Foundations of Science (2006) 11:349–368
DOI 10.1007/s10699-005-3196-x

W. S. PARKER

UNDERSTANDING PLURALISM IN CLIMATE MODELING

ABSTRACT. To study Earth's climate, scientists now use a variety of computer simulation models. These models disagree in some of their assumptions about the climate system, yet they are used together as complementary resources for investigating future climatic change. This paper examines and defends this use of incompatible models. I argue that climate model pluralism results both from uncertainty concerning how to best represent the climate system and from difficulties faced in evaluating the relative merits of complex models. I describe how incompatible climate models are used together in 'multi-model ensembles' and explain why this practice is reasonable, given scientists' inability to identify a 'best' model for predicting future climate. Finally, I characterize climate model pluralism as involving both an ontic competitive pluralism and a pragmatic integrative pluralism.

KEY WORDS: climate change, computer simulation, models, pluralism, uncertainty

1. INTRODUCTION

Scientists now use a variety of computer simulation models to study Earth's climate. It is noteworthy that many of these are designed to be models of the climate system as a whole, rather than complementary models of different components of that system. This raises the question: why so many models? After all, there is but one terrestrial climate system.

One reason is that models of different complexity are useful for different modeling tasks. For example, simple climate models, some of which represent only one spatial dimension of the climate system and only a few physical processes, are often used when computational expense is a constraining factor and when global average parameters are to be predicted. Complex models, on the other hand, are relied upon in studies of regional climate change, which are thought to require representation of all three spatial dimensions of the climate system in some detail. While the ability of complex models to simulate regional climate change is still being investigated, it is agreed that very simple models probably will not be of much use for such studies. Thus one dimension of model pluralism in climate science crosses levels of model complexity and exists primarily because, at a given time, different modeling tasks may be best undertaken using different types of climate models.

A second dimension of climate model pluralism seems more puzzling. There are not just simple and complex models – there are many simple models and many complex models. A recent chapter on climate model evaluation, for example, identified more than 30 complex climate models (see Houghton et al., 2001, Chapter 8). Why does this pluralism within levels of model complexity exist? In this discussion, the focus will be on complex model pluralism in particular. In part, complex model pluralism reflects the fact that there are different mathematical techniques available to climate modelers. For instance, the atmosphere can be represented as a grid of points corresponding to volumes of atmosphere or in terms of a series of waves of differing frequencies, known as a spectral representation (see, e.g., Holton, 1992, p. 450). Either technique can be used, so one finds some complex models with grid-point representations of the atmosphere and some with spectral representations of the atmosphere. These models often incorporate many of the same assumptions about large-scale atmospheric dynamics, even if they handle the mathematical treatment of physical equations in different ways. But this explains only part of the diversity, since complex models also typically differ in some of their assumptions about climate system processes and hence in some of the predictions and retrodictions that they make. That is, complex climate models generally are physically incompatible with one another – they represent the physical processes acting in the climate system in mutually incompatible ways and produce different simulations of climate.

What explains the persistence of this plurality of incompatible models? The explanation given for pluralism across levels of model complexity does not really apply here; it is not that the different complex models are each developed and used for quite different purposes. On the contrary, although a variety of studies are carried out with complex models, most complex models are at least sometimes used to investigate how Earth's climate might change in the future in response to increasing greenhouse gas emissions. Section 2 will discuss several reasons why climate scientists have been *unable* to further narrow the field of complex models used to investigate future climate: (1) there is genuine scientific uncertainty about how to best represent the climate system; (2) there are difficulties in testing model predictions and retrodictions; (3) it is difficult to define an overall 'figure of merit' for climate simulations; (4) no model is clearly superior to the rest with respect to measures of simulation quality currently in use.[1] Examining (1)–(4) will give some sense of why neither scientists nor philosophers find model evaluation to be a straightforward matter.

