The Role of Models in Prediction for Decision

Roger A. Pielke, Jr.

Discussion Draft of 28 February 2001

Paper prepared for

Cary Conference IX: Understanding Ecosystems:

The Role of Quantitative Models in Observations, Synthesis, and Prediction

Abstract

The processes of science and decision making share an important characteristic: success in each depends upon the researcher or decision maker having some ability to anticipate the consequences of their actions. The predictive capacity of science holds great appeal for decision makers who are grappling with complex and controversial environmental issues, by promising to enhance their ability to determine a need for and outcomes of alternative decisions. As a result, the very process of science can be portrayed as a positive step toward solving a policy problem. The convergence – and perhaps confusion -- of prediction in science and prediction for policy presents a suite of hidden dangers for the conduct of science and the challenge of effective decision making. This paper, organized as a set of inter-related analytical vignettes, seeks to expose some of these hidden dangers and to recommend strategies to overcome them in the process of environmental decision making. The analytical vignettes are titled: Modeling for What?, Importance of Uncertainty, Communicating Uncertainty, Understanding Predictability, What is a "Good" Model?, and the paper concludes with a recommendation: For Better Decisions, Question Predictions. In particular, this paper seeks to distill some of the lessons gleaned from research on modeling, prediction, and decision making in the earth and atmospheric sciences for quantitative modeling of ecosystems. One clear implication of the few lessons presented in this paper is that the belief that modeling and prediction can simultaneously meet the needs of both science and decision is untenable. For ecosystem science, there fortunately exists a body of experience in understanding, using and producing predictions across the sciences on which to build, to the potential benefit of both research and policy.

The Role of Models in Prediction for Decision

Roger A. Pielke, Jr.

Discussion Draft of 28 February 2001

Paper prepared for

Cary Conference IX: Understanding Ecosystems:

The Role of Quantitative Models in Observations, Synthesis, and Prediction

Introduction: Prediction in Science and Prediction for Decision

The processes of science and decision making share an important characteristic: success in each depends upon the researcher or decision maker having some ability to anticipate the consequences of their actions. On the one hand, "[being] predictive of unknown facts is essential to the process of empirical testing of hypotheses, the most distinctive feature of the scientific enterprise" (Ayala 1993). Of course, in science the "unknown facts" in question could lie in the past or the future. "Decision making," on the other hand, "is forward looking, formulating alternative courses of action extending into the future, and selecting among alternatives by expectations of how things will turn out" (Lasswell and Kaplan 1950).

The predictive capacity of science holds great appeal for decision makers who are grappling with complex and controversial environmental issues, by promising to enhance their ability to determine a need for and outcomes of alternative decisions. As a result, the very process of science can be portrayed as a positive step toward solving a policy problem. The appeal of this "two birds with one stone" line of reasoning is obvious for decision makers who would place the onus of responsibility for problem solving onto the shoulders of scientists. But this reasoning is seductive as well for scientists who might wish to better justify public investment in research and for a public that has come to expect solutions as a consequence of such investment (Sarewitz and Pielke 1999).

The convergence – and perhaps confusion -- of prediction in science and prediction for policy presents a suite of hidden dangers for the conduct of science and the challenge of effective decision making. This paper, organized as a set of inter-related analytical vignettes, seeks to expose some of these hidden dangers and to recommend strategies to overcome them in the process of environmental decision making. In particular, this paper seeks to distill some of the lessons gleaned from research on modeling, prediction, and decision making in the earth and atmospheric sciences for quantitative modeling of ecosystems, the focus of Cary Conference IX. The background materials for the conference noted that

> Recent years have seen dramatic advancements in the computational power and mathematical tools available to modelers. Methodological advances in areas ranging from remote sensing to molecular techniques have significantly improved our ability to parameterize and validate models at a wide range of spatial scales. The body of traditional, mechanistic, empirical research is also growing phenomenally. Ecosystem science is ripe for major gains in the synthetic and predictive power of its models, and that this comes at a time of growing need by society for quantitative models that can inform debate about critical environmental issues.[1]

This background indicates that the community of ecosystem scientists is following other fields – particularly the atmospheric, oceanic, and earth sciences -- down a path of using integrative environmental modeling to advance science *and* to generate predictive knowledge putatively to inform decision making. This paper distills some of the most important lessons from these other fields that have journeyed down this perilous path, focusing on the use of models to produce predictions for decision.

Modeling for What?

Bankes (1993) defines two types of quantitative models, consolidative and exploratory, differentiated by their uses (cf. Morrison and Morgan 1999).[2] A consolidative model seeks to include all relevant facts into a single package and use the resulting system as a surrogate for the actual system. The canonical example is that of the controlled laboratory experiment. Other examples include weather forecast and engineering design models. Such models are particularly relevant to decision making because the system being modeled can be treated as being closed, i.e., one “in which all the components of the system are established independently and are known to be correct”

[1] http://www.ecostudies.org/cary9/Conference_Background_and_Goals.htm
[2] Of course, not all “models” are quantitative or described in computer code, see, e.g., Morgan and Morrison (1999).

(Oreskes et al. 1994, 642). The creation of such a model generally follows two phases: first model construction and evaluation, and second operational usage of a final product. Such models can be used to investigate diagnostics (i.e., "what happened?"), process ("why did it happen?"), or prediction ("what will happen?).

