Whose Probabilities? Predicting Climate Change with Ensembles of Models

Wendy S. Parker[†‡]

Today's most sophisticated simulation studies of future climate employ not just one climate model but a number of models. I explain why this "ensemble" approach has been adopted—namely, as a means of taking account of uncertainty—and why a comprehensive investigation of uncertainty remains elusive. I then defend a middle ground between two camps in an ongoing debate over the transformation of ensemble results into probabilistic predictions of climate change, highlighting requirements that I refer to as ownership, justification, and robustness.

1. Introduction. There is now a broad scientific consensus that Earth's climate has warmed significantly over the last century and that much of this global warming can be attributed to greenhouse gas emissions associated with human activities (IPCC 2007). Less clear are the details of how climate will evolve in the future if greenhouse gas emissions increase or decrease in particular ways. This is currently a major focus of climate research, undertaken in part to provide input to policy decisions.

The hope is that changes in climate with dire human impacts can be largely avoided by implementing international agreements to reduce greenhouse gas emissions. But by how much do emissions need to be reduced, and by when do the reductions need to take place in order to keep global warming under some specified threshold? What will the climates of particular regions be like if emissions continue accelerating? Policy makers look to climate scientists to provide this information. Climate scientists, in turn, look to their computer simulation models.

Today, the most sophisticated climate simulation studies employ col-

†To contact the author, please write to: Department of Philosophy, Ohio University, Ellis Hall 202, Athens, OH 45701; e-mail: parkerw@ohio.edu.

‡Thanks to Lenny Smith and Reto Knutti for helpful discussion. This material is based on work supported by the National Science Foundation under grant 0824287.

Philosophy of Science, 77 (December 2010) pp. 985–997. 0031-8248/2010/7705-0021$10.00
Copyright 2010 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

lections—or *ensembles*—of climate models, each of which is used to produce one or more simulations of future climate under a chosen emission scenario (e.g., Stainforth et al. 2005; Tebaldi et al. 2005).[1] In what follows, I explain why this ensemble approach has been adopted—namely, as a means of taking account of uncertainty—and why a comprehensive investigation of uncertainty remains elusive. I then defend a middle ground between two camps in an ongoing debate over the transformation of ensemble results into probabilistic predictions of climate change, highlighting requirements that I refer to as *ownership, justification*, and *robustness*.

2. Climate Models, Uncertainty, and Ensembles. A computer simulation model is a computer-implemented set of instructions for repeatedly solving a set of equations in order to produce a representation of the temporal evolution (if any) of selected properties of a target system. In the case of global climate modeling, the target system is Earth's climate system—encompassing the atmosphere, oceans, sea ice, and land surface—and the equations of interest are ones that represent the transport of mass, energy, moisture, and other quantities by processes within the system. These processes include the formation of clouds and precipitation, the movement of large-scale weather systems, ocean currents, the melting of sea ice, the absorption and emission of radiation by atmospheric gases, and numerous others.

There is considerable uncertainty, however, about how these global climate models should be constructed. In part the reason is that some climate system processes are not well understood from a theoretical point of view. At least as significant, however, are constraints due to limited computing power. In order to complete simulations of twenty-first-century climate (and beyond) in a reasonable amount of time, global climate models need to have rather coarse spatial resolution: solutions to modeling equations are typically estimated for grid points whose spacing corresponds to ~100–200 kilometers in the horizontal and ~100 meters in the vertical. Yet many processes that significantly influence climate, such as the formation of clouds and precipitation and the turbulent transfer of momentum, take place on much smaller scales. These subgrid processes must be parameterized, that is, represented in terms of the values of

1. An emission scenario specifies how greenhouse gas emissions and other climate forcing factors evolve over a time period.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

variables whose evolution the model does simulate explicitly.[2] In most cases, it is not clear how this parameterization can best be accomplished.

The upshot is that a number of climate models, which differ from one another in various ways—in some of their equations, in their spatiotemporal resolution, in their mathematical techniques—can nevertheless have approximately equal prima facie plausibility as tools for predicting future climate change on global and continental scales (see also Parker 2006). At present, there are around two dozen state-of-the-art climate models in development at modeling centers around the world that are viewed in this way (Meehl et al. 2007). These models are considered the best-available climate models for these predictive tasks.

