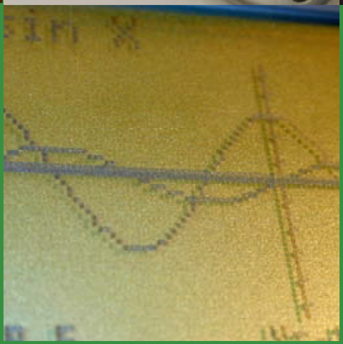





Characterizing And Communicating Scientific Uncertainty

A Report of the Aspen Global Change Institute
Elements of Change Series

Susan Joy Hassol
John Katzenberger
Editors

ASPEN GLOBAL CHANGE INSTITUTE

Characterizing and Communicating Scientific Uncertainty

A report on the Aspen Global Change Workshop
July 31 - August 8, 1996
Aspen, Colorado USA

CO-CHAIRS:
Richard Moss
Stephen Schneider

WRITER-EDITORS:
Susan Joy Hassol

ELEMENTS OF CHANGE SERIES EDITOR:
John Katzenberger

Furthering the understanding of Earth Systems and global environmental change

FUNDING
The Aspen Global Change Institute gratefully acknowledges support for its 1994 summer science sessions
PROVIDED BY THE
National Aeronautics and Space Administration
Grant Number NAGW-3583,
National Science Foundation
Grant Number OCE-9417138,
Environmental Protection Agency, and interagency support of the Subcommittee on Global Change Research of the U.S. Global Change Research Program

VIDEO AND HARDCOPY
Archival videos of the presentations from this session are available through Aspen Global Change Institute's website, www.agci.org. Hardcopy versions of the report are also available on a cost of reproduction basis. Contact AGCI for more information. Publications of AGCI are available on-line in cooperation with the Global Change Research Information Office (GCRIO) at http://www.gcrio.org/agci-home.html

THE RECOMMENDED CITATION:
Hassol, S.J., and J. Katzenberger, eds. 1997, electronic edition 2008. Natural Hazards and Global Change. Proc. of an Aspen Global Change Institute Workshop 10-20 July 1996, Elements of Change series, AGCI.

DESIGN AND PRODUCTION*:*
Original hardcopy edition: Kelly Alford
Electronic edition: Susannah Barr

ASPEN GLOBAL CHANGE INSTITUTE
100 EAST FRANCIS STREET • ASPEN COLORADO 81611
970 925 7376 • agcimail@agci.org • www.agci.org

Table of Contents

Acronyms

DOC: Dissolved Organic Carbon

EMF: Energy Modeling Forum

EN: El Niño

ENSO: El Niño/Southern Oscillation

FCCC: Framework Convention on Climate Change

GCC: Global Climate Coalition

GCM: General Circulation Model

GDP: Gross Domestic Product

GHG: greenhouse gas

GWP: global warming potential

IAM: integrated assessment model

IMAGE: Integrated Model for Assessment of the Greenhouse Effect

IPCC: Intergovernmental Panel on Climate Change

ITC: induced technological change

MSU: Microwave Sounding Unit

NAS: National Academy of Sciences

NGO: non-governmental organization

ODP: ozone depleting potential

OECD: Organisation for Economic Cooperation and Development

ppmv: parts per million by volume

PSHA: Probabilistic Seismic Hazard Analysis

SAR: Second Assessment Report

SO: Southern Oscillation

SOI: Southern Oscillation Index

SSHAC: Senior Seismic Hazard Analysis Committee

SST: sea surface temperature

TFI: Technical Facilitator/Integrator

WG: Working Group (of IPCC)

Session Synthesis Essay: Characterizing and Communicating Scientific Uncertainty: Building on the IPCC Second Assessment

Richard Moss and Stephen Schneider, Session Chairs

> How can science be most useful to society when evidence is incomplete or ambiguous, the judgments of experts vary, and policymakers seek guidance and justification for courses of action that could cause significant societal changes?

Uncertainty, or more generally, debate about the level of certainty required to reach a "firm" conclusion, is a perennial issue in science. The difficulties of explaining uncertainty become increasingly salient as society seeks policy prescriptions to deal with global environmental change. How can science be most useful to society when evidence is incomplete or ambiguous, the subjective judgments of experts about the likelihood of outcomes vary, and policymakers seek guidance and justification for courses of action that could cause significant societal changes? How can scientists improve their characterization of uncertainties so that areas of slight disagreement do not become equated with purely speculative concerns, and how can individual subjective judgments be aggregated into group positions? And then, how can policymakers and the public come to understand this input and apply it in deciding upon appropriate actions? In short, how can the scientific content of public policy debates be fairly and openly assessed?

The Case of the IPCC Second Assessment Report

Interest in climate change, potential impacts, and adaptation /mitigation policy options increased dramatically during 1995-96. While there are many explanations for this, one contributing factor was the conclusion, reached by the Intergovernmental Panel on Climate Change (IPCC) in its Second Assessment Report (SAR)(IPCC 1996a-c) that, even considering remaining uncertainties, "the balance of evidence suggests that there is a discernible human influence on global climate." This conclusion, which was negotiated over a period of a year by hundreds of scientists and policymakers, acknowledged the strong belief of most experts that human modification of atmospheric composition has led to noticeable climatic effects and likely significant climate change in the decades ahead. Though not especially controversial in the scientific community, the statement created a maelstrom of both support and criticism from a variety of interest groups who seemed confused by the different ways in which uncertainties and knowns were explained in the technical chapters and the Summary for Policymakers (e. g., see Science, Nature, 1996).

The most recent IPCC report and its predecessors provided "best estimates" of possible climatic futures, as well as a broad range of plausible outcomes, including possible "outlier" events. The implications encompass consequences of climate change that range from mildly beneficial to potentially catastrophic changes for ecosystems and human activities such as water management, development of coastal margins, and agriculture. Although more confidence was naturally expressed in outcomes near the centers of those wide ranges between the high and low outliers, some interest groups understandably focused on possible extreme outcomes, which sharpened the debate and created substantial contention.

The purpose of the IPCC and other assessments of scientific research is to convey to interested publics, including decision-makers, advisors, the media, private-sector businesses, and environmental /community groups, the most up-to-date information available. One of the major challenges is that the assessments necessarily must present a snapshot of information which is continuously evolving. At the time of preparation of the SAR, the uncertainties included, for example, the possibilities of large and/or abrupt climate changes and/or technological breakthroughs that could radically reduce emissions abatement costs in the future. Given the use of the IPCC reports in policy making, and the need of decision-makers to determine their response to the risks of climate change before all uncertainties can be resolved (even in principle) to the satisfaction of every interest group, the available information, imperfect as is may be, must be synthesized and evaluated at periodic intervals.

Thus, a great deal of importance is attached to the need to assess and explicitly distinguish which aspects of technical controversies that affect our understanding of these uncertainties are well understood and enjoy strong consensual support, which aspects are somewhat understood, and which are highly speculative. Unfortunately, in the media and political debates, such degrees of certainty and uncertainty often become blurred. As a result, the nuanced characterization of uncertainty that might occur in professional assessment is often mis - translated into the appearance of scientific cacophony in the public arena.

The nuanced characterization of uncertainty that might occur in professional assessment is often mis-translated into the appearance of scientific cacophony in the public arena.

At the same time, scientists themselves struggle with the highly subjective and qualitative nature of the assessment process, preferring, by tradition, to study individual components of problems that can be tested, rather than the necessarily more difficult synthesizing of these components of relevance to decision-makers and the public. Qualitative descriptions of the level of certainty attached to a particular finding terms such as “almost certain,” “probable,” “likely,” “possible,” “unlikely,” “improbable,” “doubtful,” “almost impossible” mean different things to different people and hence are not precise descriptors, and they are sometimes used rather inconsistently or uncritically in assessments (let alone by the general public and the media). Individuals and groups often use simplifying assumptions or heuristic procedures for making judgments about uncertainty. The consequence of this can be a tendency towards overconfidence in the likelihood of median outcomes and a tendency to underestimate the probability of outlier events or surprises.

In order to examine previous practice in characterizing uncertainties, as well as opportunities for improving the assessment process, a group of researchers and analysts met for nine days at the Aspen Global Change Institute (AGCI) in August 1996 to investigate ways of communicating uncertainty within specific content areas of climatic assessment. The group included independent scholars, statisticians, decision analysts, media and policy analysts, and a number of Lead Authors from all three working groups of the IPCC SAR, including researchers from the physical, biological and social sciences. Our overall goal was to examine possibilities for achieving greater consistency in evaluating the judgments of scientific experts and in communicating these judgments to non-specialists. Serving as Co -Chairs of the workshop, we facilitated discussions on four basic sets of questions:

1 What approaches to establishing uncertainty ranges and confidence levels were used in the preparation process for the SAR (IPCC 1996)? How did these approaches and other factors affect the conclusions and ranges presented?

2 What approaches could be used to represent the center, the body, and the range of informed technical opinion in future assessments (including quantification of uncertainties)?

3 How do uncertainty and the presentation of levels of confidence translate into media coverage and policy debates? What approaches are used in other fields or could be developed to communicate more effectively the nuances of uncertainty as understood by climate experts, impacts specialists and policy analysts, to the media, policymakers and the public at large?

4 What recommendations for improving future IPCC or other international or national assessments could be developed by authors and experts?

Our overall goal was to examine possibilities for achieving greater consistency in evaluating the judgments of scientific experts and in communicating these judgments to non-specialists.

Our overall goal was to examine possibilities for achieving greater consistency in evaluating the judgments of scientific experts and in communicating these judgments to non-specialists.

The workshop took advantage of the progress being made in decision sciences, in which the ability to deal formally with uncertainty is improving. These formal approaches to incorporation of new information into the policy process via adaptive strategies (e. g., see Schlesinger, in this report) provide a promising framework that involves not only provision of new data, but also changes in model structure, validation and aggregation. The analyses and methodological alternatives and suggestions for improvements in the assessment process from the participants make up the most substantial portion of this report (see presentation summaries which follow).

This summary essay synthesizes the analyses and suggestions made during the AGCI meeting. In the next several sections we will:

1 review briefly the recent IPCC process and several other climate change assessments as examples of how formal assessments have treated uncertainty,

2 describe how estimates of central tendencies and ranges have been made in previous assessments and reports,

3 briefly review alternatives for more formal and consistent definitions of outlier uncertainties, and

4 discuss the potential applicability of these methods to future assessments, in particular the Third Assessment Report (TAR) of the IPCC, scheduled to be completed during 2000-2001.

IPCC: Background and Case Studies from the SAR

The IPCC was established by the United Nations Environment Programme (UNEP) and the World Meteorological Organization (WMO) in 1988, to assess scientific information about climate change relevant for international and national policy formulation. The IPCC produced its first full assessment report in 1990, a supplementary report in 1992, a special report on radiative forcing of climate in 1994, and its Second Assessment Report in 1995. It is also producing a series of Technical Papers and Special Reports in 1996-98.

The IPCC operates at two levels: as a formal intergovernmental body, and as a scientific and technical assessment body.

Intergovernmental body: Representatives of each government involved in the climate change negotiations meet in formal plenary sessions to approve the topics for assessment and the workplans for preparation of the reports. They also review and "accept" (as a whole, rather than on a line-by-line basis) the detailed reports, as well as "approve" (on a line-by-line basis) the Summaries for Policymakers that highlight the policy implications of the reports.

Scientific/technical assessment body: The body of scientists and technical experts involved as lead authors, contributors, and reviewers of IPCC assessments represent academia, governments, industry, and environmental organizations around the world.

The process used to prepare IPCC reports can be quite lengthy, often spanning two or more years. The process is initiated by national governments which commission the reports, either through the IPCC itself or through the subsidiary bodies to the U. N. Framework Convention on Climate Change (UNFCCC), approve preliminary outlines which set out the topics to be addressed, and nominate experts as potential lead authors, contributors or reviewers. Non-governmental organizations (NGOs) from both the environmental and business sectors also participate in this process. Lead and contributing authors are chosen for each chapter from among the nominations by the Bureau of the Working Groups responsible for overseeing the report. Authors and reviewers come from many countries and are trained in disciplines ranging from atmospheric chemistry to zoology and engineering to economics. The Bureau works to achieve substantive and geographical diversity. Insofar as practicable, the composition of a group of Lead Authors reflects a balance among different points of view, and includes experts from both industrialized and developing countries.

The IPCC operates at two levels: as a formal intergovernmental body, and as a scientific and technical assessment body.

During the preparation of reports, authors are charged with reviewing the most up-to-date scientific information, reconciling competing views where possible, and characterizing the range of perspectives and the reasons for disagreements within author teams when consensus does not exist. Lead authors are instructed by the IPCC not to advocate particular points of view, but rather to characterize the range of scientific information available; however, no formal roles (e. g., mediators, evaluators, proponents, etc.) are assigned. Text is prepared by Lead Authors or solicited by the authors from other contributors; unprompted submissions are also accepted for consideration. As will be discussed below, the process followed by different writing teams to arrive at an agreed-upon text which describes areas of both agreement and disagreement varies considerably, from Working Group to Working Group, and issue to issue.

The IPCC documents typically undergo two stages of peer review, first by experts and subsequently by governments and accredited organizations. The expert review is open to those who have established research or technical credentials in a field related to the chapters of the report being reviewed. Effort is made to provide opportunities for input from scientific skeptics. The documents are made freely available to the full range of stakeholder groups, including environmental and industry NGOs. The Working Group Bureaus have been responsible for ensuring that comments received from reviewers are considered by the authors (a proposal under consideration for the next assessment would give this responsibility to editorial boards). Authors are requested to document their responses to comments, and these are kept on file in the relevant Working Group Support Units. The government review is open to all participating governments, as well as accredited NGOs and experts who participated in the first round of review.

The final stage of the report process involves review of the report by an intergovernmental meeting of the Working Group and/or IPCC plenary, where representatives speak for their

governments or organizations. Most chapter authors usually attend these sessions as well, to interact with the government representatives over the report. The plenaries are responsible for ensuring that the full reports have been prepared according to established IPCC procedures that ensure openness and balance. They are also responsible for "approving" on a line-by-line basis the report's Summary for Policymakers, which is constructed by consensus in a manner that will make it consistent with the underlying technical report, representing a balance among viewpoints or interpretations of the science.

Making Collective Judgments

Collective judgments about the degrees of certainty associated with particular conclusions have been an important part of the IPCC assessment process. This is because there are no precise, mathematical approaches for simply combining the varying conclusions that different experts have reached regarding such exceedingly complex issues as the sensitivity of climate to increases in greenhouse gas (GHG) concentrations, the amount by which observed temperatures have already changed at a global mean level, or the reductions to per capita GDP that might result from climate changes associated with doubled concentrations of GHGs. Collective judgments were reached in a variety of ways, depending on the subject domain and the particular set of lead authors. The following examples from each of the three IPCC working groups Working Group I (The Climate System), Working Group II (Impacts, Adaptation, and Mitigation) and Working Group III (Social and Economic Dimensions of Climate Change) (IPCC 1996 a, b and c, respectively) illustrate the diversity of approaches used.

Authors are charged with reviewing the most up-to-date scientific information, reconciling competing views where possible, and characterizing the range of perspectives and the reasons for disagreements.

The fact that there is variation in approach across issues and chapter writing teams does not mean that the judgments provided in the assessment are irreproducible, of poor quality, or misleading IPCC authors developed qualitatively reasonable ranges and mean values through open, iterative discussions. Given similar information, these or very similar estimates would likely have been reached by different teams of Lead Authors acting independently. Here we demonstrate that improvements to the assessment process are possible, based in part on recent advances in statistics and methods for subjectively estimating probabilities of different events or conclusions. In our view, these improvements could be usefully applied to estimation of particularly important parameters or judgments regarding key issues. The techniques could be especially useful to improve the clarity and consistency of the estimates of the outliers or end points of the ranges of estimates generated by the writing teams. We also argue that further research into additional methods for developing collective judgments of certainty is needed in order to expand the range of approaches available to scientific assessors.

Statistics Refresher

1 The mean is the arithmetic average of a variable.

2 The median is the value beneath which half of the observations for a variable fall.

3 The mode is the most common value for a variable.

4 The mean, median and mode are called measures of location or measures of central tendency.

5 The standard deviation is a measure of how heterogeneous (spread out) the values of a variable are. It is the square root of the sum of the squared deviations around the mean. A “deviation” is the arithmetic difference between each value and the mean of the variable. One can think of the standard deviation as roughly (not literally) the average disparity between the mean and each value.

6 The range is another measure of heterogeneity and is the difference between the maximum and minimum value.

7 The interquartile range is another measure of heterogeneity and is the difference between the 75th percentile and the 25th percentile. It is the interval in which the middle 50 percent of the values fall.

8 The standard deviation, range, and interquartile range are called measures of variability or spread.

9 A probability distribution function is the set of probabilities, each associated with a discrete value of the variable in question.

10 A probability density function contains the same information as a probability distribution function but for variables whose values are continuous, not discrete.

11 Commonly used probability distribution functions are the binomial, multinomial, and the Poisson. Commonly used probability density functions are the normal, T, Chi-squared, and F.

12 A joint probability is the probability of two or more events occurring. A joint probability distribution is the set of such probabilities for two or more events. For example, suppose in the future the climate will either have more precipitation or less and be either warmer or colder. There are four combinations, then, with a probability attached to each. Those four probabilities, one attached to each combination, constitute a joint probability distribution.

13 A cumulative probability is the sum (or integral) of the probabilities (or densities) less than or equal to some value. For example, one might compute the probability that the temperature increase due to global warming is less than or equal to 3°C.

When comparing values from two different variables or the same variable measured on two different objects, we often cannot compare the values directly. For example, how could one compare a temperature of 30°C to rainfall of half a meter? Or how can one compare a hot day in Los Angeles to a hot day in Seattle? So, we standardize. One common way is with percentiles. We might say that a high temperature of 90°F is at the 75th percentile for Los Angeles summer days while a high temperature of 90°F is at the 95th percentile for Seattle summer days. We can then compare percentiles.

Another way is to compare z-scores, which are in standard deviation units. Thus, a 90°F temperature in Los Angles might have a z-score of around 1.5 which means it is 1.5 standard deviations above the mean. Since a standard deviation is roughly the average distance from the mean, Los Angeles temperature of 90°F is a bit more than an average departure from the mean. A 90°F temperature for Seattle might have a z-score of around 2.0, which means that it is about 2 standard deviations above the mean. So, it’s about twice the average distance from the mean. So, z-scores and percentiles get at much the same thing and in fact are very closely related. For example, when one can work with the normal density function for the problem at hand, one can move easily between z-scores and percentiles: z-scores are monotonically related to percentiles.

Case 1: "Climate Sensitivity"

For almost 20 years, a significant reference point in scientific assessments of climate change has been the estimate of "climate sensitivity" the globally averaged surface temperature response that eventually can be expected to occur (i. e., in equilibrium) if CO 2 were to double from pre-industrial levels and remain at this concentration indefinitely. This factor, "T2x," was first identified in a 1979 National Research Council report (NRC 1979) as 1.5-4.5°C warming (for a discussion of how this value was determined, see Schneider, in this report). This estimated range for climate sensitivity has remained essentially invariant through various assessments over the years, not because the studies that have taken place every few years revealed the same exact quantitative judgments, but largely because the relative changes from assessment to assessment were too small for scientists to justify changing this range on any plausible scientific grounds. While some changes in the underlying science have occurred (e. g., new formulations of cloud or biophysical parameterizations), small (±0.5°C) changes to the outliers (end points) of the ranges have seemed unjustified even frivolous to many assessors, in view of the absence of fundamental new data or tested theory. Thus, the range has remained the same.

Of course, subjectivity is inherent in complex systems analysis, but the process of achieving more consistent aggregate scientific judgments is critical to establishing more meaningful and credible ranges of potential outcomes like T2x. More consistent estimates of the endpoints of a range for any variable would minimize misunderstandings of the intent of the Lead Authors of assessments such as IPCC 1996a, and would, in turn, reduce the likelihood that interest groups could misunderstand or occasionally, misrepresent the scientific assessment process. However, such attempts to achieve more consistency in evaluating the subjective judgments of appropriate experts in specifying outlier values (or the distribution of subjective probabilities within and beyond the outlier estimates) have not received as much attention as we believe they deserve. Assessors should be encouraged to state whether the range of estimates provided for T 2x (or other parameters) is within some percentile limits (e. g., 10 and 90 percent) or within one, two or three standard deviations of the mean (and, if possible, to provide the shape of the subjective probability distribution). This would make the ranges more meaningful to other scientists, the policy community and the public, and would help to increase the likelihood that the assessors themselves all have had the same level of probability in mind for the estimated range.

IPCC authors developed qualitatively reasonable ranges and mean values through open, iterative discussions. Given similar information, these or very similar estimates would likely have been reached by different teams of Lead Authors acting independently.

Case 2: Observed Changes in Surface Air Temperature

Another example from Working Group I is the estimate of 0.3° - 0.6°C surface air temperature warming for the world since the late 19th Century. This range was maintained in the SAR (1996a) assessment at the same value as in the 1990 first assessment, even though the years 1990, 1991 and 1995 were among the highest in the record. The AGCI Workshop participants familiar with the process suggested that it was more a matter of strong personalities not wishing to defend to their colleagues any change than a collective vote of the knowledgeable group which led to the range remaining unchanged. It is our understanding that a number of scientists in the process may have preferred changing the warming trend estimate to 0.4°- 0.7°C, but a sense of collegiality, combined with the difficulty of scientifically justifying changes to a range which was not initially chosen on consistent criteria for picking outlier endpoints, simply resulted in the range remaining constant. This is another example where more formal analytic methods may have led to a slightly different outcome.

Case 3: Detection/Attribution

One of the major areas of advance in IPCC 1996a came in the area of detection and attribution of climate change. "Detection" is the process of demonstrating that an observed change is highly unusual in a statistical sense (analogous to measuring one's body temperature and determining that it is significantly higher than the normal). "Attribution" is the process of establishing cause and effect (analogous to explaining why one's body temperature is elevated once its baseline, "normal," has been established). While major scientific uncertainties still exist with regard to estimates of the magnitude, pattern and evolution of different forcings, IPCC 1996a concluded that the Earth is warming and that its mean temperature is warmer than it has been since at least the 15th Century; that observed geographical and vertical patterns of temperature change are becoming increasingly similar to model predictions that incorporate combined GHG warming and sulfate aerosol cooling effects; that the observed patterns are different enough from those due to natural variability; and that observations and model predictions generally agree in overall magnitude and timing of change.

The process used by the authors of WG I Chapter 8 on Detection of Climate Change and Attribution of Causes was statistically rigorous in fact, it was more rigorous than that used in many other chapters or most previous assessments. In addition, the key conclusions were supported by multiple lines of evidence and were not particularly surprising to most members of the informed scientific community given increases in atmospheric concentrations of GHGs. Yet the chapter's findings resulted in what one of the Lead Authors referred to as "a nightmare media aftermath" of controversy.

The key conclusions were supported by multiple lines of evidence and were not particularly surprising to most members of the informed scientific community. Yet the chapter's findings resulted in what one of the Lead Authors referred to as "a nightmare media aftermath" of controversy.

Why did the chapter produce such a commotion, and what, if anything, should we learn about the use of more consistent procedures for estimating levels of confidence and confidence ranges? Extracting lessons from the case is difficult due to its uniquely charged nature. Not only does detecting a human influence on climate cross a psychological/cognitive barrier (i. e., humans can and are changing the planet on a global scale), but this conclusion also raises the issue of a need for potentially significant changes in economic and technological patterns that could affect many important interests. In addition, the case is clouded by perceived although not actual violations in the rules of procedure usually followed to prepare and edit chapters. The attempt to be as up-to-date as possible also meant that some of the newest materials had not been vetted for very long in the scientific literature and the community; in addition, permissible editorial changes were made following the final plenary session at which the chapter was discussed.

A subtle but critical distinction in the IPCC rules between "acceptance" of chapters (which does not require line-by-line approval) and "approval" of the Summaries for Policymakers (which does require line-by-line agreement to text) also contributed to the confusion. Some critics were either unaware of or chose to overlook this distinction, and criticized changes that were made to the text in the late stages of preparing the report. These changes were completely permissible, however, since the chapters are not "approved" on a line-by-line basis (a practical impossibility for a report of many hundred pages). (Edwards and Schneider, 1997, have noted that scientists typically bestow greater levels of trust in their colleagues to respond to critiques than is typically tolerated in legal or political proceedings an epistemological professional difference that contributed to political attacks on the IPCC process after the SAR was written.) Moreover, the conclusions were first communicated to the public in The New York Times , months ahead of the official release of the report, and this may have induced suspicions of procedural irregularity. This situation led to a well-orchestrated attack on the IPCC process

by a number of skeptics, including some scientists, who believed (or at least stated) that their views were not adequately represented in the report. The IPCC responded vociferously to these charges, perhaps exacerbating the problem by calling more attention to the attacks than they otherwise may have received (for further discussion, see MacCracken, Edwards and Schneider, in this report).

From the point of view of lessons to be learned in shaping treatment of uncertainties and outlier ranges, this case seems to illustrate the fact that some conclusions, no matter how statistically rigorous, will produce controversy in the political arena. Clarifying procedures for preparing and revising IPCC reports so that all stakeholders are familiar with them, and developing clear procedures within the IPCC for responding to media controversies, may do just as much to avoid such controversies in the future as would employing advanced techniques for aggregating and communicating to general audiences the opinions of experts.

From the point of view of lessons to be learned in shaping treatment of uncertainties and outlier ranges, this case seems to illustrate the fact that some conclusions, no matter how statistically rigorous, will produce controversy in the political arena.

Case 4: Impacts of Climate Change on Vegetation and Ecosystems

The primary controversies in projections concerning the future dynamics of terrestrial ecosystems in response to climate change involve (1) the decline of terrestrial biodiversity and (2) the role of the terrestrial biosphere in the global carbon cycle. (Examinations of climate impacts on animals are in their infancy e. g. , Root in this report, Root and Schneider, 1993, and will be a stronger focus in future reports.) Impacts on ecosystems and the ability of various species to migrate in response to climate change have been investigated through models which attempt to determine the implications of a number of competing processes that affect the rates of change to which species can adapt, and hence to determine the sorts of ecosystems which potentially could persist in different locations. In the case of the carbon cycle, uncertainties preclude, at present, a confident assessment of the current or future role of the terrestrial biosphere as either a source or a sink of carbon. Some evidence suggests that the biosphere has been a source of CO_2 emissions during natural warmings of the past 250,000 years, producing a positive climate feedback leading to further warming. Other evidence suggests that, since CO_2 stimulates growth of vegetation and phytoplankton, an increased CO_2 concentration will induce an increased carbon sink, providing a negative feedback on the global carbon cycle by moderating the projected growth in the CO_2 concentration. Depending on the interaction of effects in the transition from current to future climate, there is also the potential for a “pulse” of carbon to be released from dying forests (but not to be fully reabsorbed from the atmosphere in new forests for many decades or centuries), creating a positive or amplifying feedback.

In IPCC 1996a and b, Working Group I and II authors writing chapters covering terrestrial ecosystems contended with a number of these issues. Both groups estimated the potential effects of climate change on the carbon cycle, while Working Group II authors estimated, inter alia, the percentage change in vegetative cover for specified climate change scenarios. These latter estimates were based on group discussions of model results and shared agreements regarding their interpretation, rather than any formal statistical procedure for aggregating subjective probabilities (again, the lack of formal methods for objectively assessing outlier ranges limits the techniques available for application). While those involved in the discussions were familiar with the literature on both modeling and observations, the ranges ultimately chosen were not based on specified subjective probabilities for outliers. Methods to develop a judgment as to whether the outliers were within some percentile range or within one, two or three standard deviations of the mean (or to depict the shape of the subjective probability distribution) would have added clarity to the assessment.

Case 5: Impacts of Climate Change on Agriculture and Food Security

Another controversial aspect of the impact assessment concerned the adaptive potential of agriculture to climatic change and the potential implications for regional and global food security. The assessment of these issues involved a wide range of controversies, from crop/ climate /insect interactions and carbon dioxide fertilization to the socioeconomic conditions under which hunger occurs and the economic modeling of agricultural trade. IPCC 1996b reached the conclusion that "global agricultural production can be maintained relative to baseline production under climate change as expressed by general circulation models (GCMs) under doubled CO_2 equilibrium climate scenarios." However, in reaching this conclusion, the authors disagreed over whether net global effects really had meaning, or were in fact meaningless and should not be reported. There were also concerns regarding both over- and under-estimation of damages in this level of aggregation. The final wording reflected a careful if imprecise compromise.

The authors disagreed over whether net global effects really had meaning, or were in fact meaningless and should not be reported. There were also concerns regarding both over- and under-estimation of damages in this level of aggregation.

Estimates of the impact of equilibrium doubled-CO_2 climate conditions were made using crop-climate models, and the assessment indicated a wide range of yield changes, from large and positive to large and negative, when compared to results under current climate conditions. Some of the variation in results is accounted for by differing assumptions in different studies regarding the extent to which farmers would be able to adapt their crops or farming techniques to the new climate, thereby offsetting some of the damages or even, in some cases, taking advantage of opportunities created by expected alterations in weather conditions. A significant debate exists in the impacts community over the rate, costs, and levels of adaptation that are possible. Furthermore, assuming that some adaptation would eventually occur, others have questioned whether it would be delayed by decades because of the high natural weather variability that would mask any anthropogenic climate signal (e. g., Schneider, 1996, Morgan and Dowlatabadi, 1996).

Another debate occurred over how to operationalize one of the key concepts employed in Working Group II "vulnerability." Vulnerability could be defined for different aspects of agriculture: yield, farmer income (affecting regional, national or global economic vulnerability), or broader issues of food availability and security. While the writing team developed a precise definition of this term based on the probability density function for climate change and a damage function that relates impacts to varying levels of climate change in practice, specific climate model results played little role in the determination of vulnerability. These determinations were based on other factors, such as identifying those populations judged as being vulnerable to hunger or famine, or identifying potential thresholds for different crops.

Another difficulty was incorporating the potential for surprises that are possible and could be important (e. g., spread of a plant disease to new regions or crops). A "linear" model of food supply (which assumes that supply will grow to meet demand) may have provided a reasonable gross approximation for the recent past, but that does not mean that future assessments should not systematically consider other possibilities. As in other areas, an important next step in improving assessments would be to reflect the potential for such surprises in outlier ranges.

Case 6: Aggregate Economic Impacts of Climate Change and Emissions Abatement

Of all the cases examined, conducting a quantitative assessment of the economic impacts of climate change presented the most challenges (see Chapter 6, The Social Costs of Climate Change, IPCC 1996c). The IPCC writing team was faced with an overwhelming amount of

information on the potential for a diverse set of impacts in many sectors and regions, and there was no robust methodology for assigning values to each impact, aggregating them, and presenting the sum in monetary terms or as a percentage of GDP.

Valuation was difficult enough for market sectors such as agriculture or human infrastructure in coastal zones, for which markets help to establish monetary values; it was an especially serious challenge when it came to valuing impacts on non-market sectors such as ecosystems and the goods and services that they provide. Due to the absence of such a methodology, the chapter authors restricted their work to a review of the existing literature on the topic, without much discussion of alternative assumptions. Valuation was also extremely controversial when applied to the question of valuation of a human statistical life. Some argued that this value was higher in developed than developing countries, while others refused to accept this result on moral grounds; more fundamentally, some objected to even attempting to assign a monetary value to a human life. However, in spite of these and additional uncertainties associated with estimates of economic impacts, there is great pressure from the policy community to provide such estimates because (1) they deliver a single indicator of climate change impact, and (2) they make impacts directly comparable to the costs of emission control (for further discussion, see Tol in this report, and Fankhauser, Tol and Pearce, 1997).

The team was faced with an overwhelming amount of information on the potential for impacts in many sectors and regions, and there was no robust methodology for assigning values to each impact, aggregating them, and presenting the sum.

Working Group III estimated damage costs resulting from climate change impacts to range from $5-$125 per ton of carbon emitted, the range stemming from differences in estimation techniques as well as different assumptions about the appropriate “discount rate” to use in assessing future impacts in current monetary terms (note that the choice of discount rate is a source of significant dispute, but not a source of uncertainty, because it is a political/ethical choice, and the effects of different discount rates on the estimates of climate impacts can be calculated with precision). The assessment presented a range of best guess estimates of aggregate damages. Globally, annual impacts under doubled CO_2 equilibrium climate were estimated to range from 1-2 percent of GDP; regionally, the annual impacts on GDP were estimated to range from slightly positive (for the former USSR and perhaps China), to mildly negative (for OECD countries), to as much as negative 10 percent (for Africa).

It is extremely important to note that ranges associated with these impacts estimates simply represented the range of best guesses of the authors, not their estimates of the full range of potential damages from low to high. No estimation of uncertainties with regard to the full social costs of climate change was made, and in fact, no systematic calculations are available in the literature as yet, as far as we know; but generally, the uncertainties are known to be large and difficult to quantify formally because of such issues as valuing non-market impacts in monetary terms (e. g., Nordhaus, 1994). The lack of estimating techniques made the existing uncertainties difficult to communicate. In the opinion of one lead author, Chapter 6 of WG III conveys a message that knowledge is better developed than it, in fact, is, and that uncertainties are smaller than they actually are. It does seem likely that the ranges would not be nearly as small as portrayed in the chapter had procedures to tap outlier opinions been more formal, thus allowing lower probability minimum and higher maximum range outliers to be included in the estimates provided.

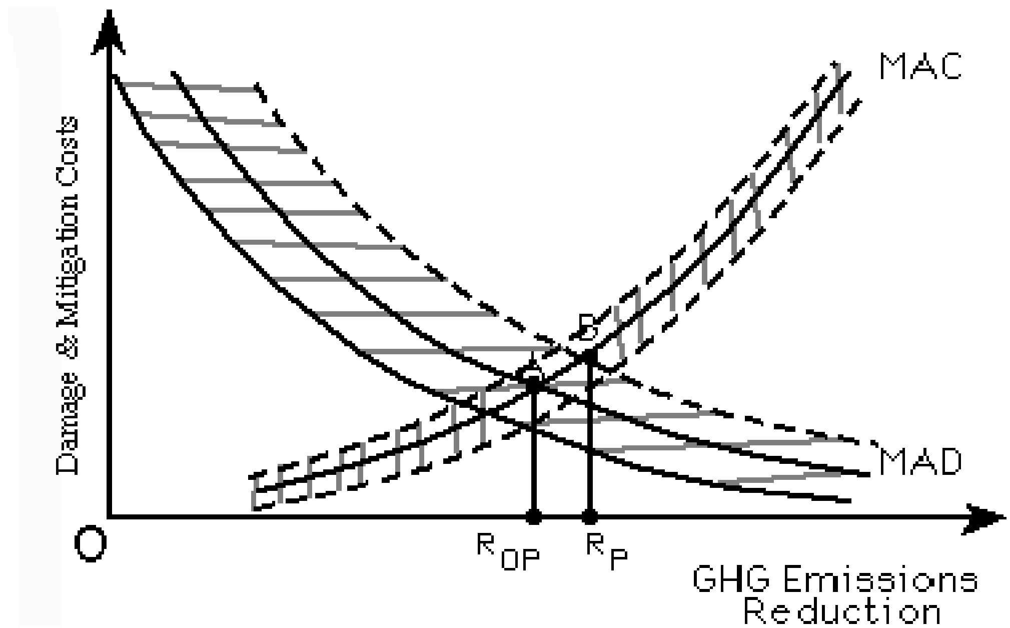


Figure 2.1
Marginal Avoided Damage Costs (MAD) and Marginal Abatement Costs (MAC)

An important practical handicap in performing cost-benefit analyses for mitigating damages due to climate change is the likelihood that both the cost and benefit curves may exhibit great uncertainty. Despite the greater uncertainty implied by this figure's shaded area for climate damages than for mitigation costs, consideration of all the factors not included in most conventional mitigation cost estimates (see accompanying text) could well make mitigation cost uncertainties far wider than conventional thinking presupposes.

Source: Munasinghe, M., P. Meier, M. Hoel, S. W. Hong and A. Aaheim, 1996. "Applicability of Techniques of Cost Benefit Analysis to Climate Change," IPCC, 1996c, Social and Economic Dimensions, chapter 5, Cambridge Univ. Press, UK.

Also reproduced in C. J. Jepma and M. Munasinghe, 1997. Climate Change Policy, Cambridge Univ. Press, Cambridge, UK, chapter 3.

Globally, annual impacts under 2xCO2 climate were estimated at 1-2 percent of GDP; regionally, the annual impacts on GDP were estimated to range from slightly positive, to mildly negative, to as much as negative 10 percent (for Africa).

One of the important estimates to emerge from Working Group III is a figure (Figure 2.1) that contains a relatively wide range of climate damage estimates as a function of degree of carbon dioxide abatement. Plotted on the same axes is a narrower range of estimates of the costs to the world economy of emissions abatement (e. g. , carbon taxes). With regard to the relatively tight uncertainty range that Working Group III experts offered for costs of mitigation, one of the criticisms of economic analyses of such costs (e. g., Grubb et al., 1994, Goulder and Schneider 1997) is that the range of uncertainty associated with these estimates should probably be broader, reflecting uncertainties in the studies, including assumptions about the efficiency of the baseline scenario, rates of economic growth, the existence of market failures which may limit diffusion of efficient or low-emissions technology, trade patterns, energy prices, and resource availability. Another factor which has not received adequate attention is the issue of induced technological change (ITC). That is, increased prices of conventional energy would stimulate research, development and deployment of alternative energy technologies, as well as improved efficiency of all energy use, which would over time reduce costs of achieving certain abatement targets relative to those calculated in the absence of such induced technological change (see discussion in Schneider, 1997). Had such issues and uncertainties been more firmly debated and more formal procedures applied to estimate outliers, it is likely that the range given for mitigation costs would have been wider.

Application of Techniques for More Consistent Quantitative Subjective Estimation of Outliers

In the above brief set of examples from the three IPCC Working Groups, the outlier ranges which are used to characterize the level of uncertainty associated with a particular estimate often imply potential risks or benefits with costs or values very different from best guess or central tendency estimates. Thus the values assigned to outlier outcomes are important to policymakers in that they can be taken as an indicator of the sorts of low-probability but potentially high consequence risks that are associated with different outcomes (the basis for most insurance activities). The examples outlined in Cases 1 -6 indicate that, while IPCC ranges are qualitatively reasonable and have been developed through open discussions of authors and reviewers, more attention needs to be paid to the process of developing more consistent and methodical evaluations of outlier outcomes with an explicit statement of the subjective nature of the probabilities attached. We next turn our attention to the question of what approaches might be available to develop more consistent and clear estimates of ranges and levels of certainty, and to what extent any of these techniques would be potentially applicable to the IPCC process.

While IPCC ranges are qualitatively reasonable and have been developed through open discussions of authors and reviewers, more attention needs to be paid to the process of developing more consistent and methodical evaluations of outlier outcomes.

While IPCC ranges are qualitatively reasonable and have been developed through open discussions of authors and reviewers, more attention needs to be paid to the process of developing more consistent and methodical evaluations of outlier outcomes.

Eliciting Expert Opinion

There is a rich literature in the decision-analytic and social -psychological disciplines of methods for eliciting subjective opinions of expert groups. To gain some insight into their potential usefulness in the IPCC, we will provide a few examples from the area of climate analysis and comment on their strengths and weaknesses. To our knowledge, there are four formal elicitations published in the area of climate change, and one of us (SHS) has been a participant in all four (NDU, 1978; Morgan and Keith 1995; Nordhaus, 1994; Titus and Narayanan, 1995).

The National Defense University's elicitation

The first took place in 1978, when the National Defense University undertook a survey requesting various scientists to present their opinions as to whether global temperatures would increase or decrease. Given the controversy at the time, the heterogeneous backgrounds of the experts selected and the ambiguity of the questions asked, it is not surprising that the opinions varied widely. The NDU study authors averaged the opinions from the participating experts, some of which leaned toward warming and some toward cooling, producing very low aggregate estimates of future climate change projections a point that was criticized later (Schneider, 1985, and Stewart and Glantz, 1985). For example, Schneider (1985) argued that it is the standard deviation or spread of the experts' opinions that has most policy relevance, since that provides a subjective metric of how different the climate of the near future could be from the current (based on the premise that difference from present is likely to cause impacts). Moreover, he questioned whether all experts surveyed were comparably knowledgeable about the underlying science of climate projections. This leads to the very controversial (and still relevant) question of devising techniques which are appropriate for aggregation of opinions from experts with heterogeneous skills, a matter we discuss below.