A closer look at (1)–(4) will also be preparation for examining another surprising feature of complex model pluralism: these incompatible models are used together as *complementary* resources for investigating future climate change. By contrast, when two incompatible theories are available, they typically are viewed as competitors, and scientists seek evidence that disconfirms one theory and supports the other. Why is it different in the case of complex climate models? As Section 3 will argue, it is not because climate scientists consider the models to be purely instrumental tools (in which case their incompatibility might not matter, so long as they helped scientists to accomplish their goals). Rather, scientists' attitudes toward complex climate models typically involve both instrumentalist and realist components. Section 4 will show that, given this mixed status of complex models and the difficulties faced in evaluating them, it makes sense for scientists to use incompatible climate models together to investigate climate change. The same section will explain why the 'multi-model ensemble' approach currently used is advantageous and why scientists must nevertheless be careful in interpreting results obtained via this approach. Finally, in Section 5, it will be suggested that this interesting use of incompatible models involves two kinds of model pluralism: ontic competitive pluralism and pragmatic integrative pluralism.

2. EXPLAINING THE PERSISTENCE OF A PLURALITY OF INCOMPATIBLE MODELS

This section examines (1)–(4) in some detail to see how they together promote the continued existence of a plurality of incompatible complex climate models.

2.1. *Scientific Uncertainty with Respect to Climate System Representation*

Some of the differences among complex climate models reflect scientists' uncertainty about the nature of processes acting in the climate system and about how such processes should be represented in climate models. Some physical processes are still poorly understood, and some occur on spatiotemporal scales smaller than those resolved by the models. In either case, while it may be believed that the processes influence climate to a degree that merits their inclusion in the models, it may be unclear how this can best be done, leading to the development of incompatible representations of those processes in different climate models.

The case of clouds provides a good illustration of this. Clouds play an important role in shaping climate. Initial warming due to increased greenhouse gas emissions might lead to changes in the amount and types of clouds that form, thereby enhancing or offsetting the initial warming. Although scientists would like their climate models to incorporate some representation of the effects of clouds, individual clouds occur on scales that the models do not explicitly resolve. In addition, there is genuine uncertainty about how clouds interact with larger-scale dynamical processes in the climate system and hence about how the effects of clouds can be best parameterized.[2] As a result, several different parameterizations of clouds have been developed, reflecting different approaches to representing clouds within the bounds of present scientific uncertainties.[3]

These parameterizations disagree in some of the assumptions that they make and generally give somewhat different predictions about the effects of clouds on climate. In fact, clouds are an especially problematic case: as of the last major review of climate change research, models incorporating different cloud parameterizations did not even agree on whether changes in cloud formation due to increased greenhouse gas concentrations would have a net warming effect or a net cooling effect (Houghton et al., 2001, pp. 427–431).

Despite remaining uncertainties, the collection of incompatible climate models might be reduced in size if scientists could test the models in a straightforward way; as the rest of this section will illustrate, however, climate model evaluation is a messy and complicated undertaking.

2.2. *Difficulties in Testing Model Predictions and Retrodictions*

Especially given recent concern over global warming, climate models are of interest as resources for making predictions of future climatic conditions. In order to identify models that are most promising as predictive tools, scientists consider (among other things) the models' histories of predictive successes and failures. Unfortunately, for today's climate models, there are virtually no such predictive track records. Today's models make predictions about what might happen 10 or 50 or 200 years from now under conditions that may or may not actually obtain during the intermediate years.

This is why such predictions are typically referred to as 'projections' instead; they are projections of what would happen if greenhouse gases were to be emitted at particular rates over the course of decades or centuries. Weather forecasting models, by contrast, make predictions about what will actually happen over time periods of hours, days or weeks. Scientists can and do compile much information about the predictive strengths and weaknesses of these models. But for climate models, there is almost no such information, since the observational data that is needed in order to assess the quality of their predictions will not be available, even in principle, for quite some time.

Given this situation, simulations of past and present climate conditions (i.e., 'retrodictions') have become a focus of climate model evaluation. Scientists attempt to compare such simulations with available observational datasets. One serious problem, however, is that data are available for only a few quantities (e.g., temperature, pressure, precipitation), for only relatively recent time periods, and primarily for land locations and near-surface locations, and even these records are incomplete and of variable quality. Scientists lack a solid observational foundation against which to compare even the retrodictions of climate models.

