Risk Analysis, Vol. 19, No. 1, 1999

Mixed Levels of Uncertainty in Complex Policy Models

Elizabeth A. Casman,[1] M. Granger Morgan,[1] Hadi Dowlatabadi[1]

The characterization and treatment of uncertainty poses special challenges when modeling indeterminate or complex coupled systems such as those involved in the interactions between human activity, climate and the ecosystem. Uncertainty about model structure may become as, or more important than, uncertainty about parameter values. When uncertainty grows so large that prediction or optimization no longer makes sense, it may still be possible to use the model as a ''behavioral test bed'' to examine the relative robustness of alternative observational and behavioral strategies. When models must be run into portions of their phase space that are not well understood, different submodels may become unreliable at different rates. A common example involves running a time stepped model far into the future. Several strategies can be used to deal with such situations. The probability of model failure can be reported as a function of time. Possible alternative ''surprises'' can be assigned probabilities, modeled separately, and combined. Finally, through the use of subjective judgments, one may be able to combine, and over time shift between models, moving from more detailed to progressively simpler order-of-magnitude models, and perhaps ultimately, on to simple bounding analysis.

KEY WORDS: Uncertainty; model uncertainty; epistemic uncertainty; integrated assessment.

1. INTRODUCTION

The past two decades have witnessed substantial progress in the way in which routine quantitative policy analysis deals with uncertainty. From a norm of single-value-best-estimate analysis, with sporadic discussion of sensitivity, the field has now progressed to the point where the use of probability distributions to describe uncertain coefficients and the use of methods such as stochastic simulation to propagate that uncertainty through policy models have become the norm in engineering safety analysis and common in health and environmental risk assessment. Of course, there are still holdouts, particularly among the biomedical community,[1–3] but continuing progress is apparent.

Uncertainty about coefficient values can arise both because the world is full of variability and random processes, and because our understanding of how it works is incomplete.[2] Sometimes it is important to distinguish between these two sources of uncertainty. However, recent emphasis on the distinction,[5] particularly by EPA[2] has sometimes resulted in the distinction being overdrawn.

While an adequate treatment of parameter uncertainty is important, in many domains of risk and other forms of policy analysis, uncertainty about coefficient values is swamped by uncertainty about the appropriate functional form of the models that should be used. Model uncertainty is frequently im-

[1] Department of Engineering & Public Policy. Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.

[2] These two sources of uncertainty are sometimes referred to as ''aleatory'' and ''epistemic.''[4] While we have no disagreement with this classification, we avoid the use of the terms simply because we and many others have difficulty remembering what they mean, and which is which!

0272-4332/99/0200–0033$16.00/1

portant when the system involved is sufficiently complex that key influences have not yet been identified, or have been intentionally omitted or simplified to make the model computationally tractable. It can also be important when the causal signals are so weak or buried in so much noise that system structure can not be readily inferred. Dealing adequately with uncertainty about model structure can be difficult.

Evans and his colleagues constructed an array of plausible alternative models to describe low-dose carcinogenic potency. They elicited subjective probability distributions from experts to characterize the coefficients in these models, and then combined the models using probability trees, in which the weights were based on their own careful reading of the literature,[6] or the elicited judgments of experts.[7] In the area of seismic analysis, Budnitz *et al.*[8] took this approach a step further, using teams of experts working collaboratively to thoughtfully ''weigh'' (not simply attach weights to) alternative models and collectively produce a probabilistic consensus judgment of how the models should best be combined. Paté-Cornell[9] describes this process as involving ''gathering a group of well-informed and socially adjusted individuals,'' who construct '' a complete set of hypotheses'' and then assess ''axiomatically correct probability distributions [across these hypotheses] based on all scientific evidence.''

However, as the quality of scientific understanding becomes poorer, developing meaningful probability judgments to combine alternative models of the world becomes increasingly more difficult. In such circumstances, many Bayesian theorists would advise the analyst to specify the (perhaps infinite) set of all priors and models which fit the constraints imposed by whatever limited knowledge one has. Probability weights (which might all be equal) should then be applied across this set, and the problem should be solved for all cases. While we have no basic theoretical disagreement with such an approach, we also know from experience that a prescription that one's analytical formulation should grow in complexity and computational intensity as one knows less and less about the problem, will not pass the laugh test in real-world policy circles.

2. INTEGRATED ASSESSMENT OF CLIMATE CHANGE

Our interest in the problems of dealing with uncertainty in very poorly understood systems has been stimulated in recent years by our work on integrated assessment of climate and other types of global change.[10] The idea of integrated assessment is to use the various tools of policy analysis to try to better interpret available knowledge, identify and explore the implications of alternative policy options, and identify future research priorities which can best serve the needs of policymakers. Since understanding is incomplete and uneven across the problem, integrated analysis must typically include elements which are incompletely understood.