An exploratory model – or what Bankes (1993) calls a "prosthesis for the intellect" -- is one in which all components of the system being modeled are not established independently or are not known to be correct. In such a case, the model allows for experiments with the model to investigate the consequences for modeled outcomes of various assumptions, hypotheses, and uncertainties associated with the creation of and inputs to the model. These experiments can contribute to at least three important functions (Bankes 1993). First, they can shed light on the existence of unexpected properties associated with the interaction of basic assumptions and processes (e.g., complexity or surprises). Second, in cases where explanatory knowledge is lacking, exploratory models can facilitate hypothesis generation to stimulate further investigation. Third, the model can be used to identify limiting, worst-case, or special scenarios under various assumptions of and uncertainty associated with the model experiment. Such experiments can be motivated by observational data (e.g., econometric and hydrologic models), scientific hypotheses (e.g., general circulation models of climate), or by a desire to understand the properties of the model or class of models independent of real-world data or hypotheses (e.g., Lovelock's Daisyworld).

Both consolidative and exploratory models have important roles to play in science and decision settings (Bankes 1993). However, the distinction between consolidative and exploratory modeling is fundamental, but rarely made in practice or in interpretation of research results. Often, the distinction is implicitly (or explicitly) blurred to "kill two birds with one stone" in modeling and predicting for science and policy (Pielke and Sarewitz 1999). Consider, for example, the goal of the U.S. Global Change Research Program, from 1989:

> To gain an adequate predictive understanding of the interactive physical, geological, chemical, biological and social processes that regulate the total Earth System and, hence establish the scientific basis for national and international policy formulation and decisions (CES 1989).[3]

And following from this blurring, most presentations by scientists and the media of the results of national and international climate assessments have sought to imbue the imprimatur of consolidative knowledge upon what are inherently exploratory exercises.[4] Those who conflate the science and policy roles of prediction and modeling trade short-term political or public gain, with a substantial risks of a more lasting loss of legitimacy and political effectiveness (Sarewitz et al. 2000).[5]

Thus, one of the most important lessons to be learned from the experiences of other scientific endeavors in which modeling has a potential role to play in research and decision is: *be clear about the purposes for which the modeling is to be used and carefully examine any assumption that presumes isomorphism between the needs of science and the needs of decision.*

Importance of Uncertainty

Uncertainty, in the view of economist John Maynard Keynes, is the condition of all human life (Skidelsky 2000). Uncertainty means that more than one outcome is consistent with our expectations (Pielke 2001). Expectations are a result of judgment, sometimes based on technical mistakes and interpretive errors, and shaped by values and interests. Because uncertainty is a characteristic of every important decision, it is no surprise then that society looks to science and technology to help clarify our expectations in ways that lead to desired outcomes.

Because decision-making is forward-looking decision makers have traditionally supported research to quantify and even reduce uncertainties about the future. In many cases, particularly those associated with closed systems – or systems that can be treated

[3] On the USGCRP see Pielke 2000a and 2000b.
[4] On the interpretation of climate model results see Trenberth 1997, Edwards 1999, IPCC 2001; On the media's presentation of climate research results see Henderson-Sellers 1998; On their role in decision see Sarewitz and Pielke 2000 and Shackley et al. 1998.
[5] Of course, quantitative models have uses beyond simply producing "predictions" (see Bankes 1993).

as closed -- understanding uncertainty is a straightforward technical exercise; probabilities in a card game are the canonical example. Two real-world examples include error analysis in engineering and manufacturing and the actuarial science that underlies many forms of insurance. But in many other circumstances – particularly those associated with human action -- systems are intrinsically open and cannot be treated as closed, meaning that understanding uncertainty is considerably more challenging. In recent decades, many scientists have taken on the challenge of understanding such open systems, e.g., global climate, genetic engineering, etc. And in the process of securing the considerable public resources to pursue this challenge, scientists often explicitly promise to "understand and reduce uncertainties" as input to important societal decisions.

Conventional wisdom holds that uncertainty is best understood or reduced by advancing knowledge, an apparent restatement of the traditional definition of uncertainty as "incomplete knowledge" (cf. Cullen and Small 2000).[6] But in reality, advances in knowledge can add significant uncertainty. For example, in 1990 the Intergovernmental Panel on Climate Change (IPCC) projected that a doubling of CO_2 would result in a 1.5º to 4.5º C mean global temperature change. In 2001, after tens of billions of dollars of investment in global-change research, the IPCC now concludes that a doubling of CO_2 will result in a 1.5º to 6.0º C temperature change. Even as the IPCC has become more certain that temperature will increase, the uncertainty associated with their projections has also increased. Why? Researchers have concluded that there are many more scenarios of possible population and energy use than originally assumed, and have learned that the global ocean-atmosphere-biosphere system is much more complex than was once thought (IPCC 2001). Ignorance is bliss because it is accompanied by a lack of uncertainty.

The promise of prediction is that the range of possible futures might be narrowed in order to support (and indeed to some degree determine) decision making. By way of contrast, in his Foundation series, science fiction writer Issac Asimov introduced the notion of "psychohistory." Asimov's psychohistorians had the ability to predict the future with certainty based on complex mathematical models. We know that Asimov's

[6] Uncertainty is also defined as a decision making bias and a psychological perception, see Weber (1999).

characters lie squarely in the realm of science fiction – there can be no psychohistory. The future, to some degree, will always be clouded. But experience shows that this cloudiness is variable, we *can* predict some events with skill and the promise of prediction can be realized. Understanding, using and producing predictions depends upon understanding their uncertainty. What is it that leads to the uncertainty of earth and environmental predictions? What are the prospects of knowing the uncertainty of specific predictions?