The Morgan-Keith study

The second study, carried out by Granger Morgan and David Keith, from Carnegie-Mellon University (CMU), was very different from the NDU study. In this case, the authors first visited each of the respondent scientists, provided in advance a background paper on climate change to each scientist, and requested comments and corrections from those scientists on the concepts of that paper. Morgan and Keith then revised their questionnaire to be sure that the content about which expert opinion was elicited would be meaningful scientifically. The authors then personally visited each of the respondents, asking each one to draw cumulative density functions displaying their subjective probabilities for a number of outcomes in the global warming debate. We concentrate here on the respondents' estimates for the climate sensitivity parameter, T2x (see Figure 2.2). These tentative initial cumulative probability functions were then discussed by each respondent and the visiting study authors, with the latter pointing out to the former precisely what the hand-drawn sketches translated into: that is, the respondent implied an x percent probability that climate sensitivity would be less than 1°C, y percent probability it was less than 4°C, and z percent probability it was less than 10°C. This feedback helped to ensure that each scientist's sketch in fact conformed with his own internal beliefs. Not only did this feedback help respondents to represent their own thinking more consistently, but was the opening exercise of a long interview.

Through this interactive process, scientists were able to rethink, reexamine and arrive at a refined estimate of their own true subjective opinions of probability in formal terms.

Later in the interview, the CMU team asked each scientist to specify the components of the climate problem that they believe created the climate sensitivity and led to its uncertainty distribution. Problems such as cloud feedbacks, ocean fluxes, etc., were addressed by respondents. The study authors then checked for consistency between the range of uncertainty elicited from the scientists on the outcome (i. e., T2x on Figure 2.2), and the degree of uncertainty associated with processes such as cloud feedbacks. This helped the CMU authors determine the relative internal consistency of each scientist by comparing their reasons for uncertainty in the outcome and their actual subjective probability functions as displayed in Figure 2.2. These results were then returned to the scientists, who, witnessing their own possible inconsistencies, had the opportunity, if they wished, to redraw their cumulative probability functions on outcomes such as climate sensitivity. Through this interactive process, scientists were able to rethink, reexamine and arrive at a refined estimate of their own true subjective opinions of probability in formal terms.

While it is true that the elicitation process took four hours of interview time on each occasion, and subsequent revisits to the task, the product of this study is a remarkably consistent set of cumulative probability functions (see Figure 2.2). Interestingly, although most scientists interviewed had largely overlapping ranges of subjective probability, with means and medians for T2x close to the conventional wisdom of 2-3°C (also the best guess of the IPCC) nearly all sixteen analysts also allowed somewhere between a 5 and 20 percent chance of small climate change (less than 1°C). Moreover, most of the sixteen allowed a significant (greater than 10 percent) chance of large climate change (greater than 4°C). The two columns on the right of Figure 2.2 show the mean and standard deviation of the cumulative density functions of each scientist.

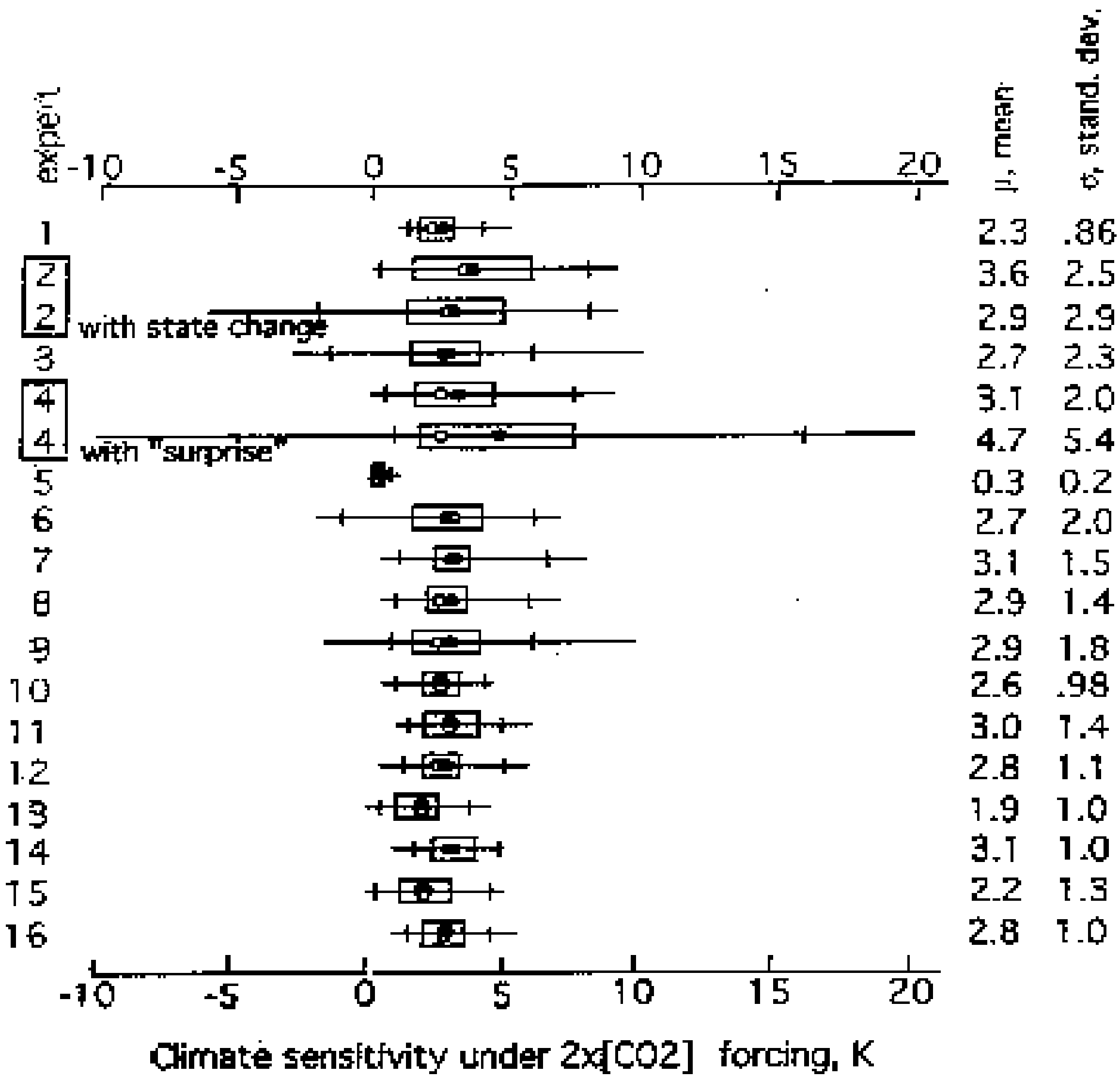


Figure 2.2
Climate Sensitivity under 2xCO2 forcing, °K

Box plots of elicited probability distributions of the change in global average surface temperature resulting from a doubling of CO2. The horizontal lines denote the range from minimum to maximum assessed possible values. Vertical marks indicate the locations of the lower 5th and upper 95th percentiles. The boxes indicate the interval spanned by the 50% confidence interval. The solid dots indicate the mean and open dots, the median. Source: Morgan and Keith, 1995.

One scientist, number 5 on Figure 2.2, had a radically different distribution from the rest of the group. Scientist 5 essentially viewed climate sensitivity via a different paradigm than the others. Granger Morgan has argued that it is inappropriate to aggregate the opinions of experts into a group mean and standard deviation when experts appear to hold to different paradigms. For example, had half the experts agreed with Scientist 5 that climate sensitivity is negligible, while the other half had estimates that lay in the center of the standard range, aggregating the two would create a bimodal distribution. This distribution would have peaks around 0.5°C and

2.5°C, but its mean would be about 1.5°C. Because the mean of a bimodal distribution will often be a value that is itself unlikely, the net effect of aggregating in this case would place the group average sensitivity in between the two paradigms where neither group believes the best guess belongs. Under such conditions, Morgan and Henryon (1990) have argued that it is best to simply exhibit the two groups of elicitations in separate blocks. There is, therefore, no group average given to accompany Figure 2.2. Given the summary nature by which aggregate distributions in that study were presented, aggregating the distribution would have been misleading to some readers. Aggregating may be appropriate in other studies where readers are unlikely to be misled, such as studies that present probability density functions where the bimodal nature of the results would be evident, or Monte Carlo modeling assessments where the model employed can make efficient use of the input information even if it includes bimodal distributions.

Because the mean of a bimodal distribution will often be a value that is itself unlikely, the net effect of aggregating in this case would place the group average sensitivity in between the two paradigms where neither group believes the best guess belongs.

The Titus-Narayanan EPA sea level rise study

The third study in the case of global climate change was undertaken at the Environmental Protection Agency (EPA) by Jim Titus and Vijay Narayanan (1995). This study was explicitly designed to improve upon the existing IPCC integrated assessment framework. The authors started with the set of models that IPCC uses to project emissions, concentrations, transient temperature change, and sea level rise resulting from thermal expansion and glacial melting. IPCC's projections require different groups of experts to specify low, medium, and high assumptions for various key processes (a wide variety of processes involving many different types of expertise) and the end results are the official IPCC low, best estimate, and high projections of future temperatures and sea level. Titus and Narayanan instead sought to estimate a probability distribution for temperature and sea level rise by soliciting different groups of experts to specify probability distributions for key coefficients and the likelihood of particular models.

Like the NDU, CMU, and Nordhaus (1994) studies, Titus and Narayanan also included an effort to elicit subjective cumulative probability distribution functions of subjective probability from various experts, but in this case the focus was on obtaining probability distributions of model parameters. Only one of the parameters dealt with outcomes (i. e., T2x) while the others dealt with values of partly known (but critical) parameters used in climate systems models (e. g., mixing rates between upper and deep oceans, polar amplification of warming, etc.), rather than climate change outcomes. (The distinction between outcome and parameter in a model is relative. For example, the T2x parameter is the outcome of an equilibrium run of a 3-D GCM, but it is a parameter in the 1-D ocean models used by IPCC and EPA to project temperatures and sea level.)

This study not only solicited distributions for parameter ranges from climate modelers, but also treated all aspects of the problem with distribution functions. Thus, the authors began with a subjective statistical distribution of carbon emissions, fed those emissions into a carbon cycle model, fed those results into the climate modeling section just described, then fed those results into ocean mixing models, and fed those results into glacier melt models to calculate eventual probabilities of sea level rise. At each step, subjective probability distributions for key parameters were fed into the models; in some cases, reviewers also attributed probabilities to alternative models. For example, while the IPCC assessment did not include any models that allowed for a disintegration of the West Antarctic Ice Sheet, the EPA authors included some models that did so. In a few cases, reviewers even suggested models that had not been within the original model framework. Monte Carlo techniques were used to generate final probability distributions.

Titus and Narayanan only sought opinions on processes for which individual scientists were recognized experts, and allowed anyone to opt out of specifying a probability distribution for any particular parameter where the scientist lacked expertise. The reviewers fell broadly into three groups: climate and ocean modeling, polar amplification, and glacial modeling, but not every expert expressed an opinion on every parameter related to his or her field. The rationale for this approach was (1) not to ask reviewers to estimate parameters on which they have little expertise, and (2) this approach is consistent with the IPCC assessment, which had different committees responsible for emissions, concentrations, ocean modeling, and glacial processes. The questions posed to the experts concentrated on estimating the values and uncertainties in the process parameters, letting the models convert these uncertainties into outcome probability distributions for global temperatures and sea level rise.

Like the Morgan and Keith study, there was one scientist whose estimate of climate sensitivity was an order of magnitude less than the eight other scientists involved in the EPA climate model elicitation. There was also one glaciologist whose suggested parameters and model probabilities implied a 5 percent chance of a major Antarctic contribution to sea level rise while the other glaciologists provided parameters suggesting the probability to be only 0.5 percent. In both cases, the EPA authors weighted all of the scientists' opinions equally in calculating the combined distribution, to ensure that the one "contrarian" scientist's response, which resulted from his fundamentally different view of climate sensitivity, was given full effect. These aggregations were criticized in editorial comments by Keith, 1996, on the grounds that this is an inappropriate procedure when an outlier scientist clearly adheres to a different paradigm than the other scientists interviewed. Titus and Narayanan considered arguments against this approach, but, they argue, to simply ignore, for example, the outlier glaciologist's opinion because it is a minority view would be to disregard the possible risk of a large rise in sea level from rapid Antarctic disintegration. They also noted that, unlike solicitations that focus on outcomes, a modeling study in which different experts estimate different parameters must devise a means of combining opinions, and that Keith's criticism applied to all aggregation procedures. (For further discussion, see Titus, in this report.)

Titus and Narayanan sought to estimate a probability distribution for temperature and sea level rise by soliciting different groups of experts to specify probability distributions for key coefficients and the likelihood of particular models.

Despite some debate over methodologies, this pioneering study did produce a wide distribution of probabilities for the outcome of concern (sea level rise), ranging from a small probability of no rise, to a small probability of greater than a meter rise by the end of the next century. The overall shape looks somewhat like a normal distribution except it is skewed to the right allowing a small chance of a very large rise in sea level, and, as such, it encompasses a somewhat wider range than that resulting from the outlier ranges chosen by much less formal means by IPCC 1996a.

The Nordhaus study

In order to produce integrated assessments or cost-benefit analyses, attempts are made to weigh the impacts of anthropogenic climate change on environment and society (i. e ., climate damage) against the costs of mitigation activity, such as carbon taxes. In order to perform such optimal cost-benefit analyses, scientists need probability distributions of both costs and benefits. Nordhaus (1992) provided such a calculation based upon an assumed cost to the world economy of about 1 percent of GDP from a CO2 doubling. This climate damage estimate was criticized as arbitrary and too small by a number of scientists, which led Nordhaus (1994, the fourth study considered here) to perform an elicitation of the opinions of 19 economists and other scientists regarding the damage to the world economy from several scenarios of climate change: (a) a warming of 3°C by 2090, (b) a warming of 6°C by 2175, and (c) a more rapid warming of

6°C by 2090. The probability distribution elicited for each scenario required the respondents to provide 10 percent probability, median, and 90 percent probability estimates for the likely damage to the economy from these climate scenarios. Figure 2.3 shows this result for the 6°C warming by 2090. Nordhaus performed his elicitation by a combination of formal written questionnaires and follow-up telephone interviews. Like Morgan and Keith, he gave authors a chance to reconsider their own judgments based upon the elicitations from fellow respondents.

Nordhaus performed his elicitation by formal written questionnaires and follow-up telephone interviews. Like Morgan and Keith, he gave authors a chance to reconsider their own judgments based upon the elicitations from fellow respondents.

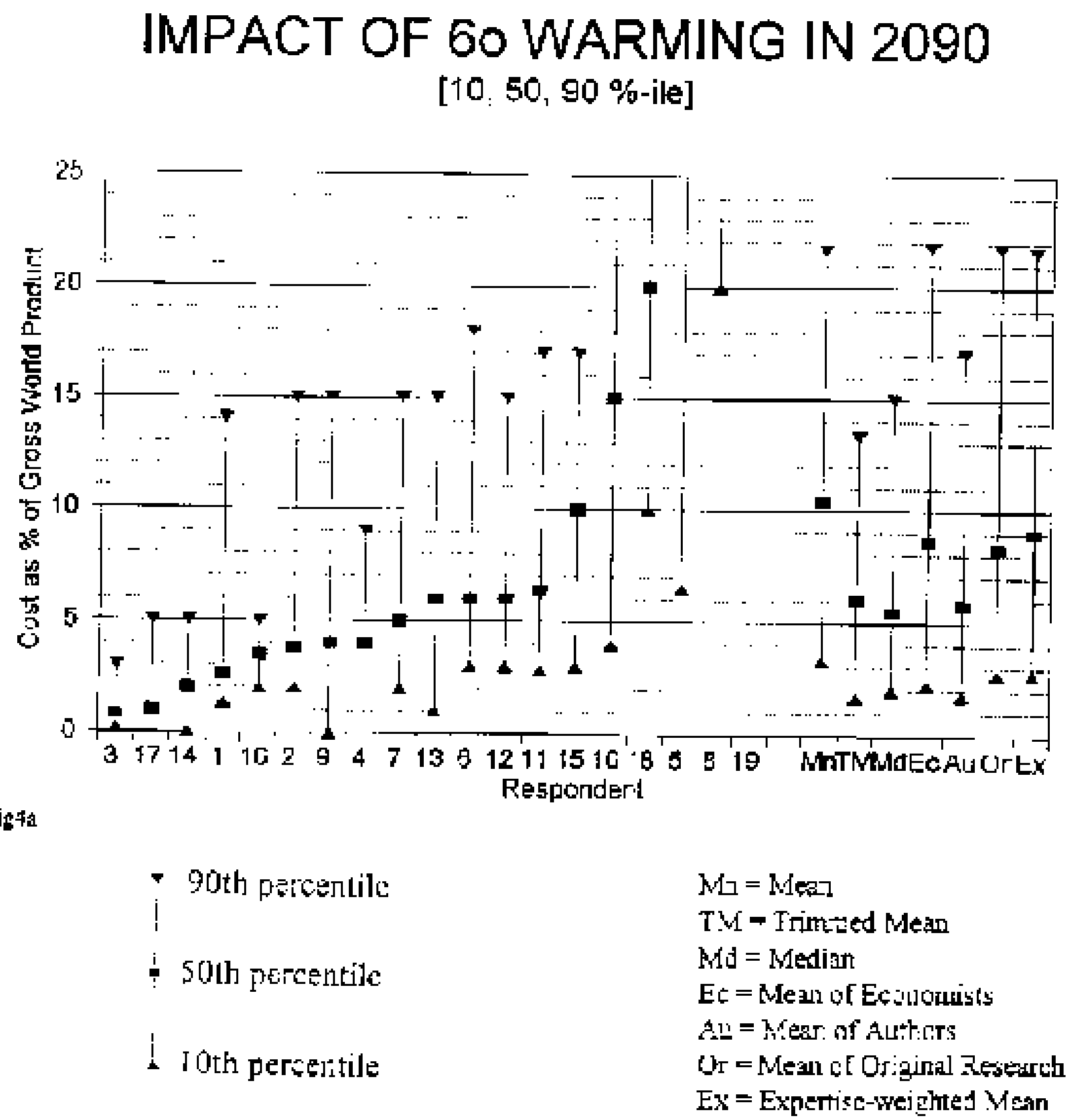


Figure 2.3
Impact of 6° Warming in 2090 (10, 50 and 90 percentile)

Opinions of various experts on the impact of a 6°C warming by the year 2090 as a percentage loss of gross world product. Each respondent's best guess of impacts is shown as the 50th percentile. Source: Nordhaus, W. D., 1993, Survey on Uncertainties Associated With Climate Change, Yale University, New Haven, Connecticut.

Two features are apparent in Nordhaus' results. First, in addition to very widespread ranges, in general, there are clusters near the ends of the range, at low and high damage levels. Nordhaus classified these according to the profession of the respondents, concluding that mainstream economists are a factor of twenty less worried than natural scientists, even about a 6°C warming in a century. (At a meeting where the results were discussed, Nordhaus quipped that those who know the most about the economy are least worried, to which Schneider counter-quipped that those who know the most about nature are most worried (e. g., Schneider, 1997b, chapter 6)). Differences in training, world views, and the relative amount of damages attributed to non - market sectors (e. g., biodiversity loss) contributed to these paradigm differences among groups of different professionals.

Interestingly, despite the order of magnitude difference in their absolute GDP damage estimates, which correlated with disciplinary backgrounds, both natural scientists and economists drew non-linear damage functions. That is, both felt that the 6°C warming to 2090 AD would have damages more than twice the 3°C warming to 2090 case. It is interesting to contrast this with an elicitation by Berk and Schulman, 1995, of citizens in California's San Fernando Valley, who were asked how much they were willing to pay to avoid climate changes that were 5, 10, 15 or 20°F warmer or colder than the summertime averages in that location. These citizens were willing to pay more for larger deviations from the mean, but achieved a diminishing return in their willingness to pay, relative to the elicitation of the experts. Few said they would be willing to pay as much as $500 to prevent a climate like that of Death Valley (mean temperatures of 120°F) from occurring in the Los Angeles Basin! This is in sharp distinction to the non-linear (more like temperature difference squared; see Roughgarden and Schneider, 1997) views of the experts, whether economists or natural scientists. Moreover, Roughgarden and Schneider's analysis of Nordhaus' survey data reveals that there was a positive correlation between respondents who believed there would be large climate damages and the proportion of damages assigned to non-market sectors. Clearly, such formal techniques are useful for revealing different categories of subjective probability among various groups, but it remains unclear how to develop a widely accepted conclusion.

Nordhaus concluded that mainstream economists are a factor of twenty less worried than natural scientists, even about a 6°C warming in a century.

Elicitation of expert judgments from other fields

The AGCI summer session also considered approaches not yet tried in the climate-change arena, including an additional kind of formal procedure for evaluating and aggregating subjective expert opinion, based on experience with earthquake hazard assessments. Probabilistic seismic hazard analysis is a methodology for estimating the likelihood that various levels of earthquake-caused ground motions will be exceeded at a given location in a given future time period. Due to large uncertainties in the geosciences data and in their modeling, multiple models and interpretations are the norm, leading to disagreements among the experts. A process of aggregation of expert opinions, both social and scientific, was designed by the Senior Seismic Hazard Analysis Committee (SSHAC) funded by the Department of Energy, the Nuclear Regulatory Commission and the Electric Power Research Institute and chaired by Robert Budnitz. It has been reviewed by a National Research Council committee (NRC, 1997). The objective is to obtain a single mean estimate of the annual frequency of exceedance versus peak ground acceleration at a given site, and composite uncertainty bands reflecting the informed scientific community's current uncertainty in that estimate. The process represents an alternative way of dealing with the problem of aggregating expert opinion in a particular scientific area for purposes of policy making. It has been applied recently to other subjects such as volcanic risks. (This process is discussed by Cornell, in this report.)

The first stage of the SSHAC process involves discussion by a large group of experts

that reviews and weights different approaches to modeling, dropping some approaches as demonstrably wrong or as not worth pursuing further in depth and promoting others. The second stage involves the careful selection of experts for the remainder of the process who are asked to wear different hats, some as proponents of particular models or approaches and others as evaluators, who play the major role in the process by evaluating the form of uncertainty distributions on parameter values and/or alternative models. The ultimate responsibility for the success of this process lies with the Technical Facilitator/Integrator (TFI). The TFI is a single entity, sometimes an individual, but preferably a small team that is responsible for identifying the key issues and components of the analysis, structuring and directing the interaction and debates among the proponents and evaluators, conducting any necessary numerical analyses, and documenting the process followed and the results obtained.

The objective of the process is to allow researchers to be in a position to give an aggregate, composite range of the group's uncertainty.

The SSHAC process provides for clarification of objectives, consensus, and the roles of experts. It also involves numerical or mechanical aggregation schemes, and social integration procedures. The objective of the process is to allow researchers to be in a position to give an aggregate, composite range of the group's uncertainty. The process involves very strong interaction, with the TFI like a panel of court judges (see original proposal for a science court by Kantrowitz, 1967 and 1976) holding the final responsibility for defending the result as an accurate characterization of the group's collective judgment. The cost of conducting this process well for any given issue may be up to $1 million.

Despite the cost obstacle, some argued that an independent panel might be more credible in the policy community than experts who might be perceived as having vested interests in any aspect of their respective elicitations. On the other hand, AGCI participants recognized that such vested experts may be uniquely qualified to estimate the subjective probabilities of outcomes they specialize in studying relative to more "neutral" judges who are less vested, but also less knowledgeable. Of course, large committees of experts, acting as both assessment witnesses and peer reviewers, can minimize the likelihood of vested interest bias progressing very deeply into the aggregated assessments. This is another area in which different degrees of peer trust among professional communities (i. e., scientific versus legal versus political) may strongly influence what each community believes to be appropriate assessment process design (see discussion in Edwards and Schneider, 1997).

Options for Improvements

We considered three basic steps that assessment bodies such as the IPCC should consider taking in order to achieve more consistent and clear subjective opinions of participating experts.

1 Improving consistency in estimating ranges and levels of confidence

Most straight forward would be the presentation of consistency standards that are qualitative, but could nonetheless provide more systematic definitions of levels of confidence based on degrees of evidence and consensus (Table 2.4, Moss 1996 and Table 2.5, Trenberth 1996, see also National Academy of Sciences report, 1987).

At a minimum, employing such consistency tables would force participants to think more carefully and consistently about their subjective probabilities, and help to translate words like high, medium, and low confidence into reasonably comparable probability estimates. This step would be relatively straightforward to implement, and could improve the consistency of the subjective estimates in future assessments.

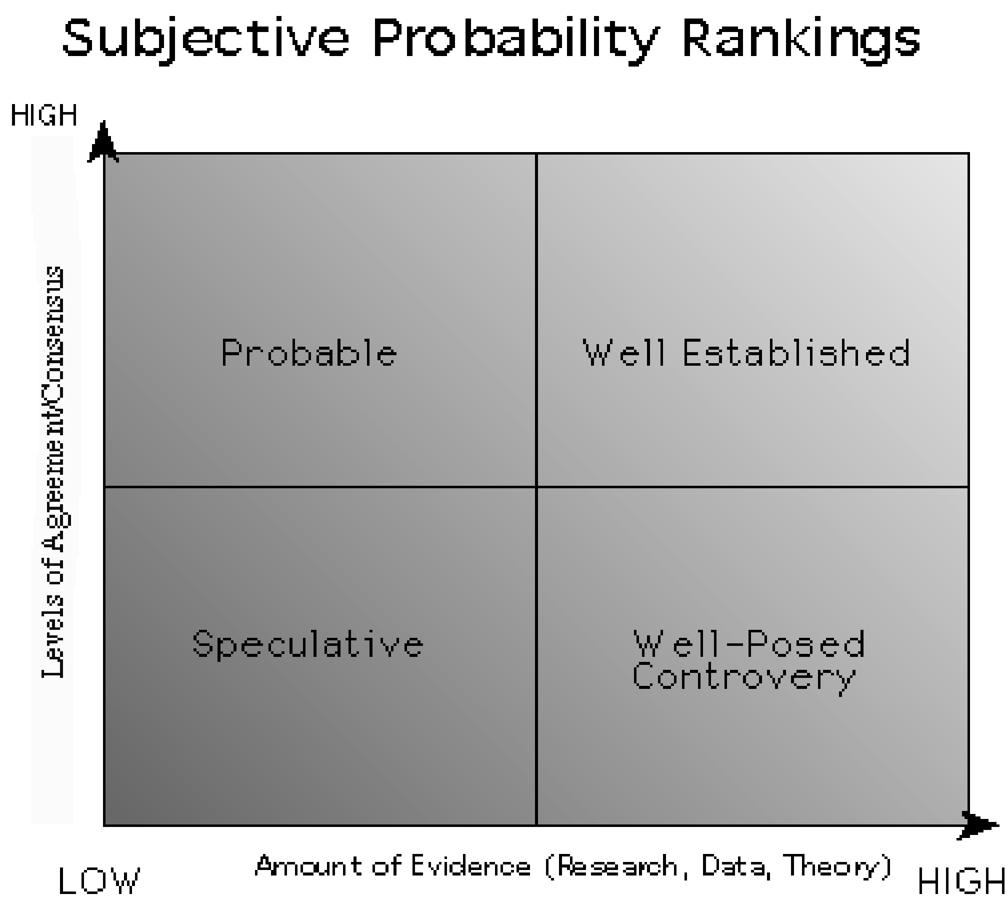


Table 2.4

Moss, 1996, Subjective Probability Rankings

Well-Established This category denotes wide agreement, based on multiple findings through multiple lines of investigation. A finding could be removed from this category not by a single hypothesis, observation or contention, but only by a plausible alternative hypothesis, based on empirical evidence or explicit theory, and accepted by a substantial group.

Well-Posed Controversy A well-established finding becomes a well -posed controversy when there are serious competing hypotheses, each with good evidence and a number of adherents.

Probable This category indicates that there is a consensus, but not one that has survived serious counter-attack by other views or serious efforts to "confirm" by independent evidence.

Speculative Speculative indicates not so much "controversy" as the accumulation of conceptually plausible ideas that haven't received serious attention or attracted either serious support or serious opposition.

Table 2.5
Trenberth, 1996, Common Approaches to Ranking Levels of Confidence

Given a cumulative probability distribution, there should be a standard terminology adopted to translate that into plain English for communication with policymakers, media and the public.

Percent		Confidence	Confidence Bands		
99 %		Extremely High			Extremely Confident (1 %–99 %)
95 %	+2	Very High		Very Confident (5 %–95 %)	
75 %	Inter Quartile Range (25 %–75 %)	High	Confident (Low–High)		
50 %		Medium			
25 %		Low			
5 %	-2	Very Low			
1 %		Extremely Low			

Other terms
Certain
Highly probable
Probable
Likely

The problem is then to assign a probability distribution (may be a band rather than a line, see figure below). Generally this will be empirical or subjective in origin, even if based in part upon objective procedures.

Problem

Divide into	Random component (cf. random errors)
	Systematic component (bias)
Beware	Dependencies (correlations) among ingredients
	Conditional probabilities
Issues	Confidence in assumptions
	Confidence in model, given assumptions
	Confidence in parts of model
	Top down (model validation)
	Bottom up (components uncertainties)
	Relevance for particular question

2 Elicitation of subjective probability distributions

The second technique, following Morgan and Keith, would frame a number of explicit questions in which ranges of estimates of outcomes or parameters would be called for from authors or other participants. Formal one-on-one interviews, group interviews or mail-in questionnaires (perhaps with follow-up contacts) could be used to establish subjective probability distributions for a few key parameters or issues. In addition to soliciting outcomes (e. g., estimates of climate sensitivity), processes and major uncertainties in data or theory could also be elicited (e. g., Titus and Narayanan, 1996). Furthermore, questions could be designed to improve the consistency between outcome estimates and process uncertainty estimates made by each individual or group.

These probability distributions would not necessarily be published as part of the final report, but could be used in the process of preparing and revising drafts at successive author meetings. Examination of these different formal estimates of uncertainty by each individual, particularly in a group setting, could well help to sharpen the intuitive scientific judgments and not only add consistency, but enhance the quality and quantitative reliability of the range limits provided by the assessors. Not only might a single elicitation activity be undertaken, but successive meetings of the working groups could each devote a short period in which re-elicitation is attempted, based upon what each participant learned from the previous elicitation and the research that they each undertook in-between meetings, partly stimulated by the differences in each respondent's elicitation and those of his/her colleagues. Such differences are likely to cause discussions to be sharpened and focused on points of uncertainty among the participants. The privacy and formality of such procedures would also help to overcome the difficulty of dominant personalities having undue influence in small groups of people from different cultures and personalities.

Examining these different estimates of uncertainty by each individual could help to sharpen the intuitive scientific judgments and not only add consistency, but enhance the quality and reliability of the range limits provided by assessors.

3 Assessment Court

Finally, an assessment court, or panel of independent judges, could be selected to listen to the debates, to review the elicitations of authors or other experts, and to provide an independent evaluation of the experts' assessment. While such procedures are more formal and cumbersome (as well as costly) than the previous two, if they resulted in higher credibility for the assessment report or more consistency in the estimation of outlier events whose damages could be nonlinearly more significant than smaller outlier ranges, then the extra effort might be worth the trouble. Even if the formal process were not applied, sensitizing authors to the various roles identified (proponents, evaluators, etc.) could help establish more systematic guidelines for participating in assessments.

What are the pros and cons of these three techniques? Under which circumstances could they be applied? No universal recommendations can be developed, as the purposes and context of each assessment varies, and techniques appropriate to one context may not be appropriate to others. However, the clarification of language, as in Tables 2.4 and 2.5, would minimize the likelihood that different participants meant different quantitative probabilities while using the same words. That strikes us as an absolute minimum requirement for future assessments.

Formal elicitation would have several benefits. It would force people to confront their doubts and reexamine their relative uncertainties. It would provide assessors with a preliminary estimate of comparability to their colleagues thinking in quantitative terms, which, in turn, might lead to debate that could result in different (presumably better) subjective estimates than would take place in the absence of intercomparison of various authors' formal elicitations. This formal procedure certainly would help to get the opinions of more reserved participants counted equally with more extroverted ones.

The assessment court procedure would bring to the process the ostensible credibility of an independent group, eliminating the frequent criticism that the recommendations from the assessment scientists can be (even if unconsciously) motivated by self-interest, since the scientists indeed are interested parties in performing the scientific research they recommend as needed.

What are the drawbacks of such formal techniques? With regard to more systematic definitions of uncertainty via qualitative categories (e. g., Tables 2.4 and 2.5), we cannot think of any. It seems essential, we have already argued, that this be a minimum next step. With regard to formal elicitations, a number of possible drawbacks come to mind. First, for those who do not understand the nature of subjective probability, it could convey a false sense of analytic precision, at least for a reader not recognizing that the probabilities, though quantitative, are still highly subjective. Of course this objection could be mitigated by careful writing and clear explanation of the activity.

For the procedure to be most effective, it needs to operate in the learning mode, that is, learning by inter-comparing various authors' elicitations both at initial and follow-up meeting

The second difficulty is that for the procedure to be most effective, it needs to operate in the learning mode, that is, learning by inter-comparing various authors' elicitations both at initial and follow-up meetings. This puts a premium on having the same sets of authors attending most of the meetings, so that the learning process can lead to evolving and more consistent subjective probability estimates. Another potential drawback is that peer or political pressures could induce people whose elicitations might be publicly displayed to produce cumulative density functions that do not reflect their true beliefs (this objection could also be applied to the consistency table approach to assessment discussed above). However, it seems easy enough to overcome any potential problem by assigning a number to the elicitation of each respondent and displaying this number rather than the name of the scientist who offered each response. Thus, the responses would be known only to the individual scientists, and would be revealed only if a scientist chose to identify him or herself with a particular number. Another potential drawback is the (remote) possibility for political trade -offs among delegates, but once again this danger also exists in current qualitative assessments, and is best avoided by open meetings and extensive peer review practices already in place.

A difficult potential drawback is that not all authors bring the same knowledge base to bear in assessment workshops, and thus the appropriate relative weights in the aggregation of authors' opinions may not be equal. It is very divisive to try to determine whose opinion is more significant than another's, and this weighting process needed for aggregation could lead to significant internal dissension in the assessment team. However, an anonymous system of elicitation display where only the scientist knows which chart is his or hers, and vigorous discussion among participants willing to either admit their views or defend certain elicitations of others, could lead to substantial and rapid education among those less well acquainted with specific topics, thereby improving the overall quality of the final elicitations that result at the end of the iterative process. Finally, rather than averaging all results, they can be either aggregated into paradigmatically similar groups or displayed in full without averaging.

One drawback of the assessment court approach is that a search for "neutral" experts outside of the climate community might not produce higher quality opinions, and might require that group to meet for a long time to get up to speed on the expert dimensions of the very complex sets of climate issues, thereby invalidating the extra credibility that would ostensibly derive from their alleged unbiased (even if less knowledgeable) presence. However, on politically sensitive items (such as biodiversity protection or the value of a statistical human life used in cost/benefit analyses), limited use of such independent assessment teams may, nonetheless, be worth considering.

Another major drawback of this approach is the fact that court systems in general tend to create advocates who typically ignore or denigrate data and evidence that do not necessarily support their theses. In other terms, whereas the display of expert opinions (via distributions and/or aggregations) should make use of all available information, court systems tend to create incentives for the opposing sides to either suppress or exaggerate the flaws of evidence that does not support their thesis. This may result in either an artificial level of confusion or a substantial truncation of data that lie in the middle and do not support clearly extreme positions (for more, see Paté -Cornell, 1996). Such risks can be mitigated by proper training and control by the assessment judges, but it is still not clear whether scientific assessments in general should move closer toward (or even further away from) a legal model in which opposing advocates often operate on an "it's not my job to make my opponent's case" epistemology. IPCC has encouraged its scientific assessment teams to consider the full range of evidence, not to serve as advocates.

The Media and Political Context

IPCC has encouraged its scientific assessment teams to consider the full range of evidence, not to serve as advocates.

When scientists' debates regarding uncertainties and levels of confidence in findings move into the media and political contexts, many more variables come into play. From the difficulties journalists have in understanding and communicating these issues to the public, to the use and misuse of scientific uncertainties by interest groups, an already confusing set of issues is further confounded. To use one example, some media accounts portrayed the fact that the SAR's 1995 estimate of mean annual surface temperature increase for 2100 A. D. was about 1°C lower than the temperature estimate in the 1990 report as an indication that climate was much less sensitive to GHG emissions than previously thought. However, this change in the projection was due to the inclusion of the potential effects of sulfate aerosols in the models, not to a change in the sensitivity of climate to forcing, the estimate of which remained constant (see Case 1 above). While scientists understand and explained this distinction, it is often read by much of the public to mean that the problem is not as bad as was thought a serious misunderstanding of the findings, since (a) aerosols exhibit a local and short-lived cooling effect which alters regional patterns of climate change, and (b) projections of future aerosol concentrations have a very large uncertainty.

Another example is the manner in which the detection and attribution of climate change, discussed in Case 3 above, was portrayed in the media. The IPCC chapter on this subject is one of the most statistically rigorous in the entire assessment. Its conclusions are based on multiple supporting lines of evidence. Although most of the debate within the scientific community centered on narrow concepts of statistical significance, the debate was characterized in the media and political contexts as being much broader (Edwards and Schneider, 1997). In both of these examples, much of the media portrayed what were actually debates about narrow scientific concepts or minor procedural disputes as though they involved more extensive uncertainties or were based on political advocacy.

Conclusion

Assessment is not focused on original scientific results, but rather on evaluation of existing science, to address policy-related issues and to provide perspectives on the relative parts of the problem that are well understood and to distinguish these from more speculative aspects. Given the nature of the issue, climate change assessments need to encompass end-to-end coverage of a broad spectrum of issues including emissions of greenhouse gases, changes in radiative forcing, implications for global and regional climate impacts, and options for reducing emissions in energy sectors and land management. As a consequence, such assessments need to be multidisciplinary, which requires a diverse set of researchers and experts to be involved in preparing and reviewing the documents. Communication across such a broad range of disciplines about the joint probabilities of related events is a challenge; the strength of assessments can be improved by more systematic attention to the terms and approaches used in assessing and describing the degree of certainty associated with findings.

The strength of assessments can be improved by more systematic attention to the terms and approaches used in assessing and describing the degree of certainty associated with findings.

Assessments need to make estimates of the likelihood that (and time frame over which) new scientific knowledge can resolve existing uncertainties and assess the impacts of changes to physical, biological and societal conditions. They also need to assess the potential costs of those impacts relative to the potential costs of mitigation activities that would reduce the environmental changes that create the impacts (the so-called optimization process). The assessment process and its interaction with the policy community is a cyclical process over time. Since any one assessment will necessarily be incomplete in regard to all of the needed information, it is clear that rolling reassessments will be required every so often (thought to be about five year intervals by the IPCC), so that governments and other decision making groups can reassess the degree to which mitigation, adaptation, or other responses should be stepped up or delayed.

Since reassessment is needed for adaptive management, it is critical that each succeeding assessment refine the way it treats and communicates subjective estimates of uncertainty and the ranges of estimates that are provided. The proposals developed during the AGCI summer session provide useful ideas for the next step in the process of reassessing the potential implications of climate change.

Scientific uncertainty, whether about climate change policy or many other socio-technical issues, is not a problem which can be “solved.” Rather, it is an ongoing challenge and an increasingly important part of the context in which policy decisions will be made. With negotiations of the Framework Convention on Climate Change progressing, it is clear that the world will not wait for virtual certainty before making policy. It is also likely that some interests will demand precisely such certainty before agreeing to implement policies that they perceive negatively affect their interests. In this context, improving the characterization and communication of uncertainty is invaluable. The AGCI sessions’ contribution, in addition to specific recommendations contained in this report, also stems from improving understanding and dialog among disciplines from the natural and social sciences, the field of decision analysis, and the media and policy arenas.