Dowlatabadi, Morgan and co-workers have built a large stochastic simulation model called ICAM (for Integrated Climate Assessment Model)[11] in the Analytica™ software environment (formerly Demos). This environment provides a powerful graphic user interface which represents the model structure in the form of hierarchically organized influence diagrams. Users can explore the model by ''double clicking'' on various elements, moving down through the model hierarchy until they reach individual model elements, where they can observe the mathematical relationships between variables and read documentation on some of the values being used and the assumptions that have been made. Users can easily substitute alternative values or probability distributions. The model is available at http://www.hdgc.epp.cmu.edu. A demonstration copy of the Analytica™ software can be obtained at http://www.lumina.com/software.

In the current version of ICAM, the world is divided into twelve regions. Time is stepped in 5 year increments from 1975 to 2100. Demographic and economic processes lead to emissions of greenhouse gases and aerosols. These modify the composition of the atmosphere and bring about climate change. Climate change leads to various impacts which in turn can affect demographic, socioeconomic and ecological processes. It is possible to make policy interventions in energy use, in emissions management, and in adapting to impacts. In some user-selected structural variants of ICAM, economic factors, climate change, and climate impacts can influence the initiation and path of these interventions.

In developing ICAM we found that uncertainty about the appropriate functional form of different sub-models is sufficiently large, and the difficulty of constructing all plausible alternatives sufficiently great, that it is often best to report results parametrically across a set of combinations of different model structural assumptions, in much the same way that one reports the results of parametric sensitivity studies of coefficient uncertainty. For example, in an application of ICAM-2 designed to explore the probability that a specific carbon tax policy[3] would yield net positive benefits, we found that the probability ranged from 0.15 to 0.95 for the world as a whole, depending upon the structural assumptions made.[10] A more recent example, from a study of the costs of delaying mitigation activities, illustrates the effects of alternative model structures in just the energy and carbon emission control modules of ICAM-3 (Table I).

Many climate policy models are designed and solved as long-term optimization problems. In a settings with uncertainties as great as those displayed in Table I, we doubt the utility of conventional optimization formulations. As an alternative, rather than try to search for the optimal policy, we have set out to search for robust behaviors. Just as in the model environment, real-world policymakers will always face great uncertainty. They must observe the world, use what they see together with models to make forecasts, choose what they think is the best strategy at the moment, and then a few years later, repeat the entire cycle. By building simple "decision agents" we have been able to do something very similar within the world of the ICAM model environment. Then, across a range of alternative model worlds we run repeated stochastic simulations of the model and ask, among a range of plausible alternative behavioral strategies which our agents might adopt, which one does best in the face of the uncertainties about both coefficient values and model structures? In the case of the climate problem, a strategy that tracks and attempts to control atmospheric concentration of greenhouse gasses (as opposed to emissions or temperature), using a quadratic penalty function, seems to do best.[12] Of course, not all problems with high uncertainty will yield such a single general result. In some cases, even a recasting of the problem in behavioral terms is likely to lead to different behaviors for different combinations of model structure.

Table I. Illustration of the Wide Range of Results that Can Be Obtained with ICAM Depending upon Different Structural Assumptions

	Model variants								
Model components	M1	M2	M3	M4	M5	M6	M7	M8	M9
Are new fossil oil and gas deposits discovered?	no	yes	no	no	yes	yes	no	yes	yes
Is technical progress that uses energy affected by fuel prices and carbon taxes?	no	no	yes	no	yes	yes	yes	yes	yes
Do the costs of abatement and nonfossil energy technologies fall as users gain experience?	no	no	no	yes	no	no	yes	yes	yes
Is there a policy to transfer carbon saving technologies to non Annex 1 countries?	no	no	no	no	no	yes	yes	no	yes
TPE BAU in 2100 (EJ)									
Mean	1975	2475	2250	2000	3425	2700	1450	3550	2850
TPE control in 2100 (EJ)									
Mean	650	650	500	750	500	500	675	750	725
CO_2 BAU 2100 (10^9TC)									
Mean	40	50	50	40	75	55	25	73	55
SD	28	18	36	29	29	23	22	27	21
Mitig. cost (% welfare)									
Mean	0.23	0.44	0.14	0.12	0.48	0.33	0.05	0.23	0.17
SD	0.45	0.23	0.23	0.22	0.28	0.12	0.07	0.12	0.11
Impact of delay (% welfare)									
Mean	−0.1	0.2	−0.6	0.0	−1	−0.5	−0.1	−0.6	−0.4
SD	1	0.3	1	0.7	1.2	0.9	0.5	0.8	0.6

[a]TPE = Total Primary Energy, BAU = Business as Usual (no control and no intervention), Sample size in ICAM simulation = 400.

[3] In this case, the tax began in the year 2000 and increased by \$4.00/ton of carbon every 5 years through the year 2100.

