is necessary to understand prediction as a process.

Prediction as a Process

The prediction process can be thought of as three parallel decision activities:

- *Research.* Including science, observations, modeling, etc., as well as forecasters' judgments and the organizational structure—all of which go into the production of predictions for decision makers.
- *Communication.* Both the sending and receiving of information—e.g., who says what to whom, how is it said, and with what effect.
- *Use.* The incorporation of predictive information into decision making. Of course, decisions are typically contingent on many factors other than predictions.

These activities are not sequential. They are more accurately thought of as integrated components of a broader *prediction process*, with each activity proceeding in parallel, and with significant feedbacks and interrelations between them.

A robust conclusion of this book is that good decisions are more likely to occur when all three activities of the prediction process are functioning well—and research activity is often the least critical of the three. Open communication and consideration of alternative policy approaches can lead to successful decisions in the face of unsuccessful prediction products, but the opposite is unlikely to occur (see chapter 14). Consider the following examples:

- The case of the Red River flood illustrates how a technically skillful forecast that is miscommunicated or misused can result in costs rather than benefits. The overall prediction process broke down in several places. No one in the prediction process fully understood the uncertainty associated with the prediction, hence little attention was paid to communicating the uncertainty to decision makers, and poor assumptions were made about how decision makers would interpret and use the predictions. As a result poor decisions were made. Given that that region will to some degree always depend on flood forecasts, the situation can be improved in the future by including local decision makers in the research activity in order to develop more useful products (Pielke 1999).
- In the case of earthquake prediction, a focus on developing skillful predictions of earthquakes in the Parkfield region of California brought together seismologists, local officials, and emergency managers with the original goal of preparing for a predicted earthquake. A result was better communication among those groups and overall improved preparation for future earthquakes. In this case, even though the prediction product was a failure, the overall process adapted to that failure and made decisions that enhanced awareness, refocused attention on alternatives, and arguably reduced vulnerability to future earthquakes. (See box 18.5.)
- Global climate change seems to display attributes similar to the early stages of earthquake prediction. Policy making focused on prediction has run up against numerous political and technical obstacles, while alternatives to prediction are becoming increasingly visible. The prediction *process* will be said to work if it addresses the goals of climate policy—i.e., if it reduces the impacts of future climate changes on environment and society (Pielke 1998). More and better predictions are not a prerequisite for this desirable outcome.
- Nuclear waste disposal has also evolved from a situation in which the development of skillful predictions played a central role into

BOX 18.5

A User's Perspective on Earthquake Prediction and Public Policy

Shirley Mattingly, Emergency Management

As calamities go, earthquakes pose a special threat because the really disastrous events occur infrequently. Earthquakes challenge those who would predict them and those who would respond to predictions because uncertainty surrounds the science and the response and provides an excuse for no action.

Predictions, per se, can disrupt life in a city at risk, and they don't have to be valid or even scientifically based. The self-proclaimed clairvoyant Nostradamus predicted a devastating event in May 1988 in the "new city," assumed to be Los Angeles. Widespread publicity fed rumors, which led to near panic in one community. Hundreds of families took their children out of school and permanently left their homes, relocating to the relative seismic safety of Fresno or Oregon. I was disheartened that earthquake drills at school were regarded as proof that the coming catastrophe was inevitable.

Panic is not a good thing. But emergency managers took advantage of the public's heightened awareness of earthquakes to explain the science and promote simple safety measures. Nevertheless, the Nostradamus "prediction" and similar incidents negatively impact public policy makers' regard for the predictive science.

Even apparently legitimate scientific disagreements justify inaction as the preferred policy option. A so-called seismic deficit postulated in 1994 by an official science working group was believed to mean that destructive quakes were more likely to strike the region in coming decades. Subsequently, earthquake insurance rates quadrupled (Kerr 1988). Eventually, the deficit was debunked and suddenly disappeared. But for four years Los Angelenos thought that they faced either twice as many large earthquakes as normal or one huge quake many times more powerful than the last Big One. This incident didn't improve policymakers' regard for the predictive science, either.