Technical specialists should not assume they will be understood by nonspecialists unless considerable effort is invested in separating out what is speculative from what is likely in a clear and consistent manner. To do less is to invite confusion and contention that will delay the process of determining how societies intend to address the risks of climate change.

References

Berk, R. A., and Schulman, D., 1995, Public Perceptions of Global Warming, Climatic Change, 29(1):1-33.

"Climate Debate Must Not Overheat," Editorial, Nature, 381(6583): 539, 1996.

"Climate Report 'Subject to Scientific Cleansing'," Nature, 381(6583):546, 1996.

Edwards, P. and Schneider, S. H., 1997, The 1995 IPCC Report: Broad Consensus or "Scientific Cleansing?" (Ecofables/Ecoscience , Vol. 1, No. 1, in press).

Goulder, L. and Schneider, S. H., 1997, Induced Technological Change, Crowding Out, and the Attractiveness of CO2 Emissions Abatement, Resource and Environmental Economics (submitted).

Grubb, M., Duong, M. H. and Chapuis, T., 1994, Optimizing Climate Change Abatement Responses: On Inertia and Induced Technology Development. In Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change. N. Nakicenovic, W. D. Nordhaus, R. Richels, and F. L. Toth, eds. International Institute for Applied Systems Analysis, Laxenburg, Austria. pp. 513-534.

Fankhauser, S., R. S. J. Tol and D. W. Pearce, 1997, Extensions and Alternatives to Climate Change Damage Valuation: On the Critique of IPCC Working Group III's Damage Estimates, submitted to Environment and Development Economics.

IPCC, 1990a: Climate Change: The IPCC Scientific Assessment . Prepared by IPCC Working Group I [J. T. Houghton, G. J. Jenkins, and J. J. Ephraums (eds.)] and WMO/UNEP. Cambridge University Press, Cambridge, UK, 365 pp.

IPCC, 1990b: Climate Change: The IPCC Impacts Assessment . Prepared by IPCC Working Group II [W. J. McG. Tegart, G. W. Sheldon, and D. C. Griffiths (eds.)] and WMO/UNEP. Australian Government Publishing Service, Canberra, Australia, 275 pp.

IPCC, 1991: Climate Change: The IPCC Response Strategies . Prepared by IPCC Working Group III and WMO/UNEP. Island Press, Covelo, California, 273 pp.

IPCC, 1992: Climate Change 1992: The Supplementary Report to The IPCC Scientific Assessment. Prepared by IPCC Working Group I [J. T. Houghton, B. A. Callander, and S. K. Varney (eds.)] and WMO /UNEP. Cambridge University Press, Cambridge, UK, 200 pp.

IPCC, 1994a: Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios. Prepared by Working Groups I and III [J. T. Houghton, L. G. Meira Filho, J. Bruce, Hoesung Lee, B. A. Callander, E. Haites, N. Harris, and K. Maskell (eds.)] and WMO/UNEP. Cambridge University Press, Cambridge, UK, 339 pp.

IPCC, 1994b: IPCC Technical Guidelines for Assessing Climate Change Impacts and Adaptations. Prepared by IPCC Working Group II [T. R. Carter, M. L. Parry, H. Harasaw, and S. Nishioka (eds.)] and WMO/UNEP. University College London, UK, and Center for Global Environmental Research, Japan, 59 pp.

IPCC, 1995: IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by the IPCC, OECD, IEA, and WMO/UNEP. IPCC Working Group I TSU, Bracknell, UK, 3 volumes.

IPCC, 1996: Climate Change 1995: The Science of Climate Change . Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J. J., L. G. Meiro Filho, B. A. Callander, N. Harris, A. Kattenberg and K. Maskell (eds.)]. Cambridge University Press, Cambridge and New York, 584 pp.

IPCC, 1996: Climate Change 1995: Impacts, Adaptations, and Mitigation of Climate Change: Scientific-Technical Analyses . Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change [Watson, R. T., M. C. Zinyowera, and R. H. Moss (eds.)]. Cambridge University Press, Cambridge and New York, 880 pp.

IPCC, 1996: Climate Change 1995: Economic and Social Dimensions of Climate Change. Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change [Bruce, J., Hoesung Lee, and E. Haites (eds.)]. Cambridge University Press, Cambridge and New York, 464 pp.

Kantrowitz, Arthur, 1967, Proposal for an Institution for Scientific Judgement, Science, 156:763-764.

Kantrowitz, Arthur, et al., 1976, Presidential Task Force on the Science Court, The Science Court Experiment: An Interim Report, Science, 193: 653-656.

Keith, D., 1996, When Is It Appropriate to Combine Expert Judgments? An Editorial Essay, Climatic Change, 33(2):139-143.

Morgan, M. G., and Dowlatabadi, H., 1996, Learning from Integrated Assessment of Climate Change, Climatic Change, 34(3-4): 337-368.

Morgan, M. Granger and David Keith, 1995, Subjective Judgments by Climate Experts, Environmental Science and Technology, 29:468-476.

Morgan, M. Granger, and Max Henryon, with Mitchell Small, 1990, Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Policy Analysis, Cambridge University Press, New York (paperback edition 1992).

National Academy of Sciences, 1987, Current Issues in Atmospheric Change, National Academy Press, Washington, D. C., pp. 9-10.

National Defense University, 1978, Climate Change to the Year 2000 , National Defense University, Washington, D. C.

Nordhaus, W. D., 1994, Expert Opinion on Climatic Change, American Scientist, Jan.-Feb. 1994, 82:45-51.

Nordhaus, W. D., 1992, An Optimal Transition Path for Controlling Greenhouse Gases, Science, 258:1315-1319.

Panel on Seismic Hazard Evaluation, Committee on Seismology, National Research Council, 1997, "Review of Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts," National Academy Press, Washington D. C.

Paté-Cornell, M. E., 1996, Uncertainties in Global Climate Change Estimates. An Editorial Essay, Climatic Change, 33(2):145-149.

Root, T. L. and Schneider, S. H., 1993, Can Large-Scale Climatic Models Be Linked with Multi-Scale Ecological Studies? Conservation Biology, 7:256-270.

Roughgarden, T. and Schneider, S. H., 1997, (in preparation).

Schneider, S. H., 1997a. Integrated Assessment Modeling of Global Climate Change: Transparent Rational Tool for Policy Making or Opaque Screen Hiding Value-laden Assumptions? Environmental Modeling and Assessment (submitted).

Schneider, S. H., 1997b, Laboratory Earth: The Planetary Gamble We Can't Afford to Lose. Basic Books, New York, 174 pp.

Schneider, S. H., 1996, The Future of Climate: Potential for Interaction and Surprises. In Climate Change and World Food Security . T. E. Downing, ed. NATO ASI Series: Ser. I, Global Environmental Change, vol. 37, pp. 77-113, Heidelberg: Springer-Verlag.

Schneider, S. H., 1985, Science by Consensus: The Case of the National Defense University Study 'Climate Change to the Year 2000,' An Editorial, Climatic Change, 7(2):153-157.

Singer, S. F., 1996, Climate Change and Consensus, Science, 271(5249):581.

Senior Seismic Hazard Analysis Committee, 1997, "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on the Uncertainty and the Use of Experts", U. S. Nuclear Regulatory Commission, NUREG/CR-6372, Washington, D. C.

Stewart, T. R. and Glantz, M. H., 1985, Expert Judgment and Climate Forecasting: A Methodological Critique of 'Climate Change to the Year 2000,' Climatic Change, 7(2):159-183.

Titus, James G., and Vijay K. Narayanan, 1995, Probability of Sea Level Rise, Environmental Protection Agency Report 230-R-95-008, 186 pages, October.

Titus, James G., and Vijay K. Narayanan, 1996, The Risk of Sea Level Rise, Climatic Change, 33:151-212.

Bayesian Approaches to Characterizing Uncertainty

Richard Berk

University of California, Los Angeles
Los Angeles, California

There are sometimes questions of scientific uncertainty for which no useful statistical tools exist.

Berk began by noting certain limitations to what statisticians have to offer scientists grappling with issues of uncertainty. To begin, there are many "poorly posed" scientific questions that cannot be constructively addressed using statistical techniques. An instance of such a question is "is this model right?" This does not mean that the question is not important but simply that there is a lack of coherence between the tools statisticians bring to bear and what scientists want to know. There are also constraints on when available statistical tools apply. Too often, the scientific data are in a form inconsistent with what the statistical procedures require. One common case is "convenience" samples produced by unknown data generation mechanisms. And there are sometimes questions of scientific uncertainty for which no useful statistical tools exist at all. Finally, even when the scientific question is well posed, and existing statistical tools can be usefully applied, there may be practical constraints, such as insufficient computing power.

Frequentist Approach

Berk then addressed two statistical perspectives on uncertainty: frequentist and Bayesian. The frequentist approach constitutes the dominant paradigm among the fields' elders and is the one scientists most often use, though among statisticians under the age of 40, Bayesian approaches are more common. For the frequentist, probability is defined as the proportion of time something happens in a limitless number of independent, identical trials. Simply put, it is the long run relative frequency. This definition is clearly an abstraction reflecting a "thought experiment" that could never be implemented in practice. Arguably, nevertheless, the frequentist approach is often found to be scientifically useful; it is found to be an instructive way to think about uncertainty.

Implications of the frequentist approach are:

1. The values to be estimated (i.e., the parameters) are fixed; they are knowable in principle and do not change. Uncertainty comes from the data.

2. We need a very friendly world for this definition to be useful; the world really is not a place where a very large number of trials can generally be expected to be even approximately independent and identical. So ...

   a. We can make the thought experiment more credible by using randomized trials or random samples when the data are collected.

b. We can do the science to show that the thought experiment is sensible, e.g., applying the Poisson distribution to radioactive decay.

c. We can become “science fiction writers,” treating the data as “random realizations” from some hypothetical population. We really don’t have to think, we simply pretend it is true (since it is unfalsifiable). In other words, we just assume the thought experiment applies. This is an unscientific but common approach.

And as usual there are lots of caveat emptors :

a. We should not confuse the statistical technique with the model being applied. For example, the bootstrap technique can simulate the thought experiment whether or not the thought experiment really applies to the problem at hand.

b. Be careful about how confidence intervals (CI) are interpreted. For example, we might take a sample and compute the mean, the standard deviation and the CI and then ask if the parameter (e. g., average temperature) is in that confidence interval. But we can’t know that. It does not tell us if the parameter is covered by this band in this study. It only tells us that the parameter is in 95 percent of the bands we would construct if we did the study over and over using independent random samples (as in the thought experiment).

Bayesian Statistics

Bayes Theorem is not controversial in and of itself. To begin, the probability of two events equals the probability of one event times the conditional probability of the other event, given the first. After some simple manipulations we get to Bayes’ theorem. There is no dispute about the mathematics. But when we use Bayes’ theorem as a means to “learn,” serious controversies follow. In outline form, we begin with a belief about some state of the world, consider that belief in light of the data, and then revise that belief accordingly. And a key point: those beliefs are represented in probabilistic terms so that probability reflects a state of mind, not long run relative frequency. Expression in probability language becomes a means to convey what a person believes, and resides, therefore, “in here” and not “out there.” Probability reflects what a person believes about the world and not directly the condition of the world itself.

The probability of two events equals the probability of one event times the conditional probability of the other event, given the first.

Bayesian Inference

We begin with a prior probability density function representing our beliefs about the parameter of interest before looking at the data. (By parameter, we mean some quantitative feature of whatever it is that we are studying.) If there is a “tight prior,” (if the density has a small spread), it means that we have relatively clear beliefs about the likely value of the parameter. If there is a relatively “flat prior,” (if the density has a larger spread), it means that our beliefs about the likely value of the parameter are unclear.

Then we introduce the likelihood function, which represents the distribution of the parameter given the data. The prior is then revised, based on this information, and the posterior distribution that results represents what we now believe about the parameter in light of the data. This updating process can be repeated again and again; the posterior from one distribution becomes the prior for the next. But in any case, the posterior distribution contains all of the information we have about the parameter and is used to draw conclusions about it.

In order to interpret and summarize results based on the posterior distribution, we might

choose the mode of the posterior, the mean of the posterior, or the 95 percent confidence interval (in which we can say we are 95 percent certain the value falls). This is consistent with the language scientists speak.

Still, one can play this Bayesian game well and still come upon various complications. For example, what if the resulting distribution has two peaks separated by a valley; this makes the posterior very difficult to summarize. Note also that there is a big difference between the distribution of data and the distribution of the parameter of interest. The focus here is on what individuals believe about the parameter.

Extensions of this approach include multiparameter problems (which require joint probability distribution functions, and are thus much harder), non-normal probability distribution functions, and model comparisons.

One can play this Bayesian game well and still come upon various complications. For example, what if the resulting distribution has two peaks separated by a valley; this makes the posterior very difficult to summarize.

Climate Effects on Food Security

Robert S. Chen

Consortium for International Earth Science Information Network (CIESIN)
University Center, Michigan

Our understanding and ability to project future climate effects on food security are still very primitive, Chen says. There have been few international assessments, and very few coherent efforts to pull all the available information together. Only two trade models have been used to address integration of various possible impacts in an international trade framework. There is very limited research to draw upon in trying to link climate factors to impacts on the larger food system and overall food security.

Chen disagrees with the assumption that the developed world is not vulnerable to climate change impacts on food security.

Food security issues are of interest to a wide range of decision and policymakers. Climate change is but one contributing factor. What is the decision framework? What else must be considered besides climate? How do climate uncertainties compare with other uncertainties? There is a great need to better understand the context and the decision framework.

Dimensions of Food Security

Chen disagrees with the assumption that the developed world is not vulnerable to climate change impacts on food security. There is much legitimate concern about water and land resources, such as limitations on agriculture and competition for land. There are also a whole set of food access issues, such as the sociopolitical nature of many famines. Chronic undernutrition affects on the order of 1-2 billion people, and nutrient deficiencies plague many more.

Conventional projections of world hunger (not including possible climate impacts) expect some improvement in the hunger situation over the next few decades. However, in Africa, for example, even if percentages decline, absolute numbers of people in hunger may still increase. Links between hunger/food access and poverty and efforts to break these links are issues of concern, as is the basic human right to food. Even in the developed world, this right is not firmly established programs such as food stamps are always under attack. In other instances, food is used as a weapon of war.

Food utilization is an understudied area. Beyond basic questions of population and yield, there are issues of dietary preferences (such as the increasing use of animal products), cultural issues, caloric needs of people in various occupations, and rural versus urban diets. Health and disease issues also enter into the equation. For example, parasites affecting about a billion people cause losses of 10-20 percent of the calories they eat. There are also potential links to climate change of cholera and other diseases that affect nutritional status. A more general issue is the diversion of food to non-nutritional uses including pet foods, tobacco, biomass energy, and the creation of diet foods that can't be digested. Food storage and waste comprise another area of concern.

Figure 2.6 illustrates possible links between future food security and the occurrence or prevention of possible global environmental changes. Chen believes that food security is an important normative issue; the world can decide that food is a basic human right and take steps to implement that view. He also mentions the fact that there may be cultural biases in research agendas, for example, the relatively late attention given to rice which feeds a third of the world's population in terms of possible climate impacts.

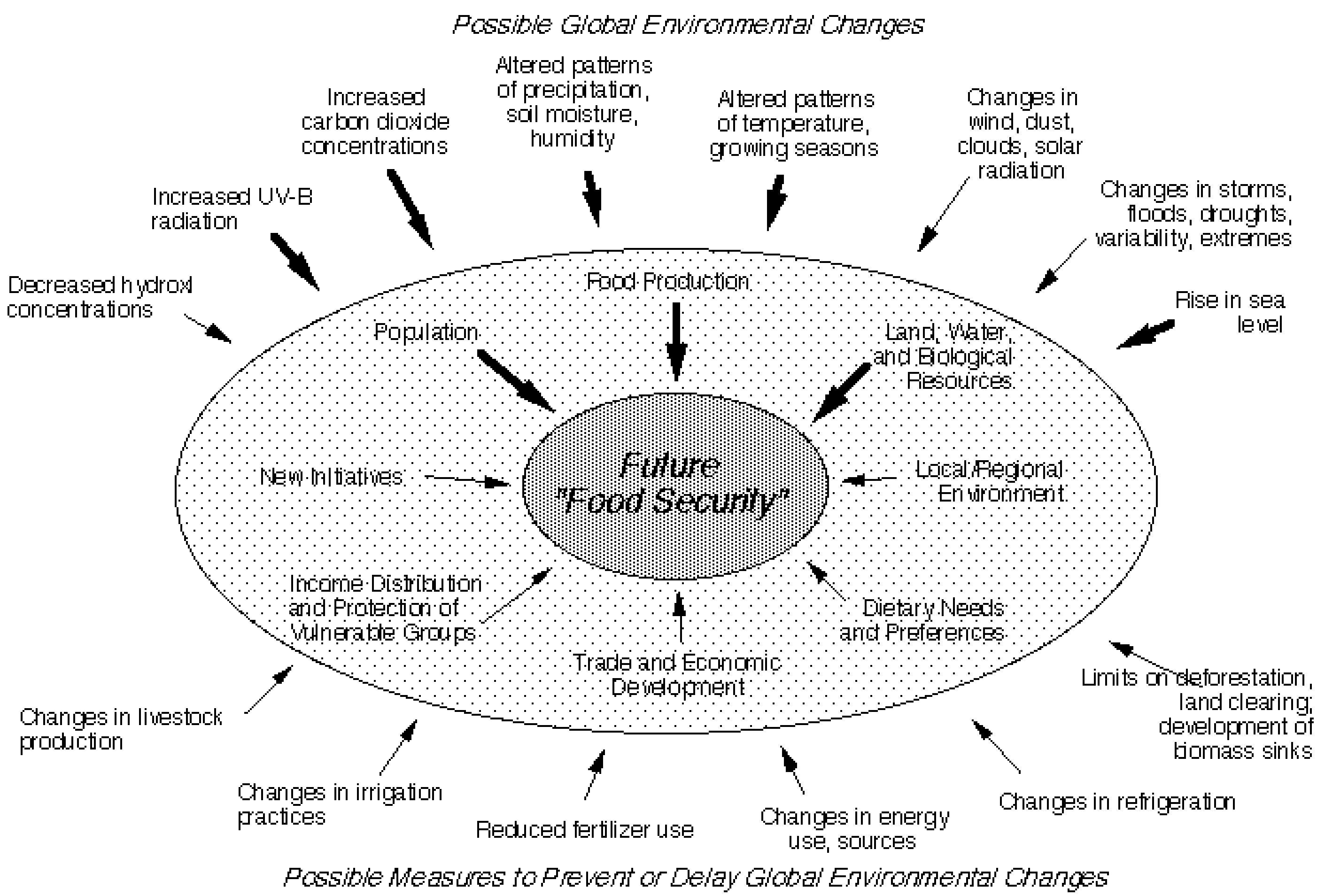


Figure 2.6

Regarding uncertainties in the decision making framework, Chen asks, "what does this all mean for agriculture now?" While there may be legitimate concerns about the potential effects of future climate change on the food system, many people are hungry now. What are the near-term food security impacts of proposed climate change mitigation strategies? For example, if a fossil fuel tax is implemented, it could change the current vulnerability of the food system and present -day access to food. This implies significant change in the system we're now trying to assess.

With regard to the decision making framework, Chen recalls the 1991-92 drought in southern Africa and the fact that it was predicted by using an El Niño/Southern Oscillation index. While information about this prediction was just beginning to surface and the drought was developing, in November of 1991 the World Bank urged Zimbabwe to make a large grain sale in exchange for a loan from the World Bank. Zimbabwe acquiesced to the Bank's request and, ironically, then used the loan to import needed food at higher prices due to the drought. The predictive information was not put into a form that was useful in a timely manner to the people who needed it.

Human population is a key issue to consider when looking at possible climate effects on the food system. It is important to remember that the United Nations' median projection is the middle of a very broad range of possible future human population sizes. Demographers make little attempt to analyze the probability distribution of the low, medium and high scenarios. Chen once tracked the history of past UN population projections for Africa and found that the current population estimate for Africa is now somewhat above the range of past projections. It is important not to take UN population projections for granted, but to study demographic models more critically. For example, one of the limitations of the UN's methodology is that the projections for future African fertility do not follow the "S" curve evident in the historic demographic transitions of other world regions.

Chen believes that if we limit ourselves to a simple "linear" perspective on food supply, we will end up having to choose the least of a number of evils. Food security is a broad issue for which there exists a wider array of options than is commonly perceived. Options include addressing food distribution, nutrient deficiencies, and eliminating war's effect on access to food. The IPCC process could begin to address this larger set of options.

In conclusion, Chen points to the continuum of decisions that are made in the food system, from local to national to global and from short to long term. There is a need to look beyond current decisions to future ones. We should also look at the robustness of decisions relative to uncertainty, and take actions that will be appropriate regardless of the outcome. Further, we must not forget about the possibility of surprises. For example, the current news about an unexpected die-off of honey bees is a reminder that we do not control everything and often cannot predict all outcomes.

While there may be legitimate concerns about the potential effects of future climate change on the food system, many people are hungry now. What are the near-term food security impacts of proposed climate change mitigation strategies?

Seismic Risk Analysis as an Example of Aggregating Expert Opinion

C. Allin Cornell

Stanford University
Stanford, California

The objective of aggregation is to represent the scientific community's composite state of knowledge on a particular issue. The process should seek to capture the diversity of interpretations.

As an example of aggregating expert opinion on a scientific issue, Cornell presented a process known as Probabilistic Seismic Hazard Analysis (PSHA) designed by the Senior Seismic Hazard Analysis Committee (SSHAC). SSHAC's PSHA is a methodology for estimating the likelihood that various levels of earthquake-caused ground motions will be exceeded at a given location in a given future time period. Due to large uncertainties in the geosciences data and in their modeling, multiple model interpretations are often possible, leading to disagreements among the experts. Cornell reported on a project, co -sponsored by the U. S. Nuclear Regulatory Commission, the U. S. Department of Energy, and the Electric Power Research Institute that was undertaken to review the state-of-the-art and improve on the overall stability of the PSHA process by providing methodological guidance on how to best perform the analysis. Cornell says that the resulting process represents an effective way of dealing with the problem of aggregating expert opinion in a particular scientific area for purposes of policy making.

The objective of aggregation is to represent the scientific community's composite state of knowledge on a particular issue. The process should seek to capture the diversity of interpretations, as opposed to the judgment of any particular expert. What should be sought in a properly executed PSHA project are: a) a representation of the legitimate range of technically supportable interpretations among the entire informed technical community, and b) the relative importance or credibility (weight) that should be assigned to the various hypotheses across that range. The type of consensus being sought, therefore, is that all experts agree that a particular composite probability distribution represents, first, them as a panel, and secondly, perhaps modified, the informed community as a whole.

The first stage of the process may involve winnowing, in which a large group of experts reviews, discusses and gives weights to different approaches to models, dropping some approaches as demonstrably wrong or as not worth pursuing further in depth because, for example, they have been replaced by later variations or extensions by the same authors. The second stage involves the careful selection of experts for the remainder of the process. These experts are asked to wear different hats; some are proponents of particular models or approaches while others serve as evaluators of different models or data sets. The major role is held by these latter experts who are charged with evaluating and passing judgment in the form of uncertainty distributions on parameter values and/or alternative models.

The ultimate responsibility for the success of the process lies with the Technical Facilitator/Integrator (TFI). The TFI is a single entity, sometimes an individual, but preferably a small team that is responsible for identifying the key issues and components of the analysis, structuring and

directing the interaction and debates among the proponents and evaluators, conducting any necessary numerical analyses, and documenting the process followed and the results obtained. The evaluator-experts should be made up of people who have been through these kinds of exercises before and have enough experience to have seen the state of the science change over time; they must have maturity and good judgment and must agree to work long and hard. They should be paid experts (which means resources are needed) who will work as "science court"-type evaluators, conducting a collegial effort to capture consensus.

At this point, participants pointed out that the IPCC process has thus far been a strictly voluntary one, and that raising the amounts of money that would be needed to follow the procedure described above could be difficult.

Like the IPCC, peer review is an important part of the process Cornell describes. He points out that it should be a "participatory" peer review, in which reviewers have full and frequent access throughout and can provide mid-course advice on both technical and procedural aspects of the process (as opposed to a late-stage peer review that occurs only after the project is almost complete).

Discussing how uncertainties are dealt with in the PSHA process, Cornell defined two classes of uncertainties. Those called epistemic are uncertainties resulting from lack-of-knowledge and which are in principle reducible through further research and gathering of more and better data. Those called aleatory or random in character are uncertainties that for all practical purposes cannot be known in detail or cannot be reduced. He mentioned using "uncertainty expression training" in which experts are taught how to express their uncertainties with more precision.

Interaction is key to the process of achieving the outcome of better science. There can be no anonymity, each evaluator must have the same information, and understand all of each other's assumptions. An important part of eliminating unintended disagreement is focused, directed interaction, which involves all participants and downplays socially dominant individuals.

In sum, the process described provides for clarification of objectives, the roles of experts, and the meaning of consensus. It also involves numerical or mechanical aggregation schemes, and social integration procedures. The objective of the process is to be in a position to simply give equal weight to each evaluator (this does not mean equal weight to each model). (Peer weighting does not seem to work unless the individuals are going to stop working in the field.) The composite uncertainty band which results is the TFI's representation of the group's uncertainty. Although he may seek the group's affirmation that his is a consensus position, the TFI holds the responsibility for defending this result as an accurate characterization of the group. The process involves very strong interaction, and should be facilitated by the TFI to exploit while not becoming the victim of scientific culture. The cost of doing this process well for any given issue may be up to $1 million. Some participants saw cost as a potential obstacle to the IPCC adopting such a process.

The evaluator-experts should be made up of people who have been through these kinds of exercises before and have enough experience to have seen the state of the science change over time; they must have maturity and good judgment and must agree to work long and hard.

What's a Journalist to Do? Challenges and Approaches to Reporting Scientific Assessment

Sharon Dunwoody

University of Wisconsin - Madison
Madison, Wisconsin

Media routines corral the behavior of journalists and mediate against the full telling of the story.

In considering how the media deal with issues of scientific uncertainty, and science in general, it is important to consider the difficulties that afflict journalists as they attempt to convey scientific issues to the public. Some of those difficulties are a product of the individual reporters themselves, of their training, and their personal opinions. Others are built into the occupation, with its deadlines and focus on events over process. A third set of difficulties comes with members of the audience, who tend to transform information to fit their existing notions. Some of the roadblocks facing journalists are the very same issues that face all people when confronted with evidentiary issues.

Journalists bring a widely inconsistent set of skills to the task of covering global warming and other scientific issues. Their profession typically values behaviors that do "end runs" around validity issues. And there is some evidence that audiences value "fair play" more than reporters' judgments about what is right or wrong. In fact, readers or viewers can react angrily to journalists who try to tell them what the "truth" is.

What can be done to help journalists develop the analytical skills needed to deal with issues of evidence and probability? Universities are now beginning to teach basic reasoning courses to undergraduates, beginning the process of familiarizing students with statistics, probability theory and other tools for evaluating evidence. In addition, there is a slow but steady accretion of statistical and other evidentiary training within journalism schools. Working journalists are increasingly products of journalism school training (the proportion of journalists who majored in journalism or communication has risen from 34 percent in 1971 to 56 percent in 1992, according to national surveys). Thus, a concerted effort to improve evidentiary training in journalism schools should have a pronounced effect on the occupation in the future.

Journalists' occupational routines and norms present another set of barriers to the accurate presentation of scientific uncertainty. Media routines corral the behavior of journalists and mediate against the full telling of the story. Journalists work under tight deadlines and cover such a variety of stories that it is difficult for them to ascertain the "truth" of any issue. In this fast-paced environment, the attempt to report "truth" has been replaced by the concept of balance: making sure your story contains the variety of available perspectives on an issue. The theory behind this seems to be: if you can't know who's telling the truth, then at least provide a plethora of views. Over time, then, the truth will emerge. As a general strategy, the concept of providing multiple views has merit, Dunwoody says. Its primary failing is that it absolves journalists from validity checks. It can also be problematic when handled simplistically; "balance" shouldn't mean finding two people on opposite sides of an issue and giving them equal space in the story.

The competition between the need for explanation and the need for brevity presents additional problems for journalists. Journalistic accounts are short and getting shorter as editors decide that consumers don't easily tolerate lengthy text. In a 200- to 300-word story, explanatory detail is the first thing sacrificed. Even when journalists put details in, editors often cut them out. To make matters worse, journalists receive little formal training in how to explain things.

In discussion, the point was raised that there are also certain cultural differences between journalists in the U. S. and those in other countries. Most American journalists eschew an investigative, critical role, seeing themselves more as interpreters and transmitters, while journalists in some European countries tend to reflect critically on and comment upon scientific subjects.

Professional subgroups of journalists also exercise some control over news making. Through interacting with one another, reporters develop a kind of shared sense of who the appropriate sources are and how to "play" a particular issue in their stories. Members of the Society of Environmental Journalists, for example, have engaged in a long -running discussion of the debits and merits of taking on an advocacy role in environmental coverage. Currently, the group is urging its members to avoid patterns of source use or topic selection that might be interpreted by others as advocacy reporting.

The competition between the need for explanation and the need for brevity presents additional problems for journalists. Journalistic accounts are short and getting shorter as editors decide that consumers don't easily tolerate lengthy text.

Elite newspapers, especially The New York Times, play an important agenda-setting role as other publications "key" on them. For example, the Love Canal contamination story originally appeared in a small regional newspaper in upstate New York but only after The New York Times picked it up, months later, did it become a nationally significant story.

The role of the audience is another important factor in how science is reported. Audiences generally do not want to be told what is true. Rather, they seem to want to read information that shows respect for all sides and allows them to make their own decisions. People don't like to be preached to. It is important not to appear to be an elitist journalist telling the audience what to think. For example, Dunwoody talked about a "haunted house" story covered by several Wisconsin journalists. Town residents, asked later about which paper provided the better coverage, selected the one that treated the frightened residents of the house and the psychic who arrived to diagnose the problem as sources with similar legitimacy to that of the skeptical scientists and police. People seemed to resent the reporting that debunked the haunted house theories. People felt the journalist who had provided all views had behaved most responsibly, informing them of other perspectives and allowing them make up their own minds.

Scientists are increasingly getting the jump on the peer review process by holding press conferences to release key results as they begin to understand the legitimizing function of the media. There is increasing evidence that having scientific research results published in elite newspapers confers benefits to the scientists involved. For example, data show that scientific papers published in peer-reviewed journals which are also featured in elite newspapers are cited more often by other researchers than those which are not reported on in newspapers.

In discussion about where reporters get their information for scientific stories, participants noted that journalists often use searches of databases such as Lexus and Nexus. These research tools point them to other newspaper stories on the subject, compounding the errors that are sometimes made in these stories, creating a feedback effect of inaccurate or inadequate news stories being relied upon as information for new news stories.

Media Coverage of Global Warming, 1985-1993

Dunwoody discussed Craig Trumbo's 1995 study (Trumbo, 1995) which examined media coverage of global warming from 1985 to 1993 in five major newspapers (The New York Times, The Washington Post, The Los Angeles Times, The Wall Street Journal, and The Christian Science Monitor), three major news magazines (Time, Newsweek, and U. S. News and World Report), television network news shows (ABC, CBS, and NBC), and the science press (55 publications including New Scientist, Science News, the news sections of Science and Nature, BioScience, Scientific American, Discover, Sierra, and Audubon). In addition, the study includes time series data for the same period on policymaker attention (gauged by the number of entries in the Congressional Record containing mention of either "greenhouse effect" or "global warming"), public opinion polls (the number of poll questions containing either of those terms), and public concern (poll respondents indicating a very high level of concern about the issue).

In 1981, only 38 percent of respondents had heard of the greenhouse effect. By 1987, that percentage had risen only to 40 percent, but over the next few years it reached near saturation at 86 percent.

When these seven variables were plotted over time, they rose and fell together, "like a school of fish," rising to a coverage peak between 1989 and 1990 (coinciding with the extreme summer heat waves of 1989) and then declining steadily through 1993. Among newspapers, The New York Times reported most frequently on global warming over the eight years (195 stories), while The Wall Street Journal came in last (49 stories). Among the magazines, Newsweek had 23 stories on the subject while U. S. News and World Report had 11. ABC World News aired 34 stories compared to 25 on the CBS Evening News and 22 on NBC Nightly News. Among the science press, New Scientist led the pack with 207 stories, followed by Nature with 174; Science News had 78 stories and Science , 75.

In 1981, only 38 percent of respondents had heard of the greenhouse effect. By 1987, that percentage had risen only to 40 percent, but over the next few years it reached near saturation at 86 percent. Level of concern rose steadily between 1987 and 1989, such that, in 1989, there were many more "extremely" concerned than "moderately" concerned individuals. By 1993, though, the number of extremely concerned had fallen so there were about the same number as the moderately concerned. Still, global warming was consistently ranked as the most important environmental problem by only 5 to 9 percent of respondents throughout the period; that percentage peaked at 12 percent in November 1989, and coincided with high levels of media coverage.

Based on the study results, Dunwoody discussed who set who's agenda on this issue. The science press led both newspapers and TV, and newspapers led TV. This inter-media agenda setting was strongest early in the time series, suggesting that the mass media seek the kind of issue legitimization that coverage by others provides. The agenda -setting power of the science press diminished as time went by. At any given time, TV coverage and public concern about global warming were closely related, but none of the media variables predicted the rise and fall in public concern over time. "If public concern is being shaped by television coverage, the effect takes place in a shorter time span than that measured here," the author of the study says. The mass media and policymakers seemed to go back and forth, influencing each other, with the media's influence fading over time. Finally, and interestingly, an increase in policymakers' attention seemed to correspond to a decrease in public concern. Colleagues in the audience cautioned that the data on this last effect should be analyzed seasonally before drawing conclusions.

Did the global warming issue peak too soon in the media? In 1988 -89, media interest peaked, long before the current "certainty peak." Now that it may be time for action on this issue in the policy arena, media coverage could be important in getting the attention of policymakers.

In closing, Dunwoody provided a few tips for scientists wishing to improve the accuracy of science stories for which they are sources of information. It is helpful to provide something in writing for the journalist, in addition to an interview. Data show that accuracy ratings go up when such written information is provided. Insist on face-to-face rather than telephone interviews, as this also improves accuracy ratings. At a scientific meeting, make time to meet with reporters and discuss your research in language they can understand. If there's an important caveat, say it in the middle of a very quotable sentence so it's less likely to be cut. Provide graphics to reporters and use familiar metaphors. Share good metaphors among scientists and think about how to convey uncertainties and other scientific issues visually.

Did the global warming issue peak too soon in the media? In 1988-89, media interest peaked, long before the current "certainty peak."

Reference

Trumbo, C., 1995, Longitudinal Modeling of Public Issues: An Applica -tion of the Agenda-Setting Process to the Issue of Global Warming. Journalism & Mass Communication Monographs, No. 152.

Emissions Mitigation and Atmospheric CO2

Jae Edmonds

Pacific Northwest National Laboratory
Washington, DC

The reason Annex I has responsibility for taking the lead in mitigation is that these are the nations which raised the level of atmospheric CO2, bringing about the climate change problem to begin with.

Edmonds examined implications of alternative actions for atmospheric CO2 concentrations. He discussed a number of proposed strategies and how they relate to the goal of the Framework Convention on Climate Change (FCCC), which is to stabilize concentrations of greenhouse gases (GHGs). The analysis he presented is based on an integrated assessment model known as MiniCAM 2.0 that resides at Pacific Northwest National Laboratory and Tom Wigley's National Center for Atmospheric Research COMMIC models. The analysis is calibrated to the IPCC IS92a scenario except that the period between 1990 and 2010 has been recalibrated to the World Energy Outlook in order to include the events that have transpired since the development of IS92a, particularly in Eastern Europe and the former Soviet Union (which followed the most successful emissions reduction strategy to date: near economic collapse!).

Annex I refers to the industrialized countries, including those of Western Europe, the U. S., Canada, Japan, Australia, New Zealand, Eastern Europe, and the European part of the former Soviet Union. The rest of the world, essentially the developing countries, are non -Annex I. Three cases were examined by Edmonds: Annex I countries acting alone, Annex I joined by non-Annex I, and atmospheric stabilization by Annex I and non-Annex I countries. Edmonds likened the Annex I countries' non-binding agreement to aspire to the goal of returning CO2 emissions to 1990 levels by the year 2000 to "aspirational pole-vaulting," which differs from the real thing in that you don't actually have to go over the bar. The question is: will the Annex I nations be real pole-vaulters? He points out that the reason Annex I has responsibility for taking the lead in mitigation is that these are the nations which raised the level of atmospheric CO2, bringing about the climate change problem to begin with.

He began by discussing the following possible Annex I actions:

· 0 percent Reduction: stabilizing emissions at 1990 levels by the year 2050

· 5 percent Reduction: stabilizing emissions at 1990 levels by the year 2010 and making 5 percent further reduction thereafter

· 10 percent Reduction: stabilizing emissions at 1990 levels by the year 2010 and making 10 percent further reduction thereafter

· 15 percent and 20 percent Reduction: stabilizing emissions at 1990 levels by the year 2010 and making 15 percent and 20 percent further reduction thereafter

The basic finding is that if Annex I acts alone, the actions can have only modest affects on global CO2 concentrations. In the reference case, CO2 concentrations in the year 2100 would be 689 ppmv. With Annex I following the 0 percent reduction strategy defined above, the CO2 concentration would be 669 ppmv, and if they followed the aggressive 20 percent reduction strategy, the concentration would be 646 ppmv. The bottom line is that no matter which of the above strategies Annex I countries follow, if they act without non-Annex I, CO2 concentrations will still be rising in the year 2100 (see Figure 2.7).

Taking the above scenarios (see Figure 2.7) except that non-Annex I nations begin to reduce their emissions by 10 percent beginning in the year 2010, again we find that there is not a great effect on atmospheric CO2 concentrations and that we do not even come close to stabilizing emissions. So even if the Annex I actions defined above are supplemented by 10 percent non-Annex I reductions in these scenarios, findings estimate CO2 concentrations of 643 ppmv for the 0 percent reduction case and 621 ppmv for the 20 percent reduction case, and again, concentrations are still rising in the year 2100 (See Figure 2.8).

The bottom line is that no matter which of the strategies Annex I countries follow, if they act without non-Annex I, CO2 concentrations will still be rising in the year 2100.

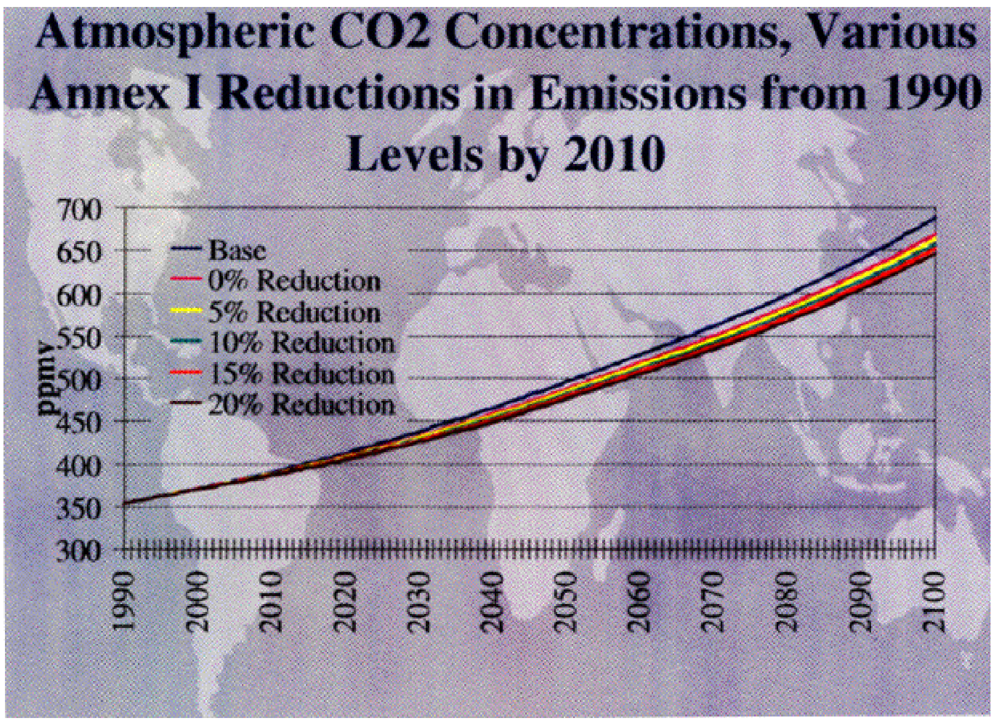


Figure 2.7

Clearly, the strategies discussed thus far do not get us close to the stated goal of the FCCC, which is atmospheric CO2 stabilization. This begs the question, what would we have to do to achieve this goal?