Recently, these best-available models were used together in a study of future climate, known as the Coupled Model Intercomparison Project, Phase 3 (CMIP3): the models were run under three different emission scenarios, representing futures with high, medium, and low levels of greenhouse gas emissions, respectively, producing a collection of simulations for each emission scenario.[3] CMIP3 is an example of a *multimodel ensemble* (MME) study—a study in which multiple models that differ in the form of some of their equations and in various other ways are each used to simulate the behavior of a target system under conditions of interest. MME studies like CMIP3 investigate the extent to which climate models with different structures differ in their predictions.[4]

A second type of ensemble study, known as a *perturbed-physics ensemble* (PPE) study, investigates the impacts of uncertainty regarding parameter values. Here, a single climate model is run multiple times under a given emission scenario, changing only the values assigned to one or more of its uncertain parameters (e.g., Stainforth et al. 2005). For instance, the model might include a parameter that represents the fall speed of ice crystals in clouds, assigning a single value to that speed as a simplification. But it might be unclear which value would be best, given the predictive

2. For instance, a very simple cloud parameterization might represent cloud formation as a function of relative humidity: at any time step in the simulation, if the relative humidity at a grid point exceeds a threshold value, then a volume of atmosphere associated with that grid point is assumed to include clouds of a given type.

3. Most models were used to produce simulations for each of the three emission scenarios; some models were used to produce simulations for only some of the scenarios (see Meehl et al. 2007, table 10.4).

4. By the *structure* of a climate model, I mean the computer simulation model itself—a physically implemented algorithm for transforming inputs into outputs—except without values assigned to parameters in the model's equations. In what follows, I will also speak of model structures in the abstract, by which I will mean possible implementations of simulation algorithms on computers with specified properties (e.g., precision). I do not address any metaphysical questions concerning these model structures.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

task at hand, so several different values might be tried on different runs. Once again, the immediate product of the study is a collection of simulations, each representing the evolution of the climate system under the same emission scenario. Some of these simulations may subsequently be excluded from the collection, however, if it is found that the model versions producing them fail to simulate past climate well enough on some chosen metric of performance.

Both MME and PPE studies sometimes incorporate another kind of ensemble, known as an *initial condition ensemble* (ICE). A chosen model or model version is run multiple times, each time from somewhat different initial conditions, as a way of exploring the implications of uncertainty about how to represent the initial state of the climate system for the simulation of interest. Often, the simulations are then averaged, with the result taken to represent the response of that model or model version to the emission scenario under investigation.

These ensemble studies explore how different sources of uncertainty affect predictions of future climate: MME studies investigate the impacts of uncertainty regarding model structure, also known as structural uncertainty; PPE studies investigate the impacts of uncertainty regarding the value of parameters within a given model structure; ICEs are used within both MME and PPE studies to investigate the impacts of uncertainty regarding initial conditions.

It is easy to imagine a grand ensemble study incorporating all three types of ensemble: for each of the two dozen or so state-of-the-art climate models currently available, a PPE study would be undertaken, and each model version in that PPE would itself be run with several different sets of initial conditions. In practice, however, this study has not been carried out, in part because of limited computing power; there are enough uncertain parameters in a state-of-the-art climate model, and the range of values that those uncertain parameters might plausibly take is large enough, that a thorough PPE study for even a single model would require many simulations and thus a tremendous amount of computing power. To date, the most ambitious PPE/ICE studies have been undertaken as part of the climateprediction.net project (Stainforth et al. 2005), which relies on donated idle processing time on thousands of home computers around the world. But even in these studies, which focus on climate models developed at the UK Hadley Centre, not all uncertain parameters in the chosen model are perturbed, and only a small set of alternative values is typically considered for any given parameter.[5]

5. More thorough PPE studies are sometimes carried out on simpler models and/or with the help of statistical emulators, but it is unclear how these results relate to the results that would be produced in a full PPE study on a state-of-the-art model.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

Even if the grand ensemble study described above could be carried out, however, it would not provide a comprehensive analysis of current uncertainty regarding how climate would change (in various respects that interest us) under a given emission scenario. The reason is that state-of-the-art climate models now in existence are not designed to span the full range of current uncertainty about model structure, that is, about the form that modeling equations should take, the numerical techniques that should be used to estimate solutions to them, the spatiotemporal resolution that models should have, and so forth. Indeed, as we will see, it turns out to be challenging to even conceptualize what a systematic and comprehensive analysis of this structural uncertainty would involve.