A simple example might prove useful. Consider the poker game known as five card draw.[7] In a standard 52 card deck there are a total of 2,598,960 possible five card poker hands. Lets assume that in your hand you hold a pair. What are the chances that by exchanging the other 3 cards that you will draw a third card to match the pair? . In this instance you can know with great precision that in 71.428 . . . % of such situations you will fail to improve your hand. Thus, when you exchange three cards you are "uncertain" about the outcome that will result, but you can quantify that uncertainty with great certainty.

This sort of uncertainty is that associated with random processes, that is, one in which each element of a set (in this case a deck of cards) has an equal chance of occurring. Because we know the composition of the deck and the set of possible events (i.e., the relative value of dealt hands), it is possible to precisely calculate the uncertainty associated with future events. Scientists call this *aleatory* uncertainty and it is studied using mathematical statistics (cf. Hoffman and Hammonds 1994, Stewart 2000). Such uncertainty, by definition, cannot be reduced. One can never divine what the next card will be, although one can precisely calculate what one's chances are of receiving a particular card. Similarly, in predictions associated with the earth and environmental sciences there is also irreducible uncertainty associated with the nature of random processes.[8]

[7] For those unfamiliar with the game, each player is dealt five cards, with the object to obtain cards ranking higher than those of the other players. After the initial deal of 5 cards per player, each player has the option to exchange up to 3 cards. The poker statistics reported in this section are taken from Scarne (1986).
[8] There are other sources of irreducible uncertainty; some of these are discussed below.

But let's take the poker example a step further. Assume that you find yourself playing cards with a less-than-honest dealer. This dealer is adding and removing cards from the deck, so that the deck no longer has the standard 52 cards. The process is no longer *stationary* – it is changing over time. If you were to know the cards added and removed, i.e., to have the ability to quantify the changing composition of the deck, to quantify uncertainty, you would simply need to recalculate the probabilities based on the new deck of cards. But if you were unaware that the deck was changing in its composition, then you could easily miscalculate the uncertainty associated with your options. Similarly, if you were aware that the deck was changing, but not privy to the exact changes, you would be unable to precisely calculate the uncertainty (but would know that the assumption of a standard 52-card deck could be wrong). This sort of uncertainty is called *epistemic* uncertainty and is associated with incomplete knowledge of a phenomenon—and incomplete knowledge of the limits of one's knowledge (cf. Hoffman and Hammonds 1994, Stewart 2000).

Unlike aleatory uncertainty, epistemic uncertainty can in some cases be reduced through obtaining improved knowledge. In the case of the changing deck of cards reduction of uncertainty could be done using several methods. For instance, one could carefully observe the outcomes of a large number of hands and record the actual frequencies with which particular hands occur. For instance, if four Aces were added to the deck, one would expect to be able to observe the results in the form of more hands with Ace combinations. Of course, the more subtle the change, the more difficult it is to detect.[9] An alternative approach to understanding uncertainty would be to build a model of the card substitution process. This would require some sort of knowledge of the underlying dynamics of the card substitution process. Such knowledge might include the use of a known scientific "law" or observed relationship. For instance, research might reveal that the number of cards added to the deck is proportional to the number of players

[9] In a very similar fashion some studies of global climate change use such a method to assess whether the storms, temperature, precipitation, etc. of one period differ significantly from that of another period (e.g., Trenberth and Hoar 1998).

at the table. With such knowledge, a quantitative model of the poker game can be created and the model can be used to generate understandings of the uncertainty associated with various outcomes. The more one understands about the card replacement process, the better understanding one can have about the associated uncertainties. But if the process of changing cards were continuous (i.e., highly non-stationary or variable), then based on the observations of dealt hand one might develop numerous equally plausible theories about the changing nature of the probabilities. Unless one could discover the pattern underlying the change process (i.e., in effect "close" the system, cf. Oreskes et al. 1994) then such theories would be subject to continuous revision as experience unfolds.

But even though epistemic uncertainty can in principle 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 90 percent to their prediction of the Parkfield earthquake, but the earthquake never occurred.[10] Were the scientists simply confounded by the unlikely but statistically explicable one-out-of-ten chance of no earthquake? Or because their probability calculation was simply wrong—i.e., because the uncertainty associated with the prediction was in fact huge? Similarly, regardless of the sophistication of global climate models, many types of unpredictable events (volcanic eruptions that cool the atmosphere; new energy technologies that reduce carbon emissions) can render today's climate predictions invalid, and associated uncertainties meaningless (see, e.g., Keepin 1986).

A central theme that emerges from experience is that important decisions are often clouded by inherent uncertainty, and in many instances, efforts to reduce uncertainty have the opposite effect (Pielke 2001).[11] Efforts to reduce uncertainty can lead to a discovery the vast complexities associated with phenomena that evolve slowly over long periods -- like earthquakes, global climate change, and nuclear waste disposal -

[10] See Nigg 2000, Savage 1991, and Sieh et al. 1989.
[11] A related consideration is that attempts to eliminate uncertainty by changing thresholds for decision, e.g., changing the wind-speed criteria for evacuation, invariably result in trade-offs between false alarms and misses (i.e., Type I and Type II errors), with associated societal costs. See Stewart (2000).

