

ISSUES

IN SCIENCE AND TECHNOLOGY

NATIONAL ACADEMY OF SCIENCES
NATIONAL ACADEMY OF ENGINEERING
INSTITUTE OF MEDICINE
THE UNIVERSITY OF TEXAS AT DALLAS
ARIZONA STATE UNIVERSITY
SPRING 2015

A New Model for the American Research University

Clean Energy Diplomacy
from Bush to Obama

Physics Envy: Get Over It

The Limitations of Climate
Models as Guides for Policy

Welcome to the Anthropocene

Empowering Social Science

An Excess of Research Space?

First Science Fiction
Contest Winner



ANDREA SALTELLI
PHILIP B. STARK
WILLIAM BECKER
PAWEL STANO

Climate Models as Economic Guides Scientific Challenge or Quixotic Quest?

The uncertainties associated with mathematical models that assess the costs and benefits of climate change policy options are unknowable. Such models can be valuable guides to scientific inquiry, but they should not be used to guide climate policy decisions.

In the polarized climate change debate, cost-benefit analyses of policy options are taking on an increasingly influential role. These analyses have been presented by authoritative scholars as a useful contribution to the debate.

But models of climate—and especially models of the impact of climate policy—are theorists’ tools, not policy justification tools. The methods used to appraise model uncertainties give optimistic lower bounds on the true uncertainty, at best. Even in the finest modeling exercises, uncertainty in model structure is presented as known and manageable, when it is likely neither. Numbers arising from these modeling exercises should therefore not be presented as “facts” providing support to policy decisions.

Building more complex models of climate will not necessarily reduce the uncertainties. Indeed, if previous experience is a guide, such models will reveal that current uncertainty estimates are unrealistically small.

The fate of the evidence

Climate change is the quintessential “wicked problem:” a knot in the uncomfortable area where uncertainty and disagreement about values affect the very framing of what the problem is. The issue of climate change has become so resonant and fraught that it speaks directly to our individual political and cultural identities.

Scientists and other scholars often use non-scientific and value-laden rhetoric to emphasize to non-expert audiences what they believe to be the implications of their knowledge. For example, in *Modelling the Climate System: An Overview*, Gabriele Gramelsberger and Johann Feichter—after a sober discussion of statistical methods applicable to climate models—observe that “if mankind is unable to decide how to frame an appropriate response to climate change, nature will decide for both—environmental and economic calamities—as the economy is inextricably interconnected with the climate.” Historians Naomi Oreskes and Erik M. Conway, in their recent book *The Collapse of Western Civilization* (2014), paint an apocalyptic picture of the next 80 years, beginning with the “year of perpetual summer” in 2023, and mass-imprisonment of “alarmist” scientists in 2025. Estimates of the impact of climate change turn out to be far too cautious: global temperatures increase dramatically and the sea level rises by eight meters, resulting in plagues of devastating diseases and insects, mass-extinction, the overthrow of governments, and the annihilation of the human populations of Africa and Australia. In

the aftermath, survivors take the names of climate scientists as their middle names in recognition of their heroic attempts to warn the world.

That the Earth's climate is changing, partly or largely because of anthropogenic emissions of CO₂ and other greenhouse gases, is not the core of the dispute. Instead, the focus has moved to the more difficult questions of "what will be the consequences?" and "what is the cost of offsetting climate change?" Opposing factions use scientific uncertainty instrumentally either to deny or to support the urgency of action to mitigate adverse outcomes. In their previous book, *Merchants of Doubt*, Oreskes and Conway talk of uncertainty amplification by unscrupulous stakeholders and draw parallels between the denial of the health effects of smoking by the tobacco lobby and the denial of climate change by climate sceptics. Conversely, climate scientist Richard Lindzen, a possible source of Michael Crichton's novel *State of Fear*, compares climate change to eugenics at the beginning of the 20th century: a wrong idea vigorously supported by scientists. In the case of eugenics, this included some of the founding fathers of statistics.

In this battlefield of norms and worldviews, what is the fate of scientific evidence?

Given the economic and societal ramifications of climate change, it is not surprising that several disciplines claim to provide certainties and solutions. Among these, computer modeling is perhaps the most visible, pervasive, and opaque.

A recent example that rests on such models is the *American Climate Prospectus: Economic Risks in the United States*, a study commissioned by the Risky Business Project, a non-profit group chaired by former New York mayor Michael Bloomberg, former U.S. Treasury Secretary Henry Paulson, and environmental philanthropist Tom Steyer. The report invokes a broad battery of models for the impact of climate change, predicting when, where, and how much temperature variation to expect and what the costs will be to address the consequences.