2.3. *Difficulty in Defining an Overall Figure of Merit*

Another difficulty in evaluating climate model retrodictions stems from the vast amount of model output produced. Climate models generate output for thousands and thousands of grid points for years and years of simulated time and for numerous variables. How should the overall quality of this output be judged? There are measures for quantitatively assessing model-data fit for individual fields (e.g., monthly maximum temperature), but there are many such measures. Even if climate scientists were to privilege a small number of complementary measures for each individual field, they would still need to decide how to combine the scores received for individual fields into an overall 'figure of merit' for each model. It is not at all obvious how this can best be done. As the most recent Intergovernmental Panel on Climate Change (IPCC) chapter on model evaluation reports, 'it has proved elusive to derive a fully comprehensive, multidimensional "figure of merit" for climate models' (Houghton et al., 2001, p. 475). Thus far, scientists have been unable to use a measure of overall retrodictive performance to further narrow the field of complex models.

One might wonder why a comprehensive figure of merit is needed at all. After all, if scientists are interested in predicting future temperature changes, why not just evaluate climate models according to how well they simulate temperature changes up until now? One reason is that models may have been 'tuned' to some degree to ensure that they do a reasonably good job of reproducing available temperature observations. Tuning involves the manipulation of adjustable parameters in a model in order to make its output more closely match observational data. Tuning need not be informed by what is known about the physics of the system being simulated. In fact, the parameters being adjusted may not have any known correlate in the represented system but rather may be included in an ad hoc fashion expressly for the purpose of improving model-data fit. If a climate model must be tuned in an ad hoc manner in order to approximately reproduce the available temperature record, there is reason to question whether (or even doubt that) it will do a similarly good job in predicting future temperatures, since its past successes may have had relatively little to do with how well it described the physical processes that shape climatic conditions.

A more general reason that scientists would like to have a comprehensive figure of merit has to do with the nature of the climate system. Climate is thought to result from the interaction of numerous physical processes acting on a broad range of spatio-temporal scales. This means that errors in simulating one process may degrade the quality of many other aspects of the simulation. Thus it is desirable, even for the sake of prediction, to have climate models that perform well in simulating a range of climatic variables.[4]

2.4. *No Clearly Superior Model Emerges from Present Evaluations*

Most recently, climate model evaluation has involved large 'intercomparison' projects. The Coupled Model Intercomparison Projects (CMIP1 and CMIP2) are perhaps the best known of these. For these projects, different modeling groups carry out comparable simulations (e.g., using the same initialization fields and for the same simulated periods of time) to produce time series data for particular climatic variables of interest. Even if no comprehensive figure of merit has been developed, the output produced for these particular variables can be quantitatively compared among the models and with available observational data as long as some measure of model-data fit is selected. To date, no single measure has emerged as the 'gold standard' for comparison, even for individual variables, and a variety of measures are in use. Still, it is possible that one complex model would outperform all others for most variables and for a wide variety of measures of model-data fit. In practice, however, it turns out that when model output is compared with available climatic datasets, no single model consistently scores best even for the limited set of variables and measures of fit that are selected. Instead, some models perform better for some fields and measures of fit, while other models perform better for other fields and measures of fit (see, e.g., Lambert and Boer, 2001; Houghton et al., 2001, p. 482).

The situation is further complicated by the fact that there are several different 'observational' datasets with which model output might be compared. Because available observations of climate are sparse and of variable quality, it is not a simple matter to produce global climate datasets for use in model evaluation. As noted by Edwards (2001, p. 61), some recent climate model evaluations have made use of data produced via 'reanalysis' projects. These projects synthesize observational data and output from weather forecasting models to produce global datasets for hundreds of variables of interest, for the entire globe on a regular grid, and for regular time intervals. Models are used to fill in gaps in datasets, to interpolate observational data to particular grid points, and to derive non-observed fields (e.g., temperature advection and momentum exchange) from observed ones. Thus, some of the reanalysis data are determined almost entirely by observations, while other data are 'completely determined by the model' (Kalnay et al., 1996). Given a particular field for comparison (e.g., global annual mean precipitation) and a particular measure of model-data fit (e.g., root mean square error), some climate models score better for one 'observationally-based' dataset while other models score better for another such dataset (see, e.g., Figure 8.4 in Houghton et al., 2001).[5] In the most recent IPCC report, the authors of the chapter on model evaluation go so far as to say that they '... do not believe it is objectively possible to state which model is "best overall" for climate projection, since models differ amongst themselves (and with available observations) in many different ways' (Houghton et al., 2001, p. 475).[6]