- were in fact previously underestimated, thereby having the effect of expanding the range of future uncertainties (Sarewitz et al. 2000). In a decision setting, this 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 goes unaddressed. No case illustrates this better than global climate change (Sarewitz and Pielke 2000).

One of the most critical issues in using models to develop information for decision is to understand uncertainty, its sources and potential reducibility. As Weber (1999, 43, emphasis) observes,

> If uncertainty is measurable and controllable, then forecasting and information management systems serve a high value in reducing uncertainty and in producing a stable environment for organizations. If uncertainty is not measurable and controllable, then forecasting and predictions have limited value and need to be understood in such context. *In short, how we view and understand uncertainty will determine how we make decisions.*

Any effort that seeks to model open systems for the purpose of informing decision, particularly through prediction, should also seek to understand uncertainty, including it sources, potential reducibility, and relevant experience, *in the context of the decision making process*. In some cases, such an effort may very lead to the conclusion that decision making should turn to alternatives to prediction (e.g., Herrick and Sarewitz 2000, Brunner 2000).

Communicating Uncertainty

Experience shows that neither the scientific community nor decision makers have a good record at understanding uncertainty associated with predictions (Sarewitz et al. 2000). Such understanding is necessary because “the decision making process is best served when uncertainty is communicated as precisely as possible, but no more precisely than warranted" (Budescu and Wallsten 1987, 76). But even in cases where uncertainty is well-understood, such as is typically the case in weather forecasting, scientists face challenges in communicating the entirety of their knowledge of uncertainty to decision makers. Often, experts place blame for this lack of communication on the perceived lack of public ability to understand probabilistic information. The resulting policy prescription is for increased public education to increase scientific literacy (e.g., Augustine 1998).[12] While improved scientific literacy has value, it is not the solution to improving communication of information about uncertainty.

Consider the following analogy. You wish to teach a friend how to play the game of tennis. You carefully and accurately describe the rules of tennis to your friend, but you speak in Latin to your English-only speaking friend. When you get onto the court your friend fails to observe the rules that you so carefully described. Following, the game, it would surely be inappropriate to criticize your friend as incapable of understanding tennis, and futile to recommend additional tennis instruction (in Latin). But this is exactly the sort of dynamic observed in studies of public understanding of scientific uncertainties. For example, Murphy (1981) documents that when weather forecasters call for, say, a 70% chance of rain, decision makers understood the probabilistic element of the forecast, but did not know whether rain has 70% chance for each point in the forecast area, or that 70% of the area would receive rain with a 100% probability, and so on.[13] Do you know?

The importance of importance of understanding and communicating uncertainties associated with a prediction product was aptly illustrated in the case of the 1997 flooding of the Red River of the North.[14] 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 49 feet, hoping 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 misinterpreted both the event being forecast and the uncertainty associated with the forecast. First, the prediction of 49 feet, rather

[12] Compare Rand (1998) and Wyat and Fox (1999).
[13] There is a considerable literature on the use of weather forecasts that supports this line of argument. See in particular the work of Murphy (e.g., 1981) and Baker (e.g., 2000).

than conveying concern to the public, instead resulted in *reduced* concern. Locals interpreted the forecast in the context of the record 1979 flood, which caused damages, but was not catastrophic. With the 1997 crest expected only a few inches higher than the record set in 1979, many expressed relief rather than concern, e.g., "We survived that one OK, how much worse can a few inches be?" Second, decision makers did not understand the uncertainty associated with the forecast. Flood forecasts are extremely uncertain, especially forecasts of record floods for which there is no experience. Forecasters issued a quantitative forecast with a simple qualitative warning about uncertainty. Hence, many decision makers interpreted the forecast uncertainty in their own terms: Some viewed the forecast as a ceiling, i.e., "the flood will not exceed 49 feet." Others viewed the forecast as uncertain and placed various ranges on uncertainty on the forecasts, ranging from 1 to 6 feet. The historical record showed that flood crest forecasts were, on average, off by about 10% of the forecast.

On April 22, 1997 at Grand Forks the Red River crested at 54 feet, inundating the communities of Grand Forks, ND and East Grand Forks, Minnesota and causing up to $2 billion in damages. In the aftermath of the flood, local, state, and national officials pointed to inaccurate flood forecasts as a cause of the disaster. With hindsight, a more reasoned assessment indicates that by any objective measure the accuracy of the forecasts was not out of line with historical performance. Instead, decision makers failed to understand the meaning of the prediction both in terms of what was being forecast and the uncertainty associated with it.

A significant literature exists on communication of uncertain information, some based on experience in the sciences and much more (it seems) from the disciplines of communication, psychology, and sociology.[15] The implications of this literature range from the straightforward: "statistics expressed as natural frequencies improve the

---

[14] For a detailed evaluation of the role of forecasts in responses to the 1997 Red River floods see Pielke (1999).