The claimed precision and resolution of these modeling efforts are staggering. The Risky Business report forecasts—at the level of individual counties in the U.S.—energy costs and demand, labor supply, mortality, violent crime rates, and real estate property prices up to the year 2100. (Evidently, these investigators were not enlisted to predict the collapse of real estate prices from subprime mortgage crisis that triggered the recent recession, or that economic catastrophe might have been avoided.) The report presents the amount of computer power and data

generated as evidence of the scientific legitimacy of the enterprise. The authors note, however, that out of an abundance of caution they did not model deterioration in cognitive performance as temperatures rise.

Other models, such as those reported in a 2013 *Nature* paper by Camilo Mora and colleagues, predict when the mean climate in various locations will move outside the range of historical variability under alternative greenhouse gas emissions scenarios, prompting the *Washington Post* headline, "D.C. climate will shift in 2047."

To quantify the uncertainty in model predictions—and especially in the response to changes in drivers such as anthropogenic CO₂—requires quantifying the uncertainty in the assumptions of the model, the structure of the model, the physical approximations in the model, the data used to calibrate the model, and the parameters of the model. As a result, forecasts of global average temperature even a few years into the future are extremely uncertain, even at the level of global averages—and likely much more so at the level of counties or cities. For example, in the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC), forecasts for the period 2071-2100 generated from 12 different climate models predict that England will experience anywhere from a 60 percent decrease to a 10 percent increase in summer precipitation. Our own technical research has shown that if one accounts for all the uncertainties built into Nicholas Stern's well-regarded and influential *The Economics of Climate Change*, the uncertainty in the cost of climate damage is so large as to preclude any meaningful conclusion as to the urgency to act at the present time to counteract it. The same critique applies to those who, like William Nordhaus, conclude from the same type of cost-benefit analyses that action can be delayed, in contrast to Stern.

Some major climate modeling exercises strive to give a level-headed assessment of uncertainty. The UK Meteorological Office attempted to quantify the overall uncertainty in climate forecasts provided by their HadSM3 model, using simple numerical approximations of the more complex, but still approximate, numerical climate model to account for the uncertainty in a wide range of model parameters. This is laudable, but such "emulators" of climate models have their own large uncertainties as approximations of the already-approximate, high-dimensional models.

Likewise, the fifth assessment report of the IPCC gives qualitative descriptions of uncertainty for forecasts and parameters. For example, the assessment

report states, “Equilibrium climate sensitivity is likely in the range 1.5°C to 4.5°C (high confidence), extremely unlikely less than 1°C (high confidence), and very unlikely greater than 6°C (medium confidence).” These qualifiers reflect the type, amount, quality, and consistency of the evidence, using probabilistic language to describe uncertainty and expert judgment. Unfortunately, the probabilistic assessments appear to rely on stylized, untestable assumptions without an empirical basis, and the expert judgments are, at the end of the day, opinions rather than facts—opinions influenced by the very models whose reliability is in question. Indeed, the 1.5°C to 4.5°C uncertainty range has persisted through 30 years of advancing climate science; as long ago as 1998, Jeroen van der Sluijs and others showed that the range was arrived at through what amounts to a back-of-the-envelope estimate whose stability over time is much more a social and political phenomenon than the result of any sound statistical method.

Although climate models have become more complex and techniques for evaluating their uncertainty have evolved in the past decades, quantifying their uncertainty usefully in climate models remains impossible. In many cases little or no information is available to justify even crude ranges of uncertainty assumed for various inputs. “Expert opinion” is not a substitute for data or tested physical theory.

When the uncertainties are uncertain

It may be possible, at least in principle, to quantify the uncertainties of inputs to the model—the range of plausible values of a given driver. But “structural uncertainty,” which concerns the choice of variables and processes to include in the model, as well as how the variables and processes are characterized mathematically and how they interact, is likely a larger source of uncertainty.

A variety of methods have been proposed in recent years that, according to their proponents, can explore structural uncertainty, thereby giving a full accounting of forecast uncertainties. The most common method, which uses the results of multiple climate models from various research institutes, is based on the assumption that differences between the results of different climate models are a reasonable approximation of the structural uncertainty of a particular model as a representation of the real world. This amounts to assuming that the biases of different models average out over the “ensemble,” whereas in fact, the models are likely to share similar biases, since they involve similar approximations,

similar algorithms, and similar training data.