3. DOMAINS OF MODEL VALIDITY

Frequently one has confidence in a model used in an application such as ICAM only within a specified domain of the model parameter space. For example, when we elicited judgments about climate sensitivity from 16 leading U.S. climate scientists,[13] we also asked for a judgment of the probability that the forcing from a doubling of CO_2 would induce an irreversible change in the state of the climate system, such that once the change had occurred the climate system would not promptly return to its previous state if the excess radiative forcing were removed. The end to deep water formation in the North Atlantic is a possible example of such a change. If warm water were no longer to circulate to high latitudes in the North Atlantic (as indeed it has not on several occasions in the geological record) dramatic and widespread impacts could be expected on the climate of Europe. We would term this change irreversible if the circulation did not promptly resume when the radiative forcing was removed.

In addition to asking for the probability that a $2x[CO_2]$ forcing would induce a climate state change, we also asked experts to estimate the concentration of CO_2 that would yield a 0.2 probability of a climate state change. Eight experts[4] answered both questions or gave a response to the first question that was 0.2 or greater. Using these expert judgments in ICAM, Fig. 1 illustrates the temperature sensitivity results for three experts whose judgments span the range of usable responses we received. For each expert, the upper curve reports the mean global temperature response to growing concentrations of greenhouse gases under a business as usual (no abatement and no intervention) economic scenario. The lower curves indicate the probability, over time, that a change in climate state will have occurred. Box plots on the curves were constructed from the full probability distributions generated through stochastic simulation in ICAM.

Figure 1 displays uncertainty at several different levels. First there is the disagreement among different climate experts about the magnitude of climate sensitivity (upper curves) and the probability of state change (lower curves) given changes in the atmospheric concentration of CO_2. Then, for any given climate expert, the box plots on the upper curves represent uncertainty about the mean global temperature as a function of time, which depends both on the expert's uncertain judgment of climate sensitivity as well as other uncertainties in the ICAM model. Thirdly, the lower curves display the expert's estimate of the probability that this model projection is not valid because of a state change, since once such a change occurs, the predictions of the ICAM model no longer hold. Fourth, the box plots associated with the lower curves represent the range of uncertainties about that probability because of uncertainties in the values of other parameters in the ICAM model (such as those determining population growth, emissions levels, etc.). Of course, we could make additional

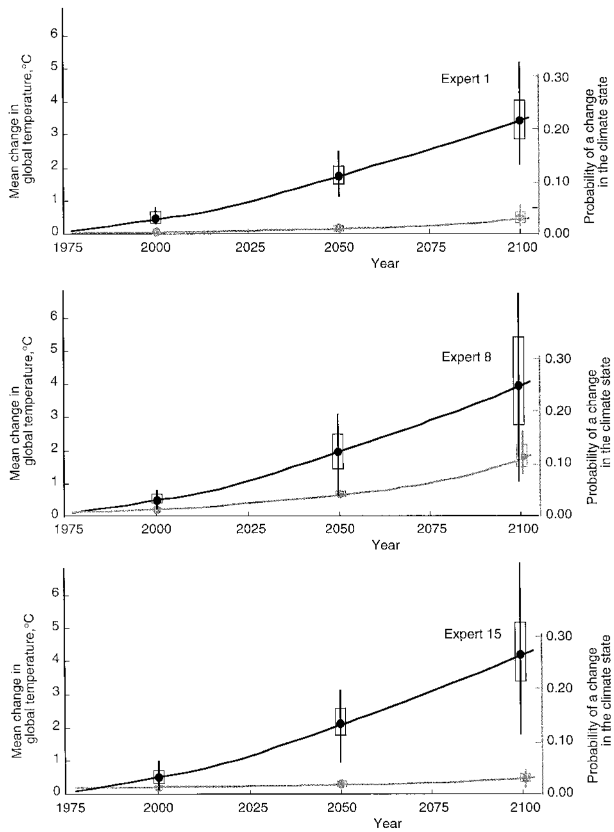

Fig. 1. Examples of predicted warming estimated via the ICAM model (upper curves), and probability that the associated radiative forcing will induce a state change in the climate system (lower curves) using the probabilistic judgments of three different climate experts. Solid dots indicate mean values (indistinguishable from medians on this scale). Boxes span the 0.25 to 0.75 and whiskers span 0.05 to 0.95 confidence intervals as deduced from the full probability distributions produced through stochastic simulation in ICAM.

[4] In addition to the eight experts discussed here, two other experts estimated probabilities of the order of 0.1 to the first question but did not answer the second. Three gave zero as their answer to the first question and did not answer the second. Two gave no answer to either question. For details see Morgan and Keith.[13]

runs of ICAM using alternative model structural assumptions (as previously illustrated in Table I) and produce an additional series of curves corresponding to a range of model structures.

When conducting our elicitations with the sixteen climate experts, we encountered two experts who chose to produce two, rather than one, estimates of climate sensitivity.[13] Expert 2 gave us a distribution similar to that of most other experts and then gave us a second distribution with much longer tails to include the possibility of a "state change." Expert 4 did much the same, but referred to the wider distribution as including "surprise." The fact that these two distributions are much wider than those of all the other experts is troubling. Our protocol asked experts to consider *all* possible eventualities, and we discussed many possible contingencies during the interviews. Nevertheless, given the striking difference between these "surprise" responses and those of the other experts, the strong psychological influence of the IPCC consensus judgments, and the strong evidence in the literature that most people are systematically over-confident,[14,15] we suspect that most of our experts truncated the tails of their distributions.