3. THE MIXED STATUS OF COMPLEX CLIMATE MODELS

Having given some reasons for the persistence of a plurality of incompatible complex models, there remains the task of explaining why these incompatible models are viewed not primarily as competitors but as a team of models to be used together to investigate how climate might change in the future. One possibility is that climate scientists understand the models to be purely instrumental tools, so that their incompatibility need not be troubling as long as each of them can be used individually in some effective manner. But this explanation seems most promising if the models are to be used for different purposes, and the situation to be explained here is one in which incompatible models are used together in tackling the same modeling tasks.

Furthermore, it is clear that climate models are not considered by scientists to be purely instrumental tools. This can be seen, among other places, in scientists' own statements concerning the basis for their confidence in climate models:

> Confidence in climate models depends partly upon their ability to simulate the current climate and recent climate changes, and partly upon the realistic representation of the physical processes that are important to the climate system. (Houghton et al., 1996, p. 274)

That is, when evaluating climate models, scientists are concerned with both the simulations that the models produce and the assumptions that the models incorporate. Contrary to what would be expected if the models were viewed purely instrumentally, it is not enough for a model to simulate with some accuracy the past and present climate; the model should give the right results (i.e., accurate simulations) for the right reasons (i.e., because the relevant physics of the situation has been accurately described). So, a model may be praised or faulted either on the basis of how well its assumptions mesh with existing background knowledge about the climate system or on the basis of the perceived quality of its simulations.

An illustration of some scientists' concern over getting the right results for the wrong reasons can be found in the case of flux adjustments. When scientists began to join complex atmosphere models with ocean models (to produce complex coupled climate models), they observed that the climate simulated by the coupled models tended to slowly drift away from an equilibrium that was expected to be maintained. This drift occurred in part due to a mismatch between the fluxes of energy at the atmosphere–ocean interface in the coupled models. To remedy the situation, ad hoc adjustments to the flux values were (and sometimes still are) made in order to keep them in line with one another. But the need for flux adjustments is thought by many scientists to indicate that the assumptions built into the coupled climate models are fundamentally deficient. Even scientists who take a somewhat more pragmatic view toward modeling seem to consider flux adjustments to be something of a 'necessary evil' and agree that it is preferable for models to perform well without the need for flux adjustments (see, e.g., the analysis in Shackley et al., 1999). If a purely instrumental view were taken, then the need for flux adjustments would not be considered problematic.

Of course, not all climate scientists have exactly the same attitude toward climate models. In Shackley et al. (1999), it is argued that among climate scientists there are at least two different epistemological approaches to modeling, which are characterized as 'purist' and 'pragmatist' approaches. In effect, purists and pragmatists differ with respect to how closely they adhere to the 'right results for the right reasons' requirement mentioned above. Pragmatists tend to be less disturbed than purists by the introduction of ad hoc adjustments whose sole purpose is to improve the fit between model output and available data. This is in part because purists often view simulation of the climate system as a scientific exercise that might advance theoretical knowledge, while pragmatists often are concerned with simulating climate for purposes of aiding practical decision-making. This does not mean, however, that pragmatists have a purely instrumental view of climate models. Rather, they simply have a somewhat greater (not infinite) tolerance for ad hoc maneuvers than purists do. For both kinds of modelers, the realistic representation of physical processes is considered important and desirable, and it seems likely that both would characterize the ideal modeling situation as one in which climate models were constructed entirely via straightforward application of well-established physical principles.

There is a thus strong realist component to the perceived status of climate models: an ultimate aim is to develop models whose assumptions are approximately true of the real climate system.[7] At the same time, because global warming is perceived by many to be an urgent environmental problem, there is pressure to work around present model shortcomings and find a way to use them to answer key questions about how climate is likely to change in the future. In other words, there is also an instrumentalist component to the perceived status of climate models. Climate models have come to have a mixed status: they should incorporate realistic assumptions insofar as this is possible, but they also should be useful tools for addressing particular problems and questions.