To understand the quantification problems posed by the structural uncertainties, let us investigate the analysis of uncertainty in a climate model given by the UK Meteorological Office in 2009’s *UK Climate Projections (UKCP09) Science Report: Climate Change Projections*. In this nearly 200-page document, the uncertainties from many sources are explained, and their effects propagated through to the model results, ending in “probabilistic” forecasts of climate change. The report states, “[although] a number of methods for probabilistic climate projection have been published in the research literature, we are not aware of any that have been designed to sample uncertainties as comprehensively as is done in UKCP09.” This suggests that UKCP09 represents the state-of-the-art in uncertainty analysis of climate models as of 2009.

UKCP09 seeks to sample both parametric and structural uncertainties. Sampling parametric uncertainties is done by constructing a “perturbed physics ensemble,” which is simply a collection of model runs of variants of the Met Office’s climate model, where each run has different, but plausible, values assigned to variables. For example, the model may be run several times with different values for temperature and water vapor in a given area to determine future cloud cover. The distribution of input values is an invention constructed by eliciting opinions from experts. Running the model many times, using many different variables results in some climate simulations that are more realistic than others, so each model variant is additionally weighted according to its performance in predicting recent climate observations and in hindcasts over the past 90 years. In the case of cloud cover, if the model output bears little resemblance to observed cloud cover in a particular area, that projection is given less weight than one that more accurately describes observed cloud cover. The perturbed physics ensemble thus produces a set of projections that have, according to the report, statistically knowable parametric uncertainties that can be described in terms of probability. However, the reporting obscures the fact that the probability distribution is an *input*, not an *output*.

UKCP09 also attempts to evaluate structural uncertainties in its model. It does so by comparing its ensemble of predictions to those of alternative climate models from other research groups. This approach is based on the assumption that “the effects of structural differences between models can be assumed to provide reasonable a priori estimates of possible structural differences between [the

Overall, there is a worrying tendency in climate science to claim that model uncertainties are quantifiable and manageable, but there is little or no statistical basis in the methodology of estimating uncertainty.

model] and the real world.” That is, it assumes that predictions from members of the ensemble somehow bracket the truth.

This assumption has no basis. UKCP09’s authors compare the projections of their ensemble with those of the other climate models and then estimate the structural uncertainty as a probability distribution of those projections, effectively treating the ensemble of models as if they were a random sample of all possible models.

For instance, for the UKCP09 to project future rainfall in a particular area, the authors use the Met Office climate models to run the perturbed physics ensemble with many different variables that might affect future rainfall to generate a range of possible outcomes. These are then judged by how well these outcomes conform to observed rainfall to produce an initial probability distribution—say, a 90 percent probability that summer rainfall in 2050 will decrease by less than 40 percent. This distribution is compared to the rainfall projections from alternative climate models, which have their own uncertainties—but which are assumed to be uncertain in their own distinct way—so that the predictions of the ensemble are, overall, unbiased. This, according to the authors, allows for the quantification of both parametric and structural uncertainty in the model projections of future rainfall. But there is no reason to think that the variability across the ensemble is a probability distribution at all, much less a distribution centered at the “real” answer.

This approach may give *lower* bounds on the uncertainty, but cannot give defensible *upper* bounds, for several reasons. First, such “ensembles” are not in any sense representative of the range of possible (and plausible) models that fit the data. Second, the structural uncertainties of the models considered are related: the models generally rely on similar physical approximations, similar numerical methods and algorithms, similar parameterizations, and similar calibration data. Climate models reflect, to differing degrees and with varying approaches, climate science’s best understanding of the processes that govern the Earth’s climate. But some processes, such as aspects of the methane cycle and the feedback effects of clouds, are poorly understood and not accounted for reliably in current climate models. So models share common errors whose magnitudes are simply not known.

The errors in the ensemble of model runs with differing variables (using the Met Office’s climate models) are not statistically independent—or even random. This fact is acknowledged in the UKCP09

report, which states, “our estimates of discrepancy [structural uncertainty] can be viewed as a likely lower bound to the true level of uncertainty associated with structural model errors.” Furthermore, the structural uncertainty is estimated, in part, by comparing the model with other models in hindcasting historical climate change. Since these models rely on similar modeling assumptions and approximations and are calibrated to the same historical observations, one would expect their output to be similar—and to have similar errors and biases. The biases of different models are unlikely to offset each other. The models do not provide independent estimates of climate change.