Science is often foreign territory for politicians, and politics is often foreign territory for scientists. Scientists and public policy setters often don't even begin to speak the same language. They generally have very different backgrounds, motivations, and aims, and they work in different milieus. So communication does not come naturally.

There's art and science in both predictions and politics. In my experience, decision makers rely on input from people they trust, people with whom they have a history. They like their advisors to do their homework, define the problem, identify alternative approaches, evaluate potential solutions, and find solid answers. And they want advice that is clear and easy to understand. Then they act based on what they've heard and on other factors we don't know about. While they are pragmatic, they can be swayed by people with passionate beliefs.

Scientists and policy makers can help each other to understand the environment that will receive—perhaps eagerly, perhaps kicking and screaming—a prediction. They should collaboratively decide how to communicate information, how to frame it, and when to release it. Both should pursue good relationships with the local media, even when there is little news, because if the media are well informed, they'll do a better job of reporting accurately when there *is* a story.

For local decision makers to take predictions seriously, they must have faith in the prediction and the predictor. Scientists must have credibility with peers and with the people they hope to influence into action. That requires sustained dialogue and mutual trust. The responsibility lies with both.

I remember one highly respected seismologist coming downtown to Los Angeles' city hall, more than once, to discuss his research findings with any public official who would listen. He came on his own initiative. He moved us, not immediately, but over time. He changed public officials' perceptions and influenced public policy. I saw it happen, and ever since, I've been trying to make it happen again, anytime and anywhere that anyone will listen.

one in which decision making focuses on actions that can achieve desirable societal outcomes under various possible futures. Initially, the success of the repository seemed to be entirely dependent on predicting the hydrologic system at the disposal site over the next ten thousand years. Unanticipated complexities associated with this natural system led to decreased emphasis on prediction and increased emphasis on designing an engineered containment system. However, while the behavior of this engineered system is likely to be much more predictable than that of the hydrologic system, it will have its own problems, and uncertainty cannot be entirely eliminated. Additional options, such as monitored, retrievable storage, may be necessary to accommodate the remaining uncertainties.

A User's Guide to Prediction and Decision Making

When to Rely on Prediction Products

The case studies in this volume provide insight into when decision makers should look to prediction products and when they should look to alternative sources of information to help make decisions.[2] The conditions under which predictions should be relied on are easy to lay out in principle but may be difficult to apply in practice. In principle, predictions should be relied on when:

1. Predictive skill is known.
2. Decision makers have experience with understanding and using predictions.
3. The characteristic time of the predicted event is short.
4. There are limited alternatives.
5. The outcomes of various courses of action are understood in terms of well-constrained uncertainties (e.g., the likelihood and effects of false positives and false negatives).

Conversely, alternatives to prediction should be sought when:

1. Skill is low or unknown.
2. Little experience exists with using the predictions or with the phenomena in question.
3. The characteristic time is long.
4. Alternatives are available.
5. The outcomes of alternative decisions are highly uncertain.

Incorporating these principles into real-world decision processes may be difficult. Organizations often choose to gather more information (even if it is useless) rather than take action (e.g., Feldman and March 1981). Political incentives favor "the basing of policy on supposedly neutral forecasts [that allow] decision making institutions to assume a cloak of objectivity" (Robinson 1982, p. 240). Rejecting that cloak in favor of the hair shirt of realism requires decision makers to:

1. Be flexible.
2. Learn from experience.
3. Search for alternatives.
4. Hedge their bets.
5. Evaluate progress with respect to goals.
6. Evaluate predictive skill with respect to decisions.
7. Focus on good decisions, not just good predictions.

Each of these guidelines might seem obvious or common sensical—but as the cases here dramatically show, they are often neglected. Overcoming this neglect requires decision makers to change their focus—from predictions as a product to predictions as a process.

Creating a Successful Prediction Process

If society is to benefit from the predictive information products of the earth sciences, scientists and decision makers should together pay attention to the broad process in which predictions are made. In particular, participants in the prediction process must take action in six areas.