How complex climate models are being used together to answer questions about future climate change is the subject of the next section.

4. THE MULTI-MODEL ENSEMBLE APPROACH IN CLIMATE MODELING

The situation in climate modeling thus can be summarized as follows. Scientists have developed a collection of incompatible complex models, each of which is intended to be a realistic representation of the climate system, insofar as current knowledge and technology permit. But even state-of-the-art complex climate models currently are constituted by a 'balance of approximations' (Lambert and Boer, 2001, p. 105) reflecting genuine scientific uncertainty, modeling preferences, and the desire to produce reasonably realistic-looking simulations of past and present climate. None of these models has emerged as clearly superior for purposes of investigating future climate change.

If no model stands out from the others as a more promising resource for predicting future climate, how are climate scientists to proceed? It would not be very sensible to pick a model randomly and then draw conclusions and choose actions based on the results given by that model alone, since it might turn out that one or more of the other models will (unbeknownst to scientists now) give much more accurate predictions of future climate. Scientists will have riskily 'put all of their eggs in one basket'. Instead, scientists are pursuing a 'multi-model ensemble' approach. The multi-model ensemble approach assumes that members of a set – or 'ensemble' – of complex models count as approximately equally plausible representations of the climate system.[8] Put slightly differently, although the climate models differ from one another in various respects, each model is assumed to be a reasonable balance of approximations, given present uncertainties in representing the climate system. The entire ensemble of models is then used in investigating future climate change. For a given greenhouse gas emission scenario, each model will be run individually to generate a projection of its own, but the product of interest from the study will be the entire collection of projections. In this way, scientists can investigate the implications of their uncertainty in representing the climate system; insofar as the ensemble of models spans that uncertainty, it will be reflected in the range of projections produced.[9]

To illustrate: suppose that scientists identify a 'most likely' greenhouse gas emission scenario and then use an ensemble of climate models to make projections of global mean temperature for the year 2050 under that scenario. It might happen that the members of the ensemble produce temperatures that span a wide range of values – the ensemble indicates that the temperature in 2050 might be somewhat cooler or somewhat warmer or just about anything in between. In this case, scientists learn that their uncertainty in representing the climate system translates into substantial uncertainty with respect to the result of interest. They must conclude (without any further information) that their present understanding of the climate system does not allow them to say with confidence what the temperature will be like in 2050. On the other hand, it might happen instead that nearly all of the projections of 2050 temperatures cluster rather tightly around one particular value. For example, perhaps nearly all of the models agree that there will be moderate warming by 2050. In this case, scientists' uncertainty in representing the climate system seems not to matter much; despite the differences in the models' assumptions, they tell a univocal story about what will happen to global temperature by 2050.

For decision-making and planning purposes, it may seem that the latter situation, in which the model-derived evidence points to a single conclusion, is preferable. However, one must proceed with caution. The fact that the models substantially agree in their temperature projections is no guarantee that the (approximately) agreed upon projection is an accurate one, even if the emission scenario is a realistic one. It is possible that the models in the ensemble all systematically underestimate or overestimate 2050 temperature. For present-day climate models, the possibility of this kind of systematic error may not be as unlikely as one might think, because the models have not been developed independently of one another. Many of today's models are descendents of a small number of climate models constructed in the early decades of computerized weather and climate modeling (see Edwards, 2000) and thus are likely to have some simplifying assumptions (and even computer code) in common. Although the models do differ from one another in important respects, their output may exhibit some of the same systematic errors because of what they have in common. In fact, recent model intercomparison projects have documented some typical systematic errors found in simulations of past and present climate (see Lambert and Boer, 2001).[10]

Despite the fact that one must be careful when interpreting the results produced by multi-model ensembles, when it comes to addressing the global warming issue, the ensemble approach seems clearly better than the two most obvious alternatives, that is, relying on a single model and/or making no use of climate models until a single 'best' one can be identified. An ensemble of models incorporating different parameterizations of key climate processes is currently being used to determine how climate might change under a variety of emission scenarios and is considered the most promising research strategy to pursue at present (see Houghton et al., 2001, p. 511). Instead of pretending that uncertainties do not exist, the ensemble approach acknowledges them and seeks to determine their implications. In the context of this approach, the fact that the models are physically incompatible need not be problematic; what is important is that each model (in conjunction with associated initial and boundary conditions) be considered a *plausible* representation of the climate system. For the reasons outlined in the preceding sections, this is the way that many scientists currently view many complex climate models.