The Met Office also performed sensitivity analysis to assess the relative impact of assumptions about various inputs on the uncertainty of model outputs. But UKCP09’s sensitivity analysis relies on the *assumed* randomness of the input parameters, among other things. As the report puts it:

The key point is that although the UKCP09 probabilistic projections provide estimates of uncertainties in future climate change, it is also inevitable that the probabilities are themselves uncertain. If the uncertainties in the probabilities are sufficiently small compared with the uncertainties quantified by the probabilities, then the UKCP09 results are likely to be sufficiently reliable to be used in support of assessments of impacts, vulnerability, or adaptation.

This neglects the fact that the “uncertainties in the probabilities” are also assumptions—not a matter of physics or measurement error.

Finally, the sensitivity analysis varies only a subset of the assumptions and only one at a time. That precludes interactions among the uncertain inputs, which may be highly relevant to climate projections. Elsewhere we have called these styles of sensitivity analysis “perfunctory.” In multidimensional spaces,

varying one factor at a time is tantamount to counting uncertainties instead of weighing them.

Overall, there is a worrying tendency in climate science to claim that model uncertainties are quantifiable and manageable, but there is little or no statistical basis in the methodology of estimating uncertainty, especially structural uncertainty. The evidence suggests that even the quantifiable uncertainties are significantly underestimated. Used in this way, models become, in the words of a classic 1993 paper by Silvio Funtowicz and Jerome Ravetz, “substitutes for disciplined thought and scientific rigor.”

Scientific insight is not policy evidence

Economist Frank Knight distinguished between risk and uncertainty in a celebrated 1921 work entitled *Risk, Uncertainty, and Profit*. According to Knight, risk corresponds to the possibility of applying probabilistic reasoning and the calculus of probabilities (such as when playing roulette), whereas uncertainty corresponds to the real-life situations where such a computation is impossible, simply because “we do not know what we do not know.” Applied to the economy, profit should accrue to one who accepts inescapable uncertainty, not to one who can compute risks. Intellectually, Knight’s position was close to that of John Maynard Keynes, in denying that our thinking can be described by a probability distribution over all possible future events. The hubris of modeling “epistemic uncertainty”—the unknown unknowns—as a probability distribution lies, in part, in assuming that we know a great deal about what we do not know.

Brian Wynne, a leading risk theoretician, expanded Knight’s framework by distinguishing between:

RISK - Know the odds.

UNCERTAINTY - Don’t know the odds: may know the main parameters. May reduce uncertainty but increase ignorance.

IGNORANCE - Don’t know what we don’t know. Ignorance increases with increased commitments based on given knowledge.

INDETERMINACY - Causal chains or networks open.

Wynne elaborates: “Science can define a risk, or uncertainties, only by artificially ‘freezing’ a surrounding context that may or may not be this way in real-life situations. The resultant knowledge is therefore conditional knowledge, depending on whether these pre-analytical assumptions might turn

out to be valid. But this question is indeterminate—for example, will the high quality of maintenance, inspection, operation, etc., of a risky technology be sustained in future, multiplied over replications, possibly many all over the world?”

Yet some experts seem convinced that they can model the structural uncertainty of climate models and refuse to make the Knightian distinction between risk and uncertainty, let alone ignorance and indeterminacy. In truth, no one knows how well climate models can forecast climate change: estimates of uncertainty are themselves almost entirely uncertain. One only has to examine the IPCC’s recent assessment report to see this: of the 114 models in the ensemble, 111 failed to predict the recent 15-year slowdown in the increase of global mean surface temperature, many by an order of magnitude. Although natural decadal variability may explain this to some extent, and the long-term trend might still be valid, it does not inspire confidence in the accuracy of model forecasts.

Research is improving our understanding of the complex processes that affect the climate. Does this, coupled with increasing computing power, mean we will eventually have models that provide accurate forecasts of climate change and the costs of its effects? No. Even proponents of model-based assessment of climate impacts admit that the more one understands climate, the more model predictions may reveal that current estimates of uncertainty are unrealistically small. The UKCP09 report states that although they hope that better modeling will eventually reduce uncertainties, “this cannot be guaranteed as the introduction of processes not yet included (for example, feedbacks from the methane cycle), or as yet unknown, could have the opposite effect.”