[15] See for example, related to the atmospheric and earth sciences, Dow and Cutter (1998), Baker (2000), Mileti and Sorenson (1988), Nicholls (1999), Glantz (2000) and from psychology, Hoffrage et al. 2000, Wallsten et al. 1993, Erev et al. 1993, Gonzalez-Vallejo et al. 1994, Konold 1989, Wallsten et al. 1986, Hamm 1991.

statistical thinking of experts and nonexperts alike" (Hoffrage et al. 2000) to the more challenging: "probability expressions are interpreted differently by speakers and listeners" (Fillenbaum et al. 1991). However, it is clear that the substantial research on communication of uncertainty has not been well-integrated with the research in the earth and environmental sciences that seeks to understand and describe uncertainties relevant to decision making.

Understanding Predictability

Consider again the poker example. With perfect knowledge of a card substitution process engineered by a less-than-honest dealer, one would thus be able to quantify completely and accurately the associated uncertainties in future hands. But this situation is quite different from most cases that we find in the real world of modeling and prediction in the environmental sciences. In the real world, systems are open and there are fundamental limits to predictability. And perhaps surprisingly many scientific efforts to divine the future proceed without an adequate understanding of the limits to predictability. In addition to the aleatory and epistemic uncertainties discussed above, there are a number of other reasons for limits to predictability, among these are sensitivity to initial conditions, complexity, and human agency.

First, predictability is limited because knowledge of the future depends upon knowing the present, which can never be completely or accurately characterized. For example, weather forecasts depend upon knowing the present state of the atmosphere and then projecting forward future behavior of the atmosphere, based on computer models. A result of the dependence on these "initial conditions" is that small changes in the initial conditions can subsequently to large differences in outcomes. Knowledge of initial conditions is obtained with instruments, in weather prediction these can include balloons, radar, satellites, and other instruments that are subject to measurement errors. But even without such measurement errors, the simple act of rounding off a decimal can lead to vastly different outcomes. Popularized as the "butterfly effect," this is the fundamental characteristic of a chaotic system with limited predictability (Gleick 1986). Scientists have established that about 10-14 days is the limit of predictability for weather forecasts.

In many other contexts the same limits hold, but are not as well understood. Meteorologists seek to understand sensitivity to initial conditions by running models repeatedly with small variations in input data (and sometimes in the model itself) to begin to understand the sensitivities of model output to initial conditions (e.g., Krishnamurti et al. 2000).

A second factor is that the environmental sciences phenomena of interest to policy makers are often incredibly complex and can be the result of interconnected human and earth processes. Consider nuclear waste disposal (Metlay, this volume). Predicting the performance of a waste facility 10,000 years into the future depends upon knowing, among a multitude of other potentially relevant factors, what sorts of 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, and social changes, and so on. What at first seems a narrow scientific question rapidly spirals into great *complexity*. One characterization of the concept holds that "a complex system is one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve" (Whitesides and Ismagilov 1999, p. 89). Scientists are just being to understand the implications of complexity for prediction (see Waldrop 1992).

A third factor is the role of human agency. In situations where human decisions are critical factors in the evolution of the future being predicted (that is to say most every issue of environmental policy[16]), the aggregate record of prediction is poor. Ascher (1981, p. 247) argues, "unless forecasters are completely ignorant of the performance record, or are attracted solely by the promotional advantages of the scientific aura of modeling, they can only be attracted to benefits not yet realized." The poor performance of predictions of societal outcomes is consistent across diverse areas that include energy demand (Keepin 1986), energy supplies (Gautier 2000), population (Cohen 1996),

[16] One exception might be the prediction of asteroid impacts on the earth (see Chapman 2000).

elections (Mnookin 2001), corporate financial performance (Dreman and Berry 1995), macro-economics (CBO 1999), and medicine (Fox et al. 1999). To the extent that modeled outcomes depend upon some degree of accuracy in predicting factors such as these, predictability will clearly be limited.

And yet, effective decision making cannot occur without some way to anticipate the consequences of alternative courses of action. The record of cases where prediction and modeling contributed to effective decisions is of course large. The practice of insurance and engineering would not be possible without predictive ability. And more relevant to present purposes, the apparently successful response to stratospheric ozone depletion would not have been possible without predictive and diagnostic modeling (Pielke and Betsill 1998). But understanding (and indeed creating) those situations where prediction and modeling serves effective decision making is not straightforward, if simply because questions about the roles of models and prediction in decision are rarely asked much less answered.

What is a "good" model?

This section focuses on model's are a means to produce predictions for decision, as well as a social and scientific mechanism that fosters integration of knowledge, and its potential use. Thus, the "goodness" of predictions produced from models can be understood from two distinct perspectives, product and process.

Prediction as Product

The first and most common perspective is to view models simply as generators of an *information product*. Often, when a model is applied to decision problems, it is used to produce a prediction, i.e., 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. From this perspective the goal of modeling is simply to develop "good" predictions (Pielke et al. 1999). Three important considerations in the production "good" predictions are accuracy, sophistication, and experience.