3. Ensemble Studies and Systematic Uncertainty Analysis. Ensemble studies of future climate might be classified as *sensitivity studies* in that they investigate the extent to which predictive results of interest are sensitive to changes in the models used to produce them. When the changes made reflect uncertainty about how to represent the target system, as they very often do, then this sort of investigation may also be referred to as *uncertainty analysis*. It is an epistemically laudable activity since instead of ignoring uncertainty about how to represent a target system, it acknowledges this uncertainty and seeks to understand its implications. But there is a significant difference between the investigation of parametric and initial condition uncertainty via PPEs and ICEs on the one hand and the investigation of structural uncertainty via MMEs on the other hand: PPEs and ICEs investigate uncertainty about numerical values, while MMEs investigate uncertainty about model structure. For the former, it is relatively straightforward to conceptualize, and then sample systematically within, the space of plausible alternatives; for the latter it is not.

For instance, recall the uncertain parameter mentioned in the last section—a parameter meant to account in a simplified way for the fact that ice crystals slowly descend in clouds. The default setting for the parameter (i.e., the value assigned in the standard version of a state-of-the-art model) might be 1 meter per second, but scientists might readily acknowledge that it is not clear that 1 meter per second is the best value; it is plausible that some other value is best. Perhaps scientists judge that the best value might plausibly be found in the range 0.5–2 meters per second; they might even specify, in the form of a subjective probability distribution, where in that range they think the best value is more and less likely to be found. Having specified this range and/or distribution, there is little puzzle about what it would mean to sample values systematically from it, running a simulation with each value chosen.

Before I contrast this with the case of structural uncertainty, it is worth making clear why the "best" parameter values, rather than the "true"

https://doi.org/10.1086/656815 Published online by Cambridge University Press

values, are of interest. First, depending on how a parameter is interpreted, there may be no true value. There is no single speed at which all ice crystals fall in clouds. But if that parameter is instead understood to represent the average speed at which ice crystals in clouds would fall during the simulated period of time under the emission scenario of interest, then it might well be meaningful to speak of its "true" value, even if we do not know (and perhaps cannot know) what that value is. The trouble is that if the model is imperfect in various other ways, then the true parameter value may not be the value with which the model performs best, relative to the predictive goals of the study. Because of simplifications, approximations and errors elsewhere in the model, not to mention omissions, there may be reason to think that the most accurate prediction of a quantity of interest will be achieved if the parameter is assigned a value other than its true value, perhaps even a value significantly different from its true value. So when specifying plausible ranges for uncertain parameters in one of today's state-of-the-art climate models—models thought to have imperfect structures—one should seek not the ranges in which the true values might plausibly fall but rather the ranges in which the best values (i.e., the values that produce the best performance) might plausibly fall, given everything else that is known about the model.[6]

So the systematic exploration of parameter uncertainty turns out to be more complicated than one might think. But at least in the case of parameter uncertainty (and initial condition uncertainty), the space of possibilities in which we are to identify plausible alternatives is clear: it is a space of numerical values. When the uncertainty under investigation is structural uncertainty, then the space of possibilities instead ranges over model structures—simulation algorithms to be implemented on machines with specified precision and so forth—and the task of systematic uncertainty analysis begins to look overwhelming indeed.

