GLOBAL CLIMATE SCIENCE, UNCERTAINTY AND POLITICS:
Data-laden Models, Model-filtered Data

PAUL N. EDWARDS

Global climate change is among today's most visible and controversial areas of science-based policy. By almost any measure—number of researchers, size of budgets, attention from the public and from policymakers—the importance of this field has grown dramatically during the last decade. The collective annual budget of the US Global Change Research Program (USGCRP), a clearinghouse organization charged with coordinating research sponsored by a dozen different US government agencies, hovers near $1.8 billion (US Global Change Research Program, 1998). Although the USGCRP covers many areas in addition to atmospheric science, a large number of these—including oceanography, ecology, agriculture, and forest studies—are linked via the question of how a changing climate may affect them.

The degree to which climate change research can or should be characterized as 'policy-driven' science remains hotly contested (see e.g. Brunner, 1996). But on two points there is no debate.

First, climate change issues have become the focus of important and far-reaching policies, both national and international. Most of the world's nations have signed the Framework Convention on Climate Change (FCCC), negotiated at the 1992 United Nations Conference on Environment and Development at Rio de Janeiro. This commits them to further negotiations on limiting the 'greenhouse gas' emissions believed to contribute to global warming; the ensuing series of Conferences of Parties to the FCCC has been closely and widely watched. A number of nations have already enacted policies restricting or taxing greenhouse-gas emissions.

Second, scientific work has played an enormous role in legitimat-

---

Address correspondence to: Paul N. Edwards, School of Information, 301D West Hall, University of Michigan, 550 E. University Avenue, Ann Arbor, MI 48109-1092, USA, E-mail: pne@umich.edu.

0950-5431/99/040437-36 © 1999 Process Press

ing this political activity and increasing its pace. The work of the Intergovernmental Panel on Climate Change (IPCC), the official scientific advisory body on climate to the United Nations Environment Programme, has been highly visible and often controversial (Edwards and Schneider, forthcoming). In the summer of 1996, the IPCC's long-awaited Second Assessment Report announced that despite continuing uncertainty, 'the balance of evidence suggests a discernible human influence on global climate' (Houghton *et al.*, 1996, p. 5). Soon afterward, US Under-secretary of State for Global Affairs, Tim Wirth, formally announced to COP-2 that the United States would now support 'the adoption of a realistic *but binding* target' for emissions. Wirth noted in his address that 'the United States takes very seriously the IPCC's recently issued Second Assessment Report'. He proceeded to quote the report at length, proclaiming that 'the science is convincing; concern about global warming is real' (Wirth, 1996, provided by USGCRP). This is just one example of the tight coupling between climate science and policy at the highest levels.

☐ *Models and politics*

Computer models are arguably the single most important tool of global climate science. They range in size from simple programs that can be run on a personal computer to ultra-complicated simulation models of the entire Earth system, which strain the capacities of even the most powerful supercomputers. Much of climate science could not exist without them, since by definition, world-scale processes cannot be studied by controlled laboratory experiments. Instead, global 'experiments' must be performed on models. Furthermore, the huge size of global data sets makes it impossible to process or understand them in any detail without the aid of computers (Edwards, in press).

This article focuses on global climate models and their role in the political issue of climate change. However, most of my analytical comments apply equally to other kinds of global models (such as ocean, ecosystem, and agriculture models). First, I will describe how climate models work and discuss some of the key problems faced by modellers. Second, I will examine the extremely fuzzy boundaries between models and data in global climate research. Finally, I will look at how these scientific problems are taken up in political debates about climate change.

To a large degree these debates are in fact *about* the model/data relationship: whether model results agree with observations, how much each of these can be trusted, and what role these model projections should play in policymaking. I will argue that some parties to these debates have relied upon a conceptual separation between models and data that is, at best, misleading. The interdependent, even symbiotic, relationship between theory and observation in global climate science requires a different conception of the nature of scientific work. Uncertainties exist not only because of quantifiable, reducible empirical and computational limits, but also because of unquantifiable, irreducible epistemological limits related to inductive reasoning and modelling. These uncertainties can be, and have been, employed as political resources.