5. PRAGMATIC INTEGRATIVE PLURALISM

The foregoing discussion has shown why a plurality of incompatible climate models persists and how these models nevertheless are being used together as complementary resources for investigating future climate. How should this model pluralism be characterized?

In the philosophical literature, two primary forms of pluralism have been identified – competitive and compatible pluralism (see, e.g., Mitchell 2002). Although it is not always emphasized, these forms of pluralism ultimately are concerned with accounts of what the world (or some part of it) is like, that is, with its ontology. Here, their labels will be expanded to ‘ontic competitive pluralism’ and ‘ontic compatible pluralism’. In the context of scientific modeling, ontic *competitive* pluralism exists when two models incorporate conflicting assumptions about the part of the world that they are intended to represent. In other words, as representations of the same target system, the models are mutually exclusive. For example, there might be one model of the solar system according to which the planetary orbits all lie in the same plane and another model according to which not all orbits lie in the same plane. Typically, when scientists have two representations that conflict in their assumptions about the world (and each representation is a candidate for belief or acceptance), the representations are viewed as competitors – it does not make sense to accept both of them as true of the world, so they compete for scientists’ belief/acceptance. By contrast, ontic *compatible* pluralism exists when there are two or more representations that can be true of the world at the same time. These representations do not conflict in their assumptions about what the world is like. For example, there might be one model of radiation transfer in the atmosphere and another model of plant respiration, which could be used at the same time in constructing a larger representation of the climate system. These can be viewed as compatible sub-models of a larger, more comprehensive model (see Bailer-Jones, 2000, for another example and discussion). Alternatively, there might be one model that describes only the aggregate features of some system (e.g., mean global temperature and precipitation) and a second model that describes the system in greater detail (e.g., temperature and precipitation on a fine spatial grid), but if the models closely agree in the nature of their assumptions about the system and in their predictions of the aggregate features, then the situation may be considered one of ontic compatible pluralism. One need not believe that either of two ontically compatible models is actually true of the world, but it is at least a logical possibility that they both are true (or approximately true) of it.

These forms of pluralism take one only so far in making sense of the situation in climate modeling. The situation does seem to be one of ontic competitive pluralism, since the climate models incorporate mutually conflicting assumptions about what the climate system is like. Ideally, scientists would like to choose from among the complex climate models that which does incorporate the most realistic assumptions about the physical processes that will shape future climatic conditions of interest (whatever those processes are). But for a variety of reasons, as shown above, scientists simply have been unable to identify such a model. The interesting feature of current climate modeling, however, is that scientists are not obsessed with paring down the collection of complex models that they have. In fact, as discussed in the last section, they are actually using the models together to investigate future climate. Are the models somehow compatible after all?

Sandra Mitchell's recent work on 'integrative' pluralism in biology (see Mitchell, 2002) may be of some help in making sense of the situation in climate modeling. Her analysis seems relevant because it is concerned with situations in which apparent competitors end up being compatible. More specifically, she shows how several idealized causal models that seem to provide competing explanations of some type of phenomenon (e.g., division of labor in social insects) can turn out to be compatible when it comes to explaining a particular, concrete instance of that phenomenon (e.g., division of labor in leaf-cutting ants). Mitchell argues that the idealized causal models are not actually in competition, because each applies in a different idealized (non-actual) situation (Mitchell, 2002, p. 64). Competition *can* occur among explanations of a particular concrete phenomenon, but each of those competing explanations typically will invoke several contributing causes. In other words, a plurality of idealized causal models will often be integrated in explaining an actual, complex biological phenomenon. There can be pluralism at the level of theoretical modeling, even though there will be only one 'true' integrated explanation of any particular, concrete phenomenon (Mitchell, 2002, p. 67). Mitchell's integrative pluralism thus seems to be a particular variety of ontic compatible pluralism – different possible accounts are brought together in producing a single actual account. This can be called 'ontic integrative pluralism'.