The IPCC constructed 10 emissions trajectories which stabilized the atmospheric CO2 concentration. These were defined by 5 different CO2 concentration ceilings, 350, 450, 550, 650, and 750 ppmv, and two timing cases, a more rapid initial mitigation effort and a slower one. Modeling results for these 10 cases reveal that all of these emissions trajectories have three phases: first, increasing emissions, second, emissions stabilization, and third, long-term emissions phase out (see Figure 2.9). All IPCC emissions trajectories exhibit this characteristic pattern for ceilings of 450 ppmv or greater. It is also clear from the results that ceilings below

450 ppmv (close to today's level) require immediate emissions reductions. Immediate and continued reductions in emissions below 1990 levels can, in theory, achieve a ceiling of less than 500 ppmv. But is this the correct policy to pursue?

Even if the Annex I actions defined above are supplemented by 10 percent non-Annex I reductions in these scenarios, concentrations are still rising in the year 2100.

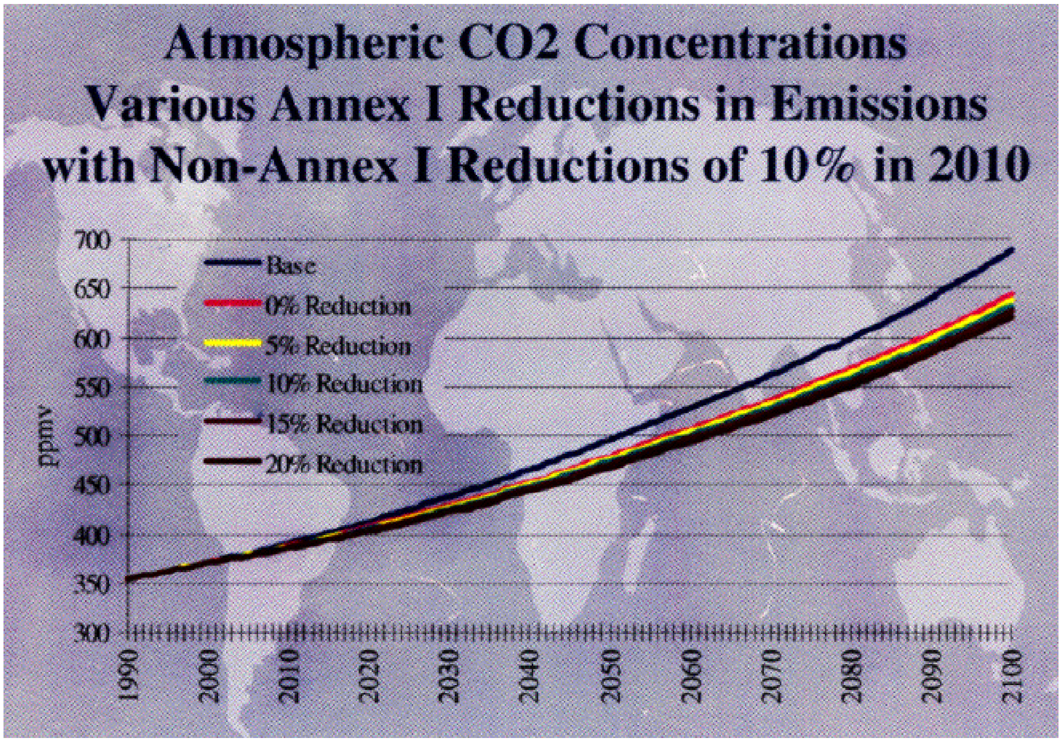


Figure 2.8

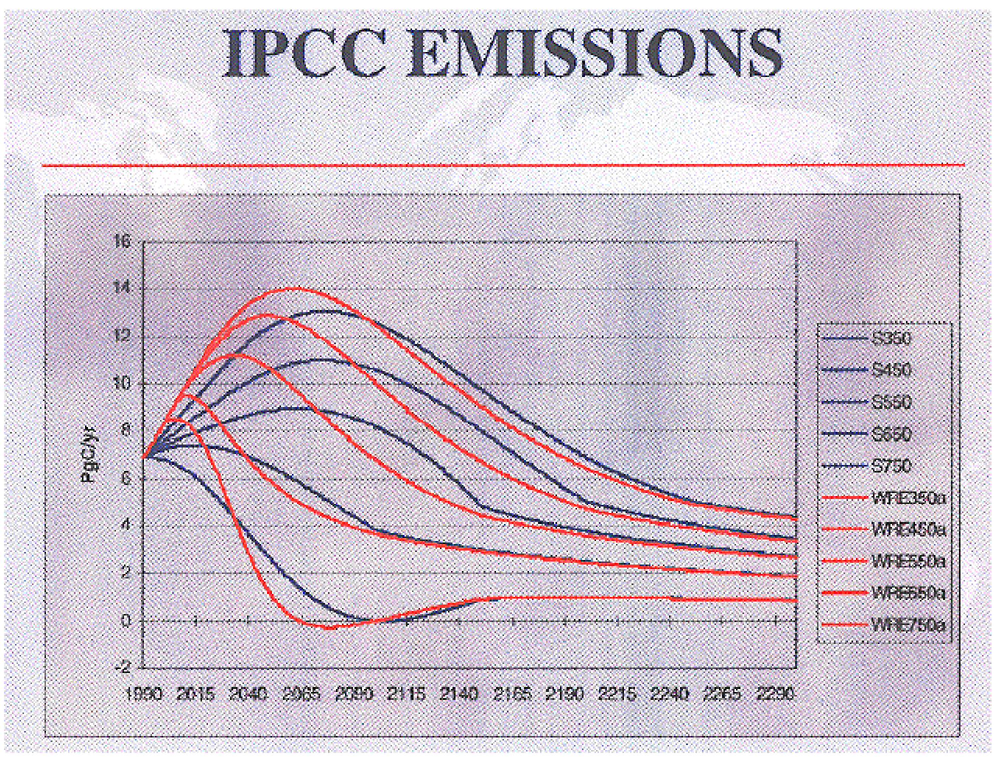


Figure 2.9

Edmonds says that by deciding on a policy of immediate and continued reductions, we would be ignoring what the optimum ceiling is. Instead, he says, we could decide what ceiling we wish to achieve and design a policy accordingly, thus working from the ceiling to the policy. However, no consensus yet exists regarding what is the right ceiling. Some argue that a prudent hedging strategy would not exceed a 400 ppmv ceiling until major uncertainties are resolved, thus keeping all options open. Edmonds believes this is an incorrect use of the analysis and could force society to more drastic and immediate action than might be called for. On the other hand, he acknowledges that once we pass a certain level of emissions, it is very difficult to go back; ceilings once reached then become floors.

Edmonds then turned to what must be done and by what dates in order to meet the target ceilings specified in the analysis. The "deflection date" is defined as the year in which emissions are 0.3 PgC/yr or more below IS92a, in other words, the year by which we would have to be significantly departing from that emissions trajectory. For a 450 ppmv ceiling, the deflection date is 2008 and global emissions would peak in 2011 at 9.5 PgC. This would necessitate a major, immediate effort in terms of initiating the institutions and technologies required to achieve this goal, and this would be costly. Edmonds says that there is a great non-linearity between the costs at a ceiling of 450 ppmv and one of 550. With a 550 ppmv ceiling, the deflection date would be 2018, and global emissions would peak in 2031 at 11.2 PgC. He believes this presents a more reasonable approach in terms of getting the entire world on track.

Immediate and continued reductions in emissions below 1990 levels can, in theory, achieve a ceiling of less than 500 ppmv. But is this the correct policy to pursue?

Modeled costs come down dramatically from the 450 ppmv ceiling to the 550 ceiling for a number of reasons, Edmonds explains. There is a premature retirement of capital stock or use of capital stock in ways not originally intended. The model indicates that there is a "carbon cycle dividend" equal to 20 years of emissions at current rates that can be released. There is also the factor of discounting, in which future expenditures are not weighed as heavily as current expenditures. And finally, there is the cost of technology development. A 450 ppmv ceiling would force society to spend large sums developing new technologies rapidly, whereas the 550 ppmv ceiling would allow normal rates of technological development to take care of that, Edmonds says. Due to having less time to build institutions and develop technologies, it would cost four times as much to achieve the 450 ppmv ceiling than the 550 ceiling (nearly 4 trillion dollars compared to less than 1 trillion), model results suggest. An important assumption in the model is that there is total flexibility in where and when emissions reductions are achieved and that they are accomplished in the lowest cost ways possible. It would be far more costly if each country were forced to achieve emissions reductions on its own.

There was discussion about these issues, with Nakicenovic saying that in the International Institute for Applied Systems Analysis models, when a learning curve was built in with the up-front investment costs to begin development of new technologies now, results indicated that it was more cost-effective to begin emissions reduction sooner.

Regarding the equity issue, Munasinghe pointed out that there may be a problem with asking the developing countries to start reducing their emissions by 10 percent in the year 2010 when their current levels of emission per capita are still quite small compared to those of the Annex 1 countries.

MacCracken pointed out that if the whole expenditure required to address this problem is well under 1 percent of GDP, why is it generally presented as being such an expensive solution? Why, given the uncertainties and potential for large impacts, would society choose not to make this relatively small expenditure of resources? Others felt that what was required to make this happen was the global political will and an institutional framework.

Models in the Policy Arena

Paul Edwards

Stanford University
Stanford, California

Climate change would not exist as a political issue without models, Edwards contends, launching a discussion of computer models, uncertainty and the politics of climate change.

Climate change would not exist as a political issue without models, Edwards contends, launching a discussion of computer models, uncertainty and the politics of climate change. Uncertainty in these models has important political uses, both positive and negative, by different parties. Models don't and probably won't ever control policy choices because there are other policy constraints that are too powerful, Edwards says.

There are a variety of uncertainties and epistemological problems with climate models including the limitations of computational methods, scaling issues, parameterizations and tuning, flux adjustments, uncertainty cascades (where uncertainties from one model feed into other models), limitations of data, blurring of model results and data (some "data" are in fact products of other models), and validation and verification problems. It is unlikely that these uncertainties will be eliminated.

Climate models are purely mathematical constructs that don't work in ways that are analogous to models in other fields. It is routinely reported in climate research that "experiments with the models" conclude ... But this is a new scientific paradigm one that excepts a mathematical model as a representation of global climate. Another problem concerns so-called "global" data. Very few data sets are actually global; most are regional or very scattered in their global coverage. Models have forced researchers to try to provide data on a global uniform grid and this has resulted in spotty data being filtered through models to produce "global" sets using techniques including smoothing, interpolating and gridding. But global data are required for validation and calibration of models. Thus, data used to validate one class of models are themselves the product of other models. Further, all data go through filters set up by humans. One such filter resulted in the famous "missed ozone hole," when the computer was programmed to reject data outside a certain range of values.

While models don't control policy, they do play an important role in politics and policy making. One role of models for climate science has been to build an increasingly large community around the climate change issue in which many groups and elements have come to play a role. Successful political processes succeed by enrolling allies. Scientists can't simply write up their results and publish them in journals. They also need to draw connections with other scientific disciplines, the fields of energy and economics, and with government agencies and politicians. Computer models, and particularly their graphics, have been a powerful tool for this kind of enrollment. The primary goal of models in politics is to play a heuristic role as opposed to a strong predictive role.

The Club of Rome Example

The Club of Rome's well-known modeling project that was published as Limits to Growth, grew from Jay Forrester's ideas about complex systems: that they are counterintuitive, nonlinear, and impossible for unaided minds to grasp. He also believed that policies often worsen problems

because "complex systems resist most policy changes." Models of such systems are insensitive to changes in most parameters. For all of these reasons, a model is needed to reveal leverage points which are likely not to be where you think they are.

Forrester saw models as the policy solution, and felt that models could serve policy purposes even without good data. He said that "the barrier to progress in social systems is not lack of data. We have vastly more information than we use in an orderly and organized way. The barrier is deficiency in the existing theories of structure." He believed models should be comprehensive. And he saw growth as a developmental phase, not a constant. He believed that continued exponential growth was impossible. Pointing to the importance of metaphor in the policy arena, Edwards recalled the exponential growth curve that became the icon for the Limits to Growth work.

In 1970 and 1971, the first modeling studies of truly global environmental problems were created: the Study of Critical Environmental Problems (SCEP) and the Study of Man's Impact on Climate (SMIC). These models raised for the first time anthropogenic global climate change as a major policy issue. Both studies recommended new methods for gathering global information in standard ways, integration of existing monitoring programs, and a global network of monitors. Following these efforts, climate simulation and global scale observation drove each other. Before computers, there was too much data. After computer models, there was not enough.

The primary goal of models in politics is to play a heuristic role as opposed to a strong predictive role.

In 1970, Forrester attended the first general meeting of the Club of Rome, having been invited via member Carroll Wilson, organizer of SCEP and SMIC. The Club of Rome was considering a computer model of world "problematique," which included issues with global dimensions such as population, resources and environment. Further developing Forrester's approach, they created a model called World 1. The world was divided into five major subsystems: natural resources, population, pollution, capital, and agriculture. A rapid work-up, with system structure and dynamics of greatest importance, was performed. In complexity, it was equivalent to a global average energy balance model. It exhibited the characteristic typical of all Forrester's models: overshoot and collapse. Edwards says that in fact, it is very difficult to produce a policy that does not exhibit this pattern.

The System Dynamics Group at Massachusetts Institute of Technology developed the model further, producing World 2 and World 3, the latter with over 120 interdependent variables. These models were calibrated to historical trends. Edwards says that the models were not particularly good but that they did provoke efforts to gather data and develop new models that were better. They concluded that exponential growth rates were unsustainable and that catastrophic collapse would come around 2050. In 1972, their results were published in Limits to Growth which sold 7 million copies in 30 languages. It was heavily criticized, especially by economists, but earned international respect for the Club of Rome in other circles. It had few direct policy impacts, despite the Club of Rome's promotional campaign to governments. It did, however, have a large impact on world public opinion, perhaps as much as Paul Ehrlich's Population Bomb, Edwards believes.

Limits to Growth was purely heuristic; rather than trying to make specific predictions, it tried to make three general points: the world is a system, exponential growth can't continue, and a comprehensive approach is necessary. It also helped establish a models-for-policy tradition of which climate scientists are the inheritors. This tradition includes a hybrid science/policy community and an interdisciplinary /transdisciplinary approach.

All of these early models confirmed what the creators already believed that the world faced huge problems that had to be addressed or we would face catastrophe. A huge modeling community was thus created and blossomed in the 1970s, including the International Institute for Applied Systems Analysis. After the 1972-73 energy crisis, a number of energy models and global economic

models were developed. The 1980s saw a renaissance of models-for-policy with models designed to simulate the enhanced greenhouse effect. It also saw the first integrated assessment model, IMAGE, which is a direct descendant of the world dynamics models.

Most policies that make it through the policy process have certain characteristics, Edwards says: they have a narrow focus, a high probability of success, short-term payoffs, they are tangible, easily perceived, have widely derived benefits, perceived affordability, and feedbacks. The climate change issue has none of these characteristics. The case for policy becomes even weaker when it is taken to the regional and local levels where we cannot even tell if policies will have any observable impact on global climate. If models progress to identifying regional winners and losers, that could make policy responses even less likely. A great advantage of the current regional uncertainty of climate models is that self-interest can't creep in; NIMBY is a robust phenomenon, Edwards says.

Limits to Growth tried to make three general points: the world is a system, exponential growth can't continue, and a comprehensive approach is necessary.

Edwards concludes that models do have a role to play in policymaking. They can be used for retrospective policy evaluation, helping to determine if a policy worked by comparing what actually happened to model results of what would have happened in the absence of the policy. Uncertainty can play a role in raising money to do further research involving models. He stresses that models-for-policy should be used heuristically, not predictively, and that modelers should take this into account when addressing the policy community.

Uncertainties in Observed Changes in Climate Variables

Tom Karl

National Climate Data Center
Asheville, North Carolina

Karl discussed uncertainties in observed changes in global temperature, precipitation, and other climate variables, and the three -star approach to ranking confidence levels used in the 1995 IPCC assessment. In general, the level of confidence has been raised by additional data since the 1992 assessment. In particular, there is now a high degree of confidence (three stars) that the global temperature increase has been between 0.3 and 0.6°C from the late 19th century to the present. On the other hand, there is less confidence in our understanding of the behavior of clouds and water vapor, and estimates in these areas were assigned one star in the confidence ranking. In Karl's IPCC Working Group I discussions, there was very little serious disagreement about what confidence levels were appropriate for the various estimates. Figure 2.10 illustrates and summarizes the basic consensus reached on stratospheric cooling, tropospheric warming, retreat of mountain glaciers, and other climate variables.

In general, the level of confidence has been raised by additional data since the 1992 assessment. In particular, there is now a high degree of confidence (three stars) that the global temperature increase has been between 0.3 and 0.6°C from the late 19th century to the present.

The uncertainty range is largely related to both systematic and random errors in data sets. Systematic errors include urban heat island effects, poor exposure of instruments (particularly for 19th century land -based data), uninsulated or poorly insulated buckets used to measure ocean temperature from ships, and differences between these data and ship engine intake measurements and hull contact thermistors. Random errors include inadequate spatial sampling, observer time changes (may be systematic or random), switch to automated instruments or other instrument changes, changes in local land use (such as effects of desertification), and instrument re-locations. While many errors have been corrected for, there is still a degree of uncertainty with regard to some data.

Figure 2.11 shows the combined land-air and sea surface temperature anomalies from 1860 to the present. This data set raises the question of how to decide when in the record we should begin having confidence in data. Older data are not as reliable for reasons including the errors mentioned above and because spatial sampling increases after 1900. Karl treats data collected in the 19th century with skepticism. It is thus unclear what beginning date should be used. Karl says that selection of the initial date should be dependent on the question(s) being asked.

An analysis of minimum and maximum temperatures at non-urban stations reveals that minimum temperatures are increasing by 1.33°C per 100 years while maximum temperatures are increasing by 0.83°C per 100 years. Comparisons between rural, urban and metropolitan stations show that the urban heat island effect is discernible even in small cities, though the effect is not large. The greatest warming has been measured in metropolitan areas, with less warming in smaller cities, and still less in rural areas.

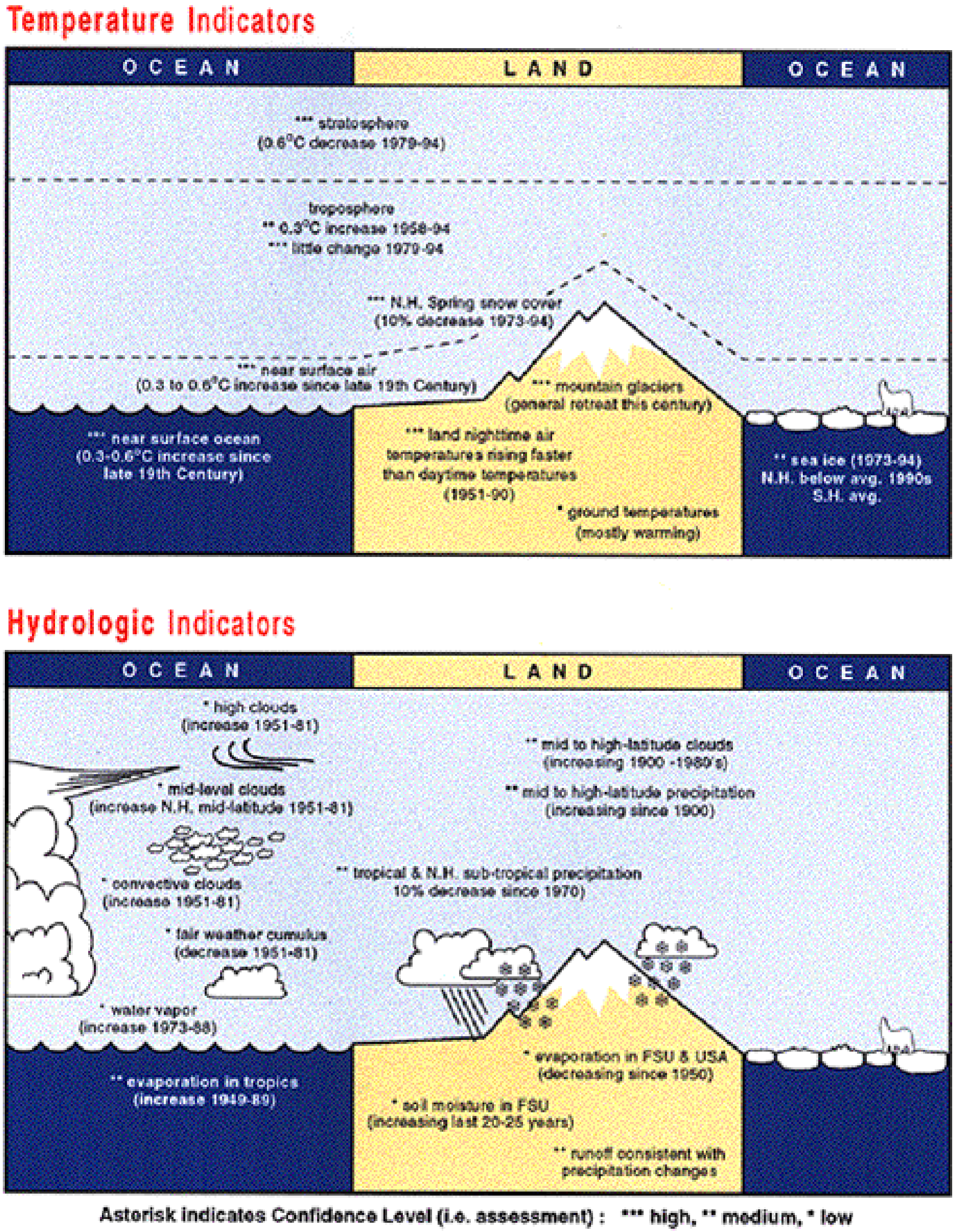


Figure 2.10
Summary of Observed Climatic Trends During Instrumental Period of Record

Source: IPPC, 1996, Climate Change 1995, The Science of Climate Change, Cambridge Univ. Press, U. K.

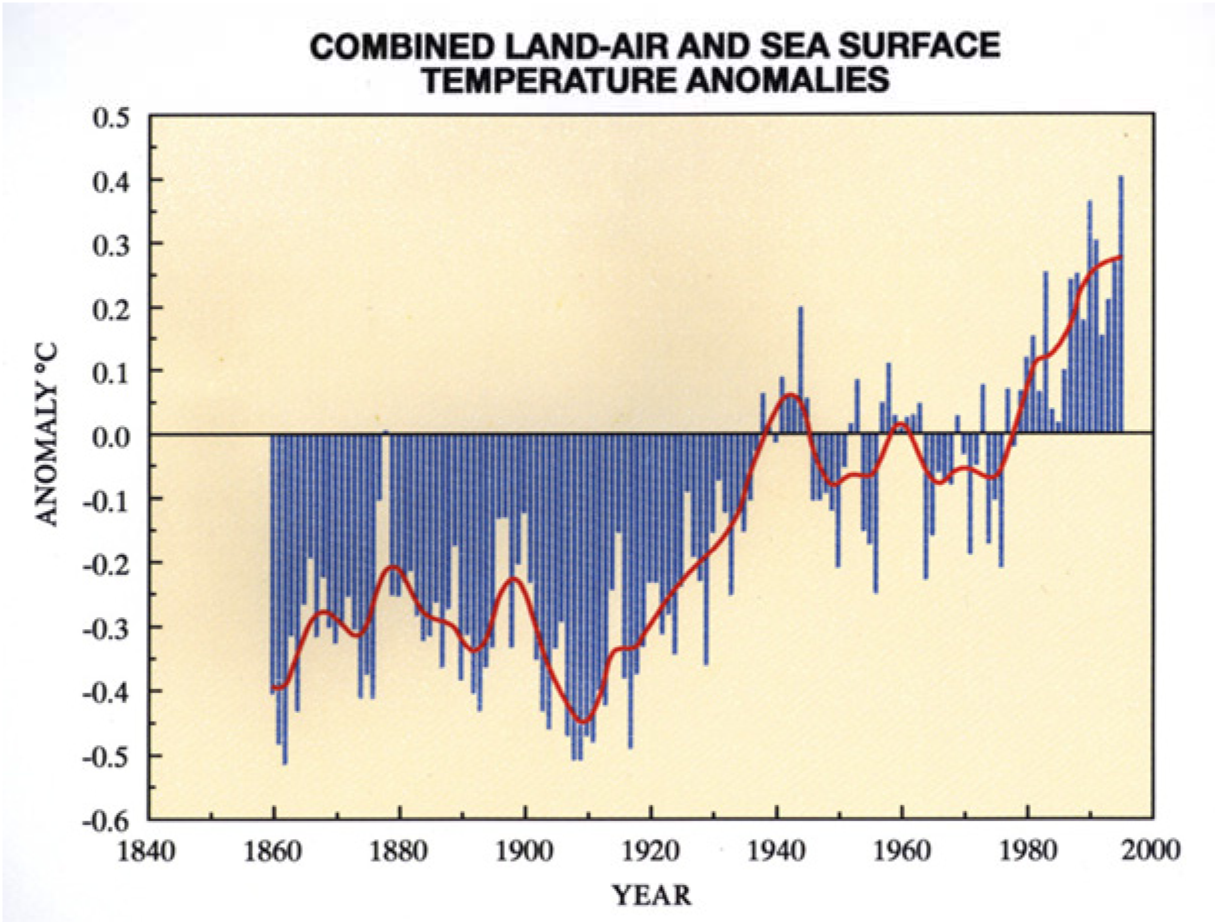


Figure 2.11
Combined Land-Air and Sea Surface Temperature Anomalies

The solid curve represents smoothing of the annual values shown by the bars, to suppress sub-decadal time-scale variations.
Source: IPCC, 1996, Climate Change 1995, The Science of Climate Change, Cambridge Univ. Press, U. K.

Assuming major errors are independent, with a 0.1°C/century error in land data, 0.1°C/century in ocean data, and another 0.1°C/century due to sampling inadequacy, the error interval is 0.17°C. Additionally, there is an urban heat island residual bias of less than 0.05°C/century. With the warming calculated as ~0.55°C/century (quoted as 0.5°C /century, after consideration of residual urban affects), the uncertainty band established by IPCC of 0.3 to 0.6°C/century in 1990 was not altered in 1995. The question was raised as to why this range includes significantly more on the down side than the up side (a 0.3 to 0.6°C range given for a calculated rise of 0.5°C). Karl expressed the belief that this range is likely to change in the next IPCC assessment but that there was not enough new data to make the changes at this time.

There is a high degree of confidence (three stars) in the data that show that alpine glacier mass is decreasing globally, and similarly, that snow cover is decreasing. Satellite data of snow cover changes over the Northern Hemisphere from 1972 to 1992 reveal a strong correlation (r=0.82) between temperature rise (in area-mean maximum) and reduced snow cover. This provides added confidence in near-surface global warming.

In regard to the question of whether the climate has become more variable or extreme, there are inadequate data and analyses to say anything about global scale changes. In some aspects, the climate has become more extreme, but in other aspects, there is evidence of little change or even a decrease in extremes. For example, there is a clear trend toward increasing precipitation rates in the U. S. and northeast Australia, while other areas (such as China) show little change. Other trends include increased intensity in extratropical cyclones in the North Atlantic since the late 1980s, and a small decrease in hurricane frequency and intensity in the North Atlantic over the past 40 years.

In the Working Group I discussions, there were some controversies about what can be said about extreme events. For temperature, the statement was made that a general warming tends to lead to an increase in extremely high temperatures and a decrease in extreme lows (e. g., frost days). It was also stated that small changes in the mean (or climate variability) can produce large changes in the frequency of extremes; a small change in variability has a stronger effect than a similar change in the mean. Karl asks, "What does this mean? The units are not even similar." A change in variability (variance) of 50 percent has a smaller effect on extremes than does a change in the temperature of 1.5°C (related to the extreme temperatures in Chicago, for example). Variability is not necessarily more important than changes in the mean with respect to extremes, Karl says, arguing that the language poorly reflects our knowledge.

There is a high degree of confidence (three stars) in the data that show that alpine glacier mass is decreasing globally, and similarly, that snow cover is decreasing.

Regarding mid-latitude storms, conclusions regarding extreme events are uncertain. The pole-to-equator temperature gradient has increased at high elevations but decreased at low elevations. There is some evidence for a recent increase in storminess around the North Atlantic (e .g., the 1988/89 abrupt increase) but a more general increase since the 1970s, implying that the North Atlantic Oscillation may not be dominant. There is not clear evidence of any uniform increase, however, and all of this led to controversy on this subject. As for tropical storms, no assessment is possible about what to expect in the future.

The number of frost days is decreasing in many parts of the world; is that an extreme event? The Chicago heat wave of 1995 was certainly an extreme event. Karl compared annual 2-day episodes of maximum temperatures over the past four decades for Chicago and found a discernible upward trend. He also used a statistical model to examine the effect of changing four variables (mean temperature, variance of daily temperature, interannual variance, and day-to-day persistence) on the occurrence of such events. Results suggest that increases in the mean and in variability increase the probability of extreme heat waves. One striking result of this research is that under present conditions, the 1995 heat wave is a 1 in 20 year occurrence for the extreme high, but less than a 1 in 1000 year event for the elevated minimum temperature during the 2 days of most intense heat. Raising the mean temperature to levels that may be realized by the end of the next century raises the probability of such an event to a 1 in 3 year event for the extreme high and a 1 in 200 year event for the elevated minimum temperature still quite unusual.

Why the IPCC Detection Chapter Was So Controversial

Michael MacCracken

U. S. Global Change Research Program
Washington, DC

Chapter 8 of the IPCC WGI report discusses the detection and attribution of anthropogenic climate change. Understanding why this chapter has raised significant attention may provide lessons for the future. Thus, MacCracken began with the rather intriguing question: Why did the IPCC chapter on detection of a human influence on climate receive so much attention when other areas of research involved even greater uncertainties? The process used by the authors of Chapter 8 was elegant and rigorous; key conclusions had multiple supporting evidence; and the result was not so surprising. So why did this chapter lead to such a commotion, and what can we learn from this?

The process used by the authors of Chapter 8 was elegant and rigorous; key conclusions had multiple supporting evidence; and the result was not so surprising. So why did this chapter lead to such a commotion, and what can we learn from this?

MacCracken suggests that the answer lies in a number of contributing factors:

1. The key conclusions of the detection chapter crossed an important psychological barrier, sometimes even nearing the "religious" belief that humans can not affect the global environment. Thus, saying that we can is an important paradigm shift. Detecting a human influence on climate crossed this line, saying that humans can and are changing the planet on a global scale.

2. Procedural issues were called into question by a number of the critics, although nothing out of the ordinary actually occurred procedurally (e. g., other chapters were also changed late in the process). The question thus arises whether such a seemingly strong conclusion might have threatened a trillion dollar industry, such that all means, even apparent procedural ones, might be used to discredit the IPCC conclusions and those who reported them.

3. There were misperceptions about whether any substantive changes were actually made in the chapter. MacCracken suggests that, contrary to claims, no changes actually were substantial if the chapter is read as a whole in contrast to focusing on specific phrases that were deleted.

4. Detection was one of the first areas in which there was an effort to be statistically rigorous. Was the jump to being quantitative, instead of just qualitative, part of the problem?

5. Some of the key critics apparently felt offended by the IPCC process; they felt insufficient attention was paid to their points of view in the chapter development process. The set of critics in this area, sometimes referred to as the "contrarians," have been well funded and have access to the media and thus were able to gain attention.

6. There was confusion on the part of some between the Summary for Policymakers and the chapter itself. Because the summary integrates across all chapters, it cannot contain all the caveats and explanations of each chapter.

7. Another problem may have been the manner in which the key conclusions of the detection chapter were first communicated to the public in The New York Times, which is perceived by some as a “pro-environment” newspaper, rather than directly from the IPCC (although the IPCC report was put out for review with a notice in the Federal Register a month before The New York Times article).

8. The lead authors of this chapter tried to be as up-to-date as possible in a rapidly changing field. Some of the newest materials hadn’t been vetted for very long in the scientific literature and community, and some of the points were not sufficiently elaborated upon or fully reconciled (e. g., the Microwave Sounding Unit data).

Contrary to claims, no changes actually were substantial if the chapter is read as a whole in contrast to focusing on specific phrases that were deleted.

9. The manner in which the IPCC chooses authors was also a possible point of controversy. The aim was to have a comprehensive review, but questions were raised about whether some of the expert authors might have slanted the content of their chapters to their own views. This is a potential issue across the IPCC in that experts in fields are selected as authors and will often cite their own work or may omit the work of others.

10. Did the reaction of the scientists involved in the IPCC exacerbate the problem (e. g., could they have chosen not to respond to the opinion pieces in the newspapers)? Would a different response to the critics have been appropriate and not ballooned the issue to as much prominence? Did the critics receive enough attention during the process? Is there a way to include them more effectively in the future?

Discussion

Discussion ensued regarding ways to try to avoid such pitfalls in the future. One suggestion was that the IPCC lead authors should not be the final authority for checking that adequate responses were made to reviewers’ comments. Instead, an appointed editor for each chapter could be responsible for overseeing comments and making sure they are adequately incorporated. It was also suggested that adequate time to study, document and incorporate comments is needed.

It was also pointed out that the conclusions of the detection chapter present a direct threat to the world view that climate change is just a theory that warrants study. The model for many powerful people is that it’s fine to study this subject but we should not take action until there is certainty. As soon as information in the climate debate approached a “smoking gun,” as it did in the IPCC chapter on detection and attribution, it became more politically charged. It was suggested that other issues might stir a similar response in the future (e. g., ecological impacts, economic costs).

Santer pointed out that there is no unique set of words that will make everyone happy, so criticism will be an ultimate consequence. How to word key conclusions will often be contentious, even among the authors. Even if there were more time and additional changes made, the authors could still iterate endlessly and still leave some people unhappy. So who should ultimately decide how to word conclusions? Would involving more individuals at

the lead author level help? Santer suggests that the scientists closest to the work are most appropriate to serve this function.

It was further pointed out that the language of the chapter was originally written as a communication from expert to expert and thus included implicit messages about statistics in phrases such as “we cannot positively attribute...” When read by those not versed in scientific subtleties, the implicit messages are missed. Thus, there seems to be a real communication problem between experts and laypeople. Having to write the chapters as well as the Summary for Policymakers for laypeople would seem to make the overall process much more difficult.

It was suggested that there will likely be topics that will become flash points in future IPCC assessments. Identifying them early in the process might lead to more focused attention on how to best deal with them.

How to word key conclusions will often be contentious, even among the authors. Even if there were more time and additional changes made, the authors could still iterate endlessly and still leave some people unhappy.

Freshwater Ecosystems, Hydrology & Water Resources

John J. Magnuson

Center for Limnology, University of Wisconsin - Madison
Madison, Wisconsin

Outcome advocacy can interfere with improving the state of the science, increasing understanding, and making appropriate responses to the findings.

Magnuson reviewed some of the controversies that arose during the development of the water-related chapters of the 1995 IPCC WG II report as well as some of the major findings of the chapters related to fresh water. To begin, he points out that fresh water is a critical human resource and essential to the functioning of all natural ecosystems, and that freshwater systems are among the most sensitive responders to climate change (McKnight, Naiman).

Controversies in the IPCC chapter on fresh water ecology began early with the question of whether such a chapter should be included in the assessment at all. Prior to a rather late review of draft chapters, it was thought that the information would come through in chapters on hydrology, oceans and large lakes, non-tidal wetlands, and water supply; but it did not. The subject of fresh water ecology sometimes falls into the cracks between oceanography and terrestrial ecology. A second controversy had to do with water supply versus water quality issues. A third controversy arose because some authors, particularly in the water supply chapter, felt that the climate scenarios were so uncertain that perhaps they should not even be used.

In the end, Magnuson said he believed that issues dominated over personalities, and the 1995 IPCC Working Group II report usefully incorporated freshwater quality and ecology into a chapter on Hydrology and Freshwater Ecology. The review system was responsive and flexible in making a place for this information late in the writing and review process. The system could have failed but did not. Magnuson said that those involved learned that the lead authors should take a broad perspective and be inclusive rather than exclusive in the materials considered, that the limited number of pages allocated to a chapter not be used as a way to exclude points of view or significant new information, and that advocacy for a particular outcome must be recognized and dealt with.

The strongest advocacy in the water area concerned water resources. One extreme was an engineering perspective (we can fix this problem) while the other was an environmentalist perspective ("the sky is falling"), Magnuson says. Outcome advocacy can interfere with improving the state of the science, increasing understanding, and making appropriate responses to the findings.

Information on aquatic ecosystems is scattered through the IPCC Working Group II report in chapters on the cryosphere (Chapter 7), wetlands (Chapters 6 & 9), oceans (Chapter 8), hydrology and freshwater ecology (Chapter 10), water resources management (Chapter 14), and even fisheries (Chapter 16). While fresh waters are sensitive to climate change, it was

not possible to model freshwater ecological impacts with the most recent models provided by the IPCC. The output of these models was available far too late to be incorporated. Also, the temperature and precipitation data provided were not sufficient to model many of the effects on aquatic ecosystems; data on humidity, cloud cover, and wind is needed, and extreme events are important.

Major results are water-related chapters indicate that as a result of climatic warming, the water supply per person and water for irrigation decline, wetlands and wetland areas decline, and cryosphere area and ice cover durations decline. At the same time, flow variability and floods increase, lake and stream temperatures increase, poleward dispersal of aquatic biota increases, and extirpations and extinctions increase. All of these were high confidence results with little uncertainty, provided that the climate changes as indicated in the climate scenarios.

Changes owing to climate warming in the per capita availability of water for direct human use is simulated to vary considerably from country to country. A key conclusion, however, is that water resources would become more critical with climate warming. Water levels of lakes and rivers are far more dynamic than those of the oceans and people have already had to adapt to such changes in the past, often at high costs.

As a result of climatic warming, the water supply per person declines, wetlands decline, and cryosphere area and ice cover durations decline. Flow variability and floods increase, lake and stream temperatures increase, and extinctions increase.

Focusing on the cryosphere, scenarios for 2050 in the Northern Hemisphere indicate a 6-20 percent decline in snow cover, a 25 percent decline in glaciers, 16 percent shrinkage in permafrost, and one month less lake and stream ice. In discussion, Tom Karl pointed out with regard to the projected 6-20 percent decline in snow cover that there has already been a 10 percent decline in snow cover observed in the last 20 years.

Magnuson says that more use of lake and stream ice data would be useful in future assessments. Ice phenologies for lakes and streams have several useful features. Many long records of “ice on” and “ice off” dates exist in northern temperate, boreal, and Arctic latitudes in North America, Europe and Asia. While direct observations are common, ice dates also can be estimated across large regions using satellite data. Ice on and off dates can be modeled well with physical process models using climatic data. Long-term changes in lake ice phenologies have common features globally, such as progressively shorter ice durations over several centuries; year-to-year variation reflects more local conditions and shorter-term phenomena such as El Niño. Recent papers pointing out some of these features are in a 1996 special issue of Limnology and Oceanography on climate change (McKnight et al., 1996).

Ecosystem effects of climate warming would be varied and touch on most physical, chemical, and biological aspects of lakes and streams. Interpretations in the 1995 assessment were from direct observations and from modeled changes forced by the GCM scenarios.

Lake and stream physics are responsive to climate change. Summarizing effects simulated under doubled CO2 conditions: stream temperatures track air temperatures, lake surface temperatures rise from 1 to 7°C, deep water temperatures range from -6 to +8°C, thermocline depth ranges from -4 to +4 meters, and the thermocline gradient becomes sharper. Some improvements can be made in the physical hydrodynamic models used in aquatic climate change analyses, but the uncertainty is less in these models than in the GCMs, Magnuson says.