As we consider what such a task would involve, we first need to clarify what the uncertainty at issue is uncertainty about. Is it uncertainty about

6. As understood here, a computer simulation model has an imperfect structure if there is no possible assignment of values to its parameters that would make it a perfect predictive model for the simulated period of time that interests us, i.e., that would make it a model for which there is some assignment of initial conditions (roughly consistent with our knowledge of the target system) such that, for the simulated period that interests us, the model delivers predictions/results that are always consistent with relevant observations. See Smith (2002) for a related perspective on perfect/imperfect model structures and the challenges that arise when a model's structure is imperfect. Note also that if a model has an imperfect structure, then unless special steps are taken (e.g., Rougier 2007), a thorough PPE study will reveal a range of predictions in which the best prediction of that model structure can be expected to fall, but this says nothing about how good that best prediction will be.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

what would constitute a perfect model structure? Presumably not, since current attempts to explore structural uncertainty include today's state-of-the-art models in the analysis, and they are believed to have imperfect structures. So perhaps it is uncertainty about what would constitute an adequate model structure.[7] After all, adequacy (not perfection) is all that is really needed, and it is plausible that the structures of today's models are adequate for some predictive purposes of interest. But care should be taken here: we do not want to set ourselves the task of identifying all plausibly adequate model structures, for this would seem to require that we survey all possible algorithms that might be implemented on today's computers (and as part of those algorithms, all combinations of all mathematical functions)—a truly mind-boggling task. Only one adequate model structure is needed. So the real task is to identify a collection of plausibly adequate structures among which we can expect to find at least one that is actually adequate.

Whether this can be done for a given set of predictive purposes, and whether the collection of structures that is identified is tiny or vast, depends on the extent of our background knowledge.[8] If we cannot specify such a collection of structures, then it seems we do not have a clear idea of the structural uncertainty that we aim to systematically investigate. And even if we can specify such a collection, then unless it includes only a small, finite number of model structures, so that we can simply try them all (i.e., carry out a PPE/ICE study for each), we will face the further challenge of determining what it means to sample systematically from such a collection of structures; this is not at all obvious.

Climate scientists have not yet attempted a systematic analysis of structural uncertainty. For many predictive tasks of interest to decision makers (e.g., prediction of regional changes in climate), current understanding of the climate system is not sufficient to pick out a small, finite collection of model structures that can be expected to include an adequate one. Again, part of the difficulty stems from computational constraints: we need climate models that can be implemented and run on today's computers, which requires parameterization of (or some other accounting for) a number of important climate system processes. Matters are further com-

7. As understood here, an adequate model structure is one that can be implemented and run on existing computers and that also has the following property: there is at least one assignment of initial conditions (roughly consistent with our knowledge of the target system) and parameter values such that the model delivers predictions/results that are accurate enough for our purposes, and not just as a matter of accident; the model captures enough of the relevant physics in an approximate way that it is "good enough" for our purposes.

8. In this discussion, it is assumed that we can individuate model structures; clearly this may be a nontrivial task.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

plicated by the fact that we typically want to predict what climate will be like if greenhouse gas concentrations are significantly different from those previously observed (e.g., much higher); we must consider whether processes and feedbacks represented poorly—or not represented at all—in today's models will become more important under these new conditions, and this can be hard to determine from our current vantage point.

In practice, investigations of the impacts of structural uncertainty on predictions of future climate have been somewhat limited. The most ambitious investigation using state-of-the-art global models is the CMIP3 study mentioned in section 2, an MME study in which two dozen or so models were used to produce simulations of twenty-first-century climate under a few emission scenarios of interest (Meehl et al. 2007). Such MMEs are acknowledged to be "ensembles of opportunity," assembled from models available at the time and only insofar as research groups are willing to participate; the "sampling" by which they are produced "is neither systematic nor random" (Tebaldi and Knutti 2007, 2068), and the set of models involved "is not designed to span an uncertainty range" (Knutti et al. 2008, 2653).[9]

In light of this, the question naturally arises: How should the results of such MME studies—and ensemble studies more generally—be interpreted? What can we infer from these studies about future climate, given that they take into account only some recognized sources of uncertainty? This turns out to be a controversial matter.

4. From Ensembles to Probabilities? Much of the controversy concerns the appropriateness of transforming collections of simulation results from individual PPE and MME studies into probability density functions (PDFs) over climate change outcomes. Some scientists are producing these PDFs from ensemble results, using Bayesian and other methods (Murphy et al. 2004; Tebaldi et al. 2005; Furrer et al. 2007), while others argue against doing so (Stainforth et al. 2007). In this section, I aim to carve out something of a middle ground: it may be appropriate for today's scientists to offer PDFs for some predictive variables, but methods currently used to transform results from individual PPE and MME studies into PDFs are problematic.