Our theoretical and observational knowledge give us strong reasons to be concerned about the consequences of climate change and to act to mitigate its future effects. Modeling the economic cost of climate change necessarily requires accurate forecasts of the magnitude and locations of its various effects (which aren’t available). But, as detailed in a recent article by Richard A. Rosen and Edeltraud Guenther, extra layers of uncertainty are added because of the need for assumptions about the future growth of the economy and population—not to mention the highly subjective “discounting” of future costs and benefits against present ones, the uncertainty in future technologies, and the uncertainty of future changes in production and consumption patterns.

A good demonstration of the volatility of these

cost estimates is that *even when using the same model*, different teams cannot agree on the urgency and severity of action. Much of the difference is due to the assumed form of the function that relates damage to surface temperature. Different research teams are apparently able to propose different functions that are equally plausible, since there is little or no evidence to constrain them, as inadvertently shown in a recent paper by Simon Dietz and Nicholas Stern that tries to show that models used by the economist William Nordaus to suggest less urgency to act on climate are consistent with an interpretation of more urgency.

Given their uncertainties, the value of climate models to the policy debate depends on the important difference between policy simulation (performed by scholars to gain insight in their discipline) and policy justification (where the same scholars or other parties produce evidence to support adopting a specific policy). The difference between scientific insight and evidence for policymaking is profound. Indeed, in policymaking, the question raised by some model-based “facts” is whether, in the face of such substantial uncertainty, they are scientific at all. Yet when these numbers enter the policy world, often through the media, they look and sound like incontrovertible facts. The *Financial Times* recently published an article by its chief economic commentator Martin Wolf, who considers as “plausible” economic models suggesting “that the aggregate loss of world output by 2030, under the low-carbon option, would be equivalent to a one-year hiatus in economic growth.” If models are judged by their plausibility—by definition a non-quantitative attribute—why should we derive from these models quantitative information? And why should this information be relevant for policy?

A parallel could be drawn here with long-term weather forecasts. A story is told by Nobel Laureate Kenneth Arrow, who was asked to provide weather forecasts one month in advance during World War II:

The statisticians among us subjected these forecasts to verification and they differed in no way from chance. The forecasters themselves were convinced and requested that the forecasts be discontinued. The reply read approximately like this: “The commanding general is well aware that the forecasts are no good. However, he needs them for planning purposes.”

Climate models, by augmenting our scientific intuition, may help us to understand climate, though we might question whether the cost in time and resources absorbed by these exercises might be deployed better elsewhere. But when we attribute to

them predictive capabilities and attempt to introduce them, through political processes, into our policy planning, the numbers pollute the debates with a spurious impression of rationality, prediction, and control. One danger for society is that we will be condemned to endless debates over uncertainties, with models deployed to support various competing positions about which policy pathways to follow. Another, just as serious, is that, with excessive confidence in our ability to model the future, we will commit to policies that reduce, rather than expand, available options and thus our ability to cope with the unknown unknowns of our future.

Andrea Saltelli (andrea.saltelli@gmail.com), William Becker, and Pawel Stano are at the Joint Research Centre's Unit of Econometrics and Applied Statistics of the European Commission. Philip B. Stark is professor and chair, Department of Statistics, University of California, Berkeley.

Recommended reading

- Frank Ackerman, Stephen J. DeCanio, Richard B. Howarth, and Kristen Sheeran, “Limitations of Integrated Assessment Models of Climate Change,” *Climatic Change* 95, no. 3-4 (2009).
- Silvio O. Funtowicz and Jerome R. Ravetz, “Science for the Post-Normal Age,” *Futures*, 25, no. 7 (1993).
- Michael D. Mastrandrea and Stephen H. Schneider, “Probabilistic Integrated Assessment of ‘Dangerous’ Climate Change,” *Science* 304, no. 5670 (2004).
- Richard L. Revesz, Peter H. Howard, Kenneth Arrow, Lawrence H. Goulder, Robert E. Kopp, Michael A. Livermore, Michael Oppenheimer, and Thomas Sterner, “Global Warming: Improve Economic Models of Climate Change,” *Nature* 508, no. 7495 (2014).
- Richard A. Rosen and Edeltraud Guenther, “The Economics of Mitigating Climate Change: What Can We Know?” *Technological Forecasting and Social Change*, 91 (2015).
- Andrea Saltelli and Beatrice d’Hombres, “Sensitivity Analysis Didn’t Help: A Practitioner’s Critique of the Stern Review” *Global Environmental Change* 20 (2010).
- Andrea Saltelli and Silvio O. Funtowicz, “When All Models Are Wrong: More Stringent Quality Criteria Are Needed for Models Used at the Science-Policy Interface,” *Issues in Science and Technology* 30, no. 2 (Winter 2014).