Stream export of various compounds into a lake is strongly associated with an observed warm, dry period for a lake in northwestern Ontario monitored over 20 years, the Experimental Lakes Area (ELA). Dissolved organic carbon (DOC) declines because as water levels decline in the surrounding wetlands, there is less decomposition and hence less export of DOC via the streams to the lake. The water flowing into the lake becomes progressively clearer and clearer, thus the lake gets clearer and clearer, and light transmission increases. This changes the mixing depth and the depth of photosynthesis. Changes in runoff also caused changes in the export of phosphorous, silica, and other compounds. In a warmer, drier climate, less primary production takes place in the lake, owing to these changes. Many simulations of effects on lakes or streams place the boundary of the system at the edge of a water body and do not account for how climate change would affect land-water interactions and thus the export of materials from terrestrial to aquatic portions of the landscape.

In a warmer, drier climate, less primary production takes place in lakes due to a variety of changes. And clouds dramatically reduce primary production in lakes.

Primary production, or the total amount of organic carbon fixed, would be responsive to changes in cloud cover. This is apparent from measurements at the North Temperate Lakes Long-term Ecological Research Site in Northern Wisconsin. Photosynthetically active radiation was monitored, lab incubations were done at different temperatures, and the carbon fixed was calculated on a daily basis for Crystal Lake, Wisconsin. A key expectation from the results is that clouds dramatically reduce primary production in lakes, underscoring the importance of the role of clouds and the need for a better understanding of how they can be expected to change if we are to simulate ecological effects. On the other hand, if primary production increases as has been simulated in some studies, dissolved oxygen would more likely be depleted in the deeper, colder waters of the lake and result in the loss of cold water fishes like lake trout.

The thermal niche of a fish species can be defined by lethal, controlling, and directive (thermal preference) criteria. The criterion with the broadest limits is the lethal criteria: the fish can survive within an upper and lower temperature limit. The optimum range, in which the fish's growth rate, swimming speed and other factors are greatest, is more narrow. The most narrow range is specified by behavioral thermoregulation of the fish; fish move to better temperature habitats in much the same way as we do when we move out of or into the sun or wind. These ranges are used to calculate how much more space there is in a lake for fish before and after climate warming and whether it increases or decreases on a annual basis. Simulations to date suggest that warming may benefit fishes in large deep lakes but be detrimental in streams, shallow lakes and in the shallow areas of large, deep lakes.

In conclusion, Magnuson reiterates that freshwaters are critical to the human condition, are sensitive to climate changes, interact with other influences, have heterogeneous spatial impacts, and have many uncertainties associated with them. A rich array of these potential effects of climate change are in the IPCC chapters.

References

McKnight, D., D. F. Brakke, and P. J. Mulholland. eds., 1996. Freshwater Ecosystems and Climate Change in North America. Limnology and Oceanography, 41(5).

Naiman, R. J., J. J. Magnuson, D. M. McKnight, J. A. Stanford, 1995. The Freshwater Imperative. Island Press, Washington, D.C., USA. xvi + 165 pp.

Quantitative Expert Subjective Judgment: Does It Have a Role in Future IPCC Assessments?

Granger Morgan

Carnegie Mellon University
Pittsburgh, Pennsylvania

There are two traditional approaches to uncertainty: ignore it or describe it qualitatively using words such as “unlikely” or “doubtful.” In the case of a problem such as climate change, ignoring uncertainty is obviously not acceptable. Thus, in the past, many treatments have used qualitative descriptions. However, Morgan noted that the experimental social science literature indicates that qualitative descriptions of uncertainty can mean very different things to different people. For example, when a 1986 study by Wallstein, Budescu and Rapoport asked subjects for a subjective judgment of the chance that an event termed “unlikely” would in fact occur, responses ranged from 0 to almost 40 percent!

Qualitative descriptions of uncertainty can mean very different things to different people.

To avoid this problem, one can adopt a Bayesian perspective, using quantitative statements of probability to convey one’s degree of belief. In at least limited ways, scientists have long engaged in such practices (e. g., “Our best estimate of the speed of light including all sources of error is 299,775 km/sec ± 5 km/sec.”) In order to illustrate how expert judgments about uncertainty can be formalized, Morgan asked one participant to consider how long it would take him to drive from his office to the airport at a given time on a given morning. They began by refining and clarifying the question. It was quickly shown that developing a well-posed question requires a lot of clarification in order to be sure of getting the answer to the question one thinks was asked. Once that is accomplished, standard methods can be used to elicit the subjective probability distribution, typically starting with questions designed to identify the extremes of the distribution.

This basic idea of expert subjective judgment has been a main-stay in the field of decision analysis for many years. By way of background, before turning to the issue of applying such judgments in the IPCC process, Morgan discussed:

- the role of cognitive heuristics which people use in making judgments under uncertainty and the biases to which they can lead; and

- previous experience with the use of expert subjective judgments in related fields.

He went on to argue that in designing a procedure for use in the IPCC, it was important to focus on eliciting the reasons for the results (the whys are important, not just the numbers) and to pay careful attention to how to display the results and under what circumstances it does and does not make sense to combine judgments from different experts.

Morgan explained that people use a variety of heuristic procedures, or subconscious rules of thumb, when making judgments that involve uncertainty. One is availability, which results in probability judgments being driven by the ease with which people can think of previous occurrences of the event or can imagine such occurrences. Another is anchoring and adjustment, which results in probability judgments being overly influenced by the starting point which becomes an "anchor." A third is representativeness, which results in people judging the likelihood that an object belongs to a particular class in terms of how much it resembles that class. Often these heuristics serve us well, but in some situations they can lead to significant bias in judgments.

Over a wide range of circumstances respondents are systematically overconfident. Interestingly, the harder the question, the greater the overconfidence.

One consequence of the operation of these heuristics, particularly that of anchoring and adjustment, is a strong tendency to overconfidence. Experimental studies reveal that over a wide range of circumstances respondents are systematically overconfident. There is a very large literature on calibration which shows consistent overconfidence. Most of this literature is related to judgments by laypeople. Scientists may be different from laypeople in some respects, for example, in the greater amount of substantive knowledge they have to fall back on. But even among expert physical scientists, there is strong experimental evidence of frequent overconfidence, Morgan reported. Calibration does depend upon how hard the questions are. Interestingly, the harder the question, the greater the overconfidence. But calibration can get quite good with feedback and practice (as in the case of precipitation forecasts by weather forecasters). Calibration also improves when respondents are asked to give reasons for their answers and address counter-factual examples.

Morgan briefly described a number of examples in which subjective expert judgments about uncertainty have been used in addressing a variety of environmental problems. These included the Rasmussen, et al., 1975 study of nuclear reactor safety, the 1979 National Academy of Sciences (NAS) study of ozone depletion and CFCs, his own 1984 study of sulfur long-range transport and resulting health impacts, and the NAS 1986 study of radon from uranium mill tailing piles. He then focused on work he and David Keith performed in 1995, which involved detailed interviews with 16 U. S. climate scientists.

In the Morgan and Keith research, the interviews were far more technically detailed, and used a much more complex interview protocol than any previous expert elicitation on any topic, Morgan says. Morgan and Keith's interview protocol involved six parts:

1. introductions and an explanation of what the researchers were trying to accomplish;

2. general discussion in which the experts were asked to critique the researcher's background paper and discuss how they think about the problem;

3. judgments on a small number of policy-relevant global variables elicited in the form of full subjective probability distributions; judgments elicited about how well knowledge of global-scale climate change allows predictions on smaller scales;

4. disaggregated sources of uncertainty in global average temperature change dealt with by asking experts to systematically identify key contributors to overall uncertainty, and the extent to which the three global-scale variables were separable;

5. experts asked to discuss factors that should be considered in designing a national R&D program and then make a series of judgments about how resources should be allocated in such a program;

6. experts asked to talk about how research programs can lead to surprising results and then asked to make judgments about the likely state of their knowledge after 20 years of the research program they designed in part 5.

Morgan displayed and discussed some illustrative results of the study, including the projected climate sensitivity, pole to equator temperature gradient with doubled CO_2 forcing, zonally averaged precipitation, and radiative forcing due to anthropogenic aerosols. He also discussed a task in which the experts were asked to sort a set of cards on which possible causes of uncertainty were identified. Each expert thus ranked which uncertainty terms contributed the most uncertainty to his answers. The top five uncertainty terms identified were, in order, cloud optical properties, convection/water vapor feedback, CO_2 exchange with terrestrial biota, CO_2 exchange with the oceans (including ocean biota), and ocean convection.

Morgan then discussed a similar elicitation procedure he is developing to be administered to a group of terrestrial ecologists. They will be asked a series of questions designed to assess:

In the Morgan and Keith research, the interviews were far more technically detailed, and used a much more complex interview protocol than any previous expert elicitation on any topic, Morgan says.

1. What factors and processes are most important in determining the responses of forests and grasslands to climate change?

2. What will be the likely change in standing biomass in tropical and northern forests and the extent of grasslands, under fairly gradual and modest climate change?

3. How rapidly can trees and grasses in the Northern Hemisphere migrate in the face of climate change?

4. How might the southern boundary for the region of Arctic tundra move under fairly gradual and modest climate change, given that the change in regions of permafrost is known?

5. questions related to species loss and the introduction of non -indigenous species;

6. questions about how the mix in biota in two specific northern forests might change over time in the face of fairly gradual and modest climate change;

7. questions similar to those above, but for fairly abrupt and substantial climate change;

8. questions about research needs and priorities for impacts of climate change and interaction between climate and natural ecosystems.

Finally, Morgan addressed how an elicitation procedure of this kind might work for the IPCC. He argued that someone with previous experience conducting elicitation studies should work with the authors to define the questions to be asked. Care should be taken to explicitly elaborate all the conditioning assumptions. The procedure should place emphasis on collecting information about the reasoning that underlies the probabilistic judgments obtained, not just probability distributions. The group of respondents should be chosen by the IPCC review experts.

The objective should be to include all major points of view, not to produce a “statistically representative” sample. The initial judgments should be collected individually. Once the results are all in, they should be summarized, and then shared among the respondents. The results and associated reasons should be discussed at a workshop and respondents given an opportunity to revise their views.

For any particular question, there are a number of possible outcomes. For some questions, it may be suitable to combine the assessments into one summary distribution. For others, it might be more appropriate to combine them into several distributions reflecting different “schools of thought” on an issue, always conveying the reasons for the differences. In other cases, it might be best to summarize but not combine the assessments, again, conveying reasons. Morgan illustrated what some of the resulting displays might look like and stressed that understanding why the judgments are what they are is critical before combining them in any way.

The procedure should place emphasis on collecting information about the reasoning that underlies the probabilistic judgments obtained, not just probability distributions.

Some discussion took place around the issue of how the group of experts should be chosen and how to ensure their independence. Ideally, Morgan thinks, the set of experts should be chosen to describe the range of seriously held views and the current state of knowledge including the diversity of opinions. There is a need, he says, to display the range of respectable opinion. It was also pointed out that it may be optimal to go to different experts for different questions.

Morgan concluded by pointing out that the point of this exercise is to summarize the state of knowledge and belief, not to achieve a false consensus. Masking disagreement or different perspectives will help neither scientists nor policymakers in the long run, he cautions. He suggests that the first IPCC report conveys a greater degree of consensus than does the results of the Morgan and Keith research. Finally, he stresses again that it is of utmost importance to convey the reasons which underlie any differences in views whenever expert opinions are combined.

Reference

Morgan, M. Granger and David W. Keith, 1995, Subjective Judgments by Climate Experts, Environmental Science & Technology, 29:468-476.

Subjective Probability Rankings and the IPCC Process

Richard Moss

IPCC Technical Support Unit
U. S. Global Change Research Program
Washington, DC

The issues of representation of uncertainty, how to characterize outlier events in assessments, and how to combine expert judgments in very diverse fields, have been around for a long time. Moss' interest is focused on how we can do a better job next time than was done in the latest IPCC process. A lot of progress has already been made, including in the area of characterizing uncertainty, but much more remains to be accomplished. How to best rank confidence levels is still a difficult issue and so far the effort to deal with it has not been very systematic.

How to best rank confidence levels is still a difficult issue and so far the effort to deal with it has not been very systematic.

Moss presented a framework he and colleagues developed as one possible way of standardizing subjective probability rankings in the IPCC process. The draft framework consists of placing conclusions into one of four categories:

Well-established This category denotes wide agreement, based on multiple findings through multiple lines of investigation. A finding could be removed from this category not by a single hypothesis, observation or contention, but only by a plausible alternative hypothesis, based on empirical evidence or explicit theory, and accepted by a substantial group.

Well-posed controversy A well-established finding becomes a well-posed controversy when there are serious competing hypotheses, each with good evidence and a number of adherents.

Probable This category indicates that there is a consensus, but not one that has survived serious counter-attack by other views or serious efforts to "confirm" by independent evidence.

Speculative Speculative indicates not so much "controversy" as the accumulation of conceptually plausible ideas that haven't received serious attention or attracted either serious support or serious opposition.

As shown in Figure 2.12, there are two dimensions to this framework with regard to confidence: amount of evidence, and level of agreement. This framework was developed too late in the process to be used in the 1995 IPCC assessment just completed, so for this version, they fell back on using high, medium and low confidence levels, designated by one, two or three asterisks, which is a less precise framework.

A more systematic process and tools for ranking confidence levels are needed. It is both particularly important and difficult to do this in an environment that is becoming increasingly political. We must also pay attention to how uncertainties are characterized in the media and in the political process.

We must also pay attention to how uncertainties are characterized in the media and in the political process.

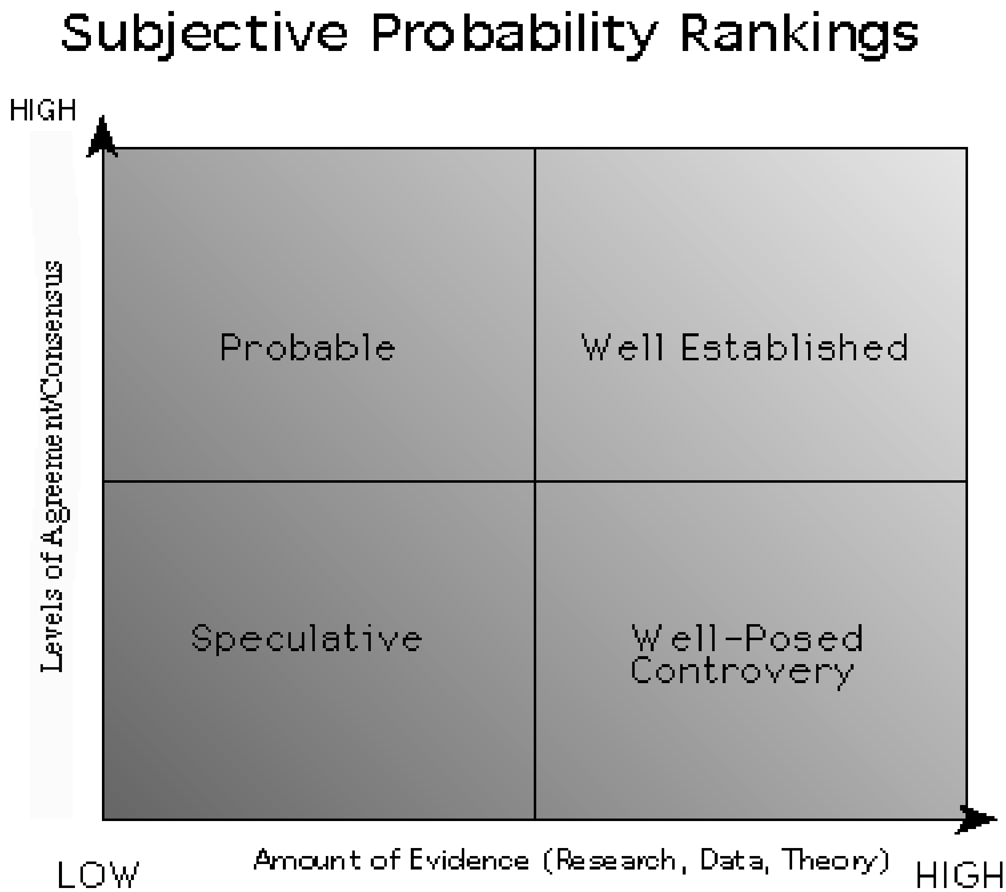


Figure 2.12

Estimating Damages and Costs of Climate Change

Mohan Munasinghe

University of Colombo
Colombo, Sri Lanka

Compared to Working Group I's effort on the science of climate change, the development of Working Group III's material on the economic and social dimensions of climate change is relatively new. The 1995 IPCC WG III assessment should thus be seen as a "first cut" at the question of responses to climate change and as such, does not have the depth of material nor sufficient time to do the probing analysis of confidence intervals that we see in the Working Group I report.

Predictions about climate change, its impacts, and the costs of mitigation are important for the policy making dimension. We are not seeking numbers for their own sake. Questions about climate change reside within broader questions about sustainable development. The objective of human development is sustainability, and the pursuit of greater precision in climate prediction can help with progress toward this goal.

While climate change has an important bearing on sustainable development, it is crucial to recognize that, especially for the developing countries, there are a number of other priorities. Hunger and malnutrition, poverty, and pressing local environmental issues present far more immediate needs. For example, particulates from burning fuelwood in traditional stoves expose large numbers of women and children to the equivalent of smoking 20 packs of cigarettes a day. And in middle income countries, sulfur dioxide pollution, especially in capital cities, is a much higher priority than climate change.

While climate change has an important bearing on sustainable development, it is crucial to recognize that, especially for the developing countries, there are a number of other priorities.

In terms of priority setting and getting the attention of decision makers in low and middle income countries, it is important to have information emerge from the IPCC process to show how climate change is related to these other priorities. For example, the most attractive climate change policies would be so-called "win-win" options, such as energy conservation, which provide significant in-country benefits as well as reductions in greenhouse gas emissions. It is also important to consider equity issues, like the huge disparity in per capita carbon dioxide emissions between the high and low income countries, when decisions are made concerning future international burden-sharing to mitigate climate change.

Some of the less developed countries are beginning to pay attention to the issue of climate change however, and Munasinghe believes things have progressed considerably in this regard. His own country, Sri Lanka, has at least three reasons to be concerned about climate change, he explains. In the short term, the UN Framework Convention on Climate Change (FCCC) will involve joint implementation projects that could earn the country foreign exchange currency. In the middle term, as the FCCC comes increasingly into force, it could affect decisions regarding

the building of power plants in Sri Lanka and have impacts on other aspects of sustainable development. In the long term, future generations will be affected worldwide. How we pose the problem and solutions is important. To get attention, we must focus on certain near-term issues, embedding them in a long-term context.

Regarding key uncertainties in Working Group III's assessment, Munasinghe says that the range of confidence in the predictions is very loose. Mainly, the results consist of "best guess" estimates based on the judgments of the so-called experts who wrote the chapters. No error bands can be assigned and there are no probabilities attached to particular outcomes.

The chain of causality from emissions to impacts to responses is summarized in Figure 2.13. Hopefully, through the FCCC, the allocation of rights and responsibilities will close the loop in a way that reduces net emissions and also controls impacts and damages.

It is important to consider equity issues, like the huge disparity in per capita carbon dioxide emissions between the high and low income countries, when decisions are made concerning future international burden-sharing to mitigate climate change.

The decision framework roughly breaks down into three major components, which are not mutually exclusive:

1. **global optimization** is the most important component, which looks at costs and benefits and averages and aggregates them to find some optimal level for reduction of GHG concentrations, adaptation responses, and so on;

2. **collective decision making,** which covers international bargaining, equity issues, and other such matters that cannot be captured in global optimization models;

3. **mechanisms and procedures,** which deals with process issues, including creating or adapting international bodies, developing monitoring mechanisms, etc.

There are many complications in the analysis of global climate change. First, the tools for such an analysis have never been used on such large spatial and temporal scales. Similarly, the complexity of this issue is far greater than any other tried previously. The issue of "irreversibility" causes another complication we may lose future options which have value. What would we be willing to pay to have the flexibility of exercising certain options in the future? Non-linearity, or the possibility of catastrophic outcomes, adds another complication to the analysis. And, directly on the subject of this meeting, uncertainty is a key complication. We have, in this analysis, socioeconomic uncertainties that overlay all the scientific uncertainties, compounding the problem. Finally, we have equity and social issues, which have a strong influence on the negotiating process and the final outcome, but are very difficult to factor into the analysis especially in the decision making stage.

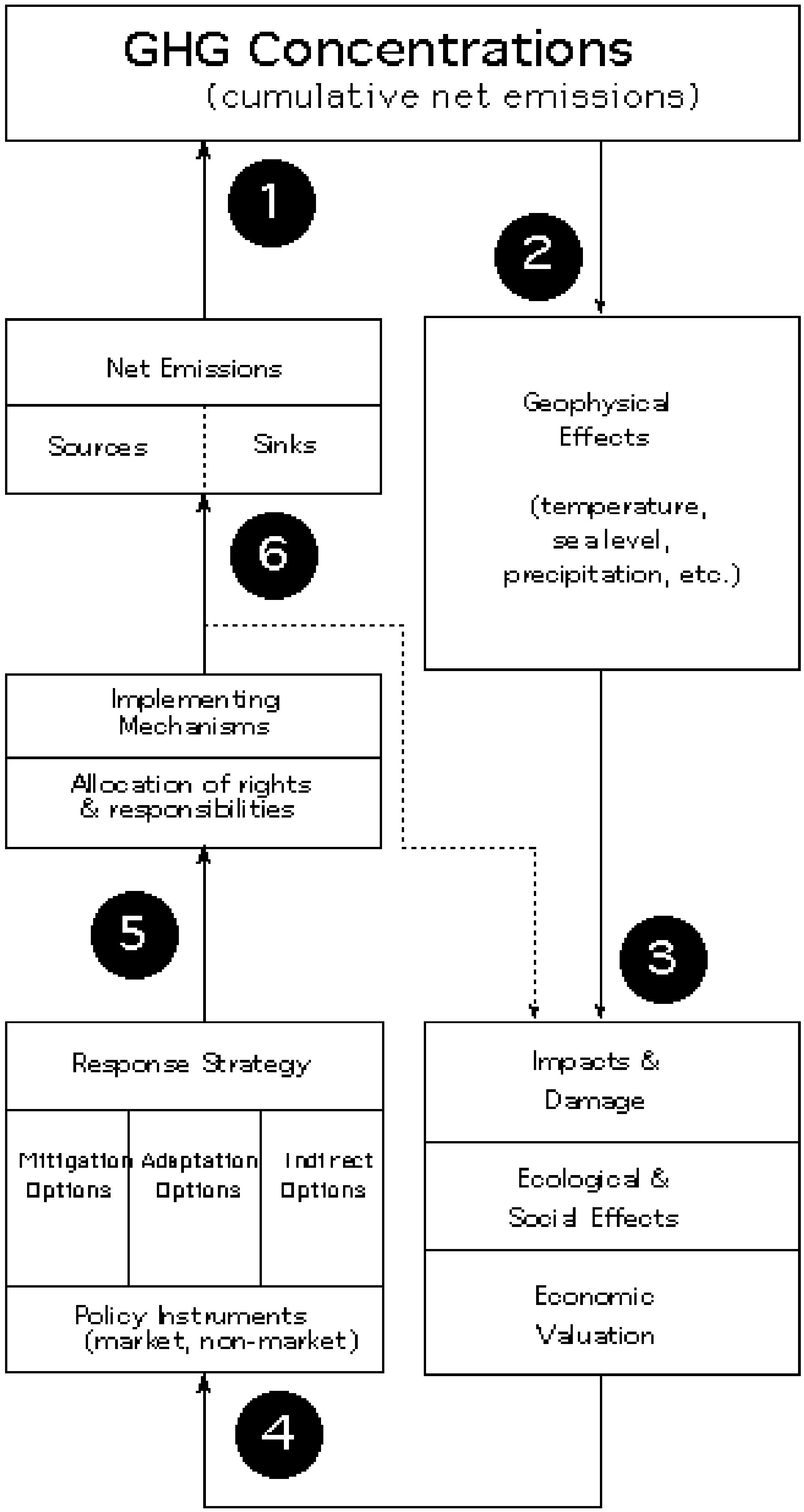


Figure 2.13
Chain of Causality

The most attractive climate change policies would be so-called “win-win” options, such as energy conservation, which provide significant in-country benefits as well as reductions in greenhouse gas emissions.

In the cost/benefit approach used to determine the globally optimal level of emissions, cumulative emission reduction is plotted against the cost of removing each ton of carbon from the atmosphere. Analogously, the marginal benefit curve is drawn based on the damage one avoids by removing one ton of carbon from the atmosphere. Theory says that we should carry out emissions reduction up to the point where the cost of removing the last ton of carbon is equal to the damages avoided (where the two lines cross). The problem is that we do not know the shape of these curves.

Suppose the damage curve was poorly defined, but we still had some idea of the costs of mitigation. Then, instead of the optimal reduction level approach mentioned above, we could use the "affordable standard" (which implies that we reduce emissions up to the point that we can afford to, say 1 percent of GDP per year). A more stringent criterion may be based on the "absolute standard" approach, which indicates that there is some benchmark or point beyond which any transgressions will result in very steeply rising costs and high risk of some catastrophic outcome.

We have, in this analysis, socioeconomic uncertainties that overlay all the scientific uncertainties, compounding the problem.

Munasinghe and colleagues feel that damages involve many more uncertainties and are far more difficult to estimate than costs, and as such, they assigned wider error bars to their damage estimates. Damages involve valuation of ecological impacts such as biodiversity loss which are very difficult to estimate and have a much higher order of uncertainty. Also, damages involve potential social impacts which may be so severe in some places that they damage the whole social fabric causing mass migrations, civil strife and conflicts that are impossible to quantify. For these reasons, Munasinghe feels that damage estimates are inherently less precise than mitigation costs estimates (which are technology dependent).

Table 2.14 summarizes the monetary damage estimates for a doubling of CO_2 on an annual basis. Ranges are upper and lower bounds based on the existing literature; no error bars are assigned here. Assumptions are for 560 ppmv CO_2 equivalent and a 2.5°C climate sensitivity.

Table 2.14
Annual Economic Damage for Doubling of
Pre-industrial CO_2 Concentrations
(560 ppmv CO_2 equivalent, 2.5°C climate sensitivity)

Region	Nominal Percentage of GDP
OECD	1 - 2 percent
Others (developing and transition	2 - 9 percent
World	1.5 - 2 percent

The damage costs range from $5 to $125 per ton of carbon emitted. Different values of discount rate (based on the notion that costs in the future should carry less weight than costs borne today), which is used in various calculations, cause damage costs to vary substantially. If a uniformly high discount rate is used, say 5 percent, the range of damage costs narrows to $5 to $12 per ton of carbon. However, environmen -talists and others argue that with a problem like climate change, we should use a very low discount rate implying that future damages should weigh heavily in the calculus today. What discount rate to use is the source of great controversy. While the choice of a discount rate is a source of dispute and not another cause of uncertainty, it does influence estimated damage costs significantly.

Regarding mitigation costs, some experts believe that CO_2 could be reduced to 1990 levels at negligible costs, using “bottom up” models which deal with individual behavior. Conversely, the top down, macroeconomic models project the costs of mitigation to be much higher, up to 1-2 percent of GDP. Nevertheless these top-down models give answers for different countries that vary quite widely. To hazard a guess from these results is a perilous task indeed. However, for decisions that must be made today, we should strive to give decision makers information on what is currently known, and at the same time seek to improve the state of our knowledge for future assessments. Climate change decision making is not a single event, but rather a process based on a heuristic approach which makes use of continuing improvements in human knowledge.

Some experts believe that CO2 could be reduced to 1990 levels at negligible costs, using “bottom up” models which deal with individual behavior. Conversely, the top down, macroeconomic models project the costs of mitigation to be much higher, up to 1-2 percent of GDP.

Technological Potential for Mitigation

Nebojsa Nakicenovic

International Institute for Applied Systems Analysis
Laxenburg, Austria

More efficient conversion of fossil fuels is one of the greatest areas of mitigation potential. About 30 percent efficiency gains are possible in the short run, and up to 60 percent in the long run, at relatively low or even zero cost.

Mitigation options are discussed primarily in the Working Group II IPCC assessment, in the chapters that deal with the energy sector (including transport) and the other consuming sectors. It is important to recognize that the energy system involves more than just the energy sector. One typical energy chain goes from crude oil to gasoline and through the consuming sectors to the energy services provided (distance traveled, illumination, heat, etc.). There was a high degree of agreement among the authors of these chapters that one cannot look at the energy sector in isolation, but must instead look at the whole energy system and its major driving forces, namely higher quality of final energy and energy services delivered.

Currently, the world requires about 9 gigatonnes of primary energy, used at an overall efficiency of about 70 percent which leaves 6.4 gigatonnes of final energy delivered. The carbon flows from this part of the system are about 2.3 gigatonnes of carbon released in the conversion from fossil fuels to final energy, with an additional 3.7 gigatonnes released in the conversion from final energy to useful energy and energy services. The mitigation potential, from the technological point of view, involves structural change in the future as we go to cleaner and higher quality fuels, so the share of emissions from the energy sector is likely to increase; therefore the importance of mitigation in this sector may be amplified.

Technical mitigation potentials were classified into the following groups:

1. More efficient conversion of fossil fuels

This one of the greatest areas of mitigation potential. About 30 percent efficiency gains are possible in the short run, and up to 60 percent in the long run, at relatively low or even zero cost, if one allows sufficient time for technological development. There was very broad consensus on this point.

2. Switch to lower carbon fuels

Typically, this involves a shift from coal to natural gas (a 40 percent lower carbon fuel which can also be more efficiently converted) which results in about a 50 percent mitigation. That there is large mitigation potential in this area was also a non - controversial point.

3. Decarbonization of fossil fuels

Technologies to remove CO2 from power plant flue gases include amine absorption, membrane separation, and centrifugal separation (which is very expensive and not yet feasible). Today, all of the decarbonization technologies are very costly, about $100 per ton of carbon removed. Taking carbon out of flue gases also reduces the efficiency of the power plant and raises the cost of electricity by approximately 50 percent. Also discussed were technologies for removing or reducing carbon from the fuels themselves. These include gasifying coal and steam reforming natural gas. These technologies are quite expensive, in the range of $200 to $400 per ton with current economics. And once carbon is separated it must be stored. One project in Norway involves injecting CO2 into underwater aquifers in the North Sea, which is cost effective given Norway's carbon tax. There are also projects which use CO2 for enhanced oil recovery or deposit it into depleted natural gas fields. One more controversial option is to deposit CO2 into the deep ocean; there are great concerns about the environmental impacts and costs of this. Sequestering carbon, particularly by biomass was also discussed by the group and there was agreement that it is probably better to use biomass to produce energy than to sequester carbon.

4. Switching to zero-carbon options such as nuclear and renewable sources

The general perception in the literature was that the costs of nuclear power can be expected to increase in time due to many problems. The conclusions was that, at least in the short term, the potential for mitigation from nuclear power is not very large. With regard to renewables, there was a very high variation of costs, in the $50-$500 range, making these sources very attractive in some parts of world and very expensive in others.

Energy Efficiency

The potential for increases in energy efficiency is enormous, since the current end use technologies are very inefficient, in the range of 10-20 percent efficiency. The main issues are the timing, cost development of new technologies, and consumers' behavior. A surprising result of the research was that Eastern Europe used to have higher energy efficiency due to structural components, such as more people traveling by bus and using large scale systems such as district heating. The shift from collective to individual transport, etc., has brought energy efficiency down.

Over what time scale could we expect a doubling of energy efficiency? Historical data of U. S. energy intensity reveals that it took about 70 years to double the energy efficiency of the economy. As far as carbon intensity, OECD economies evolved such that it took 40 to 50 years to halve carbon intensity. There is a learning curve; the more we emit, the better we know how to emit less, Nakicenovic says. We won't run out of potential; the question is how fast we can achieve part of the potential.

Flagging the carbon intensity of primary energy in several countries, France halved its carbon intensity from 1960 to 1990 through its increased reliance on nuclear power and Japan reduced its carbon intensity by using more methane. The rapidly growing, extremely carbon -intensive economies of the large developing countries, China and India, are cause for great concern. There was a consensus that a shift away from fossil fuels will not happen due to resource constraints. There is plenty of inexpensive fossil fuel. The shift must occur for other reasons.

Today, all of the decarbonization technologies are very costly, about $100 per ton of carbon removed. Taking carbon out of flue gases also reduces the efficiency of the power plant and raises the cost of electricity by approximately 50 percent.

There was a fair amount of controversy over the potential contribution of renewable energy sources, but it appears that renewables could contribute about half of the today's total global energy use by 2025, and in the long run, gigantic potential was identified. The issue is not so much absolute potentials but rather the transition times and dynamics. The large diversity in opinions made it impossible to state specifically which technologies offered what potentials to achieve carbon reductions. The best the group could do was come up with a statement that carbon emissions could be reduced by 2 gigatonnes per year by 2100 through the use of renewable energy.

Regarding the dynamics of the energy system, Nakicenovic points out that the entire system is very interrelated. For example, we need not only a process to produce hydrogen but also the whole system that supports it. In addition, the typical lifetimes of energy system components are in the range of 30-50 years. Thus, in the 50- to 100 -year time frame of this assessment, most of the energy system infrastructure would be replaced twice. This demonstrates that some mitigation would be achieved through natural technological evolution. If we want a much lower emission world, then stronger policies also come into play. For example, there is a strong relationship between a sustained increase in price and a decrease in energy intensity. Incidentally, the time frame for such structural changes, coming about by around 2050, is the same as the time frame predicted for major changes in the climate system.

A shift away from fossil fuels will not happen due to resource constraints. There is plenty of inexpensive fossil fuel. The shift must occur for other reasons.

A final point was that research, development and demonstration of new, lower-carbon technologies is needed. Though up-front costs will increase, this will offer long term benefits and diffusion potential. Immediate investments in research and development are needed if these technologies are to reach usefulness in 50 years.

Different Levels of Treatment of Uncertainty in Risk Analysis and Aggregation of Expert Opinions

Elisabeth Paté-Cornell

Stanford University
Stanford, California

Paté-Cornell comes from the engineering systems analysis and risk analysis tradition which uses probability more than the environmental risk analysis tradition. The level of sophistication with which uncertainty is treated should depend on the use that will be made of the information. Risk analysis is necessarily a snapshot of the information available at the time a particular decision is made. The level of treatment depends on both the importance of the risk and the costs of the mitigation measures. Sometimes, a "level zero" treatment of uncertainty is all that is needed, simply asking: is there a risk or not? At a slightly higher level, we can ask, what is the maximum loss that can occur? And if we believe that we can afford to incur this loss, the analysis can end there. The trouble is that often what we assume to be the worst case can be made still worse by adding other details to catastrophic scenarios, so this level is generally not sufficient for significant risks.

Risk analysis is necessarily a snapshot of the information available at the time a particular decision is made. The level of treatment depends on both the importance of the risk and the costs of the mitigation measures.

The most often used level of sophistication in uncertainty analysis in environmental health risk assessments is what Paté-Cornell calls "quasi-worst case" or "plausible upper bounds." This is the level of sophistication, for example, in carcinogen risk assessment by the Environmental Protection Agency. The problem with this approach, she says, is that it does not enable one to rank the risks because the probable level of overestimation is likely to vary from one risk to another. If one does not have unlimited funds to spend, and has to set priorities, it is important to be to able determine which risk is likely to result in the greatest losses. Using this method, one cannot be sure that the plausible upper bound of risk one, which may be greater than the plausible upper bound of risk two, will actually result in a greater value of losses in terms of expected values.

Therefore, there is some need for a "central value" of the potential outcome distribution. The third level of uncertainty analysis in risk assessment generally involves such "best estimates" or "central values." In many of these risk analyses there are several possible models and mechanisms, and for each of these, a spectrum of possible parameter values. The approach to a "best estimate" analysis is to pick the most likely among these possible mechanisms, but often, the most likely thing is that nothing happens. The problem with this approach is that the most likely mechanism may not be the one that would result in the kinds of losses that one might be most concerned about in designing public policies or private risk management decisions. Because of all these shortcomings, the engineering field has developed the methods of probabilistic risk analysis.

In probabilistic risk analysis, one begins by dividing a complicated system into subsystems, analyzing the functions that have to be performed in order for the system to work, evaluating external events and loads, and computing the probability of failure of the whole system based on the probability of failure of its components as well as external factors. An example of this is

the seismic risk analysis that civil engineers perform in which there are two basic problems to resolve. First, what is the seismicity and the probability that the ground will move with different levels of severity in a given time period? Second, given the seismicity, what will happen to structures. In such a problem, there are several kinds of experts involved in the assessment, and probabilistic methods are needed to combine these pieces together.

Very often fundamental uncertainties are encountered about the mechanisms themselves (for example, seismic mechanisms at a given site). Some of these are epistemological (fundamental) uncertainties and others are aleatory (arising from randomness or variability) uncertainties. When it comes to treating epistemic uncertainties in the face of limited knowledge, the heart of the problem is often how to treat the disagreements among experts. Paté-Cornell discussed several methods for aggregating expert opinions into single probability distributions for the spectrum of possible mechanisms, and for the parameter values for each model.

Very often fundamental uncertainties are encountered. Some of these are epistemological (fundamental) uncertainties and others are aleatory (arising from randomness or variability) uncertainties.

The Delphi method is an iterative method which begins by interviewing experts separately. The results are collected and fed back to the experts in order to give them the opportunity to modify their opinions based on their colleagues' answers. This method generally leads to rapid convergence and consensus, but the answer may not bear much resemblance to what one would obtain by combining all of the evidence. This simple, purely interactive method is a social mechanism whereby experts are brought together to share their evidence and mental models. This allows them to constitute the larger body of evidence and to enrich each other's knowledge with new information. They are essentially left together until they can agree upon an answer.

The purely analytical method involves gathering the opinions of experts separately and then weighting their models, for example by Bayesian probabilistic analysis methods, involving a "super-expert" who determines the probability of the phenomenon given what each of the experts says. One problem that can arise is that the experts are generally not truly independent and it is difficult in practice to evaluate their influences on one another. Another problem is that they may not share the same evidence base and have no opportunity to exchange information in this method.

Another (probably better) method is thus a variation on this purely analytical one in which models are weighted according to some more subjective means than the "super-expert." This process is both social and analytical. It is analytical in that it is based on Bayesian logic, and it is also a social process in that the experts come to a common conclusion in an interactive manner. The result (and the success of this exercise) depends upon the kind of people in the group, the way they are allowed to intervene, and in what manner they are allowed to support their views, evaluate other models, and integrate that information to come up with the group's aggregation of material.

Communicating Estimated Climate Change Impacts on Agriculture

John Reilly

U. S. Department of Agriculture
Washington, D. C.

Different perspectives across scientific disciplines give rise to differences in how to communicate the risks of climate change. Both within and among disciplines, individual researchers bring different sets of experiences with regard to how information they generate will be used. Most of this background experience and expectation of how the information will be used is never addressed directly and instead may be reflected in debates about the accuracy of estimates and what types of information to report. Some of the key background perspectives scientists bring to the table of scientific assessment but rarely discuss directly are: who is the audience, what does this audience already know, and what will this audience do with the information?

Different perspectives across scientific disciplines give rise to differences in how to communicate the risks of climate change.

For the Intergovernmental Panel on Climate Change (IPCC), the audience is obviously the world community of policymakers, yet this is a diverse group. Further, there is a general recognition that the top level ministers will not read the entire report and likely will not even read the Summary for Policymakers themselves. So in addition to different perceptions scientists have about who the direct audience is for the report, they have different views about how information in the report will reach the ultimate audience. What “policy” is, and where, when and by whom it is “made,” are also issues on which scientists bring many different perspectives.

Following are three caricatures of the views different scientists bring to the table and the implications for the type of information they see as important to represent in a document like the IPCC report.