To avoid confusion, it is important to emphasize that the PDFs under dispute are representations of epistemic uncertainty about what would occur under various emission scenarios; they should be understood as depictions of someone's (or some group's) uncertainty. They do not summarize the frequency with which different changes in climate would occur

9. Murphy et al. (2007, 1995) remark that "it is not clear how to define a space of possible model configurations of which the MME members are a sample."

https://doi.org/10.1086/656815 Published online by Cambridge University Press

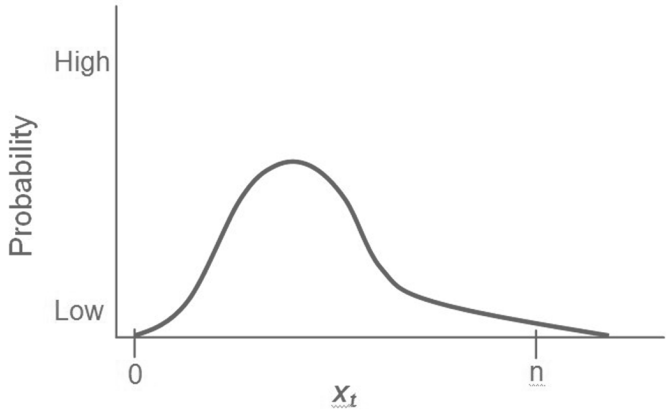


Figure 1. Schematic depiction of a PDF representing someone's (or some group's) uncertainty about the value that predictive variable x will take at time t.

if it were possible to repeat an experiment on the climate system many times. Figure 1 illustrates what such a PDF might look like.

The main motivation for producing these PDFs is the perception that they are needed, or at least desired, in policy- and decision-making contexts. Standard decision frameworks call for probabilities to be assigned to different possible outcomes of an action (e.g., in expected utility calculations). Even stepping away from such frameworks, if scientists cannot say precisely what would occur under a given emission scenario, decision makers would at least like to know the range of outcomes in which the actual outcome is expected to fall and which outcomes within that range are most plausible. So there is a perceived demand for something like a PDF.

But under what circumstances is it appropriate for scientists to offer decision makers a PDF for a predictive variable? Here I offer three requirements that seem sensible, given the far-reaching decisions that may be influenced by what is offered. (1) *Ownership*: The PDF should be one that the scientists are willing to claim as their own. That is, the scientists should take the PDF to accurately depict their uncertainty about what would happen under the given emission scenario. This means that the probabilities implied by the PDF for different outcomes are (at least very approximately) the subjective probabilities that the scientists assign to those outcomes. (2) *Justification*: The scientists should be able to justify their choice of PDF—why the PDF has the shape and width that it does rather than some other shape and width. I would argue that a PDF is not adequately justified unless, in arriving at that PDF, scientists strive

https://doi.org/10.1086/656815 Published online by Cambridge University Press

to take into account all of the available evidence regarding the predictive variable—this might include modeling results, background theory, and observational data—as well as what is known (or not known) about the limitations of those sources of evidence, their interdependencies, and so forth. (3) *Robustness*: The PDF should be robust in each of two senses; namely, it should not be highly sensitive to contentious assumptions, and it should not be expected to change substantially in the very near term as models and observing networks undergo incremental development. It would not be good, either for policy or for the credibility of climate science, to offer decision makers a PDF now with the expectation that a very different PDF—implying very different probabilities for outcomes of interest—will be produced a couple of years later.