Raising the sophistication and effectiveness of political debate about climate change requires an understanding of the many sources of uncertainty. My conclusions will avoid the pose of disinterest by using the results of my analysis to make two normative points. First, modelling is a *sine qua non* of both knowledge and policy about climate change; we cannot do without it, since the data themselves depend on modelling. Second, a responsible policy process must acknowledge the multiple forms of uncertainty inherent in both scientific and policy knowledge.

■ GLOBAL CLIMATE MODELING: CONCEPTS, TECHNIQUES, AND PROBLEMS

The Earth is bathed in a constant flood of solar energy, all of which it ultimately re-radiates into space. At the most abstract level, therefore, the Earth's temperature is a simple matter of what climatologists call 'energy balance': all the energy that goes into the system must, eventually, come out again. The atmosphere forms a blanket of gases (primarily nitrogen, oxygen, and water vapour) capable of absorbing and holding a great deal of this incoming energy as heat.[1] In theory, if the Earth had no atmosphere, then its average surface temperature would be about –19°C. Instead, the heat retained in the atmosphere and oceans maintains it at the current global average of about 15°C.

☐ *General circulation models*

Climate models are mathematical simulations, based on physical laws, of long-term atmospheric conditions as they evolve over time.

The simplest, 'zero-dimensional' models rely on the principle of energy balance (and are called 'energy balance models' or EBMs).[2] Using measured values for such factors as solar radiation and concentrations of the atmosphere's constituent gases, they compute a single global average temperature, treating the Earth as if it were a point mass. One- and two-dimensional EBMs also exist. In addition, there are two-dimensional models called 'radiative–convective models', which give a picture of the atmosphere's vertical temperature structure. In these models, temperature is computed as a function of latitude and either longitude (a second horizontal dimension) or air pressure (the vertical dimension). Such models are more complicated than zero-dimensional EBMs, yet they are still relatively simple—typically, a few hundred to a few thousand lines of computer code—compared to the *three*-dimensional models known as atmospheric general circulation models (GCMs, or AGCMs to distinguish them from OGCMs, which model the oceanic general circulation).[3]

Contemporary GCMs are typically expressed in 30–60 thousand lines of FORTRAN code. They represent the atmosphere as a three-dimensional lattice or 'grid'.[4] Typically, the grid resolution at the surface is 4–5° latitude × 7–8° longitude. (This translates into rectangles between 250 and 500 km on a side.) Between eight and 20 layers of varying depth represent the vertical dimension up to a height of 20 km. or so, with more layers at lower altitudes, where most weather (in the ordinary sense) occurs. Equations of state compute the effect of various forces (radiation, convection, etc.) on the air mass within each grid box. Equations of motion compute the direction and speed of the air's movement into the surrounding grid boxes. AGCMs usually also include representations of certain aspects of the land surface, such as elevation and albedo (reflectance), as well as some representation of the oceans.

GCMs are 'coarse-grid' versions of the models used to compute weather forecasts, known as numerical weather prediction (NWP) models. For the sake of clarity, I will distinguish NWP models from GCMs, although the latter are simply low-resolution versions of the former. The differences between them lie in how they are used, not in their structure. Three of these differences create a very important computational bottleneck for GCMs.