View One

Global negotiators on climate change, representing their own country’s interests, must make a decision about the best level of climate change control policy to undertake at this time. Once an efficient level of control is determined, negotiators can proceed to design a control strategy that addresses differential distributional consequences of damages and the costs of control across countries. From this perspective, getting an estimate of the global impact on agriculture is the first order of business for a chapter dealing with agricultural impacts. It is important that this estimate represent the best guess scientists can make at this time, even if it is highly uncertain. Moreover, from this perspective, it is useful to denominate agricultural impacts in units that can be added to or easily compared with the units being used in other sectors.

View Two

Policymakers at all levels and the public are ill-informed about the risks of climate change. Policymakers representing different interests and the general public within each country will or should have a voice in formulating a national view on the risks of climate change. The decision process is highly interactive and somewhat chaotic; thus it is not possible to identify or separate different types of policymakers who are responsible for international negotiations. Impact analysis helps determine those sectors/regions/people who are most vulnerable to climate change so that actions can be taken to reduce their vulnerability. Alerting people to the potential negative consequences of climate change is the first priority. It is important to relate impacts in terms that vulnerable groups or those responsible for them can understand. There is little benefit to seeking common units with which to report impacts.

View Three

Policy will not be made until the scientific evidence is clear. The scientific staffs of governments will review the quality of the scientific evidence presented in the IPCC process and make recommendations to policymakers on the basis of this evidence. From this perspective, there is a high priority on presenting issues of methodology and background on data quality rather than simply summarizing findings. Presentation of material in such a way that it is credible to scientific peers is essential because this is the principal audience who will interpret the material for others. The focus is on presenting variables and information that science has investigated and on which there is strong evidence rather than on answering questions that arise from outside the community of scientists investigating the problem.

These caricatures of background views on how IPCC information will be used to illustrate the considerable potential for conflicting opinions on what to include and how to present material. This conflict intensifies when it comes to writing a 10- to 20-page policymakers' summary.

These caricatures of background views on how IPCC information will be used illustrate the considerable potential for conflicting opinions on what to include and how to present material.

Points of Debate

These issues can further be illustrated with some of the specific debates that arose in preparation of the agricultural impacts chapter of the IPCC Working Group II report.

One major point of debate was how to address the issue that estimates of the net impacts of climate change on a global level were small but the regional effects could be large. Some believed that the net global effects were meaningless and should not be reported. There was skepticism about the global results from both sides; some felt they overestimated damages while others felt they underestimated them. In the end, the executive summary was worded very carefully to state that "global agricultural production can be maintained relative to baseline production under climate change as expressed by general circulation models (GCMs) under doubled CO2 equilibrium climate scenarios." The reader is left to judge whether or not this climate scenario is reasonable.

The chapters' authors made no attempt to make a statement about the dynamic path of impacts or to identify the period of time in which these GCM scenarios might occur. These studies were based on climate scenarios with global mean temperature changes of 4.0°, 4.2°, and 5.2°C. This range of temperature changes is on the high end of what Working Group I concluded to be the equilibrium climate sensitivity, and also represents global mean temperature changes that seem unlikely to occur before 2100.

If the chapter had been driven more by “View one” described above that policymakers needed scientists’ best judgment of what might happen through 2050, which was the general guidance for the Working Group II effort, the author team might well have speculated on the transient damages. Instead, the author team chose not to speculate even though this information is critical for near-term decisions. The lack of real studies of adaptation to transient climate scenarios and the lack of realistic transient scenarios (e. g. , the limited scenarios available did not include the cooling effect of sulfate aerosols) was the basis for this unwillingness to speculate.

Adaptation potential and capability was also a significant topic of controversy. This was reported as an unresolved controversy, recognizing that historically, farming systems had adapted to a variety of stresses, but recognizing that specific research addressing the future rate of climate change had not been conducted. The chapter did go on to assert that adaptation was likely but the extent depended on a number of factors. The overall wording of this section was a compromise between “View one” and “View three.”

General guidance for all chapters was to present information on thresholds, sensitivities and vulnerabilities. In early rounds of writing, there was skepticism of the usefulness of reporting information of this type.

As writing progressed, a fair amount of material was added with regard to broad temperature ranges for different crops, indicating that productivity would fall off sharply beyond these limits. Specific examples based on crop modeling studies were cited but no discussion of the extent of these thresholds or the likelihood that they would be encountered under specific scenarios of climate change were reported. In this regard, the report adopted “View two,” warning readers of potential hazards without indicating their likelihood. Technical reviewers sometimes focused on this type of reporting as a bias toward reporting negative effects.

There was a debate about whether the chapter should be about the basic agronomic relationships of temperature, precipitation, and carbon dioxide or about the results of models that simulated specific impacts. Various technical reviewers offered the idea that either (1) only the basic agronomic relationships should be reported because this was sound science whereas the models used to simulate impacts were not adequately validated under the simulated conditions or (2) only the modeling results should be reported because the basic agronomy results provided nothing new that could not be found in relatively basic texts on crop/climate interactions. In part this was a debate between those holding “View one” versus “View two,” but more fundamentally, it was a debate about what was admissible science.

There was a debate about whether the chapter should be about the basic agronomic relationships of temperature, precipitation, and CO2 or about the results of models that simulated specific impacts. Fundamentally, it was a debate about what was admissible science.

Ultimately, the chapter went relatively far in defining the concept of vulnerability. The chapter defined it as “the potential for negative consequences that were difficult to ameliorate through adaptive measures given the range of possible climate changes that might occur.” Identifying a region or population as vulnerable is thus not a prediction of negative consequences of climate change. Rather it is an indication that across ranges of possible climate changes, there are some climatic outcomes that would lead to relatively more serious consequences for that region than for other regions.

It was also recognized that vulnerability could be defined for different aspects of agriculture: yield, farmer income (in terms of regional, national or global economic vulnerability), or in terms of vulnerability to hunger. Vulnerability also depends on the scale at which one examines the problem. Here the chapter attempted to provide a relatively precise definition for a term that is used vaguely in the literature. As defined, vulnerability combines the probability density

function for climate and the damage function that relates impacts to varying levels of climate change to generate a loss function.

In practice, the chapter focused on identifying those populations that were subjectively judged to be vulnerable to hunger or famine. While in principle, the probable range of climate is important in this determination, the judgment that went into defining vulnerable populations largely viewed this probability function as reasonably flat across a wide temperature range and with roughly equal probability of precipitation increase or decrease. As such, specific climate model results played little role in the determination. As vulnerability was used in the chapter, it was roughly consistent with “View two” as described above.

A general problem the chapter authors faced was recruiting contributors. About 200 separate quantitative studies were reviewed. Given the limited participation, there was not time to critically assess the merits of individual studies. In this regard, chapter authors mainly summarized the findings and focused less on critiques of the methods and approaches. Berk suggested that there are formalized meta analysis techniques for summarizing and aggregating studies that could be helpful in validating studies.

Vulnerability combines the probability density function for climate and the damage function that relates impacts to varying levels of climate change to generate a loss function.

The chapter spanned a wide range of issues, from understanding crop/climate/insect interactions and the science of carbon dioxide fertilization to the socioeconomic conditions under which hunger occurs and the economic modeling of agricultural trade. This added to the difficulty in enlisting experts with deep knowledge in these many subject areas because in the end, an expert’s contribution would likely represent only a few paragraphs.

Reilly stressed that future research into agricultural impacts should involve the development and broad application of integrated agricultural modeling efforts, development of the capability to readily simulate agricultural effects under transient climate scenarios so as to better address the cost of adjustment, and evaluation of the effects of variability and changes in extreme events.

How to Approach Assessing Climate Impacts on Animals

Terry Root

University of Michigan
Ann Arbor, Michigan

The issue of potential effects of climate change on animals (except fish) was entirely left out of the IPCC process and Root argues that this must be remedied. Models that deal with these issues are very immature at this time. Why are we just starting to look at this problem? One reason is that there are strong controversies in ecology regarding what factors shape animal species' ranges. Root believes that both environmental factors and biotic interactions are important, but the conventional wisdom in the community for at least the last 30 years has been that biotic interactions are dominant in shaping species ranges.

Scale is another significant problem. Most ecological studies are done on areas the size of a tennis court. On this small scale, biotic interactions such as competition and predation can be seen, but not environmental factors such as climate. The dogma has been that it is impossible to understand ecology on a large scale because the biotic interactions cannot be observed at this level. It has thus been widely assumed that animal ecology can not be done on the continental scale, though this scale is needed to observe environmental factors like climate.

Bucking this trend, Root has begun a study of the effects of climate on passerine birds in North America. Using National Audubon Society Christmas Bird Count data, she has plotted maps for all species that winter in North America. These data begin in 1900 and identify both the species and the number of individuals observed by volunteers. Using ten years of these data and sixty years of climate data, Root plotted maps for all species that winter in North America. For the purposes of this talk, the edge of the distribution range for each species is of central importance. Root wanted to see how strongly environmental factors such as climate were associated with distribution edges, in hopes of challenging the notion that distribution edges are determined strictly by competition and predation. She finds that indeed, there is a strong association between climatic factors and distribution edges, particularly northern edges.

A large percentage of the 250 bird species studied have their northern ranges associated with climatic factors, most often temperature (defined as average minimum January temperature). In addition, some species ranges are largely determined by vegetation, which is itself strongly related to climatic variables. For yet other species, their ranges are associated with both temperature and vegetation. The possibility exists that with global warming, as the temperature rises, species whose ranges are related primarily to temperature are going to move north. But those related to both temperature and vegetation cannot move until the vegetation moves. So communities may be torn apart and some species may be driven to extinction. This is a concern in terms of pest control (certain bird species eat certain insects) as well as seed dispersal (some bird species transport particular seeds).

Species whose ranges are related primarily to temperature are going to move north. But those related to both temperature and vegetation cannot move until the vegetation moves. So communities may be torn apart and some species may be driven to extinction.

How to acquire animal data on a continental scale is a significant problem. With no money for such research, volunteers have traditionally been relied upon for data collection. But legislation currently before Congress could disallow the use of data collected by volunteers. This prohibition has been introduced as part of the proposed reauthorization of the Endangered Species Act (ESA) and is based on the notion that volunteers from organizations such as the Audubon Society have a vested interest in saving species, making data they collect invalid. This issue will be decided in the near future and Root makes the point that volunteer data are critically important to research such as she describes here. As further evidence of this, she points to additional volunteer data used in her most recent analysis. In this study, Root plotted 30 years of data for arrival dates for migrants and found a detectable signal suggesting that some migratory birds are returning earlier as temperatures warm.

Root plotted 30 years of data for arrival dates for migrants and found a detectable signal suggesting that some migratory birds are returning earlier as temperatures warm.

Given her conclusion that many bird species' northern ranges are associated with climate, the question arises: why are animals' ranges associated with environmental factors? Each animal species has a metabolic response curve. For any given animal, the thermal neutral zone is the temperature range in which the animal doesn't have to raise its metabolism to cool or warm itself. The animal can also live in the wings of this distribution, outside of its optimal range, where it needs to expend energy to warm or cool itself. But when the temperature gets hotter or colder than the zone of tolerance for that animal, it will die or move. This shows that temperature is very important to animal ranges. Species abundances are also highly determined by this curve, as a species will be most abundant in the optimal temperature range, less abundant in the wings, and not exist at all outside of its zone of tolerance.

For some species, the absolute minimum temperature is the key factor, while for others, it's other measures. This is species-dependent for many reasons. One reason is that microclimate effects impact species differently. Another is that one species may simply not be able to physically survive below a certain temperature, while another's survival is based upon whether it has stored up enough fat to last through a cold spell. In some species of reptiles, temperature determines sex ratios. In turtles, for example, the sex of the young is determined by the temperature at which the egg is incubated; higher incubation temperatures yield more female turtles.

Of the 250 bird species Root is looking at, winter physiology studies have only been done on 14. From these studies, she was able to find the basal metabolic rate of the species, the slope of the metabolic response curve, the temperature at the edge of the distribution, and the lower critical temperature (the temperature at which metabolism has to increase in order for the animal to stay warm). From these, she calculated the metabolic rate at the northern boundary for each species. She found that for a little over 50 percent of the species, the northern boundary metabolic rate was equal to 2.5 times the normal basal metabolic rate. This gives us a relationship that can be used in modeling studies; for those species limited by temperature, we can determine the northern range of the species if we know its metabolic rate.

Basal metabolic rate is strongly affected by food availability which in turn is strongly influenced by day length (for foraging). (Passerine birds have to forage all day to accumulate enough fat to survive through the night.) From mid-December to the first of March, Root studied birds at eight sites paired on one hand by latitude (same day length) and on the other by average minimum January temperatures. The question she sought to answer was: are individuals more closely related between sites that have the same temperature or sites that have the same day length? She found that they were more temperature related, suggesting that global warming could have important effects.

Ecological dogma is that bird distributions don't shift, but Root says this is false. Some of her work has thus focused on showing that birds can and do actually move. Temperature time series data for one year show that birds do go north when its warmer and further south when its cooler and that they are indeed shifting in response to climate. This implies that there will be changes in response to global warming. Some birds move when the temperature gets outside their zone of tolerance, while others die. This is species dependent. For example, the eastern phoebe moves, while the field sparrow dies.

Root is beginning to broaden her study to include other organisms, in particular some amphibians and reptiles. It appears that the ranges of the American alligator and the broad headed skink are also shifting. Organisms that are slow dispersers seem to be most at risk. There is a need to add the factor of land use into this study, Root says. For example, if alligators need to move due to climatic change, will changes in land use, such as farm fields, prevent their needed movement? Climate maps and land use maps are needed to further the study in this and other important ways.

A number of factors are impeding progress on the study and modeling of climatic effects on animals. There is a need for historic climate data (past 30-35 years) at a half degree grid, altitudinally adjusted (time series data). There is also a need for general circulation model data downscaled to a half degree grid, at appropriate altitudes. Abundance data on a continental scale exists only for birds at this time; distribution data is available for other animals, but no abundance data exists. Vegetation data and forecasts are needed as well, since vegetation is important to most animals. Complexity among species (many things are species specific) adds to the difficulty in modeling impacts on animals. The fact that there are no obvious feedbacks to climate from animal impacts is another deterrent. And finally, the widely held assumption that species are largely substitutable is an obstacle to furthering research in this area.

Temperature time series data for one year show that birds do go north when its warmer and further south when its cooler and that they are indeed shifting in response to climate. This implies that there will be changes in response to global warming.

Climate Change Detection and Attribution Assessment in the IPCC Proceedings

Benjamin D. Santer

Program for Climate Model Diagnosis and Intercomparison
Lawrence Livermore National Laboratory, Livermore, California

Progress in pattern-based methods involves "fingerprints" that can help discriminate between different possible causes of climate change. Different forcing mechanisms have different patterns of effects.

Santer defines climate change detection as the process of demonstrating that an observed change is highly unusual in a statistical sense, and offered the analogy of detecting a fever by measuring body temperature. Attribution of climate change is the process of establishing cause and effect, like diagnosing the cause of the fever through a complete suite of medical tests. We are now in the process of using complex diagnostics to diagnose the causes and effects of climate change.

Major advances have been made since 1990 in defining a human -induced climate change signal, defining the noise of natural climate variability, and increasing application of pattern-based methods with greater relevance for attribution. In regard to the first of these advances, identifying the human-induced climate signal, early models showed spatially coherent warming with more warming toward the poles. Then, full ocean dynamics were added, simulating the Earth's redistribution of heat. The latest advance is the addition of the direct effects of sulfate aerosols in the models. In this ongoing evolution, new anthropogenic components will be incorporated including indirect aerosol effects, tropospheric and stratospheric ozone, and aerosols from biomass burning.

Progress in defining "noise," or the natural variability of the climate system comes from new instrumental data as well as reconstructions of past climate. Progress in paleoclimate data comes mainly from recent efforts to meld information from diverse proxy sources rather than simply looking at tree rings, ice cores, corals, or borehole temperatures individually. A 1992 study by Bradley and Jones is an example of such melding. Further development of statistical models as well as physically -based climate models with full ocean general circulation are proving helpful as well. Comparing modeled estimates of the pattern and amplitude of natural variability with observations from paleoclimatic data show that model-based estimates of climate 100 years in the future are generally less than those derived from paleoclimatic data, but progress is being made. Knowledge of changes in solar forcing is now also being incorporated into general circulation models (GCMs).

Progress in pattern-based methods involves "fingerprints" that can help discriminate between different possible causes of climate change. Different forcing mechanisms have different patterns of effects. For example, tropospheric warming resulting from increases in carbon dioxide and other greenhouse gases would be accompanied by stratospheric cooling, while warming resulting from an increase in solar output would not be accompanied by stratospheric cooling. Since data show that the stratosphere has indeed cooled, this is part of the fingerprint that helps identify the cause. Another pattern-based indicator is that models which incorporate both greenhouse gases and aerosols yield a better correlation with observations, showing both

the cooling and warming pattern observed. The combined CO_2 and aerosol signal exceeds the statistical significance threshold in the last two decades, when it would be expected to, and the regional pattern is also as expected (better correlation over industrialized regions).

A methodological advance in understanding comes from new multivariable representations of the climate change signal, such Tom Karl's climate index, rather than a focus on global mean temperature alone. Karl et al., 1995, included in their climate index: asymmetric increases in maximum and minimum temperature, cold season precipitation, severe summertime drought, extreme one-day precipitation events, and day-to-day temperature variability all of which yields a more comprehensive indicator of climate change than does global mean temperature.

Santer reviewed an example of a detection and attribution study showing "coarse" (global mean retained) and "fine" (global mean subtracted) structure of an aerosol signal from Taylor and Penner, 1994. Both observations and modeled changes for increased CO_2 only (without aerosols or ozone) show both stratospheric cooling and tropospheric warming. The observed patterns of vertical temperature changes are seasonally robust. Even better agreement between the model and observation is achieved when models include combined effects of CO_2 , aerosols, and ozone. Further confidence in model results comes from the fact that there are similarities between CO_2 and aerosol signals simulated by different climate models.

A methodological advance in understanding comes from new multivariable representations of the climate change signal, such Tom Karl's climate index, rather than a focus on global mean temperature alone.

Conclusions for detection are that the Earth is warming and that its mean temperature is warmer than it has been in many centuries. For attribution, conclusions are that observed geographical patterns of temperature change at the Earth's surface are similar to model predictions that incorporate combined greenhouse gas (GHG) and sulfate aerosol effects. Observed patterns of atmospheric distribution of temperature change are also similar to model predictions that incorporate the effects of GHGs, sulfate aerosols, and stratospheric ozone depletion. These pattern correspondences tend to increase with time, and model patterns are different from those due to natural variability. Observations and model predictions generally agree in overall magnitude and timing of change as well.

Major scientific uncertainties still exist with regard to estimates of the magnitude, pattern and evolution of different forcings. GHGs, direct sulfate aerosol effects, and stratospheric ozone effects are included in current analyses, but uncertainties remain about these estimates. Indirect effects of sulfate aerosols, other anthropogenic aerosols, tropospheric ozone, and solar and volcanic influences have thus far been left out of the analysis, posing additional uncertainty. In regard to model predictions of response to forcing, key uncertainties involve the parameterization of clouds and ocean vertical mixing. Estimates of the natural variability of climate on time scales of decades to centuries is another uncertainty, as is the quality of observed data.

If climate models have errors, how can we have confidence in their results? Confidence in the emerging identification of a human-induced effect on global climate comes largely from comparisons of modeled and observed patterns of change. The results of such studies rest mainly on pattern similarities at very large spatial scales, e. g., the temperature contrast between the hemispheres and the temperature differences between the stratosphere and the troposphere. It is at these large spatial scales that we have most confidence in the reliability of model results.

Additional work is needed to give us a more complete description of human-induced climate forcing through inclusion of indirect sulfate aerosols, other anthropogenic aerosols such as those from biomass burning, and a quantification of tropospheric ozone changes. More realistic

estimates of climate response can be achieved through better representation of clouds and precipitation and the incorporation of land surface processes in models. Application of more sophisticated signal processing techniques in detection and attribution studies should involve multi-variable fingerprints (such as Tom Karl's work mentioned above) and optimal filtering techniques to enhance the signal to noise ratio. We can also achieve a better understanding of the systematic error in models through model intercomparison projects.

Major scientific uncertainties still exist with regard to estimates of the magnitude, pattern and evolution of different forcings.

Empirical Estimation of Climate Sensitivity and Anthropogenic Sulfate Forcing

and

When We Don't Know the Costs or the Benefits: Adaptive Strategies for Abating Climate Change

Michael Schlesinger
University of Illinois
Urbana, Illinois

Schlesinger discussed methods of estimating climate sensitivity (the change in equilibrium surface air temperature that results from a doubling of carbon dioxide) and sulfate aerosol forcing. He summed up the uncertainties in using general circulation models (GCMs) for estimating climate sensitivity as follows: "One can get any answer out of a GCM depending on how one treats the physical processes that are not explicitly resolved. Ergo, we cannot trust these models to correctly estimate this most important climatic quantity." To illustrate this point, Schlesinger points to the model from the United Kingdom Meteorological Office, the standard form of which produces a climate sensitivity of 5.2°C; but by changing the way it treats clouds (specifically, how fast water falls out of ice clouds relative to how fast it falls out of liquid water clouds), the sensitivity changes drastically, from 5.2°C to 1.9°C.

One can get any answer out of a GCM depending on how one treats the physical processes that are not explicitly resolved. Ergo, we cannot trust these models to correctly estimate this most important climatic quantity.

In discussing problems with trying to deduce climate sensitivity using GCMs, Schlesinger points out that the present generation of GCMs resolves only two (the planetary and synoptic scales) of the 14 orders of magnitude of physical processes that determine Earth's climate. If we wanted to increase the resolution of one of these models by one order of magnitude in scale, computing time for the simulation would increase from about 500 hours to about 50 years. But even though we can not practically resolve these processes explicitly, we can not ignore them, since many are important to determining climate. Therefore, we parameterize them, meaning that we try to determine the statistical effects of these unresolved processes on the resolved processes using information only about the resolved processes. This involves making many assumptions about how nature operates, and how we parameterize important processes (e. g., cumulus convection) largely determines the climate sensitivity produced by the models.

Schlesinger then discussed a "sequential decision analysis" he undertook with colleagues at the RAND Corporation. They compared two near-term policies for dealing with climate change: a moderate policy (conservation only, characterized by a 30 percent decrease in energy intensity over 20 years) and an aggressive policy (involving fuel switching at a rate at which half of all fossil fuel-using facilities are replaced within 40 years). They assumed that in ten years we would learn exactly what the climate sensitivity is and what the climate target should be (the maximum allowable warming we want the system to undergo). The climate policy for the long term could then be adapted as needed at the end of ten years, depending on what was learned. Looking at long -term costs, the key conclusion of this analysis is that it is most important to learn the climate sensitivity because the cost of abatement changes primarily based on climate sensitivity and the climate target, and virtually not at all on whether the aggressive or moderate policy is chosen.

So, if determining climate sensitivity is the most important thing we need to know, and the models are not reliable for estimating it, how should we do it? One approach is to use past climate data to infer the climate sensitivity. There are several problems with this approach. First, past climate data are derived from proxies (such as tree rings and ice cores) rather than direct observational measurements. Second, in addition to knowing the climate changes, we need to know the forcing mechanisms. And third, the climate sensitivity value we seek is not a constant, but rather depends upon the climate perturbed. So if using past climate data is not the best approach, we are left with using instrumental temperature observations to determine the climate sensitivity.

Given this, Schlesinger discussed a simple modeling method for using the observational record of surface air temperature and interhemispheric temperature difference reported in the IPCC 1992 and 1995 assessments. He points out that there is a larger downward trend in the interhemispheric temperature difference from the 1995 data than from the 1992 data. Schlesinger fed these data, along with specified forcing due to greenhouse gases, tropospheric ozone, and solar output variations, into a simple energy balance climate model /upwelling-diffusion-ocean model. (The calculations were done with and without solar forcing and a large difference was found.) In this simple model, the Northern and Southern Hemispheres are forced separately, all feedbacks are included by the estimated climate sensitivity, the ocean sets the time scale for the response, upwelling occurs in non-polar regions and downwelling in polar regions, and deep water formation occurs in polar regions only.

The key conclusion of this analysis is that it is most important to learn the climate sensitivity because the cost of abatement changes primarily based on climate sensitivity and the climate target, and virtually not at all on whether the aggressive or moderate policy is chosen.

Several estimates of climate sensitivities were made for the different temperatures, forcings, and sulfate emission rates to determine which of these contributes most to the uncertainties; probability distributions were also figured for these estimated quantities. (To get radiative forcing due to sulfate aerosols, sulfate emission rates are multiplied by a model -determined parameter.) The simple model is run with a number of different specified sulfate forcings and climate sensitivities and the best fit between the observed climate data and the model is determined to be the solution. Because sulfate forcing is negative (it partially compensates for greenhouse-gas forcing), the higher it is, the higher the climate sensitivity. The sulfate forcing is determined by using data on the interhemispheric temperature difference. The climate sensitivity is determined by the global-mean temperature.

This study produced four sets of results: one for the 1995 IPCC data without solar forcing, one for the 1992 IPCC data without solar forcing, and one for each of these data sets with solar forcing. As far as the interhemispheric difference in warming, the model reveals the same trend as that observed. Because sulfate aerosols are more concentrated in the Northern Hemisphere, it will warm less rapidly than the Southern Hemisphere. This becomes more true as time goes on and emissions increase. The instrument observations used in this study begin in 1856.

Results using the 1995 data without solar forcing show a climate sensitivity of over 5°C, while results using the 1992 data without solar forcing show a sensitivity of about 3°C. The significant change in results can be explained by two factors. First, substantial corrections were made to 19th century measurements since 1992 (the effect of three additional years of measurements was negligible). Secondly, the 1995 data show a much bigger downward trend line, roughly by a factor of three, in the interhemispheric temperature difference and to explain that, we need more sulfate forcing, and thus a larger climate sensitivity, to explain the global mean.

The results also make clear that including the sun in the calculation makes an even larger difference. Solar irradiance variations have been in phase with the temperature record during

the 20th century, so solar forcing is positive during most of the instrumental temperature record, and adding a positive forcing brings the sensitivity down. When we add this, sensitivity comes down by factor of two (adding a factor of two uncertainty to our estimate of climate sensitivity). We don't know what the sun's role in climate change is; this is a problem that is not going to go away, and it makes a significant difference in climate sensitivity.

All of the variations in the climatic record cannot be explained wholly by the forcings we are aware of, which brings us to a discussion of natural variability or "noise." Schlesinger and colleagues performed a singular spectrum analysis on the temperature record which reveals an oscillation about half the length of the record. They tried to reproduce this and found that they could mimic one hemisphere at a time but not both simultaneously. They then developed a "bootstrap" resampling technique to generate 900 alternate (surrogate) instrumental observation records. For each of these simulated realizations, they re -did the estimation problem. The bootstrap technique does well at reproducing the power spectrum.

We don't know what the sun's role in climate change is; this is a problem that is not going to go away, and it makes a significant difference in climate sensitivity.

Schlesinger also showed marginal and cumulative probability distributions for climate sensitivity and sulfate aerosol forcing. The 50th percentile for climate sensitivity is about 3.0°C. The 50th percentile for sulfate forcing is about &endash;0.76 watt per square meter.

Adaptive Strategies for Climate Change

Finally, Schlesinger returned to further discussion of the analysis performed with colleagues at the RAND Corporation detailed in a paper titled, "When We Don't Know the Costs or the Benefits: Adaptive Strategies for Climate Change" (Lempert et al ., 1996). While most quantitative studies of climate-change policy attempt to predict the greenhouse-gas-reduction plan that will have the optimum balance of long-term costs and benefits, Schlesinger and colleagues believe that the large uncertainties associated with the climate change problem can make the policy prescriptions of this approach unreliable. In this study, they construct a large uncertainty space that includes the possibility of large and/or abrupt climate changes and/or technological breakthroughs that radically reduce abatement costs. They perform computational experiments on a linked system of climate and economic models to compare the performance of an adaptive strategy (one that makes midcourse corrections based on observations of the climate and economic systems) and two commonly advocated "best-estimate" policies (described earlier in this summary) based on different expectations about the long-term consequences of climate change.

In this study, they construct a large uncertainty space that includes the possibility of large and/or abrupt climate changes and/or technological breakthroughs that radically reduce abatement costs.

They find that the best-estimate policies perform well in the respective regions of the uncertainty space where their estimates are valid, but can fail severely in those regions where their estimates are wrong. In contrast, the adaptive strategy can make midcourse corrections and thereby avoid significant errors. The adaptive strategy doesn't do as well as either prescriptive policy IF that prescriptive policy was right, but the penalty for being wrong is much, much smaller. Schlesinger explains that if we knew the climate sensitivity, the damage that could be expected to result from it, and the rate of technological innovation, we could choose the optimum policy. But without knowing these things, we must create a strategy that allows adaptation based on new learning. Such a policy is analogous, he says, to sending a probe to Jupiter with the ability to make course corrections en route rather than just sending it off and saying, "bon voyage, call us when you get there."

The authors of this study believe that its results suggest that society might usefully recast its view of the climate-change-policy problem. Currently the debate focuses on targets and timetables for the optimum level of near-term GHG reductions that should or should not be set. The research community views its task as improving the accuracy of the predictions of the future which will provide policymakers with better estimates of the optimum level of emissions reductions. These authors suggest that climate change be viewed instead as a problem of preparing for unpredictable contingencies. Society may or may not need to implement massive reductions in GHG emissions in the next few decades. The problems for the present include developing better options for massive reductions than those currently available and determining what observations ought to trigger their implementation.

Reference

Lempert, R. J., Schlesinger, M. E., and Bankes, S. C.: 1996, When we don't know the costs or the benefits: Adaptive strategies for abating climate change, Climatic Change 33: 235-274.

Expressing Uncertainty and Subjective Probabilities
and
The Role of the Scientist-Advocate

Stephen Schneider

Stanford University
Stanford, California

Schneider tackled a number of challenging issues: How can we take experts' subjective probabilities about climate change into account in a consistent manner? What is our level of confidence in various conclusions? Are our assessments objective or subjective? If they're subjective, can we have a consistent set of criteria and an explicit set of values apply? How can we consistently say precisely what our own uncertainties mean? How do these questions relate to the climate change debate and political context? Do we all mean the same thing when we say 1.5 to 4.5°C warming as the climate sensitivity range?

How can we bring subjective probabilities into all phases of assessment? And how can we get the media and political establishment to understand what subjective probabilities mean?

Schneider traced the history of this 1.5 to 4.5°C range projected for climate sensitivity from its inception in the Charney report, National Academy of Sciences (NAS), 1979. Before 1979, General Circulation Models (GCMs) projected ranges from 2 to 3.5°C. The authors of the Charney report decided to allow 0.5°C as an additional margin of error on the low side (since there was no evidence of negative cloud feedbacks) and 1°C on the high side (due to possible positive cloud feedbacks). Thus the range of 1.5 to 4.5°C was born, and became an "anchor" which is now widely accepted and difficult to change. Such anchoring devices can be helpful in coalition building (see Shackley in this report).

Then, at the second Villach conference in 1987, Jager, Dickenson and Clark presented a chart of projected mean global temperature rise that included uncertainties not only in physical factors but in social and biological factors (e. g., what energy technologies would be used in the future, and biogeochemical feedback processes) and such factors expanded the range greatly to 0.5 to 6-8°C. At this point, the word "scenario" began to be used for such projections, to distinguish them from predictions. In a sense, these were early, mental, partially integrated assessments. Despite the wider ranges emerging from analyses that took a wider array of uncertainties into account, the accepted range of 1.5 to 4.5°C for climate sensitivity remained in place, as it does today. Though some models may indicate slightly higher or lower values, Schneider believes this range to be a reasonable representation of the scientific consensus with regard to likely (but not low probability extreme) outcomes. The stability of this range has arisen largely because there have been no compelling data or theories to modify it especially since it was somewhat arbitrarily defined in the first place.

The Political Context

A report by Accu-Weather, Inc., prepared for the Global Climate Coalition (GCC), purported to be a comprehensive, global climate analysis, but actually used data from just three cities, Schneider reported. Tom Karl, on the other hand, used data from several thousand stations in his comprehensive analysis of global climate change. However, the GCC's public relations effort to get their side of the story into the media led to their press conference receiving much

more attention than Tom Karl's vastly superior scientific efforts. This relative attention is due, Schneider quips, to a "one fax-one vote syndrome." We must realize that this is the context in which we live and work, he says. Scientists tend to shy away from the media, but we must find ways of bringing what scientists know to a wider audience in terms they can fathom.

Scientists have generally been uneasy about offering their subjective probabilities for use in the political arena. There has, however, been a social change in the receptivity of many in the scientific community over the past couple of decades to the need to perform subjective assessments and many scientists now believe it is their responsibility to give their best subjective views on climate issues and indicate their levels of confidence in a way that policymakers and the public can understand.

Many scientists now believe it is their responsibility to give their best subjective views on climate issues and indicate their levels of confidence in a way that policymakers and the public can understand.

Two studies were briefly discussed that use elicitations of subjective expert judgments with regard to climate change: the Morgan and Keith study, and the Titus and Narayanan EPA study on sea level rise estimates (see also, the Morgan and Titus essays, in this report). With regard to such studies, key questions are: who are the experts, what is the process for choosing them, what questions should be asked, and how should the answers be aggregated? Schneider points out that in studies of this kind, there is always one "contrarian" included on purpose and that usually over-represents that point of view. How, he asks, can we bring subjective probabilities into all phases of assessment? And how can we get the media and political establishment to understand what subjective probabilities mean?

Conveying Confidence Levels

A report of the conclusions of a 1987 NAS climate workshop used a set of terms to convey levels of confidence in various findings. Large stratospheric cooling is called "virtually certain," global mean surface warming is called "very probable," etc. In the 1990 IPCC assessment, a 5-star system is used to rank confidence levels: five stars indicate virtual certainties, one star indicates low confidence. It is explicitly stated that these degrees of confidence were "determined subjectively from the amount of agreement between models, our understanding of the model results and our confidence in the representation of the relevant process in the model."

Discussing confidence levels inevitably leads to the question: confidence for what? The level of confidence needed to make global policy decisions is not the same as the level of confidence needed for other purposes, e. g., establishing statistical confidence in certain associations.

The Role of the Scientist-Advocate: Is the Scientist/Advocate an Oxymoron?

Schneider next turned to a discussion of the role of the scientist in responsible advocacy. Advocacy can be personal interest advocacy for a theory, model, measurement, crony, institution, nation, or epistemology. There is also world view advocacy for such things as entrepreneurial rights transcending commons protection; that the present is worth more than the future; that commons protection justifies global-scale rulemaking; risk aversion; or the idea that other species can be as valuable as human pursuits.

Schneider believes that scientists can practice responsible advocacy. Everyone has value positions (consciously and unconsciously); the keys to being a responsible advocate are:

a. making your value positions conscious,

b. making them explicit,

c. not letting your value positions distort your "subjective priors" (see Berk in this report) on issues of fact (a community is needed to do this well because no one can do this perfectly alone), and

d. defending your value positions separately from debates over probabilities and consequences.

Some scientists think it is impossible to do these things, and instead say they will be neutral, but no one is, Schneider says. Among the positions one can take is being an advocate of science itself, in which one argues for a rational world view and has faith that science has something constructive to contribute to decision making for the future. Even that is a value position, even if many who adhere to this position don't see it as such.

The public, media and politicians misread normal scientific contention as a lack of consensus, which leads to confusion and policy gridlock.

On the role of scientists in popularization/advocacy of scientific issues, Schneider says the context is that:

a. Scientists are contentious, but people listen.

b. The media love a fight and often set up stories as "dueling scientists."

c. The public, media and politicians misread normal scientific contention as a lack of consensus, which leads to confusion and policy gridlock.

d. Communication requires simplification but audiences deserve respect; therefore, use familiar metaphors (e. g. , cards, dice, insurance) that don't do violence to the truth. These metaphors cannot completely define the problem but the alternative is to not be in the game (or think you're not and allow journalists or the opposition to define you).

e. Sound bites are selective information transfer and present a double ethical bind. A story will go unreported if there are no sound bites, so scientists should learn to craft good ones that convey both urgency and uncertainty, if that is what the issues represent (e. g., global change).

Some scientists oppose popularization/advocacy no matter how credibly it is done. Skepticism is appropriate but some opposition is pathological and may take the form of:

1. **Elitism** "Simplification is vulgar and we do not play this game. It is not possible to communicate simply what we know to the public so we shouldn't even try."

2. **Jealousy** "A scientist's reputation comes from toiling in the lab, not from press coverage." A way to minimize this problem is to discuss colleagues' work and attribute it to them (but unfortunately, journalists generally attribute it to you anyway).

3. **Special Interests** These ideological opposites/hired guns will try to discredit the science (usually proclaiming to be defending sanctity of science), or discredit the scientist (redefine terms of debate from the science to the scientist or the assessment process, find a trivial error and expand, etc.). They count on the public and media not to check up on the credibility of their attacks, i.e., they exploit the media's "balance" doctrine.

f. Policy analysis (i. e., probability/consequences) is professional science and is amenable to expert judgments.

g. Policy choice (advocacy) is about personal values.

Schneider's Rules for Popularizing

Sound bites are selective information transfer and present a double ethical bind. A story will go unreported if there are no sound bites, so scientists should learn to craft good ones that convey both urgency and uncertainty, if that is what the issues represent.

Schneider's reaction to this context is to suggest that scientists popularize their own work or have someone else do it. He suggests these rules:

a. Use familiar metaphors that convey both the urgency (to get in press) and the essence of the uncertainties surrounding the issue (e. g., "loading dice" conveys urgency and uncertainty).

b. Back up sound bites with articles and other products (e. g., full length books, so one's full views are on record for those few who want to check up).

c. Separate expertise (probabilities/consequences) from values.

d. Enjoy the process but don't be naive about the risks.

e. Don't counterpunch with ad hominem attacks and polemics, despite what the opposing advocates do.

f. Remember that long term success is built on establishing the credibility of the scientific community over time.

Journalists in the group pointed out the social legitimization function the elite press plays in science. For example, if a paper published in the New England Journal of Medicine is reported on in The New York Times, it is cited twice as often in the medical community as a New England Journal paper which is not reported on in The Times.

Schneider believes that it is best to talk to the press often or not at all, as there are risks involved in only one or two stories which by chance could be excellent or terrible. Not all scientists should be popularizers or advocates. We should let those who are good at it do it without disdain from their peers if they tell the story straight. Scientists can serve as a resource to raise reporters' level of knowledge on important scientific issues.

Schneider says that if scientists fail to convey probabilities of particular outcomes, they fail to convey the basic information that the public needs (e. g., what is the probability of significant climate change versus the probability of an invasion by aliens from space) to make any kind

of decision (health, safety or economic). On the other hand, when one gives a subjective assessment with probabilities, there is always the danger that others can grab the tails of the distribution and run with them to suit their own agenda. A community effort is needed to overcome the "dueling scientists" problem, as well as tackle the other issues presented here.

On the other hand, when one gives a subjective assessment with probabilities, there is always the danger that others can grab the tails of the distribution and run with them to suit their own agenda.

A Social Constructionist View of Scientific Objectivity

Simon Shackley

Lancaster University
Lancaster, United Kingdom

Studies have illustrated the importance in successful environmental policymaking of policy coalitions between environmental groups, industry, scientists, and others with a stake in the outcome.

It is assumed by many that greater precision, quantification and clarification of uncertainty will create more persuasive science for policymaking purposes and will help advance the climate change policy agenda. While there is much value in this, Shackley questions whether it is as important a priority as commonly assumed. Empirical studies have illustrated the importance in successful environmental policymaking of policy coalitions between environmental groups, industry, scientists, and others with a stake in the outcome. The parties in these coalitions generally have substantial differences in their outlooks, aims and objectives, but they revolve around a core of consensus that might be thought of (following Hajer) as a "storyline."

A storyline is an account of an issue in terms of its causes and effects, the impacts of remedial actions and assumptions about the motivations and behavior of the key policy actors involved. Storylines simplify arguments and interpretations over a wide range of scientific, economic and policy issues. A specific storyline serves to hold the policy coalition together despite substantially different interpretations, values and objectives. It is simply not possible to encapsulate all the complexity surrounding many environmental issues in any one definitive, agreed-upon, account. Such "closure" is not, fortunately, required for effective political and policy action. And as is discussed later, ambiguity and uncertainty can actually help hold the coalition together.