Similar requirements are implied in a recent discussion by Kandlikar, Risbey, and Dessai (2005), who provide a method for choosing among several ways of communicating uncertainty in climate science. They argue that in the best case one can convey uncertainty via full PDFs, but in other cases it may be appropriate only to offer a range in which one expects the value of a variable to fall with some probability, or to identify the expected sign of a change without assigning a magnitude, and so forth (see also Risbey and Kandlikar 2007). When it comes to offering full PDFs, Kandlikar et al. advise, "Justify your choice of 5th and 95th percentiles. Are there any processes or assumptions that would cause them to be much wider than you have stated? If so, describe and revise. If you cannot provide justifications for why you consider the 5th and 95th percentiles to be fairly robust, then move to a lower precision category. Justify your choice of shape for the distribution. Why did you reject alternative shapes? If you are not confident of the shape of the distribution, consider specifying bounds only" (2005, 449–50).

Do PDFs produced so far meet the three requirements specified above—ownership, justification, and robustness? In most cases, they do not. First, PDFs are typically presented as preliminary "estimates" of uncertainty (Meehl et al. 2007, 809), not as depictions that any particular scientist is claiming as her own. Second, PDFs are usually based on simulations produced in a single MME or PPE study, without taking into account results from other studies, limitations of today's models, or other relevant considerations. It is implausible that the complicated process of assessing and combining the available evidence that is called for is captured in the algorithms/methods that these studies use to produce PDFs.[10] Finally, in

10. So far, the most comprehensive methodology for producing PDFs seems to be the Bayesian approach outlined by Rougier (2007), which factors in an estimate of the errors that will occur even in the best simulations in an ensemble. However, it is not clear that this estimate of residual error (known as the "discrepancy") can itself be well justified and robust.

https://doi.org/10.1086/656815 Published online by Cambridge University Press

many cases PDFs are not robust; PDFs produced in different ensemble studies but for the same predictive variable can differ markedly from one another (see, e.g., Meehl et al. 2007, fig. 10.28; Tebaldi and Knutti 2007), and contentious assumptions are often involved in their production. For instance, some methods for producing PDFs rest on the contentious assumption that current state-of-the-art models can be treated as a group of models sampled independently from a population of models distributed around a perfect model (e.g., Tebaldi et al. 2005; Furrer et al. 2007; see also Tebaldi and Knutti 2007).[11]

But even if most PDFs produced in recent ensemble studies fail to meet at least one of the three requirements (ownership, justification, and robustness), are there some predictive variables for which today's scientists are in a position to offer PDFs that meet them? While it is beyond the scope of this discussion to address this question in detail, perhaps there are a few promising candidates, including near-term changes in highly aggregate quantities, such as global decadal mean surface temperature.[12] By contrast, for long-term changes in regional climate, it seems that PDFs that meet the three requirements are generally out of reach at present; these changes are often highly sensitive to factors that are themselves not well simulated by today's models, and there are few other avenues of investigation.

Nevertheless, for both global and regional scales, there seems to be a clear trend toward presenting the results of MME and PPE studies in terms of PDFs. Would it be better not to offer these "estimates" of uncertainty? After all, even if multiple estimates for the same predictive variable, when they are available, are displayed together for decision makers, this creates the second-order problem of how to decide which estimate is best, whether any are trustworthy, and so forth. A criticism leveled against those scientists who deny that PDFs should be produced in ensemble studies is that they leave to decision makers "the challenging task of deciding how to infer things about the real world from imperfect models" (Frame et al. 2007, 15). But showing multiple, contentious PDFs from ensemble studies seems to leave decision makers in a position that

11. The concerns of scientists who refuse to produce PDFs from ensemble results might well be understood as concerns about justification and/or robustness. For instance, Stainforth et al. (2007, 2154) insist that "given nonlinear models with large systematic errors under current conditions, no connection has been even remotely established for relating the distribution of model states under altered conditions to decision-relevant probability distributions."

12. The theoretical quantity known as *climate sensitivity* is also among the more promising candidates (see Roe and Baker [2007] for an interesting related analysis).

https://doi.org/10.1086/656815 Published online by Cambridge University Press

is not much better and may even provide a false sense of security (e.g., that it is almost certain that such and such will not occur).