In order to explore this issue more explicitly we asked another climate expert (Expert 17) to describe his best estimate of global average temperature response to various "imaginable surprises."[16] He took imaginable to mean that while plausible cause and effect scenarios have been offered, the events would be a surprise if they occurred because they are not thought to be likely outcomes of increased CO_2 forcing. We elicited a probability distribution in climate sensitivity for a $2x[CO_2]$ forcing, asking the expert not to include any surprise outcomes in estimating this distribution. Then we asked the expert to give quantitative estimates to describe the effects that surprise scenarios could yield.

The expert chose to consider three "imaginable surprise" scenarios:

1. Changes to the radiative properties of clouds that would produce a net negative feedback to warming. The expert assigned a probability of 0.2 to this outcome, given a change in the global average temperature of $\geq$2°C;
2. Changes in the radiative properties of clouds that produce a net positive feedback to warming. The expert also assigned this a probability of 0.2 given a change in global average temperature of $\geq$2°C; and
3. The shutdown of the oceanic thermohaline circulation. For a change in global average temperature of $\geq$2.5°C the expert assigned this a probability of 0.1, increasing to 0.2 as the global average temperature change were to increase to more than 5°C.

The first two scenarios are mutually exclusive, but the third scenario could conceivably occur in combination with either of the first two. However, the expert felt that the interactions were likely to be so nonlinear as to preclude description of their combined effects until actual modeling studies could be conducted.

Expert 17 was unwilling to provide enough judgments to allow a systematic combination of the three imagined surprise scenarios. He did estimate that the combined probability of surprise from these and other sources was about 0.3 at $2.5x[CO_2]$, increasing to about 0.5 at $4x[CO_2]$.

This expert did not feel confident in his (or anyone's) ability to produce a complete, consistent pdf for the probability of surprise, and it is no error that some of the conditional probabilities are contingent on temperature change and others on reaching various CO_2 concentration thresholds.

This set of judgments was used to examine the implications for an ICAM temperature trajectory as shown in Fig. 2. Clearly for this expert, the uncertainty around the "no surprise" condition is greater than the conventional confidence interval.

As we noted above, our theoretical Bayesian friends tell us that the proper way to deal with a situation like this is to obtain probability judgments for all possible contingencies, and then combine the results. However, when dealing with a problem as complex as the climate system, even very cooperative scientists, such as Expert 17, typically reach a point at which they refuse to go further. Figure 2 provides a more detailed description than Fig. 1 of some of the uncertainties as the atmosphere experiences increasing radiative forcing. However, given the current state of knowledge, and the high levels of ignorance about some important issues, it is unlikely that any true expert can or will be willing to go much further.

4. PROBABILISTIC MODEL SWITCHING

Models describe the behavior of a system in some domain of the system's phase space. In many