1. Above all, users of predictions, along with other stakeholders in the prediction process, must *question predictions*. For this questioning to be effective, predictions should be as transparent as possible to the user. In particular, assumptions, model limitations, and weaknesses in input data should be forthrightly discussed. Institutional motives must be questioned and revealed. Especially in cases where personal experience may be limited (such as asteroid impacts and global warming), both scientific rigor and public confidence in the validity of the prediction will benefit from this open questioning process. "Black boxes," i.e., closed processes, generate

public distrust, especially when a prediction can stimulate decisions that create winners and losers. They can also foster complacency among those doing the predicting. Even so, because of limited experience, many types of predictions will never be understood by decision makers in the way that weather predictions are understood. Table 18.1 lists seven general questions that can be asked about predictions and gives accompanying guidelines for seeking answers.

2. If users are to question predictions, then *the prediction process must be open to external scrutiny*. This means that policy makers must give procedural aspects of democratic openness, evaluation, and accountability the same priority as issues that may seem more directly connected to policy goals (e.g., funding predictive research or establishing environmental standards). Openness is important for many reasons, but perhaps the most interesting and least obvious is that the technical products of prediction are likely to be "better"—both more robust scientifically and more effectively integrated into the decision process—when predictive research is subjected to the tough love of democratic discourse. Scientists may reasonably fear that such attention could lead to politicization of research agendas, but many of our case histories show the opposite—that, in the absence of public openness, predictive science tends to converge on results that support the tacit assumptions of the administering organizations or policy regimes (see chapter 17). External scrutiny helps to reinvigorate the healthy skepticism that is supposed to be a part of the scientific process. Consider scientists working on the Yucca Mountain nuclear waste repository who converged on a predictive product that was consistent with their institutional interests. The presence of two oversight bodies provided the additional, outside scrutiny necessary to expose the technical flaws in those predictions. Similarly, Moran (chapter 9) shows that, in the effort to predict the environmental effects of mines, informed public scrutiny of environmental impact statements is necessary to ensure that significant uncertainties are brought to light. Pilkey (chapter 8) describes a failed process for making decisions about beach nourishment that is, perhaps predictably, neither open nor subject to evaluation. And as Gautier (chapter 11) explains, when the U.S. Geological Survey opened up its oil and gas assessment program to a range of interested customers, it improved both its own technical capability, and the utility of its prediction products.

3. In this same context of openness, *predictions must be generated primarily with the needs of the user in mind*. Television weather

TABLE 18.1 Questioning Predictions.

Questions to Ask	Guidelines to Follow in Seeking Answers
What are the policy goals (i.e., outcomes) that prediction is intended to achieve?	Specify the purposes of the prediction.
How does the process of developing predictions influence the policy process (and vice versa)?	Consider alternatives to prediction for achieving the purpose. Maintain flexibility of the system as work on predictions proceeds. Recognize that a choice to focus on prediction (as well as the choice of the specific predictive technique) will constrain future policy alternatives.
What are the direct societal impacts of the prediction?	Consider alternative societal impacts that might result from the prediction (including the different roles played by prediction). Evaluate past predictions in terms of impacts on society. Recognize that the prediction itself can be a significant event. If possible, assess the impacts of inadequate predictions relative to the impacts of successful ones.
What are the scientific limitations and uncertainties of the prediction?	Evaluate past predictions in terms of scientific validity. Recognize that different approaches can yield equally valid predictions. Recognize that prediction is not a substitute for data collection, analysis, experience, or reality. Recognize that predictions are always uncertain; assess the level of uncertainty acceptable in the particular context. Beware of precision without accuracy. Recognize that quantification and prediction are not (a) accuracy, (b) certainty, (c) relevancy, or (d) reality. Recognize that computers hide assumptions; computers don't kill predictions, assumptions do. Recognize that the science base may be inadequate for a given type of prediction.
What factors can influence how a prediction is used by society?	Recognize that prediction may be more effective at bringing problems to attention than forcing them to effective solution. Recognize that perceptions of predictions may differ from what predictors intend and may lead to unexpected responses. Recognize that the societal benefits of a prediction are not necessarily a function of its accuracy. Recognize that there are many types of prediction, and their potential uses in society are diverse.
What political and ethical considerations are raised by the generation and dissemination of a prediction?	Pay attention to conflicts of interest among those soliciting and making predictions. Understand who becomes empowered when the prediction is made. Who are the winners and losers? Pay attention to the ethical issues raised by the release of predictions.
How should predictions be communicated in society?	Make the prediction methodology as transparent as possible. Predictions should be communicated (a) in terms of their implications for societal response, and (b) in terms of their uncertainties.