Accuracy. A critical criterion to be used to assess a prediction is *accuracy* (Ascher 1979). Accuracy is important because "on balance, accurate forecasts are more likely than inaccurate forecasts to improve the rationality of decision making" (Ascher 1979, 6). With a few exceptions, once a forecast is produced and used in decision making, few ever look back to assess its skill (Sarewitz et al. 2000). Measuring the skill of a prediction is not as straightforward as it might seem. Consider the case of early tornado forecasts. In the 1880s a weather forecaster began issuing daily tornado forecasts in which he would predict for the day "tornado" or "no tornado." After a period of issuing forecasts, the forecaster found his forecasts to be 96.6% correct — a performance that would merit a solid "A" in any grade school. But others who looked at the forecaster's performance discovered that simply issuing a standing forecast of no tornadoes would result in an accuracy of 98.2%! This finding suggested that in spite of the high degree of correct forecasts, the forecaster was providing predictions with little *skill* – defined as the improvement of a forecast over some naïve standard -- and in fact could result in costs rather than benefits.[17] Simply comparing a prediction with actual events does not provide enough information with which to evaluate its performance. A more sophisticated approach is needed. *Thus, predictions should be evaluated in terms of their "skill," defined as the improvement provided by the prediction over a naïve forecast, i.e., such as that which would be used in the absence of the prediction.*[18]

[17] For example, see Pielke et al. (2000) which discusses a methodology to evaluate catastrophe models used by the insurance industry.

[18] The term "skill" is jargon, however the notion of evaluating predictions against a naïve baseline is fundamental to the evaluation of weather forecasts and financial forecasts (such as expected mutual fund performance). For forecasts that are probabilistic, rather than categorical, the evaluation of skill can be somewhat more complicated, but adheres to the same principles. See Murphy (1997) for a technical discussion of the many dimensions of predictive skill. As well, there are other dimensions of predictive "goodness" that are central to evaluation of its role in decision making – including comprehensibility, persuasiveness, usefulness, authoritativeness, provocativeness, importance, value, etc., foe discussion see Ascher (1979), Armstrong (1999) and Sarewitz et al. (2000).

Sophistication. Decision makers sometimes are led to believe that sophistication of a prediction methodology lends itself to greater predictive skill, i.e., given the complexity of the world a complex methodology should perform better. In reality, the situation is no so clear-cut. An evaluation of the performance of complex models has shown that "methodological sophistication contributes very little to the accuracy of [predictions]" (Ascher 1981, p. 258, see also Keepin 1986). A lesson for decision makers is that a sophisticated prediction methodology (or by extension, the resources devoted to development of predictions) does not necessarily guarantee predictive success. Because complex models often require significant resources (computation, human, etc.), a trade-off invariably results between producing one or a few realizations of the complex model and many runs of a simpler, less intensive version of the model. For instance, the U.S. National Assessment of Climate Change used only two scenarios of future climate due to computation limitations (NACC 2000). For many decision makers, having an ability to place modeled output into the context of the entire "model-output space" would have been more useful than the two products that were produced, largely without context. This is an example of confusion between consolidative and exploratory modeling.

Experience. In weather forecasts, society has the best understanding of prediction as a product. Consider that in the United States the National Weather Service issues more than 10 million predictions every year to hundreds of millions of users. This provides a considerable basis of experience on which users can learn, through trial-and-error, to understand the meaning of the prediction products that they receive. Of course, room for confusion exists. People can fail to understand predictions for record events for which the is no experience, as in the Red River case, or even a routine event being forecast (e.g., 70% chance of rain). But experience is essential for effective decision making, and most decision makers have little experience using models or their products. Erev et al. (1993, 92) provide a useful analogy:

> Consider professional golfers who play as if they combine information concerning distance and direction of the target, the weight of the ball, and the speed and direction of the wind. Now assume that we ask them to play in an artificial setting in which all the information they naturally combine in the field is reduced to numbers. It seems safe to say that the numerical representation of the information will not improve the golfer's performance. The more similar are the artificial

> conditions we create to the conditions with which the golfers are familiar, the better will be their performance. One can assume that decision making expertise, like golf expertise, is improved by experience, but not always generalized to new conditions.

The importance of experience does not necessarily limit the usefulness of models and their products in decision making, but its does underscore the importance of the decision context as a critical factor in using models (cf. Stewart et al. 1997).

A range of experience illustrates misunderstandings of prediction as products are fundamental to decision makers' efforts to effectively use predictions. Considering the following:

- Global climate change (Rayner, 2000). 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 global average temperature has no actual meaning, and thus policy advocates have sought to interpret that "event" in different ways, ranging from pending global catastrophe to benign (and perhaps beneficial) change. The issue of uncertainty compounds the issue. As a result, predictive science has been selectively used and misused to justify and advance the existing objectives of participants in the process (Sarewitz and Pielke 2000).
- Asteroid impacts (Chapman, 2000). 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 earth – and its prediction seemingly straightforward, uncomplicated by human agency. But scientific overreaction to the discovery of 1997 XF11 and the associated prediction that it could strike the Earth on 26 October 2028 illustrates that understandings of uncertainty are critical (Chapman 2000). In this case, hype might have damaged future credibility of scientists who study this threat.

These examples, and others, each illustrate the difficulties associated with understanding prediction as a product. At the same time, the cases also illustrate that to improve the use of prediction it would be insufficient to simply develop "better" predictions, whether more precise, e.g., a forecast of a 49.1652 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 predictions are in many cases more desirable, better decisions require attention to the broader prediction process. From this standpoint, better predictions 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 Process

A second perspective is to view modeling as part of a broader *prediction process*. This includes the participants, perspective, institutions, values, resources, and other factors that together determine policies for the prediction enterprise and how the prediction enterprise contributes to public demands for action or tools respect to the issues that they bring to the attention of decision makers. From this perspective the goal of the prediction enterprise is good decisions. Modeling, due to its (potentially) integrative nature, is an important element of the prediction process.