Climate model pluralism is also integrative, but in a different way. Several mutually incompatible but individually plausible climate models are used together not in order to construct one 'true' description of the climate system or to identify one 'true' projection of future climatic conditions, but rather to estimate how uncertainty in representing the climate system translates into uncertainty about future climate change. That is, scientists use the models together in order to better gauge their current epistemic situation. While the models are indeed incompatible with respect to ontology, their results are integrated in practice, as scientists pursue a particular modeling methodology to probe the implications of their uncertainty. Thus, the situation in climate modeling might be characterized as one of 'pragmatic integrative pluralism'.

Does pragmatic integrative pluralism require that the models involved be viewed as purely instrumental tools? It is suggested here that it does not. Even if complex climate models are not thought to be 'perfect' or 'true' descriptions of the climate system, it is largely because these models are believed to incorporate relatively realistic assumptions that they are used together at all. As indicated above, this is a scientific arena in which it is considered important to get the right results for the right reasons; a climate model's perceived plausibility – regardless of whether the model is being used in a multi-model ensemble study – is usually determined to a significant degree by the extent to which representations of important climate-shaping processes are included in the model and are grounded in well-established physical principles.

It thus appears that two different types of pluralism coexist in the case of climate modeling: an ontic competitive pluralism and a pragmatic integrative pluralism.

6. CONCLUSIONS

State-of-the-art climate models conflict with one another in some of their assumptions about the climate system, but scientists have been unable to select from among the models a single 'best' one for purposes of investigating future climate change. They are prevented from doing so both because uncertainty remains about the nature of the processes that shape climate and because of difficulties faced in evaluating complex climate models.

Yet climate scientists have not given up on models as a resource for investigating how climate might change in the future. Instead, they are pursuing a multi-model ensemble approach in which mutually incompatible but individually plausible models are used together to produce (for each emission scenario) a range of projections that reflect scientific uncertainty concerning how to best represent the climate system. Of particular interest is how broad the range of projections turns out to be, since the broader the range, the less confident scientists can be about how climate would change in response to a given emission scenario. (As indicated above, however, a narrow range does not necessarily mean that scientists *can* be confident.) As scientists learn more about the climate system, what counts as a plausible model will be further constrained. In the meantime, the multi-model ensemble approach will help scientists to take into account current uncertainties when providing information to decision makers about how climate is likely to change in the future.

Reflecting on the situation in climate modeling, it is possible to distinguish two dimensions along which models can be either compatible or competing. The ontic dimension concerns the compatibility of models' assumptions about what the world is like, while the pragmatic dimension concerns the compatibility of models in practice, as scientists pursue a particular modeling methodology. Pluralism in climate modeling can be characterized as combining an ontic competitive pluralism with a pragmatic integrative pluralism.

ACKNOWLEDGEMENTS

Thanks to John Earman, Francis Longworth, Sandra Mitchell, and John Norton for helpful comments on earlier versions of this paper. Remaining shortcomings are the responsibility of the author.

NOTES

1. The present discussion will be concerned with epistemic reasons, rather than social ones, although surely there are also some social reasons for the persistence of so many models.