- **Weather vs climate.** *Weather* refers to particular events and conditions over hours, days, or weeks. Because of the inherently chaotic nature of weather, NWP methods are unable to make useful predictions beyond 2 or 3 weeks, even in principle. *Climate*, by contrast, describes the average condition of the atmosphere over long periods: seasons, decades, centuries. To simulate climate, then, GCMs must be run over long periods. A typical 'run' today is anywhere from 20 to 100 model years, or even more.
- **Predictive modelling vs simulation.** NWP models are *predictive*. They are initialized with observational data, such as temperature, humidity, and wind speed, from a wide range of sources, including surface weather stations, satellites, and radiosondes (weather balloons). The models then calculate the likely evolution of this observed initial state over short periods (hours to days). By contrast, when used for climate research, GCMs generally are not initialized with observational data. Instead, they *simulate* climate, starting—in principle, at least—with only a few empirically derived inputs such as solar radiation, gaseous composition of the atmosphere, sea surface temperatures, and orbital precession. The models may take several simulated years to reach equilibrium, the point at which they settle into their own 'preferred' climate.
- **High vs low resolution: the problem of scale.** The grids of NWP models are far finer than those of GCMs. The best NWP models today use grid scales below 1 km on a side, compared with the 250–500 km grids of most GCMs. This matters because it allows NWP models to resolve relatively small-scale processes such as the formation and motion of clouds. GCM grids cannot reproduce these directly; climate scientists refer to such phenomena as 'sub-grid-scale' processes.

GCMs recompute the state of the entire atmosphere on a regular 'time step', usually 15–30 simulated minutes. This process is extremely computationally intensive. At each time step, hundreds to thousands of complex calculations must be performed on each of the tens of thousands of grid boxes. This consumes vast quantities of supercomputer time; a typical state-of-the-art GCM currently requires tens to hundreds of hours for a full-length 'run' of 20–100 simulated years. In principle, climate modellers could achieve far better results with high-resolution NWP models. But the number of model calculations increases exponentially with higher resolutions. This creates a computational bottleneck, forcing GCM builders to make trade-offs between a model's resolution and its complexity.

'Complexity' here refers to two related things: the number of phenomena simulated, and the level of detail at which they are modelled. Existing models do not directly simulate a vast number of basic atmospheric events. The most important of these is the formation of clouds, which typically occur on scales of 0–10 km. Clouds are believed to play many key roles in climate, such as trapping heat at night or reflecting it back into space during the day. These phenomena are notoriously difficult to study empirically, and their role in climate remains controversial. Clouds are not yet perfectly modelled even with NWP techniques. Other phenomena not well captured at GCM resolutions are the activity of the planetary boundary layer (the layer of air nearest the Earth's surface) and many factors relating to the land surface, such as its roughness and elevation. For example, many current models represent the entire region between the Sierra Nevada and the Rocky Mountains as a single plateau of uniform elevation.

Their low resolution is one reason for the high levels of uncertainty surrounding climate models. (I will discuss some other, epistemological reasons below.) Techniques for getting the most out of these low-resolution models have improved them, but have also been intensely controversial. The next section reviews some of these techniques and the associated problems and controversies.

☐ *Techniques and problems: parameterization*

Most of the major issues in climate modelling stem from the problems of *scaling* described above. All sub-grid-scale processes must be represented parametrically, or *parameterized.* For example, rather than represent cloud formation in terms of convection columns, cloud condensation nuclei, and other direct causes, a GCM typically calculates the amount of cloud cover within a grid box as some function of temperature and humidity. This approach embodies what is known as the 'closure assumption'. This is the postulate that small-scale processes can ultimately be represented accurately *in terms of the large-scale variables available to the models.*

Parameterization is controversial, and its effects on the activity of models are not entirely encouraging (Shackley *et al.*, 1998). For example, some cloud parameterization schemes in early GCMs resulted in cloud 'blinking', an oscillation between the presence and absence of cloud cover in a given grid box at each time step when certain variables happened to be just at the critical threshold. Real clouds do not, of course, behave like this. The question is whether and how unrealistic behaviour of this sort in one element of the model affects the quality of overall model results.