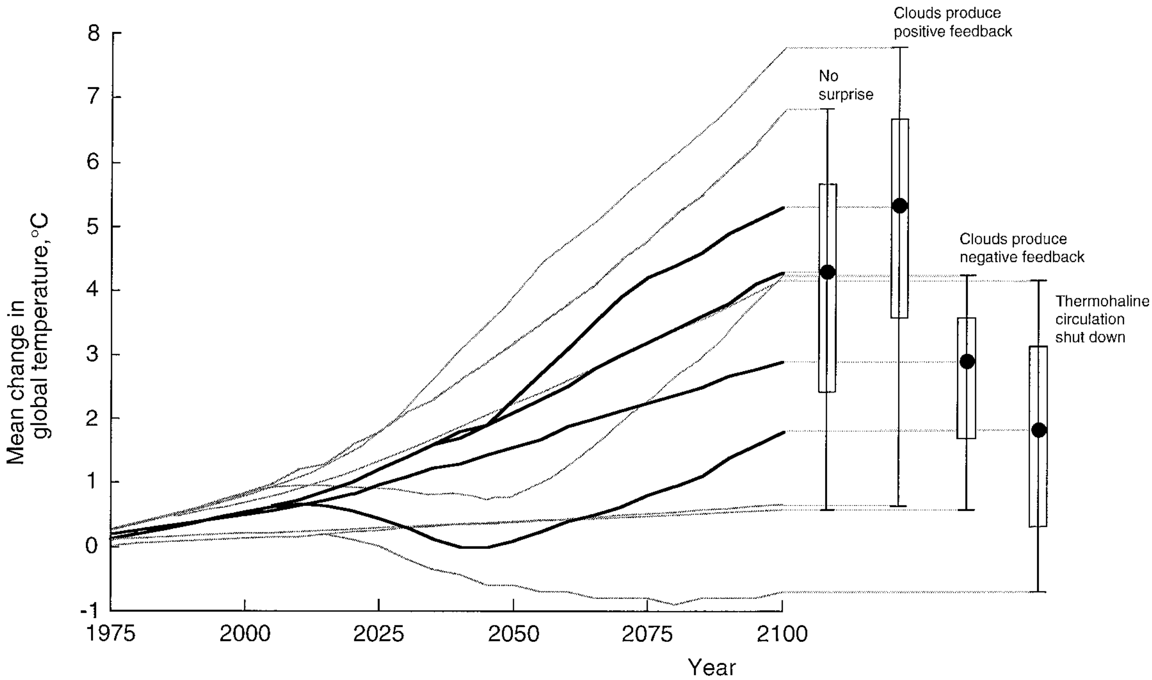

Fig. 2. ICAM simulation results for Expert 17 for the case of "no surprises," and for three "imaginable surprises" which involve the mutually exclusive possibilities of positive and negative feedback from clouds, as well as the possibility of a shutdown of thermohaline circulation in the North Atlantic. Solid curves and dots report means. Light curves report 5th and 95th percentiles. Box plots show additional information on distributions in the year 2100. (Based on a no-abatement scenario with the same energy module configuration as in Fig. 2).

cases, a model provides a realistic description across only a subset of the domain of interest. Thus, for example, in Fig. 1 we saw that the ICAM model could be taken as applicable with high probability only so long as the level of radiative forcing remained within some limit. In a complex model, composed of various submodels, one's confidence in the validity of the different sub-models is likely to change as one traverses different regions of the phase space.

The problem is easily illustrated in terms of time. So long as concentration of CO_2 stays at or below a doubling of pre-industrial levels, we have about the same confidence in the validity of the climate submodel in ICAM whether it is run for 50 years or 200 years into the future. The same is not true for submodels which deal with demographic or socioeconomic variables. In these cases, one can extrapolate from current trends and structures for perhaps a decade, or in the case of demographics, several decades, but then, our confidence in the validity of the submodels declines rapidly because they have moved into a region of phase space for which previous experience may no longer be relevant.

While a number of integrated assessment models[17,18,19] build highly detailed descriptions of economic and other social variables and then simulate the model for a hundred years or more into the future, we doubt the reliability of much of the detail after even a decade. We believe that a more defensible strategy is to deal with the more distant future (or, to be more general, with the less well understood region of the model phase space) by building very simple models, which are based on order-of-magnitude estimates when possible, and use bounding considerations such as material and energy balance and carrying capacity, when best estimates are no longer meaningful.

Figure 3 illustrates the general strategy we are proposing. One starts with a detailed model that is likely to only be reliable for a few years. Gradually one moves over to a much simpler model based on order of magnitude considerations. Finally, in the long term, one can only bound the result, without giving best estimates. The weighting functions that are used to combine models, and make the switch from one model to another over time, must, of course, be based on subjective judgment. While the illustration shows three models over time, there is no

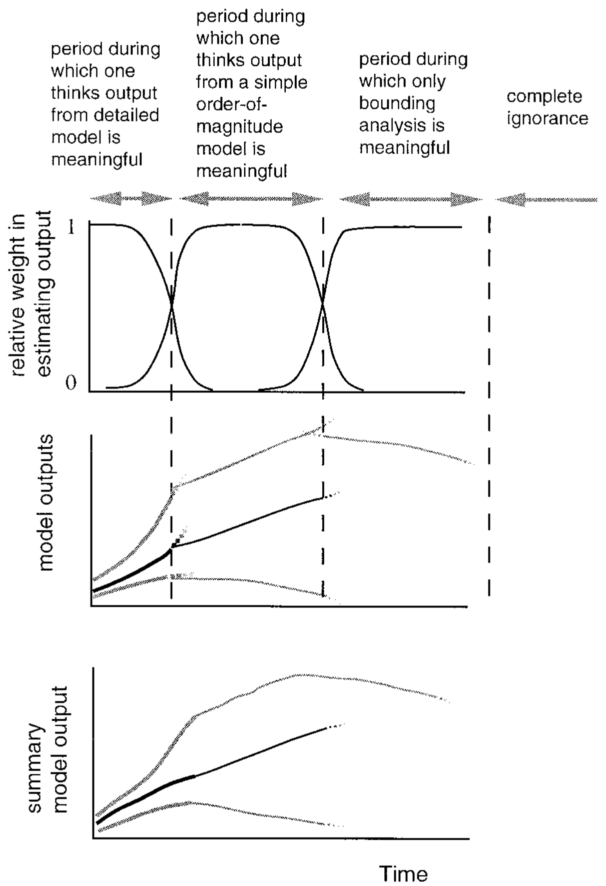

Fig. 3. Schematic illustration of the strategy of switching to progressively simpler models as one moves into less well understood regions of the problem phase space, in this case, over time. One starts with a detailed model that is likely to only be reliable for a few years. Gradually one moves over to a much simpler model based on order of magnitude considerations. Finally, in the long term, one can only bound the result, without giving best estimates. The weighting functions for combining the models are based on subjective judgment. While the illustration shows three models over time, there is no reason why the number can not be more or less than three. The worked example in Fig. 5 shows only the switch from a simple model to a bounding estimate.

reason why the number could not be more or less than three.