2. Parameterization involves representing the effects of processes whose spatiotemporal scales are too small to be resolved by the model in terms of larger-scale variables that are resolved.
3. While some differences among cloud parameterizations do reflect scientific uncertainty, some also reflect what might be termed 'engineering uncertainty', that is, uncertainty concerning how to best incorporate, using large-scale resolved variables, what *is* known about clouds. Different parameterizations often are successful in different ways. For instance, one parameterization might give rather accurate values of average cloudiness for one geographical region, while another is more accurate for mean global cloudiness.
4. Of course, scientists do know something about what it is most important to 'get right' in order to make accurate predictions of particular climatic variables, but such knowledge remains incomplete.
5. The use of reanalysis data in climate model evaluation may be of interest to philosophers of science. Weather forecasting models have many core assumptions in common with climate models, so a dataset whose content is determined in part by such weather forecasting models may not be an appropriate resource for evaluating the quality of climate model simulations. The apparent model-data fit may be artificially inflated as a result of the shared assumptions of the weather and climate models. The issue of possible circularity in the testing of climate models will not be examined here.
6. It appears that the intended claim is that they do not believe it possible to *identify objectively* which model is best overall for climate projection. Subjective model assessment is common in climate science. For examples of subjective assessments, see Dai et al. (2001, p. 515) and Houghton et al. (2001, p. 479). Oreskes and Belitz (2001) note the prevalence of subjective assessment of numerical models more generally (beyond climate modeling).
7. The reader's intuitive understanding of 'approximate truth' is relied upon here.
8. There are several variations on the multi-model approach; just one of them is described here.
9. It is possible to use ensembles of initial and boundary conditions along with multi-model ensembles. This is currently done in weather forecasting, and the results can be compared with observed conditions. These ensembles do not always span the true uncertainty in representing the atmosphere, as indicated by the fact that the range of atmospheric conditions predicted by the ensemble does not always include the observed atmospheric conditions. It is difficult to carry out similar tests of the adequacy of climate model ensembles, at least for predictions of long-term climate, since the observations of climatic conditions with which the ensemble of predictions is to be compared will not materialize for many years.

10. It would seem an important next step for scientists to investigate to what extent these known, shared systematic errors, which may be only a subset of all of the systematic errors in the simulations, are likely to impact projections of global and regional climate.

REFERENCES

Bailer-Jones, D.: 2000, Modelling Extended Extragalactic Radio Sources. *Studies in the History and Philosophy of Modern Physics* 31: 49–74.

Dai, A., T.M.L. Wigley, B.A. Boville, J.T. Kiehl, and L.E. Buja: 2001, Climates of the Twentieth and Twenty-First Centuries Simulated by the NCAR Climate System Model. *Journal of Climate* 14: 485–519.

Edwards, P.N.: 2000, A Brief History of Atmospheric General Circulation Modeling. In D.A. Randall (Ed.), *General Circulation Development, Past Present and Future: The Proceedings of a Symposium in Honor of Akio Arakawa*. New York: Academic Press, 67–90.

Edwards, P.N.: 2001, Representing the Global Atmosphere: Computer Models, Data, and Knowledge about Climate Change. In C.A. Miller and P.N. Edwards (Eds.), *Changing the Atmosphere: Expert Knowledge and Environmental Governance*. Cambridge, MA: MIT Press, 31–65.

Holton, J.R.: 1992, *Introduction to Dynamic Meteorology*. San Diego, CA: Academic Press.

Houghton, J.T., L.G, Meira Filho, B.A. Callender, N. Harris, A. Kattenberg, and K. Maskell (Eds.): 1996, *Climate Change 1995: The Scientific Basis*. Cambridge, UK: Cambridge University Press.

Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu (Eds.): 2001, *Climate Change 2001: The Scientific Basis*. Cambridge, UK: Cambridge University Press.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph: 1996, The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society* 77: 437–471.

Lambert, S. and G. Boer.: 2001, CMIP1 Evaluation and Intercomparison of Coupled Climate Models. *Climate Dynamics* 17: 83–106.

Mitchell, S.D.: 2002, Integrative Pluralism. *Biology and Philosophy* 17: 55–70.

Oreskes, N. and K. Belitz: 2001, Philosophical Issues in Model Assessment. In M.G. Anderson and P.D. Bates (Eds.), *Model Validation: Perspectives in Hydrological Science*. Chichester: Wiley & Sons, 23–41.

Shackley, S., I. Risbey, P. Stone, and B. Wynne: 1999, Adjusting to Policy Expectations in Climate Change Modeling: An Interdisciplinary Study of the Use of Flux Adjustments in Coupled A/O GCMs. *Climatic Change* 43: 413–454.

Science Studies Program
University of California, San Diego
9500 Gilman Drive, Dept. 0104
La Jolla, CA 92093-0104
USA
E-mail: wparker@ucsd.edu