Referring to a theory from Ted Porter's book Trust in Numbers , Shackley says that the desire for greater precision through quantification may be symptomatic of "weak" institutions which feel distrusted. It is possible that weak institutions collectively attempt to re-build trust through using storylines which rely on quantification. The latter have an appearance of objectivity and rationality which assists in holding policy coalitions together and is difficult to argue against by those not involved (even though quantification does not automatically imply less ambiguity). However, Shackley says that it is questionable whether this sort of quantification will secure political consensus in the way that it might have done 50 or even 20 years ago. Public alienation and distrust of governments has increased and policy making institutions frequently send out conflicting signals (environmental protection versus increased consumption, for example). In this new political context, it is not clear that quantified storylines will succeed in re-building trust between governments and the public.

Political Influences on the Scientific Process

Shackley then turned to examples of where science is being "constructed" in advisory contexts in order to illuminate how non -scientific factors enter into the formulation of scientific advice. Shackley began by discussing changes in the global warming potential (GWP) for methane

in the IPCC Working Group I 1994 assessment. The 1992 IPCC report expressed reduced confidence in the 1990 estimates and declined to estimate indirect effects. Following the 1992 report, two new three-dimensional atmospheric chemistry models were developed which accounted for indirect effects of methane. The two models did not produce the same value but the researchers were able to agree on a new GWP for methane with an uncertainty of plus or minus 35 percent, based on the range of the models' results and the modelers' interpretation.

At a meeting of lead authors in Geneva in July 1994, several authors (notably Robert Watson) expressed concern that the new GWP value for methane would have a great deal of political significance because it meant that on a 10 to 20 year time frame, methane might be a more important contributor to climate change than CO2. Because these authors could foresee a controversy over this new value, they stressed the importance of clarifying how the new GWP was calculated, and in particular, how the uncertainty range was determined. The point Shackley makes here is that the anticipated political implications had a direct effect on how the GWP and the uncertainty associated with it was represented in the report.

The anticipated political implications had a direct effect on how the GWP and the uncertainty associated with it was represented in the report.

Later that year, in the plenary session for Working Group I, methane's GWP was one of the most controversial issues. Some government delegates, particularly New Zealand's government scientists, raised concerns about this and an ad hoc meeting was held to discuss the issue. (New Zealand has very high methane emissions, primarily due to the large numbers of sheep.) The New Zealand government scientists expressed their belief that a 35 percent uncertainty was too tight a range, representing more certainty that we actually have about this value. They argued that policy might be made based on what they saw as provisional figures policies such as trading emission reductions in methane for increases in CO 2. This argument was rejected, and one of the key people in this rejection was Robert Watson, who said it was not up to scientists to tell policymakers how they could use GWPs. Shackley calls Watson a classic example of a "gatekeeper:" a scientist and policymaker who is very influential in deciding how science will be expressed and transferred to the policy arena.

Shackley alluded to the analogy with Ozone Depletion Potentials (ODPs), in which case the negotiators requested a single value for each ozone-depleting chemical. Yet, if advisory scientists accede to this request, is there not the danger that over time policymakers come to have greater confidence in the certainty and precision of the science than the scientists do themselves? Given their confidence, might policymakers not also come to believe that such precision and certainty can be expected and requested of advisory science? By contrast, if advisory scientists were to be more forthright in expressing the uncertainties, then policymakers might come to accept that it was not possible to obtain precise values.

The point of discussing this example is not that the judgment of the IPCC scientists was wrong, but that it was more than simply a scientific judgment. It involved judgments about how policy works or should work, and what policymakers need. A small number of "gatekeepers" now make these kinds of decisions. Shackley argues that more explicit discussions of these non-scientific judgments are needed. The way that scientific uncertainty is represented in advisory reports is influenced by advisory scientists' judgments about the needs of policymaking.

Another example of the construction of science involves the representation in the IPCC of flux adjustments in coupled ocean atmosphere General Circulation Models. A dominant theme of industry comments during the Madrid IPCC WG I Plenary in November 1995 was to emphasize flux adjustments as a reason climate model results should be seen as conditional. The advisory

scientists' response to such criticism is colored by recognition of the politically-inspired nature of such attacks. In such a context, who (scientists or skeptics) introduces the uncertainty into the debate is key in the wider perception of whether the advisory process has integrity.

But Shackley also argues that the actual substance of the science, as well as its portrayal, is influenced by policy considerations. As evidence of this, he describes his analysis of why flux adjustments have been used in some climate modeling centers but not in others. Most of the debate is scientific in character, but since there are good scientific reasons for both using and not using flux adjustments, more than just science is involved, he contends. Shackley and colleagues believe that factors such as the institution's mission and funding, policy roles and relations, and relations with the climate-change-impacts community can enter into the decision as to whether or not the institution uses flux adjustments in its models. Policy implications influence the debate because without flux adjustments, GCMs cannot be used to make long-term projections, which is one of the key requests made by policymakers of climate scientists. Shackley argues that the use (or non-use) of flux adjustments therefore emerges from the complex interplay of scientific debates with institutional and policy commitments and preferences.

Instead of seeing ambiguity and uncertainty as something to avoid or purify from the climate debate, it could actually be a benefit.

The Role of Ambiguity in Coalition Formation

Shackley says that both scientific arguments, and the non-scientific context of advisory science which comes to influence advisory science, contribute to the potential ambiguity of science for policy. He believes that a certain degree of ambiguity in scientific knowledge is useful and perhaps even necessary in producing a "storyline." So instead of seeing ambiguity and uncertainty as something to avoid or purify from the climate debate, it could actually be a benefit. Some of the necessary actors in the policy coalition on climate change might drop out if the uncertainties cleared up too much. Among the ambiguities he believes are important in holding together an emergent policy coalition are those in the physical science realm, such as ambiguities regarding climate sensitivity and global warming potentials.

Ambiguities in the meaning of climate sensitivity include the rules by which GCM results are adopted in the IPCC assessment, the domain of uncertainty accounted for, changes in the "best guess" to accommodate new knowledge, and changing from a narrow definition which includes CO2 only to a wider definition which includes all GHG forcings. Ambiguity in the meaning of GWPs includes the choice of GHGs for which GWPs are calculated, which indirect effects are included and how they are included, the time horizons and discount rates used, in what parameter of climate change the GWP is being measured, in what atmospheric residence time is used for CO 2, and whether GWP is calculated with sustained or pulse emissions.

The ambiguities around climate sensitivity help to hold together the GCM community, as well as helping those using simple models and those assessing impacts to calibrate their models to the GCMs. It further allows policymakers to have easy access into the climate issue, providing an "anchoring devise" and an entry point for many different communities. It also justifies ongoing modeling activities to continue research to hone these values. The ambiguities in GWPs serve similar functions, for example bringing together research communities to address problems which had not previously been tackled such as those relating to the global carbon cycle. These ambiguities also help to accommodate the preferences of policy communities, including the desire of many governments for a comprehensive approach to all GHGs, not just CO2.

Shackley presents a model of the chain from physical effects to human impacts of climate change. He calls the physical processes “up stream” and the effects on human systems “down stream.” As one moves down stream in this model, and includes more knowledge, debates and uncertainties, there is an opportunity for the whole to be destroyed in intellectual, policy and political discussions. Therefore, he says, up stream processes, and ambiguities in these processes, may serve in practice as better anchoring devices. Down stream processes are less able to generate a clear storyline and to hold the coalition together.

Discussing the significance of policy coalitions for the climate issue, Shackley says that there is a need to link climate change to other policy agendas and issues in a meaningful way. Here a dilemma emerges because in order to secure these effective cross-issue linkages, the knowledge needed is of the down stream variety which is less able to secure an effective storyline. It may be that we have to revise downwards our expectation that scientific knowledge is able to hold together policy coalitions in the climate case and look to other shared knowledge and values to provide the rationale instead.

It may be that we have to revise downwards our expectation that scientific knowledge is able to hold together policy coalitions in the climate case and look to other shared knowledge and values to provide the rationale instead.

Terrestrial Ecosystems: Controversies in Assessing Climate Impacts

Allen M. Solomon

U. S. Environmental Protection Agency
Corvallis, Oregon

Solomon discussed the primary controversies in predictions concerning the future dynamics of terrestrial ecosystems in response to climate change. Included in this discussion were controversial elements of the IPCC Second Assessment Report from Working Group II on forestry (Chapter 15) and the differences in the global carbon cycle to be expected when one uses vegetation simulation models that deal with responses to climate change by individual species dynamics versus by whole ecosystems grouped into 15 or 20 primary global biomes.

The fundamental issues in terrestrial biospheric responses to changing climate involve (1) the decline of terrestrial biodiversity and (2) the role of the terrestrial biosphere in the global carbon cycle. Biodiversity was discussed in this AGCI session by Terry Root. The carbon cycle is a problem because too little is known at present to define either the current or future role of the terrestrial biosphere as that of a carbon source or a carbon sink. Based on ice core data and the greater activity of respiration than of photosynthesis with warmer temperatures, there is good reason to suspect the terrestrial biosphere has been a source of CO_2 during warmings of the past 250,000 years, producing a positive climate feedback leading to further warming. Yet, the physiological basis for this logic is weak at best. Equally compelling (and uncertain), Solomon says, is the view that CO_2 limits growth of plants under current conditions so that increased CO_2 concentrations will “fertilize” plants, inducing an increased carbon sink with coincident warming, providing a negative feedback on the global carbon cycle.

The carbon cycle is a problem because too little is known at present to define either the current or future role of the terrestrial biosphere as that of a carbon source or a carbon sink.

Since the middle 1970s, the “missing fraction” of CO_2 that calculations of fossil fuel burning and ocean uptake predict in the atmosphere, has been attributed to the terrestrial biosphere, with CO_2 fertilization being the hypothesized mechanism for carbon uptake. Although recent calculations by Dai and Fung (1993, Snowmass Carbon Cycle Meeting) and by Keeling et al. (1996, Nature) indicate that seasonal warming at high latitudes could account for much of the missing fraction, CO_2 fertilization is still the preferred mechanism of most oceanographers and others who study the global carbon cycle (e. g., Tans, Peng, Schimel et al. as IPCC SAR from WG I, etc.).

This dependence on the CO_2 fertilization mechanism is not unlike a religious belief; in the absence of either indicative or definitive evidence, one must posit faith that the mechanism seen in herbaceous annuals and tree seedlings in greenhouse pot experiments will also be functioning (and doing so in the same way) across the process scales that link:

1. population dynamics of seedlings to survival, growth and reproduction of mature trees,

2. the interactions among mature trees with one another through competition for light, water and nutrients in forest stands,

3. the interaction of stands via exchange of species and materials to form ecosystems, and

4. the interactions among forest ecosystems through groundwater, and

5. through propagules, to form landscapes and biomes.

Forests are the focus of the carbon fertilization controversy because most of the terrestrial biospheric carbon is tied up in forest trees and soils.

The necessary long-term (multiple decade) fumigation experiments have not been done to demonstrate such a robust process transfer. Indeed, data from the very few field situations in which mature trees have been fumigated for their lifetimes by adjacent CO 2 fumaroles (experiments in Italy and Iceland) suggest that the excess carbon uptake measured in seedlings in greenhouses also occurs in nature, but that it does not cross the scale to affect processes which are important in mature trees. Instead, the measured advantage declines continuously such that no “fertilization” effect is visible after one or two decades.

Though seedlings in greenhouses fumigated with CO 2 absorb extra carbon, this does not cross the scale to affect processes which are important in mature trees. Instead, the measured advantage declines continuously such that no “fertilization” effect is visible after one or two decades.

Whether or not rising atmospheric CO2 is being increasingly sequestered in terrestrial vegetation, a second group of processes of change may release a significant amount of carbon from terrestrial vegetation during the coming decades to a century or two. These processes are the “transient” lags in response of forests to rapid climate change. First, forest succession (the developmental sequence of forest species from those which grow most quickly and are least tolerant of shade, through those which grow most slowly and are most tolerant of shade) is a slow process. It requires 50-500 years in contrast to the 1 - 10 years required for dieback of trees, such as may be expected from a chronically changing climate.

Second, continuous climate change can be expected to eventually render local tree species “climatically obsolete” requiring reforestation with species not initially present. The immigration of trees (including the transport of seeds, establishment of seedlings, growth of trees to maturity, and transport of seeds from the new trees) has been measured at 15-40 km/century during the last 10,000 years in temperate areas, although occasional “bursts” of speed beyond 100 km/century have been recorded. In contrast, the mean July isotherm in temperate areas is simulated to migrate 200-400 km/century. The lags induced in carbon uptake generated by the lack of suitable trees, and by slow forest development processes, suggests a “pulse” of carbon released from dying forests, but not retrieved from the atmosphere in new forests for many decades or centuries.

The means by which these processes (carbon fertilization, transient responses) may control the role of the terrestrial biosphere in the global carbon cycle has been characterized in simulation models designed to determine the implications of those multiple competing processes. One set of models which treat the terrestrial biosphere as a “big leaf” assume the religious tenant of a constant relationship between carbon concentration and carbon uptake, while taking no cognizance of forest “infrastructure” changes (e. g., TEM, GEM, BIOME-BGC). Other more recent models also simulate competing processes of photosynthesis and respiration rates and differences among different physiognomic biomes (e. g., temperate deciduous forests, tropical savannas, etc.), which are in turn defined by climate (BIOME2, DOLY, MAPSS). All of these

models either do not redistribute biomes' geography with climate change, or else they do so instantaneously, as though the transient processes do not exist.

Models which simulate life histories of individual tree species or species groups may be more accurate in defining the impacts of transient processes on the temporal dimension of terrestrial carbon. Models which include forest succession (gap models such as JABOWA, FORET, etc.) thus far have not been applied to the globe as a whole because the natural history information needed to parameterize them is lacking in poorly studied regions, such as many tropical areas. Tree migration models are currently under development in several research groups but none are available to apply at the global scale required to calculate implications of these processes on global carbon cycle dynamics. It seems probable that credible, globally-comprehensive vegetation models which simulate transient process mechanisms will be available within the next 5 to 10 years.

We must confront the fact that our understanding of the terrestrial carbon cycle is data-limited, not model-limited, Solomon says.

Hence, there is considerable optimism that enough is known to construct accurate carbon cycle models, and that the models will soon reveal the critical answers concerning the role of the terrestrial biosphere in the global carbon cycle. However, the answers may come from new ecological data rather than from how we manipulate the data we have. For example, those data may reveal that the processes modeled are not the processes driving the system; perhaps we will learn that the rhizosphere contains the missing fraction of atmospheric carbon, and that carbon allocation to leaves, stems, and roots within individual plants is a central process defining biospheric carbon content. Whatever the result, we must confront the fact that our understanding of the terrestrial carbon cycle is data-limited, not model-limited, Solomon says.

How Journalists Deal With Scientific Uncertainty

S. Holly Stocking

Indiana University School of Journalism
Bloomington, Indiana

Stocking discussed observed patterns of content and influences on this content with regard to how the media cover scientific uncertainty. Studies of observed patterns are limited. Most of the evidence is from studies conducted for other purposes, and all but a few studies focus on mainstream U. S. "elite" newspapers (The New York Times, The Washington Post, The Wall Street Journal, and The Los Angeles Times). There has been little attention to television, letters to the editor, editorials and op-eds, films, books, and new technologies.

The majority of journalists who wrote about science did only single-source interviews; they thus failed to tap sources who might have volunteered caveats about the research.

The evidence, such as it is, supports scientists' frequent complaints that the mainstream news media make science appear more certain than scientists think it is. At the same time, though, evidence also exists to support the view that the media sometimes make science appear more uncertain and baffling than it in fact may be.

When the mainstream news media make science appear more certain than scientists think it is, they do it in a number of ways. These include:

Downplaying caveats In simplifying science, journalists often downplay the limitations and inherent uncertainties of studies. In accuracy studies (in which scientists are asked about the accuracy of a story for which they were interviewed) omission of caveats are a frequently-mentioned problem. One important study found that the majority of journalists who wrote about science did only single-source interviews; they thus failed to tap sources who might have volunteered caveats about the research.

Offering little context Little attempt is made to link stories to previous research or background on a topic, often making science news appear disembodied and timeless.

Emphasizing product over process Journalists tend to pay more attention to findings and less to the interpretive work that goes into developing the findings.

Presenting science as a "triumphant quest" When journalists do write about the process of science, they often present scientists as "warriors" or as "detectives" trying "to unlock mysteries." This assumes that knowledge and certainty will "win the war" against ignorance and uncertainty that answers will be found and mysteries solved.

Downplaying fringe or dissenting scientists Stocking says that some of the best studies have found a high correlation between visibility of scientists in news accounts and visibility of scientists in the scientific literature as measured by citation counts. This would seem to suggest that fringe or dissenting scientists don't get undue attention in the news as some scientists complain. (The same is true for politics; fringe figures generally get little attention in the

mainstream press). However, it is quite possible that such findings would not hold up for some media or for some issues, particularly those that are highly charged, with important political and economic interests at stake.

When journalists make science appear more uncertain and baffling than it in fact may be, as is sometimes the case, they seem to do it in a number of ways, by:

Not explaining sudden changes Journalists may fail to explain why today's study disagrees with yesterday's, leaving the public to question the validity of all science. Also, they may fail to explain that two seemingly contradictory conclusions can be true (e. g., aspirin is good for the heart but bad for the stomach).

Not explaining disagreements Journalists may play up the disagreements between scientists without sorting out the differences, explaining the reasons for them, or noting which claims come from the majority and which from the minority.

In the highly-charged climate debate, some journalists (particularly on the op-ed pages) play up the views of contrarians, thereby raising an undue level of uncertainty about conclusions that are actually matters of scientific consensus.

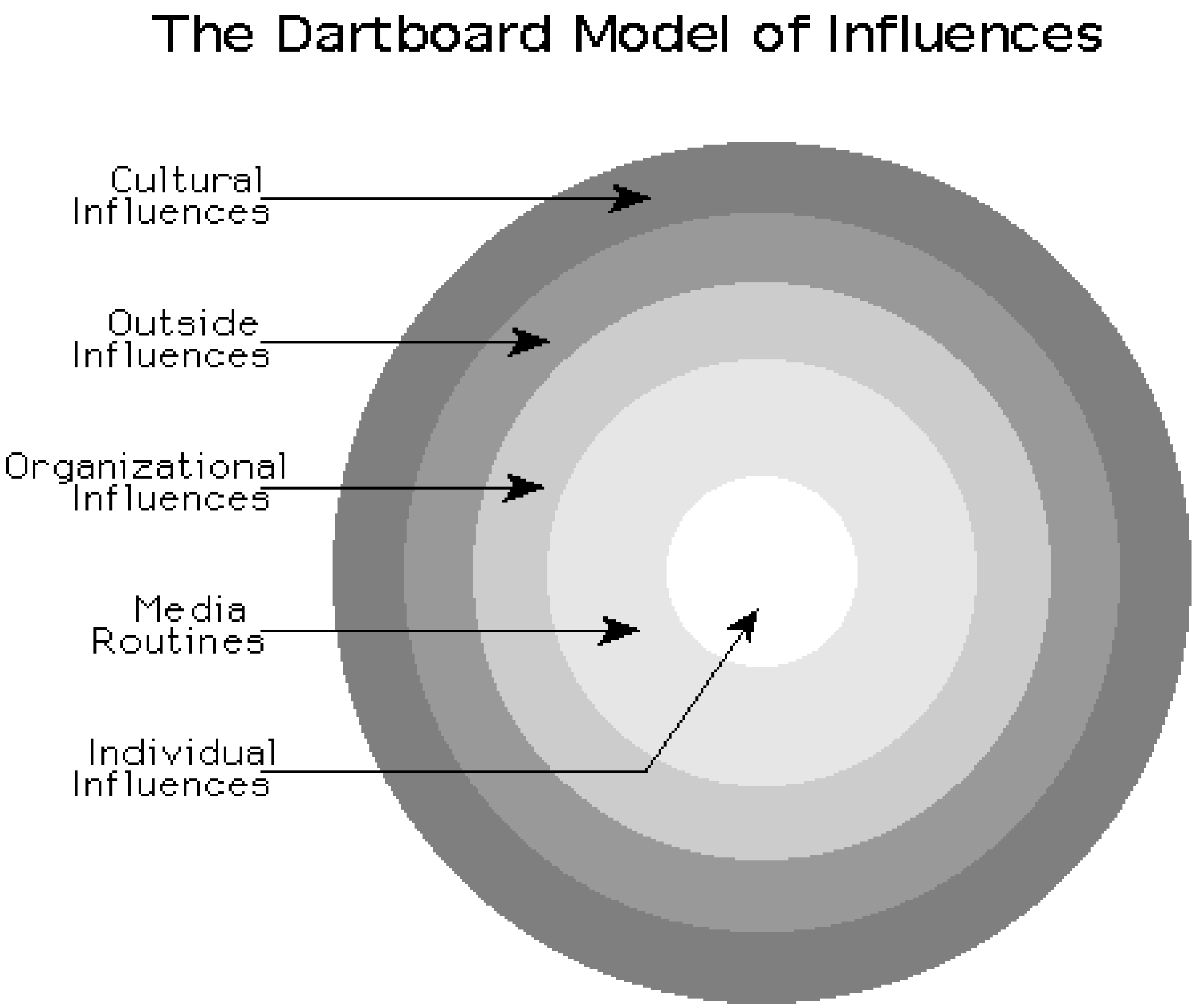
Playing up fringe or dissenting scientists This has been seen in the highly-charged climate debate, where some journalists (particularly on the op-ed pages) play up the views of contrarians, thereby raising an undue level of uncertainty about conclusions that are actually matters of scientific consensus.

Social scientists have observed all of these patterns in mainstream news media, Stocking said. And they also have observed exceptions to almost all of these patterns. Occasionally, for example, caveats and context are included, process is stressed over product, and certainty is portrayed as elusive or impossible.

If we have little direct data concerning patterns of uncertainty in journalistic accounts, we have even less direct evidence concerning the factors that go into creating these patterns. In spite of this fact, Stocking speculated on the factors that might be implicated. In doing so, she drew from what she called the "dart board" model of media constraints. This model includes all the influences on journalists, from those most distant from the individual journalist (on the outside ring of the dart board) to those most immediate (in the bull's-eye); it includes, in other words, influences ranging from cultural factors and outside influences like public relations activities (in the outer rings), to organizational influences, media routines, and in the center, individual influences such as the level of scientific knowledge (see Figure 2.15).

Stocking expressed her view that scientists' criticisms tend to focus too much on the ignorance of the individual journalist in the bull's-eye. Most social scientists who study science news, she said, assume that factors in the outer rings of influence matter more. For example, as discussed in depth in this session by Sharon Dunwoody, journalists' occupational routines and norms present many barriers to the accurate presentation of scientific uncertainty.

On all of these subjects, a great deal of discussion ensued. How do journalists define news? It must be timely, involve important people, have a huge impact, have close-to-home proximity, and/or have appeared in The New York Times (which often sets the agenda for other papers). There is a tendency for journalists to go for their stories to scientists who "give good quote," can speak in "sound bites," can speak journalists' language, and use metaphors that people can relate to.



Stocking feels that scientists' criticisms tend to focus too much on the ignorance of the individual journalist in the bull's-eye. Most social scientists who study science news, she said, assume that factors in the outer rings of influence matter more.

It is also important to realize that most news comes from public relations efforts. Press releases from organizations like the Global Climate Coalition represent a professional effort to get their side of the story out. Most scientists, on the other hand, do not like to think about public relations. They prefer to think the science will speak for itself, but this is not so.

One participant raised the problem that journalists do little background research, and what little they do is usually of other journalists' stories. This may perpetuate errors and inadequacies that appear in the popular press. Systematic fact-checking is also not common in newspapers. In-depth reporting yields much better results than the crisis reporting mode most journalists usually operate in, but it is increasingly rare. Some journalists will allow a scientific source to see the story prior to publication, but Stocking stressed that it is but a small minority of journalists who will submit to pre-publication review.

Stocking pointed out that another problem is that pictures and headlines are out of the control of the journalist who writes the story. Increasingly, with the help of new technologies, journalists are suggesting headlines, but the time pressure is extreme and editors don't have much time to edit and think about titles.

The point was also raised that when journalists go to scientific meetings for information, the scientists there are discussing the cutting edge of science. This is appropriate for the scientists, but journalists often misunderstand such debate and think that nothing is agreed upon. Thus journalists get mostly misconceptions from such meetings. There are steps scientists can take to address this problem, such as making themselves available to talk with journalists at scientific meetings and explaining things with appropriate language and context.

Finally, Stocking pointed out that journalists are quite sensitive to feedback, both positive and negative. Journalism periodicals (e. g., The Columbia Journalism Review) define the quality and ethics expected of journalists, and can be expected to constrain journalists' behavior, at least to a degree. Professional organizations and awards for science journalism can also be expected to work in this way.

Scientists have a responsibility to learn how to communicate through the news media. And journalists also have a responsibility to learn more and improve their understanding and reporting of scientific uncertainty.

Scientists, particularly those whose work has ramifications for public policy, have a tremendous responsibility to learn how to communicate through the news media. And journalists also have a responsibility to learn more and improve their understanding and reporting of scientific uncertainty.

Probabilities in Sea Level Rise Projections

Jim Titus

U. S. Environmental Protection Agency
Washington, DC

Along the Atlantic and Gulf Coasts of the United States, sea level is rising about 3 mm per year or about 1 inch every decade. For people living near the coasts and near sea level, this is significant. The natural consequences of this rise are that ocean beaches are eroding about a meter per year, floods are gradually coming further inland, and water tables are rising.

Though sea level has been changing naturally for thousands of years, what presents the current problem is the interaction between human activity and the consequences of the rise in sea level. In many coastal resorts, beaches are disappearing because houses are being built where the dunes used to be; if the houses weren't there, the dunes could move inland. Beaches are also getting narrower because bulldozers are excavating the coastline, often to build dunes. In many low-lying barrier island areas, whenever it rains at high tide the drainage doesn't work because the third of a meter (one foot) rise in sea level since these communities were built leaves the water table a third of a meter higher, thus slowing the drainage rate. Along estuarine shores like Chesapeake Bay, the natural shoreline is being replaced by walls of concrete, rock and wood. Public access to beach areas is lost, birds can't eat along the shoreline, boats can't land there; in short, much is being lost.

Addressing these and other consequences of rising sea level depends on factoring in existing sea level rise, how much it will accelerate due to climate change, and how we choose to handle the trade offs between risks and costs. In many cases, we are not yet even dealing with the fact that sea level is rising at all. For example, wetlands protection laws and other U. S. policies are based on the assumption that the sea isn't rising and shores aren't eroding, even though we know that they are.

Addressing the consequences of rising sea level depends on factoring in existing sea level rise, how much it will accelerate due to climate change, and how we choose to handle the trade offs between risks and costs.

There are cases where sea level rise is being recognized. Along ocean shores that have not yet been developed, at least half the states in the U. S. have set-back requirements, in many cases, based on current erosion rates. This shows a recognition, at least with regard to new development, that the shores are dynamic. In some cases, there is even a recognition of the acceleration of sea level rise. For example, the Dutch are incorporating accelerating sea level rise into new dike construction plans. San Francisco, Hong Kong, and a few other cities are addressing accelerating sea level rise with regard to newly reclaimed land. The state of Maine has restricted development for any area subject to a one meter (~3-foot) rise in sea level (which was the best estimate when the regulation was put in place).

There are many decisions that are sensitive to future sea level rise. Among them are determining minimum surface elevation for land reclamation, coastal setbacks, the size of urban drain pipes, and the valuation of contingent interests in land for taxation purposes and for the purposes of determining how much to reimburse land owners if their property is "taken" by

the government. Decisions such as these often deal with risks versus costs. For example, in building a new urban street drain system, it only costs 1-3 percent more to use the larger pipes that would make the system work well if there is a third of a meter rise in sea level. Current decision makers have to choose whether to spend this extra money now to prevent a possible future problem. Such decisions depend on factors including what discount rates are applied and how likely it is that the sea is going to rise a third of a meter before the system would have to be rebuilt anyway.

Regarding valuation of contingent interests in land, one way to deal with the wetland problem is to create a "rolling easement," in which the government buys an option to take property if sea level rises a certain amount. Essentially, the government is purchasing a contingent property right to ensure that nature can move inland if necessary. In calculating the value of such a rolling easement, we can use a probability distribution regarding sea level rise. In valuing these properties, a court would need expert testimony as to the probability of sea level rise. In general, Titus argues, there is a need to find ways of incorporating sea level rise into rational decision making.

In order to calculate a probability distribution for sea level rise, Titus brought together many existing models and added what was lacking, which was a way of capturing smaller probability events.

In order to begin doing this, Titus says, it is necessary to calculate a probability distribution for sea level rise. For this reason, Titus brought together many existing models and added what was lacking, which was a way of capturing smaller probability events, such as a deglaciation in Antarctica. The IPCC model assumes no contribution to sea level rise from Antarctica, but most glaciologists believe there is at least a small possibility of such a contribution. Figure 2.16 illustrates how the models used in Titus' study were brought together.

It was then necessary to specify probability distributions for all the input parameters. The IPCC and other previous assessments used high, medium and low scenarios, but did not specify probability distributions. This difficult task was accomplished by bringing in a large number of expert contributors. Each expert went "on the record" by name which allowed for probing and disclosing potential biases, etc. Experts were divided into three categories based on their having opinions on: climate parameters, glaciology parameters and precipitation parameters. In addition to the aggregations, each contributor's individual answer on any topic can also be found in the report.

Combining or aggregating their answers was a challenge. How should outlier opinions be used? How should each experts' opinion be weighted? Titus chose to alter the mainstream probability distributions based on outliers' views, and he placed a weight on them, which he chose for each in an ad hoc fashion. There was some discussion on how this should best be accomplished; the view was expressed that it is crucial to understand the underlying reasons for the differences in opinions before choosing if and how to aggregate them.

Figure 2.16
Relationship Between the Various Models Used to Project Sea Level

From Titus, J. G. and V. Narayanan, Probability of Sea Level Rise, 1995, U. S. Environmental Protection Agency Report 230-R-95-008, 186 pp.

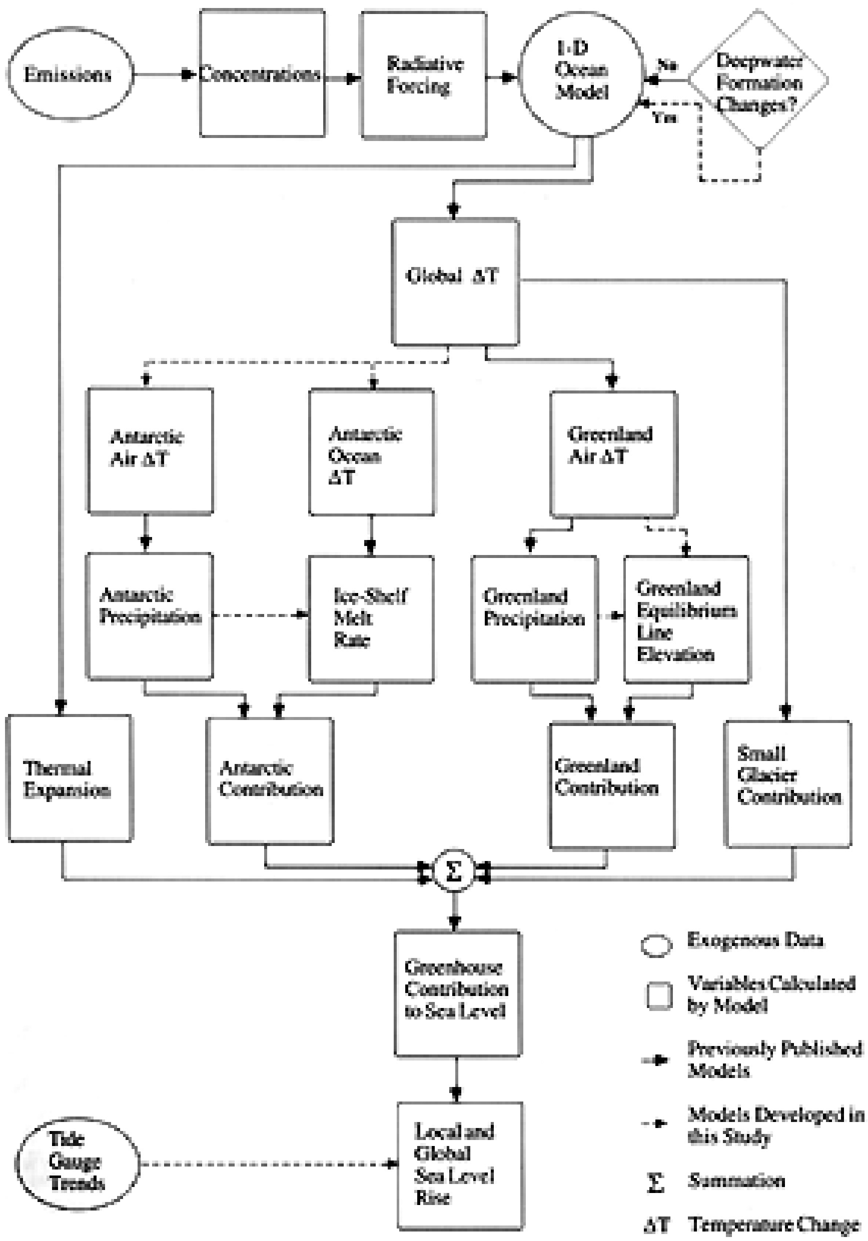


Figure 1-1. Relationship Between the Various Models We Used to Project Sea Level.

The manner in which the EPA authors chose to aggregate the scientists opinions was criticized in editorial comments by Keith, 1996, on the grounds that this may be an inappropriate procedure when an outlier scientist clearly adheres to a different paradigm than the other scientists interviewed. Titus counters that not only would it be wrong to disregard the opinion of a respected scientist simply because it was an outlier, but that the suggested courses of action recommended by Keith do not solve the problem. Keith suggested only reporting nonaggregated results, similar to the approach of Morgan and Keith. That suggestion, however, ignored the practical realities of a modeling study where different groups of experts express opinions on different subsets of parameters. In such a study, there is a vertical aggregation of opinions (combining one expert's opinion of ocean parameters with another's opinion on glacial parameters) as well as the horizontal aggregation involved in specifying a probability distribution based on alternative expert opinions. As a result, Keith's suggestion would have required the reporting of dozens of different probability distributions, for every combination of particular climate modelers with particular glaciologists with particular polar meteorologists. Paté-Cornell suggested that it might have been better for Titus and Narayanan to simply specify composite probability distributions themselves, after duly considering the various reviewer opinions. Titus and Narayanan declined to follow that course, both because the entire objective was to use the experts' opinions, and because such a procedure would undermine the modeling efforts to preserve any functional relations between parameters.

Combining or aggregating answers was a challenge. How should outlier opinions be used and weighted? Titus chose to alter the mainstream probability distributions based on outliers' views, and he placed a weight on them, which he chose for each in an ad hoc fashion.

One of the contributions of the Monte Carlo method used in the EPA study was the ways in which the possible correlations between different parameters were dealt with. Even when only one expert is providing probability distributions for different model parameters, care must be taken to avoid unreasonable combinations of parameter values (what might be called the "forbidden parameter space" problem). That is, although each expert scientist may believe that any one parameter has values that fit into a certain probability range, the same scientist might also believe that parts of that probability distribution are impossible or highly unlikely when other parameters take on certain values. In other words, certain combinations of the values of that parameter and the values of other parameters could be forbidden. Thus, a joint probability distribution needs to be constructed from subjective opinions of experts.

The EPA study did attempt to deal with this problem, both by specifying joint probability distributions and by preserving the consistent visions implied by the alternative specifications by various reviewers. In the first case, for example, Titus and Narayanan specifically asked the experts whether particular parameters were correlated. One scientist suggested that the Greenland temperature amplification factor is positively correlated with changes in the rate of deepwater formation (a 1-D ocean model parameter), because deepwater formation drives the Gulf Stream which warms Greenland, and another respondent suggested that there may be a correlation between the dynamic adjustment time of the circumpolar ocean to global warming, and both T2x and the rate of greenhouse gas buildups.

In the second case, Titus and Narayanan attempted to ensure that "internally consistent visions" of the different scientists were maintained, by requiring that each simulation sampled all the climate parameters from the probability distribution specified by one expert, all the glaciological parameters from one expert, etc. For example, although many of the reviewers did not specify a correlation between Greenland amplification and deepwater formation, a correlation across reviewers still existed because those that tended to expect the greatest decline in deepwater formation also projected the least polar amplification.

There was more variation in opinion regarding Greenland warming than global average

temperature warming and other parameters. In general, the paleoclimatologists expected more Greenland warming than did the modelers. Regarding forecasts for sea level rise, the answers were similar to those of IPCC. The one exception is that the 1 percent probability on the high end is a few meters of sea level rise in the year 2100, which is higher than the IPCC estimate. This is driven by the possibility of a large warming combined with a surprising level of melting in Antarctica. A one meter rise by 2100 is estimated at about a 50 percent probability in the EPA study.

How important are probabilities? Why not just take our best estimates of each parameter and thus come up with our best answer to a given question? Titus says that probabilities don't really matter, if you have a low discount rate. Take, for example, the case of determining a present value of an interest in land that vests when the sea rises one meter. Using a full probability distribution rather than a median sea level rise estimate, one comes up with a higher monetary value for the land; there is a non-linear valuation function because interest compounding is non-linear.

How important are probabilities? Why not just take our best estimates of each parameter and thus come up with our best answer to a given question? Titus says that probabilities don't really matter, if you have a low discount rate.

There is also the question of why the government should have to pay (with tax dollars) for the rolling easements. Why not just tell developers not to develop the coastal zone? In some cases, this is unconstitutional, as explained in Titus (1997). Even if it is legal, it is politically infeasible. For our purposes here, the important point is that these rolling easements cost about 1 percent of the cost of the land whether or not the government or the landowner takes the loss. But whether the true figure is 0.1, 1 or 5 percent depends on probabilities and discount rates, and hence, probability distributions are necessary. A one meter rise would inundate some 18,130 square kilometers or 7,000 square miles of land in the U. S. alone. Titus says that the land use conventions we adopt today will determine whether 100 to 200 years hence our coastal zones look like the Netherlands or pretty much the same as they do today.

References

Keith, D., 1996, When Is It Appropriate to Combine Expert Judgments? An Editorial Essay, Climatic Change, 33(2):139-143.

Titus, J., 1997, Rising Seas, Coastal Erosion, and the Taking Clause, South Carolina Law Review, in press.

Social Costs of Climate Change

Richard S. J. Tol

Institute for Environmental Studies, Vrije Universiteit Amsterdam, The Netherlands

An enormous controversy arose over the subject of the valuation of a human statistical life, whether such a thing is possible, and whether or not it should differentiate between income groups.

Tol discussed Chapter 6 of Working Group III, on the social costs of climate change. It is a great challenge to make a comprehensive assessment of the impacts of climate change from an economic point of view. Economists try to overcome the problem of the overwhelming amount of information on the diversity of the potential impacts of climate change by assigning values to each impact, adding them, and then presenting the sum in monetary terms or as a percentage of GDP. Advantages of such an approach are that (1) it delivers a single indicator of climate change impact, and (2) it makes the impacts of climate change directly comparable to the costs of emission control, and can thus lead to an informed notion of an appropriate level of spending on emissions control.

Controversies

Regarding the work of Tol's group, an enormous controversy arose over the subject of the valuation of a human statistical life, whether such a thing is possible, and whether or not it should differentiate between income groups. The controversy took the shape of personal attacks on lead authors in the popular press regarding alleged racism and blasphemy; it reached its summit in a short occupation of the London offices of the convening lead author's headquarters. The background of this controversy stems from (1) valid concerns about equity and justice aspects of valuation, particularly regarding human mortality; (2) less valid concerns of physical scientists who thought that the damage estimates were too low (the value of a statistical life is one of the most influential parameters in total damage estimates); and (3) the fact that the controversy gave some people the opportunity to shed a nasty light on the IPCC and try to break down at least one of its Working Groups.