We provide a simplified illustration of this approach using the demographic module from ICAM. Since the relevant sub-model in ICAM is only moderately better than "order-of-magnitude" in character, we skip the first stage and focus just on illustrating the switch from an order-of-magnitude model to a bounding argument.

In constructing the demographic submodel in ICAM, Dowlatabadi used recent regional demographic and socioeconomic data to develop linear relationships between the dependent variables: total fertility rate, infant mortality rate, and life expectancy and the regional aggregate socioeconomic indicators: disposable income, female participation in the formal workforce, access to medical care, and degree of urbanization, each of which in ICAM changes over time with economic development. The resulting linear regression models, estimated with data for the past 18 years, are described elsewhere.[11] The economic growth rates for the 12 regions were based on UN projections through the year 2100. The economic growth rates for 2100–2300 are extrapolations of the UN trends plus a stochastic component. We recognize that these variables are not independent of one another, that the direction of dependency may not be as we have presumed, and that some input variables may have been defined differently by different nations. Nonetheless, these are the best available data, and integrated assessment is all about doing the best one can with the data that are available.

The ICAM model estimates a global population of 12 ± 3.4 billion in 2050, compared with U.S. Bureau of Census estimates of about 9 billion. The latter does not provide probability estimates for alternative population projections, so it is not possible to compare the range of ICAM projections with official projections. In a recent paper, Lutz *et al.*[20] developed a probabilistic model based on expert judgments. In their model, in 67% of cases the world population in the year 2100 did not exceed 12 billion. In ICAM, for 50% of the cases, the year 2100 population never exceeds 12 billion.

Demographer Joel E. Cohen has recently explored the question "How many people can the earth support?"[21,22] by developing a very simple model that he terms a "mathematical cartoon." This model employs a carrying capacity whose magnitude evolves over time as human technical capabilities change. The population at any given time, $P(t)$, is determined as the result of a balance between human reproduction and the current carrying capacity. In the discrete time version of the model, earth's human carrying capacity, K, is defined as:

$$K(t + \Delta t) - K(t) = Lr[K(t) - P(t)]\,\Delta t \qquad (1)$$

and global population, P, as:

$$P(t + \Delta t) - P(t) = rP(t)[K(t) - P(t)]\,\Delta t \qquad (2)$$

P, K, L, and r are all positive. $K(0)$, the original carrying capacity, must be greater than $P(0)$, the original population. The constant $r > 0$ is termed the Malthusian growth rate (after the Reverend Thomas R. Malthus, who predicted that population growth would outpace growth in food supply). The ratio $L/P(t) = c$ is termed the Condorcet parameter (after the Marquis de Condorcet who ''saw the human mind as capable of removing obstacles to human progress). It can be negative, zero or positive. Cohen explains that ''when $c > 1$, each additional person increases the human carrying capacity enough for her own wants plus something extra.'' In this case, population can grow at a rate that is faster than exponential. He explains that ''when $c = 1$ each additional person adds to carrying capacity just as much as she consumes. . .$P(t)$ grows exponentially. . . . When $c < 1$, $P(t)$ grows logistically.'' Finally, Cohen terms L the Mill parameter after John Stewart Mill ''who foresaw a stationary population as both inevitable and desirable.''

Cohen did a least squares calibration of his model by hand, using point estimates of population over the past 2000 years. This led to parameter estimates of $P(0) = 0.252$ billion people, $K(0) = 0.252789$ billion people, $r = 0.0014829$ per billion people, and $L = 3.7$ billion people. These equations produce a sigmoid population curve which describes the data well and reaches an asymptote at a population of just under 20 billion people in about the year 2300.

Cohen assembled seven sources of population data, which were in turn based on hundreds of different sources of evidence. Only one of the sets of estimates he used included any indication of associated uncertainty. We conducted a careful review of these data and conferred with several prominent demographers. On the basis of these consultations, observing the range of the admittedly correlated estimates, and considering the strong experimental evidence of consistent expert over-confidence,[15] we assigned confidence intervals that range from ±17.5% in the year 0 (between 150 million and 290 million people) to ±5% in 1980 (between 4 billion and 4.9 billion).

Using these uncertain estimates of past population we then re-estimated Cohen's model using a Bayesian windowing technique, a method that assures that the uncertainty distribution of model output is consistent with the range of observed data for that output by rejecting inconsistent simulation results from the prior of sample outcomes.[23] The result is a probabilistic version of the Cohen model, which, while it should not be viewed as a reliable estimate of future world population, provides a more defensible upper bound on global population ≥100 years from now than the demographic model used in ICAM.

Following the strategy illustrated in Fig. 3, we performed a weighted combination of our probabilistic version of Cohen's model with the outputs of the demographic model in ICAM, using the weighting function shown in the upper part of Fig. 4.