In the end, perhaps the best option would be to step back and reconsider what sort of depiction of uncertainty is most appropriate for each predictive variable of interest. Though PDFs may be desirable, further reflection may reveal that the requirements of ownership, justification, and robustness can rarely be met when it comes to the predictions of greatest interest to decision makers. If so, then one of the "less precise" means of depicting uncertainty discussed by Kandlikar et al. (2005)—offering a range into which the outcome is expected to fall with (approximately) given probability, or an estimate of the order of magnitude of a change, or some other option—will typically be more appropriate. For these "less precise" depictions too, however, analogous requirements of ownership, justification, and robustness should be met. That is, whatever depiction of uncertainty is offered, it should be justified and reasonably robust, and it should be one that scientists are willing to claim as their own.

5. Conclusions. Ensemble studies are an important means of investigating the impacts of uncertainty about how to represent the climate system for particular predictive purposes. However, because today's ensemble studies investigate only some recognized sources of uncertainty, it can be difficult to interpret their results. Particularly contentious is whether ensemble results should be transformed into PDFs and offered to decision makers as such. The position taken here is that a PDF—or any other representation of uncertainty—should be offered to decision makers only if the requirements of ownership, justification, and robustness are met. For many of the predictive variables of greatest interest to decision makers, the most appropriate means of representing current uncertainty may well be something less than a full probability density function.

REFERENCES

Frame, David J., et al. 2007. "Probabilistic Climate Forecasts and Inductive Problems." *Philosophical Transactions of the Royal Society* A 365:1971–92.

Furrer, Reinhard, et al. 2007. "Multivariate Bayesian Analysis of Atmosphere-Ocean General Circulation Models." *Environmental and Ecological Statistics* 14:249–66.

IPCC (Intergovernmental Panel on Climate Change). 2007. "Summary for Policymakers." In *Climate Change 2007: The Physical Science Basis; Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. Susan Solomon et al., 1–18. New York: Cambridge University Press.

Kandlikar, Milind, James Risbey, and Suraje Dessai. 2005. "Representing and Communicating Deep Uncertainty in Climate Change Assessments." *Comptes Rendus Geoscience* 337:443–55.

Knutti, Reto, et al. 2008. "A Review of Uncertainties in Global Temperature Projections over the Twenty-first Century." *Journal of Climate* 21:2651–63.

Meehl, Gerald A., et al. 2007. "Global Climate Projections." In *Climate Change 2007: The Physical Science Basis; Contribution of Working Group I to the Fourth Assessment Report*

https://doi.org/10.1086/656815 Published online by Cambridge University Press

of the Intergovernmental Panel on Climate Change, ed. Susan Solomon et al., 747–845. New York: Cambridge University Press.

Murphy, James M., et al. 2004. "Quantification of Modelling Uncertainties in a Large Ensemble of Climate Change Simulations." *Nature* 430:768–72.

———. 2007. "A Methodology for Probabilistic Predictions of Regional Climate Change from Perturbed Physics Ensembles." *Philosophical Transactions of the Royal Society* A 365:1993–2028.

Parker, Wendy S. 2006. "Understanding Pluralism in Climate Modeling." *Foundations of Science* 11:349–68.

Risbey, James, and Milind Kandlikar. 2007. "Expressions of Likelihood and Confidence in the IPCC Uncertainty Assessment Process." *Climatic Change* 85:19–31.

Roe, Gerard H., and Marcia B. Baker. 2007. "Why Is Climate Sensitivity So Unpredictable?" *Science* 318:629–32.

Rougier, Jonathan C. 2007. "Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations." *Climatic Change* 81:247–64.

Smith, Leonard A. 2002. "What Might We Learn from Climate Forecasts?" *Proceedings of the National Academy of Sciences* 99:2487–92.

Stainforth, David A., et al. 2005. "Uncertainty in Predictions of the Climate Response to Rising Levels of Greenhouse Gases." *Nature* 433:403–6.

———. 2007. "Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions." *Philosophical Transactions of the Royal Society* A 365:2145–61.

Tebaldi, Claudia, et al. 2005. "Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach." *Journal of Climate* 18:1524–40.

Tebaldi, Claudia, and Reto Knutti. 2007. "The Use of the Multi-model Ensemble in Probabilistic Climate Projections." *Philosophical Transactions of the Royal Society* A 365: 2053–75.

https://doi.org/10.1086/656815 Published online by Cambridge University Press