Uncertainties

No assessment of uncertainties with regard to the social costs of climate change has been made as yet, but generally, the uncertainties are believed to be large. The chapter presents a range of best guess estimates. On global aggregate, impacts are expected to amount to 1 - 2 percent of GDP in a doubled CO_2 climate, on an annual basis. Regionally, this varies from slightly positive impacts for the former USSR and, some believe, China, to mildly negative impacts in OECD countries, to as much as a 10 percent negative impact in Africa. The ranges do not reflect uncertainty, however; rather, they reflect different analysts' best guesses.

These conclusions are based on only five studies; two others were ignored, a third was unheard of at that time, and a fourth appeared too late to be taken up. Tol says it is remarkable that the referees did not notice the omission of these four other studies, which are not unknown in certain circles. Three of the five authors of the studies which were used were involved in

the writing of the chapter and this may reflect certain biases. All of the five original analysts also drew heavily on the same set of underlying estimates in their work. The already wide range of best guesses mentioned above largely reflects difference in interpretation of the same "empirical" material. This, plus the small amount of information available and the small number of analysts interpreting it, leads Tol to believe that the uncertainty is quite large (but unknown). Nordhaus' 1994 poll of experts supports this notion (see Session Synthesis Essay, The Nordhaus Study in this report).

Most of the studies on which the chapter relies are based on one or two case studies which are then extrapolated to a whole nation or even to the whole world. For example, all the heat stress effect estimates were based on New York data. Agricultural impacts were based on 60 case studies spread over the world and used to obtain a global aggregate. Each of the case studies has its own uncertainty, and extrapolation brings additional uncertainties to the analysis. It is difficult to capture extrapolative uncertainties, but attempts should be made. The best strategy to reduce uncertainties is to expand the number of case studies. Tol believes that, in general, the chapter conveys a message that knowledge is stronger and uncertainties smaller than is actually justified.

The ranges of economic impacts estimates do not reflect uncertainty; rather, they reflect different analysts' best guesses.

Personalities

The main body of the first draft was prepared by one individual. This was revised and extended by another individual. These two (of seven) authors thus had the largest influence on the actual wording of the chapter, and they are known to be somewhat more pessimistic about the impacts of climate change than most mainstream economists.

Outliers

Regarding outliers, the group ignored two studies of comprehensive damage estimates, one because it was somewhat outdated and in the "gray" literature, and the second for the same reasons and also because it took two debatable views: one is that human statistical lives were valued equally and at the OECD level, and the second is that it included the impacts of potential increases in hunger and famine (whereas the five studies used in the report did not). The reasons for not considering hunger in the chapter (at least, quantitatively) include the fact that it would have dominated impact estimates if it was taken into account while the authors did not find the estimates of the increase of hunger attributed to climate change to be reliable. Agronomists have not reached beyond the stage of food availability; most people feel food accessibility is at least as large a driver of hunger, Tol explained.

However, a different outlier was included (perhaps because the author of this study was on the lead author team). That study involved the marginal benefits of GHG emission abatement. The "marginal benefit" of a tonne (metric ton) of carbon abatement is the difference in damage from emitting one tonne of carbon less than otherwise would have been emitted. The range of marginal damage estimates reported was $5 per tonne up to $125 per tonne, but most estimates were concentrated between $10 and $25 per tonne. The $125 figure comes from using a pessimistic estimate of the impact of climate change, expressing this as a function of the magnitude of climate, projecting it far into the future with no change in society's ability to adapt, and using a very low discount rate. This figure was an outlier and should have been excluded (at least, from the policymakers summary), Tol believes, but was not.

Valuation

The primary aim of the social cost chapter was to derive a comprehensive, monetary estimate of all impacts of climate change. The estimated monetary impact is the estimated impact multiplied by its estimated price. A monetary analysis is easy to do if there is a market impact, land loss due to sea level rise for example, because the price is more or less known, but much harder to do for things not traded in markets. There are basically three methods to value goods or services not traded in markets. First, is the travel cost method. People are willing to spend a certain amount to go to a particular place and that is one way of measuring the value of that place. Second, is the Hedonic method. A house on a beautiful site is worth more than the same house in an industrial area; the price difference is an indication of the worth of that beautiful site. Both of these methods are subject to measurement and value uncertainty, however, and do not capture total value. A third method is to directly interview people as to their preferences. People are asked, "Would you be willing to spend a certain amount to preserve a certain environmental good?" This presents other difficulties as most people do not have clearly expressed preferences in their minds.

The primary aim of the social cost chapter was to derive a comprehensive, monetary estimate of all impacts of climate change. The estimated monetary impact is the estimated impact multiplied by its estimated price.

Most importantly, all three of these methods are highly context specific. There are strong theoretical and empirical indications that we cannot take a measure of a non-market item from one place or one issue and use it to predict the value of a similar item in another place for another issue, but that is exactly what was done for this social cost assessment. Thus, the uncertainty is large but, again, unknown and perhaps unknowable, Tol says.

The most difficult aspect of this valuation is that monetary values are always income dependent. A poor person would be willing to spend less in absolute terms than would a rich person. This brings us back to the problem of the value of a statistical life. Based on income differences, the authors ended up saying that an American is "worth" $1.5 million while a person from a low-income country such as India or China is "worth" $150,000, which led to the controversies mentioned earlier. There was some discussion around this issue, with various suggestions from the group about using value units other than dollars for human lives. Tol pointed again to the usefulness of being able to value all things in the same units in order to enable comparison of costs and benefits.

Equity

Economics is a science based on the presumption that there is a strong nation state which clearly defines and enforces property rights and which also guarantees some form of compensation between losers and winners of certain policies. The general approach taken to issues such as the valuation of lives is that one takes an average value within one nation. This approach tacitly assumes compensation between citizens. At the international level, however, property rights are not clearly defined or enforced and there are no compensation rules. Thus it makes sense to differentiate values between countries. Otherwise, inconsistencies between national policies on global issues and national policies on domestic issues would arise. If one would use an OECD value globally, it would require developing countries to put more money into climate change abatement than into basic health care. This, of course, would not make sense, Tol explains, whereas differentiated values do.

Another issue that arises in the context of valuation is responsibility for climate change. Though most of the historical accumulation of GHGs is due to OECD nations, developing countries are more vulnerable to the impacts. Why should damages be ascribed to developing

countries when they are not, in fact, responsible? To some this means that higher impacts should be ascribed to the OECD countries, making them responsible for paying for the damages to developing countries which the OECD countries caused. In addition, people value things differently depending on perceived property rights and responsibilities. It makes a great difference whether an increase in mortality risk is voluntarily, or is imposed by someone else, particularly if that someone else reaps the benefits and is much wealthier. In the latter case, theoretical and empirical results support the intuitive notion that the value of a statistical life is (much) higher.

On the whole, despite the considerable attention the chapter received, the comments in both the expert and government reviews were not very numerous or very substantive. In response to the concern that comments on the chapter may not have been taken seriously and adequately addressed by the authors, Tol says that the text was revised to reflect concerns raised but the bottom line estimates were not. The justification for this was a strict interpretation of IPCC rules directing the authors to reflect what was in the published learned literature, and most of the concerns raised were not in the published learned literature (note that the interpretation may have been so strict exactly for this purpose) or in the review.

In addition, Tol points out that economic impact assessments are so rudimentary that it comes down to discussing numbers rather than the methods for arriving at them. Therefore, concerns and caveats could not be calculated into the values, at least not in a manner that the strict interpretation of the IPCC rules would allow the authors' team. In addition, tensions rose so high that the authors' team was not inclined to address everyone's concerns. Others point out that virtually nothing has been spent on thinking systematically about these issues compared to what has been spent on the physical science of climate change, and that consequently, the quality of the work should not be judged on the same basis.

Tol points out that economic impact assessments are so rudimentary that it comes down to discussing numbers rather than the methods for arriving at them.

Recent Changes in El Niño/Southern Oscillation and Climatic Implications

Kevin Trenberth

National Center for Atmospheric Research
Boulder, Colorado

The focus here is on time series analyses which attempt to quantify how unusual recent behavior in ENSO has been.

Trenberth focused on the climatic significance of and changes in the El Niño/Southern Oscillation (ENSO) phenomenon. He also discussed the "mini global warming" that accompanies ENSO and which affects the Northern Hemisphere temperature record. In reference to James Hansen's data that show 1995 to be the warmest year on record, Trenberth highlights the pattern of the warming, which came largely in the first six months of the year and largely over North America and Eurasia. This pattern suggests that the observed warming may not be entirely due to the enhanced greenhouse effect because atmospheric circulation patterns, specifically the North Atlantic Oscillation and ENSO, appear to play a contributory role.

Northern Hemisphere winter temperature data for 1977-1993 reveal an overall warming anomaly relative to a base period (1951-1989) with a characteristic pattern related to atmospheric circulation. Trenberth showed data which suggest that changes in the North Atlantic Oscillation and the Southern Oscillation account for the major portion of this observed anomaly. Why are these atmospheric circulation patterns changing as they are? Gaining a fuller understanding of natural variability modes in the climate system is of central importance in this inquiry.

Much of Trenberth's recent work has focused on analyzing how unusual the recent behavior of ENSO has been. An article entitled, "What is Happening to El Niño?" by Trenberth appears in the 1997 Yearbook of Science and the Future, Encyclopedia Britannica, and deals more comprehensively with this subject. The original scientific article in which this work is reported appears in Geophysical Research Letters , 23:57-60, 1996. The focus here is on time series analyses which attempt to quantify how unusual recent behavior in ENSO has been.

"El Niño" technically refers only to the ocean warming off South America, but it is linked to important basin-scale changes in temperature in the tropical Pacific Ocean. In one extreme El Niño event in December of 1982, the Western Pacific warm pool extended well out into the tropical Pacific. Temperature gradients are reflected in surface pressure gradients which determine wind patterns and thus rain and atmospheric circulation patterns. The Southern Oscillation is a global wave pattern that is the atmospheric part of this phenomenon. Together, these make up a whole cycle of warming and cooling which we call ENSO, a natural mode of the coupled tropical Pacific ocean-atmosphere system arising from air-sea interactions. This is basically the mechanism by which El Niño influences climate around the world.

El Niño events occur on average every 3 to 6 years; paleoclimatic data (from corals and ice cores) show that such events have been going on for millennia. Characteristically, we see a warming in the traditional El Niño region of the tropical Pacific that is connected to a

warming across the basin and to a boomerang-shaped cooling in the North and South Pacific caused directly by changes in atmospheric circulation (SO). La Niña is a cold event in which a cold tongue of water extends along the equator into the warmer water and causes changes in rainfall patterns in the tropical Pacific and changes in temperature patterns, storm tracks, and atmospheric circulation in North America and elsewhere.

Correlations associated with the SO were discussed, related to annual mean sea level pressure. When pressure is high at Darwin, Australia, it tends to be high across the entire region extending across Africa all the way to Brazil. At the same time, there are low pressure conditions across the central and eastern Pacific, north Pacific and southern oceans which lead to cloudy, rainy weather (while high pressure in other regions leads to more dry and settled conditions). This pattern leads to droughts in Australia, Africa and parts of Brazil and Columbia, and excessive rains in other regions. There is also a structure in this pattern, associated with wintertime conditions, that is called the Pacific /North American teleconnection pattern.

The two data-collecting stations that have the highest negative correlations of any in the world are Darwin and Tahiti. At these two stations, sea level pressures vary, very distinctly, in opposite directions (see Figure 2.17). The difference between these two stations is thus often used as an index of the Southern Oscillation. One way of completely recovering the information from two stations is to take both the difference and the sum. The SO signal is the difference and the sum, a measure of noise, or non-SO variations. The signal to noise ratio for these two monthly series is 2.0. An appropriate smoother is used to improve the signal to noise ratio to 6.4 and this helps to reveal all the El Niño and La Niña events that correspond to the fluctuations in the Southern Oscillation. Many noise features are simply related to individual small-scale temperature events that are not related to the see-saw of the SO events. The key in analyzing this phenomenon is to focus on the phenomenon itself, and not get hung up on all the noise inherent in the system.

El Niño events occur on average every 3 to 6 years; paleoclimatic data (from corals and ice cores) show that such events have been going on for millennia.

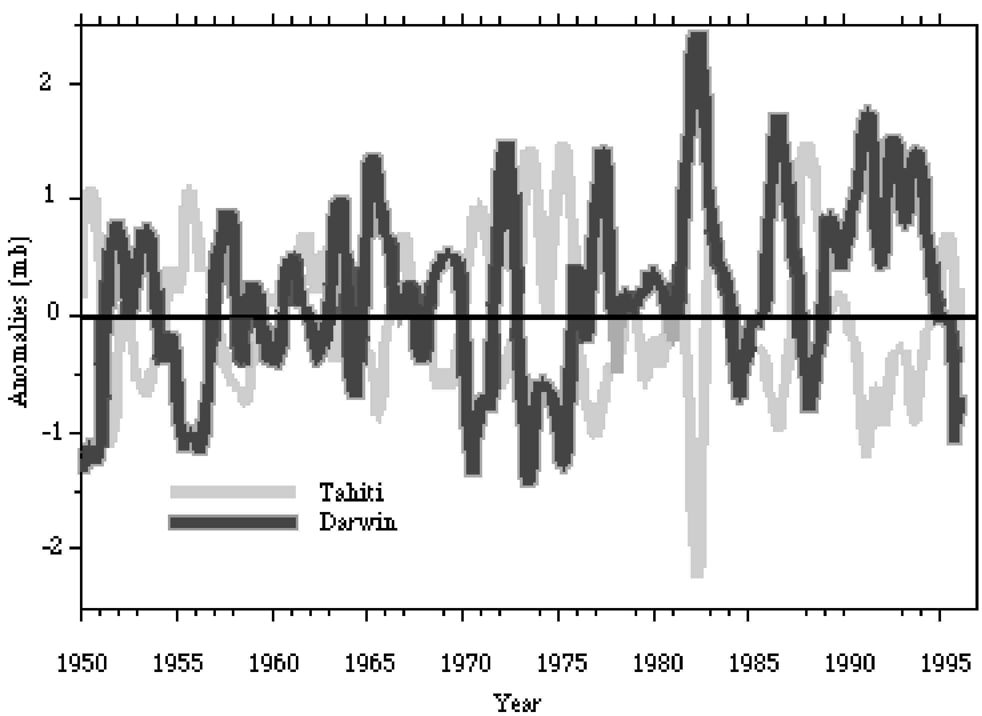


Figure 2.17
Darwin and Tahiti Pressure Anomalies

One problem is that the early record from the Tahiti station is noisy, incomplete and unreliable due to the fact that records were stored in a documentary form that was affected by bugs, mold, etc., so Trenberth doesn't use this older data. There are almost no sea surface temperature (SST) data from the tropical Pacific prior to 1950. It is thus very difficult to get reliable measurements of EN during that period of time. At the same time, it is clearly important to have a homogeneous time series of EN behavior in recent times. The SO can be reconstructed back to 1882.

After about 1976, a series of unprecedented events occurred. One La Niña and several El Niño events, including the very long 1990-1995 EN event, are unlike anything else in the record. Questions are thus raised regarding how unusual this behavior actually is and why it occurred. Prior to describing in more detail the time series analyses used to research these questions, Trenberth showed a correlation of SSTs with the SO from 1950 to 1994 in the Central Pacific which reveal the characteristic boomerang pattern and with warming along the coast of California associated with EN events (see Figure 2.18). The SO is most strongly related to SSTs in the central Pacific, not along the coast of South America.

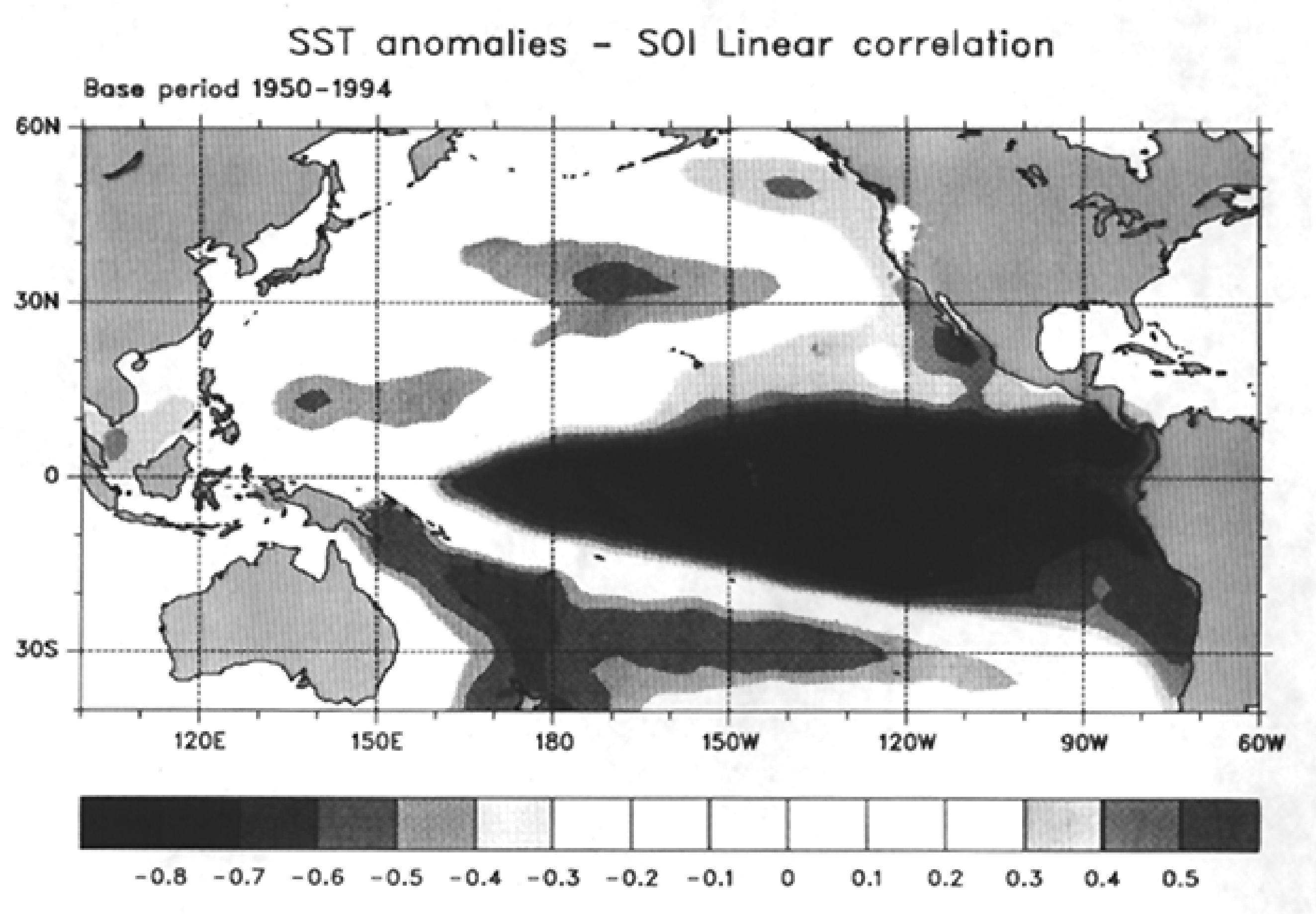


Figure 2.18

When we look at SSTs in the region Trenberth calls "Niño 3 and a half" (5°N-10°S, 180°-120°W) we see a tendency for more EN behavior recently including the prolonged event of 1990-1995, while along the coast of South America, the traditional EN region, we see more ups and downs and less tendency for typical EN behavior. Activity in this central Pacific "Niño 3 and a half" region seems to generate most of the global consequences. Looking at the recent prolonged EN event of 1990-1995, the question arises: is this a manifestation of global warming?

Looking at recent behavior in the context of the past record, we can see that the recent prolonged EN lasted from October 1989 to June 1995, 5 years and 9 months. The next longest EN in the record was from 1911 to 1915, 4.11 years. From 1906 to 1911 there was a 5-year event of opposite sign. Looking at SSTs from 1950 forward, we see that in the period from September 1989 to August 1995, sea surface temperatures were above normal. In terms of seasonal values for the Southern Oscillation, the most recent event was 22 seasons all of one sign (negative), from 1989 to 1995. The longest previous such event lasted for 15 seasons, from 1894 to 1897, and the longest such event of opposite sign lasted 13 seasons. In this context, recent behavior looks very unusual, so statistical tests were conducted to try to quantify just how unusual this behavior is.

After about 1976, a series of unprecedented events occurred. One La Niña and several El Niño events, including the very long 1990 -1995 EN event, are unlike anything else in the record.

Allowing for appropriate degrees of freedom based on the effective time between independent samples, the test results reveal a strong statistical significance of these recent events. For the 66 months from December 1989 to May 1995, a high statistical significance of t=2.82, departure 0.87 mb, significance at 0.5 percent was found. For the period of March 1977 to May 1995, t=2.68, significant at 1 percent.

Then, autoregressive-moving average (ARMA) modeling was used to better deal with persistence in the time series, using seasonal anomalies in the Darwin station data. The model was used to generate a synthetic one million-year record that contains many simulated ENSO events, averaging one about every four years, that can be tested to see how unusual the recent behavior is. Summarizing the statistical conclusions that emerge from this work: the low frequency variability and the negative trend in the Southern Oscillation Index observed in recent decades are quite unusual. Likelihood of occurrence of the 1990-1995 ENSO, given the first 100 years of the record, is once in 1,500 to 3,000 years. The change beginning about 1976 toward more frequent ENSO events is also unlikely, about a once in 2,000 year expectation. These two events are not independent and the results suggest a non-stationary influence such as would arise from a change in climate.

How will ENSO events change with global warming? ENSO involves a build up and depletion of heat as well as major redistribution of heat content in the ocean and the atmosphere during the course of events. Because GHGs trap heat, they alter the heat budget. As such, they can possibly expand the Pacific warm pool and enhance the rate of recharge of heat losses. Models which include crude simulations of ENSO show that conditions become more ENSO-like with increased GHGs. Second, the hydrological cycle is expected to speed up with increased GHGs. Increased evaporation enhances the moisture content of the atmosphere which makes moisture more available for rainfall. ENSO-related droughts are apt to be more severe and last longer, while floods are likely to be enhanced.

In discussion, MacCracken pointed out that this work demonstrates that what may seem like relatively small and subtle changes in temperature can cause dramatic changes in elements of the climate system like the ENSO phenomenon and in impacts such as precipitation. Further, he pointed out that while we often speak of phenomena like ENSO and changes in them as “natural,” it makes sense that as the climate warms, it would take advantage of such elements of the system, perturbing these natural modes of the system first and evidencing itself in this way.

Treatment of Uncertainty in Integrated Assessments

John Weyant

Stanford University
Stanford, California

Weyant focused on how uncertainty is handled in integrated assessment models (IAMs). Beginning with the general issue of uncertainty in modeling and policy analysis, Weyant points out that the formal discipline of decision analysis recognizes that uncertainty should not be used as a reason for indecision; no change in the status quo is actually a significant decision. He also distinguishes between long range projections and those useful for decision making, saying that what is important are the implications for each time frame. Conceptually, we only have to decide what to do today; over time, we can revise projections as well as decisions. Looking at the history of modeling analysis, Weyant says that the insights provided by such analysis are often a more important product than the numerical projections.

Focusing for a moment on the Energy Modeling Forum (EMF), Weyant says its primary goal has been to prove the use and usefulness of analysis. In the EMF process they have looked at why different results come from different modeling systems for the same questions, what information and insights from the collection of tools available are useful to decision makers, and what the gaps are in the information and analysis base that need to be filled.

Does explicitly adding uncertainty into models lead to better insights? Not always, Weyant says. Sometimes, simple deterministic models give very good answers. Does factoring in uncertainty lead to better numerical predictions? Weyant says it probably helps in deciding what and how to communicate to policymakers about the numerical predictions. He also points out that while we say insights are more important than numbers, in the policy process, numbers do matter. Decision makers want and need to hear at least order of magnitude trade offs when asked to make policy.

In the EMF studies, the procedure is to use standardized model comparisons to take that source of variation out of the process. Also, because model variability might depend on what scenario is used, EMF uses a standard reference case and then alternative scenarios with different variables. Then, due largely to how the results were being used, they began also doing "modelers' reference cases" in which modelers can designate preferred values for the key factors: population, economic growth, some minimal technology assumptions, and fuel prices. Doing these modelers' references cases led to a spreading out of the predictions of carbon emissions in the year 2100, almost doubling the range from the standardized reference case.

Because uncertainty is such an important issue, EMF formed an uncertainty study group to work with models that were specifically designed to explore the implications of uncertainty, rather than using the full deterministic integrated assessment models to explore these issues.

Uncertainty should not be used as a reason for indecision; no change in the status quo is actually a significant decision. ... Conceptually, we only have to decide what to do today; over time, we can revise projections as well as decisions.

Weyant says education is the main purpose of integrated assessment modeling. First, analysts need to become educated, and then they need to communicate to policymakers and the public. Modeling is primarily useful for gaining insights, not specific numbers. It can be used to identify smart and ill-advised policies and to identify previously unrecognized interactions and feedbacks. However, he reiterates, it must be recognized that numbers do matter to policymakers, whether or not they should. Models can help with numerical predictions, but we should be careful of overstating their ability in this regard. They can help to project rough trade-offs between competing objectives, using sensitivity and uncertainty analyses to check for robustness of conclusions.

The authors of Working Group III, Chapter 10 recognized that there are two distinct types of deterministic models. The first group are top -down, bottom line economic models, such as Nordhaus' DICE model, which include a simple carbon cycle, climate sensitivity and climate damages. They short circuit much of the action of a full-scale IAM regarding ecosystems, atmospheric composition, etc. The authors of Chapter 10 chose to call these "Policy Optimization Models" as they seek to balance costs and benefits (e. g., in determining the appropriate level of a carbon tax), aggregating everything to a simple dollar metric.

Education is the main purpose of integrated assessment modeling. It can be used to identify smart and ill-advised policies and to identify previously unrecognized interactions and feedbacks.

Models belonging to the second group, called "Policy Evaluation Models," include more physical properties and grow more from the physical climate modeling tradition. These models include human activities as well as more detailed representations of emissions, atmospheric composition, and climate (not incorporating full-scale GCMs, but generally two-dimensional simulations). Each modeling group specializes in a few main impact areas, some including economic evaluation and some not. These models do not attempt to aggregate all damages to the nearest dollar metric. Weyant had hoped to find a method of aggregating these two major approaches but found that deep philosophical differences between the two made that unfeasible at present.

Each of these two modeling approaches is critiqued by practitioners of the other in various ways. For example, the more complex physical modelers would say that a simple cost-benefit model tells us that a 3°C temperature rise will lead to a 2 percent reduction in global GDP, but it does not tell us whether this loss means 100,000 people drowned in Bangladesh or less sunny beach days in California. More detail is needed for policymakers to find such models useful, the argument goes.

From the other point of view, economists say that particular threshold limits suggested by the more complex models, such as the suggestion that we contain emissions enough to ensure no more than one category of change within each vegetation biome type, include no consideration of what this would cost and that it could turn out to be more than it's worth, in economic terms. Much can be gained, Weyant suggests, by going back and forth between the two approaches, using the most detailed and robust representations from each to strengthen the validity of the overall conclusions.

The desire to include uncertainty has led to two additional sets of models Weyant calls Partial and Full Stochastic Simulation Models which generate composite probability distributions. This last group, called "Decision Making Under Uncertainty Models" includes the model described in this report by Schlesinger.

Preliminary insights from Policy Optimization Models include the value of timing flexibility, which seems to suggest the value of more emissions control in the middle term and long term and perhaps more research and development in the short run, and the value of emissions trading, even if the entire world is not involved in the program.

From the Policy Evaluation Models, preliminary insights include the likelihood of sulfur aerosol policy complications, e. g., China could greatly reduce its sulfur emissions for acid rain control and other reasons only to find that doing so exacerbates climate warming. Another issue raised by these models is the land use competition implications of large - scale biomass energy development.

From the Stochastic Decision Making Under Uncertainty Models, insights include the value of hedging against bad outcomes and the robustness of research and development policies. R&D on new technologies is the one thing policymakers can't take back after the fact; there is potential for OECD countries to sponsor technologies that are salable domestically as well as transferable to the developing world.

Dealing with Uncertainty in IAMs

Weyant briefly listed approaches for dealing with uncertainty in IAMs. First, under the Stochastic Optimization approach, he mentions stochastic math programming and decision analysis, hedging calculations, partial stochastic sensitivity analysis, and value of information calculations. Under Stochastic Simulation Models he includes simulation and dynamic programming, computing probabilities of meeting stated objectives, full stochastic sensitivity analysis, and richer learning and knowledge accumulation representations. He also mentions option value approaches.

Weyant then described a simple experiment which compared a "clairvoyant case" with Sequential Decision Making Under Uncertainty (7 models could do this experiment). The numbers used are based on Nordhaus' quantification of a combination of the Morgan/Keith and Nordhaus surveys of climate experts (see Session Synthesis Essay in this report). Assuming uncertainty regarding the optimal level of emissions is resolved in 2020, and using a 5 percent discount rate, Figure 2.19 shows carbon emissions through the year 2100 based on various policies, and Figure 2.20 shows the incremental cost/value of carbon emissions through 2100. This analysis suggests that hedging is a good strategy. Doing more than would be done by following a mere cost-effectiveness test is at the core of all of their suggestions, Weyant says.

This analysis suggests that hedging is a good strategy. Doing more than would be done by following a mere cost-effectiveness test is at the core of all of their suggestions, Weyant says.

Hedging Against Bad Climate Outcomes

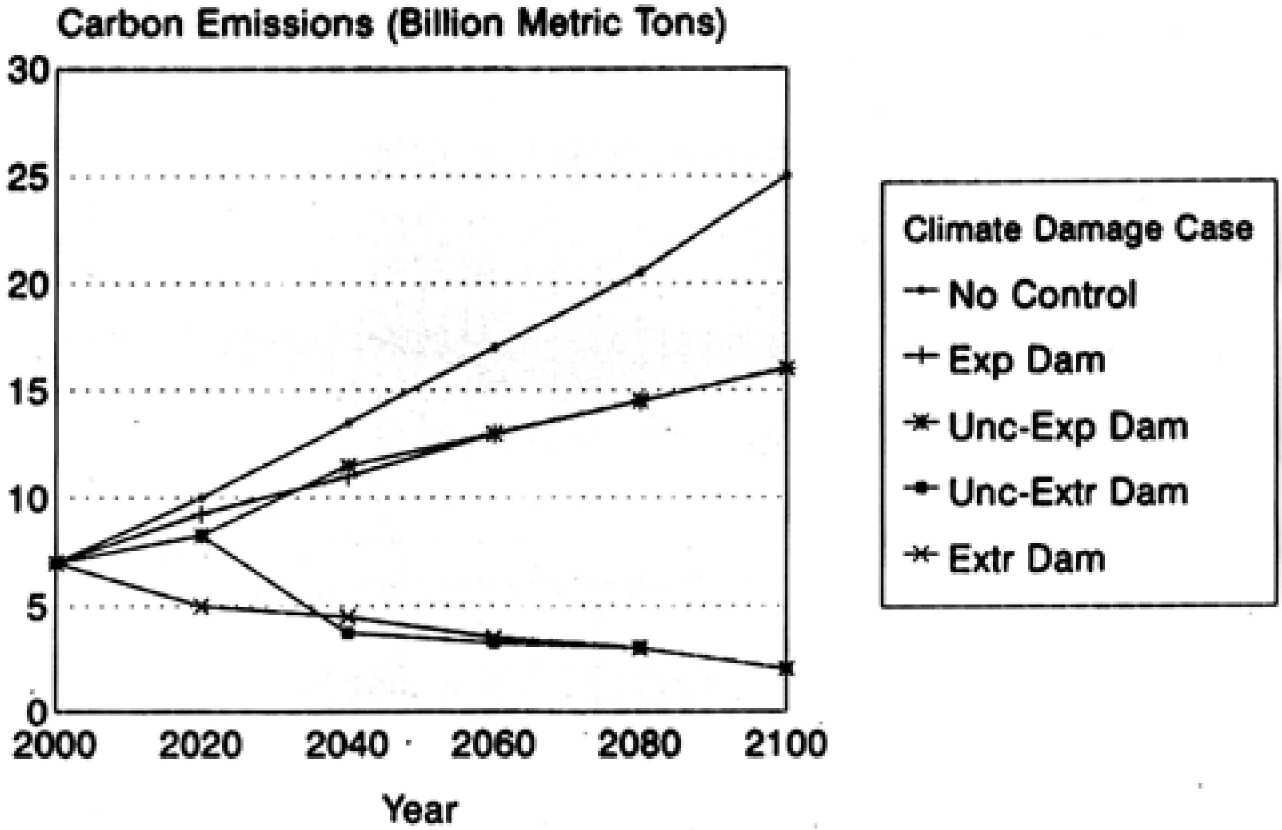


Figure 2.19
* "Exp" means expected conditions
* "Extr" means extreme conditions
* "Unc" means under uncertainty prior to 2020

Thus, Exp Damage means planning for expected damages throughout the time horizon (1990-2100), while Unc-Exp Dam means planning under uncertainty until 2020 and then learning that damages are, in fact, as expected. Unc-Extr means planning under uncertainty until 2020 and then learning that the extreme damages are in effect.

Source: Manne, A. S. (1996). "Hedging Strategies for Global Carbon Dioxide Abatement: A Summary of Poll Results - EMF 14 Subgroup; Analysis of Decisions Under Uncertainty," Energy Modeling Forum Workshop Paper 14.2, February.

Hedging Against Bad Climate Outcomes

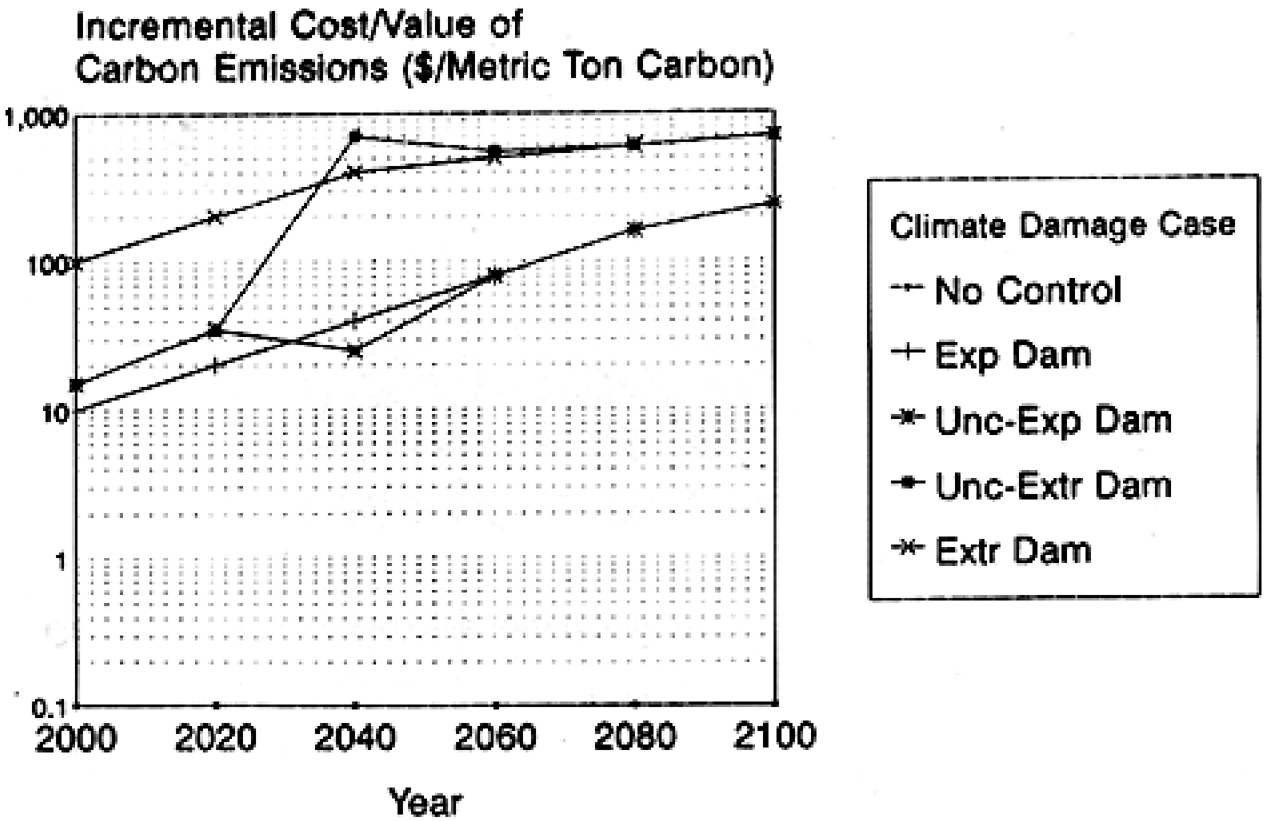


Figure 2.20
* “Exp” means expected conditions
* “Extr” means extreme conditions
* “Unc” means under uncertainty prior to 2020

Thus, Exp Damage means planning for expected damages throughout the time horizon (1990-2100), while Unc-Exp Dam means planning under uncertainty until 2020 and then learning that damages are, in fact, as expected. Unc-Extr means planning under uncertainty until 2020 and then learning that the extreme damages are in effect.

Source: Manne, A. S. (1996). “Hedging Strategies for Global Carbon Dioxide Abatement: A Summary of Poll Results - EMF 14 Subgroup; Analysis of Decisions Under Uncertainty,” Energy Modeling Forum Workshop Paper 14.2, February.

Conclusions

Many integrated assessment insights are available without much consideration of uncertainty and these can help put numbers in perspective, at least suggesting order of magnitude trade-offs. Treatment of uncertainty in integrated assessment is still in its infancy. This work should improve and expand upon available insights. Uncertainty analyses can put the numbers in perspective. More work is needed on the impact of uncertainty on economic agents. New work is also needed on multi-party negotiations and coalition formation, Weyant says. In addition, culture theory, strategic scenarios, and attempting to account for changing preferences are examples of promising approaches that are difficult to incorporate into models but could broaden their perspective.

Participant Roster

Richard Berk
Interdivisional Program in Statistics & Sociology
UCLA

Robert S. Chen
SEDAC Project Office
CIESIN

Allin Cornell
Civil Engineering
Stanford University

Sharon Dunwoody
School of Journalism & Mass Communication
University of Wisconsin, Madison

Jae Edmonds
Economics Program
Battelle Pacific NW Lab

Paul Edwards
Program in Science, Technology, and Society
Stanford University

Thomas Karl
NOAA/U. S. Department of Commerce
National Climatic Data Center

Michael MacCracken
U. S Global Change Research Program

John J. Magnuson
Center for Limnology
University of Wisconsin, Madison

Linda O. Mearns
Climate & Global Dynamics Division
NCAR

Granger Morgan
Department of Engineering and Public Policy
Carnegie Mellon University

Richard Moss
Intergovernmental Panel on Climate Change
U. S. Global Change Research Program

Mohan Munasinghe
World Bank Field Office

Nebojsa Nakicenovic
Environmentally Compatible Energy Strategies Project
IIASA

Elisabeth Paté-Cornell
Civil Engineering
Stanford University

Rick Piltz
U. S. Global Change Research Program

John M. Reilly
NRED, Economic Research Service
U. S. Department of Agriculture

Terry Root
School of Natural Resources
University of Michigan

Ben Santer
PCMDI
Lawrence Livermore National Laboratory

Michael Schlesinger
Department of Atmospheric Sciences
University of Illinois

Stephen H. Schneider
Dept. of Biological Science
Stanford University

Simon J. Shackley
Centre for the Study of Environmental Change
Lancaster University

Allen Solomon
National Health & Environmental Effects Research Laboratory
U. S. EPA, Western Ecology Division

Holly Stocking
School of Journalism
Indiana University

James Titus
Office of Policy, Planning and Evaluation
U. S. Environmental Protection Agency

Richard Tol
Institute for Environmental Studies
Vrije Universiteit

Kevin Trenberth
Climate Analysis Section
NCAR

John P. Weyant
Energy Modeling Forum
Stanford University

Aspen Global Change Institute
100 East Francis Street
Aspen CO 81611
http://www.agci.org
970 925 7376
agcimail@agci.org

Furthering the understanding of Earth Systems and Global Environmental Change