Neither of these models include explicit treatment of low probability catastrophic drops in population which might be caused by pandemic, nuclear war, or similar catastrophes, although the Cohen model is of course consistent with past catastrophic population declines such as those caused by the Black Plague in the latter half of the fourteenth century and the massive loss of indigenous populations that followed European contact in North America in the sixteenth century.[24,25] A series of order-of-magnitude calculations we performed led us to conclude that on the global scale, population decline of the order of those experienced over the past 2000 years, and of those discussed by present military strategists[24–27] can be considered second order effects on century time scales. One obvious exception would be a large meteor impact.[28] We have not included that contingency in the current work.

5. CONCLUSIONS

The characterization and treatment of uncertainty poses a number of interesting challenges when the problems involve the assessment of indeterminate or complex coupled systems such as those involving human activity, climate and the ecosystem. Uncertainty can derive from variability and random processes and from the fact that we have limited understanding of how the world works.[3] For either or both cases, one can have uncertainty about the values of key model parameters. In addition, limited knowledge about the how the world works can give rise to uncertainty in model form. When such uncertainty becomes great enough, it may no longer make sense to try to use models for prediction or optimization, though in such cases it may still be possible to use the model world as a ''behavioral test bed,'' in which one uses autonomous agents within the model world to test the relative robustness of alternative observational and behavioral strategies in the face of uncertainty.

When a complex model must be operated into a region of its phase space for which it was not de-

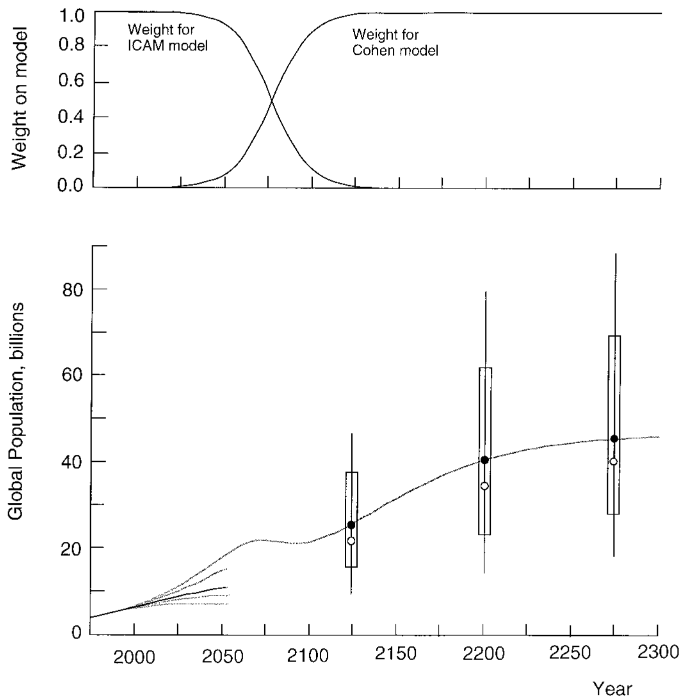

Fig. 4. Results of applying the model switch-over strategy described in the text, and shown schematically in Fig. 3, to the ICAM demographic model (until about 2050) and the Cohen upper-bound estimate of global population carrying capacity. The solid curve shows the ICAM mean projection. Shaded curves on the ICAM model show the 5th, 25th, 75th and 95th percentiles of the model projections. The 95th percentile of the ICAM model is blended over time with the Cohen upper-bound projection using the weighting functions shown above. Box plots report the uncertainty associated with the Bayesian window fit of the Cohen model to historical data. Solid dots are means, open circles medians. The boxes span the middle 50%, and the vertical lines 90%, of the distributions.

signed, different elements of the model may degrade at different rates. For example, different elements of a time-stepped model may become unreliable at different times as the model is run far into the future. One strategy is to assess the probability of model failure as a function of time, or of some endogenous model variable (incremental radiative forcing in the case illustrated here). Then one can display a time series that reports the likelihood of model failure along with the time series of model output. Alternatively, it may be possible to identify and disaggregate the various sources of model failure and model them separately. The one serious problem with this strategy is that in complex real-world situations, experts are likely to be able to identify only a portion of all the limitations or "surprises" that could be encountered, and are likely to be willing to assess probabilities for only a subset of the total. Still, identifying some and getting part way to a full treatment is clearly better than simply ignoring the possibilities.

Finally, when it is known that one portion of a model will become unreliable more rapidly than other portions of the model (e.g., over time, the socioeconomic submodel of ICAM will become unreliable before the geophysical model), it may be possible to develop much simpler order of magnitude models, or perform bounding calculations, which allow one to say something, even when detailed prediction is not possible. Through the use of subjective judgments, the results of several such analyses can be weighted and combined, in this case over time, to yield a more meaningful projection than would be obtained by running a detailed high-resolution model well past its domain of applicability.

ACKNOWLEDGMENTS

We thank the anonymous U.S. climate scientists and demographers whose judgments were used in this work as well as Joseph Kadane, Milind Kandlikar, Benoit Morel, Jerry Ravetz, James Risbey, and Patricia Steranchak for their thoughtful discussions and assistance. This work was supported by National Science Foundation grant DMS-9523602, NSF grant SBR-9521914 through the Center for Integrated Study of the Human Dimensions of Global Change, and contract WO 2141-24 from the Electric Power Research Institute.