predictions focus primarily on temperature, precipitation, and wind, rather than temperature gradients, behavior of aerosols, and barometric pressure. Scientists must understand the broader goals of the process, not the narrow goals of science; they must listen to stakeholders. Stakeholders must work closely and persistently with the scientists to communicate their needs and problems. To ensure useful prediction products, prediction research programs should be designed from their inception to include mechanisms of formal and informal, regular and frequent dialogue between prediction researchers and prediction users. More communication between producers and users of predictions always benefits the prediction process and the quest for good decisions, even if it introduces inefficiencies in the generation of prediction products.

4. *Uncertainties must be clearly understood and articulated* by the scientists, so that users understand their implications. If scientists do not understand the uncertainties—which is often the case—they must say so. Failure to understand and articulate uncertainties contributes to poor decisions that undermine relations among scientists and decision makers. *But merely understanding and articulating the uncertainties does not mean that the predictions will be useful decision tools.* For example, if policy makers truly understood the uncertainties associated with predictions of global climate change, they might decide that strategies for action should not depend only on predictions (e.g., Pielke 1998).

5. *Decision makers must realize that predictions themselves can be significant events.* Predictions can stimulate considerable action that can confer benefits or impose costs. False earthquake predictions have stimulated better earthquake preparedness, while false asteroid impact predictions have fueled needless alarm. More significantly, predictions can commit society to one course of action while foreclosing other options. The prediction of global warming, for example, has mobilized an international effort to reduce anthropogenic CO_2 emissions. Some would argue that such action is necessary to forestall disaster; others, that it is a fruitless and potentially dangerous distraction from more effective approaches to global environmental protection. In either case, the prediction itself has been a much greater catalyst for decision making than any unambiguous impact of global warming. A healthy prediction process depends on the recognition that predictions are themselves events.

6. Finally, *the quest for alternatives to prediction must be institutionalized in the prediction process*, especially when characteristic times are long, policy regimes are strong, and decision makers have limited (or no) experience with the predicted phenomenon. Alternatives to prediction should be debated and evaluated (and perhaps tried on a pilot basis) at the earliest stages of the prediction process. As our case studies show, alternatives are in fact often available. Rather than trying to predict the impacts of hard-rock pit mines on water quality as a basis for environmental regulation, spreading risk through bonding or other types of insurance might be preferable. Rather than depending on predictions of acid rain mitigation to design a regulatory command-and-control system (see chapter 12), the U.S. Congress actually implemented a system of tradable emissions permits that did not depend on predictive earth science.

In Conclusion: Question Predictions!

The emergence of an environmental challenge appears to stimulate an almost automatic call for scientific prediction as the first step toward meeting the challenge. The possible sources of this reaction range from the desire to find an objective source of information that can dictate action while protecting against political backlash, to an unquestioning modern confidence in our technological ability to control the future (e.g., Heilbroner 1959). Whatever the cause, the scientific establishment of the United States focuses a not inconsiderable proportion of its intellectual energy and technological wherewithal on predicting the future of nature in order to promote a variety of desired societal outcomes.

Considered as a whole, the cases in this book portray a pervasive and energetic societal activity—a prediction enterprise supported by substantial federal and private funds—that is unified by shared assumptions about the necessity and value of scientific foresight in environmental decision making, and rooted in a strong belief in predictability itself. The recognition that such an enterprise exists is a crucial first step toward fulfilling the goal set out at the beginning of this book: to improve environmental decision making. Only when the prediction enterprise is recognized can critical scrutiny begin.