The successful use of predictions depends more upon a healthy process than just on "good" information (Sarewitz et al. 2000). Weather forecasts have demonstrably shown value not because they are by any means "perfect," but because users of those predictions have successfully incorporated them into their decision routines. The prediction process can be thought of as three parallel sub-processes (Sarewitz et al. 2000):

Research Process	includes the fundamental science, observations, etc. as well as forecasters' judgements and the organizational structure which go into the production of predictions for decision makers.
Communication Process	includes both the sending and receiving of information; a classic model of communication is: who, says what, to whom, how, and with what effect.
Choice Process	includes the incorporation of predictive information in decision making. Of course, decisions are typically contingent upon many factors other than predictions.

Often, some mistakenly ascribe a linear relation to the processes. From the perspective of benefits to society, these three processes are instead better thought of as components of a broader *prediction process*, with each of the sub-processes taking place in parallel, with significant feedback and interrelations between them.

Peter Drucker has written an eloquent description of the modern organization that applies equally well the prediction process.

> Because the organization is composed of specialists, each with his or her own narrow knowledge area, its mission must be crystal clear . . . otherwise its members become confused. They will follow their specialty rather than applying it to the common task. They will each define 'results' in terms of that specialty, imposing their own values on the organization. (1993, p. 54)

Drucker continues with an apt metaphor.

> The prototype of the modern organization is the symphony orchestra. Each of 250 musicians in the orchestra is a specialist, and a high-grade one. Yet by itself the tuba doesn't make music; only the orchestra can do that. The orchestra performs only because all 250 musicians have the same score. They all subordinate their specialty to a common task. (1993, p 55)

In the process of modeling and prediction in support of decision making, success according to the criteria of any subset of the three processes does not necessarily result in benefits to society. Consider the following examples.

- The case of the Red River floods presented earlier illustrates that a technically skillful forecast that is miscommunicated or misused can actually result in costs rather than benefits. The overall prediction process broke down in several places. No one in the research process fully understood the uncertainty associated with the forecast, hence little attention was paid to communicate the uncertainty to decision makers. As a result poor decisions were made and people suffered, probably unnecessarily. Given that this community will to some degree always depend upon flood predictions, the situation might be improved in the future by including local decision makers in the research process in order to develop more useful products (see 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 with local officials and emergency managers (Nigg 2000). A result was better communication among these groups and overall improved preparation for future earthquakes. In this case, even though the predictions themselves could not be shown to be skillful, the overall process worked because it identified alternatives to

prediction that have led to decisions that are expected to reduce the impacts of future impacts in this region.

- The case of global climate change may be in the early stages of what was documented in the case of earthquakes (Rayner 2000). Policy making focused on prediction has run up against numerous political and technical obstacles, meanwhile alternatives to prediction have become increasingly visible. The prediction process can be said to work if the goals of climate policy – to reduce the impacts of future climate changes on environment and society – are addressed, independent of whether century-scale climate forecasts prove to be accurate (Sarewitz and Pielke 2000).
- The case of nuclear waste disposal has also evolved from one in which decision making focused first on developing skillful predictions to one in which decision making focused instead on actions that would be robust under various alternative futures (Metlay 2000). In this case, the policy problem of storing nuclear waste for a very long time (and associated uncertainties) was addressed via decision making (i.e., engineering), not prediction.

As Robinson (1982, p. 249) observes,

> by basing present decisions on the apparent uncovering of future events, an appearance of inevitability is created that deemphasizes the importance of present choice and further lessens the probability of developing creative policy in response to present problems . . . [predictions] do not reveal the future but justify the subsequent creation of that future.

The lesson for decision makers is that one is in most cases more likely to reduce uncertainties about the future through decision making rather than through prediction, again leading us back to issues associated with the broader prediction process.

The criteria for evaluating the "goodness" of a model are thus directly related to the purposes for which a model is to be used. A consolidative model will most likely be evaluated based on the accuracy of its output, whereas an exploratory model could easily succeed even if its results are highly inaccurate (Bankes 1993). Similarly, a model designed to advance understanding should be evaluated by a different set of criteria than a model designed to provide reliable products useful in decision. For society to realize the benefits of the resources invested in science and technology of scientific prediction, the entire process must function in a healthy manner, just like the sections of Drucker's orchestra must perform together to make music. Each sub-process of the broader prediction process must be considered in the context of the other sub-processes; they cannot be considered in isolation.

Conclusion: For Better Decisions, Question Predictions

The analytical vignettes presented in this paper begin to highlight some of the shared characteristics of healthy decision processes for the use of model products, particularly predictions. One characteristic is the critical importance of decision makers having experience with the phenomena being predicted, as well as experience with the predictions themselves. The less frequent, less observable, less spatially discrete, more gradual, more distant in the future, and more severe a predicted phenomenon, the more difficult it is to accumulate direct experience. Where direct societal experience is sparse or lacking, other sources of societal understanding must be developed or the prediction process will not function as effectively. Science alone and prediction in particular do not create this understanding.

More broadly, what is necessary above all is an institutional structure that brings together those who solicit and use predictions with scientists throughout the entire prediction process, so that each knows the needs and capabilities of the others. It is crucial that this process be open, participatory, and conducive to mutual respect. Efforts to shield expert research and decision making from public scrutiny and accountability invariably backfire, fueling distrust and counterproductive decisions.