REFERENCES

1. Presidential/Congressional Commission on Risk Assessment and Risk Management. *Framework for Environmental Health Risk Management, Vol. 1; Risk Assessment and Risk Management in Regulatory Decision-Making, Vol. 2* (Washington, D.C., 1997).
2. U.S. Environmental Protection Agency, *Policy for the Use of Probabilistic Analysis in Risk Assessment at the US Environmental Protection Agency* (Washington, D.C., 1997).
3. M. G. Morgan, "Uncertainty Analysis in Risk Assessment," *Human Ecol. Risk Assess.* **4,** 25–39 (1998).
4. M. E. Paté-Cornell, "Uncertainties in Risk Analysis: Six Levels of Treatment," *Rel. Engin. Sys. Safety* **54,** 95–111 (1996).
5. H. C. Frey and D. S. Rhodes, "Characterizing, Simulating and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human Ecol. Risk Assess.* **2,** 762–797 (1996).
6. J. S. Evans, J. D. Graham, G. M. Gray, and R. L. Sielken, Jr. "A Distributional Approach to Characterizing Low-Dose Cancer Risk," *Risk Anal.* **14,** 25–34 (1994).
7. J. S. Evans, G. M. Gray, R. L. Sielken, Jr., A. E. Smith, C. Valdez-Flores, and J. D. Graham, "Using of Probabilistic Expert Judgment in Uncertainty Analysis of Carcinogenic Potency," *Reg. Toxicol. Pharmacol.* **20,** 15–36 (1994).
8. R. J. Budnitz, G. Apostolakis, D. M. Boore, L. S. Cluff, K. J. Coppersmith, C. A. Cornell, and P. A. Morris, *Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and the Use of Experts* (Lawrence Livermore National Laboratory, URCL-ID-122160, Livermore, CA, 1995).
9. M. E. Paté-Cornell, "Uncertainty in Global Climate Change Estimates," *Climatic Change* **33,** 145–149 (1996).
10. M. G. Morgan and H. Dowlatabadi, "Learning from Integrated Assessment of Climate Change," *Climatic Change* **34,** 337–368 (1996).
11. H. Dowlatabadi and M. G. Morgan, *The Integrated Climate Assessment Model* (Center for Integrated Study of the Human Dimensions of Global Change, Department of Engineering and Public Policy, Carnegie Mellon University, 1999).
12. H. Dowlatabadi, "Reflections on Setting and Meeting Thresholds for Climate Change," *Climatic Change* (submitted).
13. M. G. Morgan and D. Keith, "Subjective Judgments by Climate Experts," *Environ. Sci. Technol.* **29**(10), 468–476 (October 1995).
14. D. Kahneman, P. Slovic, and A. Tversky (ed.), *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press, New York, 1982).
15. M. G. Morgan and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge University Press, New York, 1990).
16. S. H. Schneider and B. L. Turner, "Imaginable Surprise in Global Change Science," *J. Risk Res.* **1,** 165–185 (1998).
17. J. Rotmans and H. Dowlatabadi, "Integrated Assessment Modeling," in S. Rayner and E. L. Malone (eds.), *Human Choice and Climate Change, Vol. 3: The Tools For Policy Analysis,* (Battelle Press, Columbus, 1998).
18. IPCC, *Climate Change 1995, 3 Volume Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, New York, 1995).
19. E. A. Parson, "Integrated Assessment and Environmental Policy Making: In Pursuit of Usefulness," *Energy Policy* **23**(4–5), 463–475 (1995).
20. W. Lutz, W. Sanderson, and S. Scherbov, "Doubling of World Population Unlikely," *Nature* **387,** 803–804 (June 19, 1997).
21. J. S. Cohen, "Population Growth and Earth's Human Carrying Capacity," *Science* **269,** 341–346 (July 21, 1995).
22. J. S. Cohen, *How Many People Can the Earth Support*? (W. W. Norton & Co., New York. 1995).
23. K. P. Brand and M. J. Small, "Updating Uncertainty in an Integrated Risk Assessment: Conceptual Framework and Methods," *Risk Anal.* **15,** 719–731 (1995).
24. C. McEvedy and R. Jones, *Atlas of World Population History* (Penguin Books, New York, 1978).
25. R. Thornton, *American Indian Holocaust and Survival* (University of Oklahoma Press, Norman, 1987).
26. D. R. Snow, "Microchronology and Demographic Evidence Relating to the Size of Pre-Columbian North American Indian Populations," *Science* **268,** 1601–1604 (June 16, 1995).
27. Benoit Morel, personal communication (1998).
28. G. L. Verschuur, "Impact Hazards: Truth and Consequences," *Sky and Telescope* **95**(6), 26–34 (June 1998).
29. T. C. Schelling, "Intergenerational Discounting," *Energy Policy* **23,** 395–402 (1994).