Predictions—information products—lie at the heart of this enterprise. They are its rationale, its currency, its legitimacy. For this reason, any effort to assess the prediction enterprise must openly and persistently *question predictions*. This questioning has to occur on two levels simultaneously. Of course it is important to question accuracy, uncertainty, and predictability. But we have seen that prediction products mean little by themselves. Predictions must also be questioned in the context—political, cultural, economic, environmental—of the larger enterprise. Given a particular environmental problem, we need to ask: How does the enterprise operate in this case? Who are the players, and how is power, legitimacy, and participation apportioned among them? What conflicting values are hidden by debate over technical matters? What criteria should be used for judging the output of the enterprise and the outcomes of that output?

Technically "good" predictions used in a healthy decision environment can of course facilitate better decisions, as illustrated by the case of weather predictions—our only candidate for the prediction hall of fame. But the "goodness" of weather predictions arises not just from their accuracy and skill, but also from the capacity of society to make effective use of them. A pretty good flood prediction did not forestall disaster in Grand Forks, North Dakota, and a pretty bad earthquake prediction did not prevent better earthquake preparedness in central and Southern California. These types of outcomes are paradoxical or confusing only if one persists in viewing predictions as simple information products. When the whole enterprise is seen, sense and order begin to emerge.

The central issue is an uncertain future. The cause of this uncertainty is a dynamic planet, an evolving society, and the interaction between the two. Scientific prediction is one tool for coping with this uncertainty—a tool with some promise, some problems, and much unacknowledged complexity, but only one tool among many. Given the uneven performance and our lack of understanding of the prediction enterprise, a good argument can be made for the following: First, our dependence on scientific prediction has become uncritical, and at times excessive and counterproductive. Second, we need to be more careful about how and when to make prediction a central activity in addressing environmental problems. Third, as soon as new environmental problems begin to command public attention, we need to resist the urge to immediately prescribe a predictive approach and should consider instead a range of possible actions. And finally, we should worry less about making good predictions and more about making good decisions.

Notes

1. As contrasted with narrow or parochial interests, which may conflict with common interests.

2. There are other such "user's guides" for understanding predictions. One notable example, which focuses on economic and technological forecasts, is Armstrong (1999). Also see Nicholls (1999).

References

Armstrong, J.S. 1999. Introduction. In *Principles of forecasting: A handbook for researchers and practitioners,* J.S. Armstrong, ed. Norwell, MA: Kluwer Academic Publishers.

Ascher, W. 1981. The forecasting potential of complex models. *Policy Sciences* 13:247–267.

Feldman, M.S., and J.G. March. 1981. Information in organizations as signal and sign. *Administrative Science Quarterly* 26:171–186.

Fischoff, B. 1994. What forecasts (seem to) mean. *International Journal of Forecasting* 10:387–403.

Heilbroner, R. L. 1959. *The future as history: The historic currents of our time and the direction in which they are taking America.* New York: Harper and Brothers.

Keepin, B. 1986. Review of global energy and carbon dioxide projections. *Annual Review of Energy* 11:357–392.

Murphy, A.H. 1977. The value of climatological, categorical, and probabilistic forecasts in the cost-loss ratio situation. *Monthly Weather Review* 105(7):803–816.

Murphy, A.H. 1996. The Finley affair. A signal event in the history of forecast verification. *Weather and Forecasting* 8:281–293.

Murphy, A.H. 1997. Forecast verification. In *Economic value of weather and climate forecasts,* R.W. Katz and A.H. Murphy, eds., pp. 19–74. Cambridge, UK: Cambridge University Press.

Murphy, A.H., S. Lichtenstein, B. Fischhoff, and R.L. Winkler. 1980. Misinterpretations of precipitation probability forecast. *Bulletin of the American Meteorological Society* 61: 695–701.

Nicholls, N. 1999. Cognitive illusions, heuristics, and climate prediction. *Bulletin of the American Meteorological Society* 80(7):1385–1397.

Pielke, R.A., Jr. 1998. Rethinking the role of adaptation in climate policy. *Global Environmental Change* 8(2):159–170.

Pielke, R.A., Jr. 1999. Who decides? Forecasts and responsibilities in the 1997 Red River flood. *American Behavioral Science Review:* 83–101.

Robinson, J.B. 1982. Backing into the future: On the methodological and institutional biases embedded in energy supply and demand forecasting. *Technological Forecasting & Social Change* 21:229–240.