While efforts to predict natural phenomena have become an important aspect of the earth and environmental sciences, the value of such efforts, as judged especially by their capacity to improve decision making and achieve policy goals, has been questioned by a number of constructive critics. The relationship between prediction and policy making is not straightforward for many reasons, among them:

- Accurate prediction of phenomena may not be necessary to respond effectively to political or socioeconomic problems created by the phenomena (for example, see Landsea et al. 1999).

- Phenomena or processes of direct concern to policy makers may not be easily predictable. Likewise, predictive research may reflect discipline-specific scientific perspectives that do not provide "answers" to policy problems, which are complex mixtures of facts and values, and which are perceived differently by different policy makers (for example, see Herrick and Jamieson 1996).
- Necessary political action may be deferred in anticipation of predictive information that is not forthcoming in a time frame compatible with such action. Similarly, policy action may be delayed when scientific uncertainties associated with predictions become politically charged (in the issue of global climate change, for example; see Rayner and Malone 1998).
- Predictive information also may be subject to manipulation and misuse either because the limitations and uncertainties associated with predictive models are not readily apparent, or because the models are applied in a climate of political controversy and high economic stakes.
- Emphasis on predictive products moves both financial and intellectual resources away from other types of research that might better help to guide decision making (for example, incremental or adaptive approaches to environmental management that require monitoring and assessment instead of prediction; see Lee 1993).

These considerations suggest that the usefulness of scientific prediction for policy making and the resolution of societal problems depends on relationships among several variables, such as the timescales under consideration, the scientific complexity of the phenomena being predicted, the political and economic context of the problem, and the availability of alternative scientific and political approaches to the problem.

In light of the likelihood of complex interplay among these variables, decision makers and scientists would benefit from criteria that would allow them to better judge the potential value of scientific prediction and predictive modeling for different types of political and social problems related to Earth processes and the environment. Pielke et al. (1999) provide the following six guidelines for the effective use of prediction in decision making.

- Predictions must be generated primarily with the needs of the user in mind. For stakeholders to participate usefully in this process, they must work closely and persistently with the scientists to communicate their needs and problems.
- Uncertainties must be clearly articulated (and understood) by scientists, so that users understand their implications. Failure to understand uncertainties has contributed to poor decisions that then undermine relations among scientists and decision makers. But merely understanding the uncertainties does not mean that the predictions will be useful. If policy makers truly understood the uncertainties associated with predictions of, for example, global climate change, they might decide that strategies for action should not depend on predictions (cf., Rayner and Malone 1998).
- Experience is a critically important factor in how decision makers understand and use predictions.
- Although experience is important and cannot be replaced, the prediction process can be facilitated in other ways, for example by fully considering alternative approaches to prediction, such as no-regrets policies, adaptation and better planning and engineering. Indeed, alternatives to prediction must be evaluated as a part of the prediction process.
- To ensure an open prediction process, stakeholders must *question predictions*. For this questioning to be effective, predictions should be s transparent as possible to the user. In particular, assumptions, model limitations, and weaknesses in input data should be forthrightly discussed. Even so, lack of experience means that many types of predictions will never be well understood by decision makers.
- Last, predictions themselves are events that cause impacts on society. The prediction process must include mechanisms for the various stakeholders to fully consider and plan what to do after a prediction is made.

When the prediction process is fostered by effective, participatory institutions, and when a healthy decision environment emerges from these institutions, the products of predictive science may even become less important. Earthquake prediction was once a policy priority; now it is considered technically infeasible, at least in the near future. But in California the close, institutionalized communication among scientists, engineers, state and local officials, and the private sector has led to considerable advances in earthquake preparedness and a much-decreased dependence on prediction. On the other hand, in the absence of an integrated and open decision environment, the scientific merit of predictions can be rendered politically irrelevant, as has been seen with nuclear waste disposal and acid rain. In short, if no adequate decision environment exists for dealing with an event or situation, a scientifically successful prediction may be no more useful than an unsuccessful one.

These recommendations fly in the face of much current practice where, typically, policy makers recognize a problem, scientists then go away and do research to predict natural behavior associated with the problem, and predictions are finally delivered to decision makers with the expectation that they will be both useful and well used. This sequence, which isolates prediction research but makes policy dependent on it, rarely functions well in practice.

Yet once we have recognized the existence of a prediction enterprise, it becomes clear that prediction is more than a product of science. Rather, it is a complex *process*. This process includes all 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 public benefit. The value of predictions for environmental decision making therefore emerges from the complex dynamics of the prediction process, and not simply from the technical efforts that generate the prediction product (which are themselves an integral part of the prediction process). All the same, it is the expectation of a useable prediction product that rationalizes the existence of the prediction enterprise. This expectation turns out to be extremely difficult to fulfill.

This paper has presented only a few of the many considerations important to understand if scientific modeling and prediction are indeed to fulfill public expectations of the contributions of science in addressing environmental policy problems. There is considerable need for debate and discussion, supported by rigorous knowledge, on the proper role of modeling and prediction in decision, rather than simply assuming what that role should be. However, one clear implication of the few considerations presented in this paper is that the belief that modeling and prediction can simultaneously meet the needs of both science and decision is untenable. For ecosystem science, there fortunately exists a body of experience in understanding, using and producing predictions across the sciences on which to build, to the potential benefit of both research and policy.