Another example of a parameterized function is atmospheric absorption of solar radiation, the energy driver for the entire climate system. Atmospheric molecules absorb solar energy at particular frequencies known as spectrographic 'lines'. 'The contribution of each narrow absorption line must be accounted for to model the transfer of radiation. ... There are tens of thousands of such lines arising from all the absorbing gases in the atmosphere. Thus, to include all lines in a parameter of absorption would require an explicit summing over all lines at each model level and horizontal location. These types of calculations can be performed on present day supercomputers and are called line-by-line models' (Kiehl, 1992, p. 338). But such modelling is too computationally expensive.

Instead, absorption is represented in GCMs by coefficients which implicitly integrate all the absorption lines. The coefficients are approximations, derived from line-by-line (or other) models of absorption. The use of one model to generate coefficients for another is one example of the intense interplay analyzed in this essay.

In an ideal climate model, the only fixed conditions would be the distribution and altitude of continental surfaces. All other variables, such as sea surface temperature, land surface albedo (reflectance), cloud formation, etc. would be generated internally by the model itself from the lower-level physical properties of air, water, and other basic constituents of the climate system. To say that current GCMs are far from reaching this goal is a vast understatement. Instead, '[v]irtually all physical processes operating in the atmosphere require parameterization' in models (p. 336). Generating these parameters is therefore the largest part of the modeller's work.

Climate modellers do this partly by reviewing the meteorological literature and observational data to try to determine how small-scale processes and large-scale variables might be related. When they succeed in finding such relations, they call the resulting parameters 'physically based'. Often, however, they do not find direct links to large-scale physical variables. In this common case, modellers invent *ad hoc* schemes which provide the models with the necessary connections. For example, one method of cloud parameterization represents all the cumulus clouds in a given region as a single 'bulk' cloud (Yanai *et al.*, 1973). In addition, observed patterns exist which can be mathematically described, but whose physics are not understood. These, too, are represented in the models as parameters.

Another, very important part of modellers' work is known as 'tuning' the parameters. 'Tuning' means adjusting the values of coefficients and even, sometimes, reconstructing equations in order to produce a better overall model result. 'Better' may mean that the result agrees more closely with observations, or that it more closely corresponds with the modeller's judgement about what one modeller called the 'physical plausibility' of the change. In some cases parameters fit relatively well with observed data. In others—as in the case of cloud parameterizations—the connection is so uncertain that tuning is *required*. Such parameters are said to be 'highly tunable'. Since many parameters interact with others, tuning is a complex process. Changing a coefficient in one parameter may push the behaviour of others outside an acceptable range.

☐ *Techniques and problems: flux adjustment*

Today's most sophisticated climate models couple atmospheric general circulation models with general circulation models of the oceans. The latter operate on principles much like those of atmospheric GCMs. These 'coupled' models, known as OAGCMs, must somehow provide for the exchanges or 'fluxes' of heat, momentum (wind and surface resistance), and water (precipitation, evaporation) between the ocean and the atmosphere. Empirical knowledge of these fluxes is not very good, but their values have profound effects on model behaviour.

Most OAGCMs include *ad hoc* terms, known as 'flux adjustments', which modify and correct the overall model results to bring them more closely into line with observations. Without them, the models' climates drift out of line with observed values and patterns (Meehl, 1992). These adjustments are 'non-physical' model terms, in modellers' language, although they are also characterized as 'empirically determined' (Houghton *et al.*, 1996, pp. 237, 34); they are an excellent example of a 'highly tunable' parameter. Recently the National Center for Atmospheric Research has introduced the first OAGCM that does not require flux adjustments (Kerr, 1997).

Parameterization and tuning are, in effect, scientific art forms whose connection to physical theory and observational data varies widely. As one modeller told me, 'Sure, all the time you find things that you realize are ambiguous or at least arguable, and you arbitrarily change them. I've actually put in arguable things, and you do that all the time. You just can't afford to model all processes at the level of detail where there'd be no argument. So you have to parameterize, and lump in the whole result as a crude parameter'.[5]

☐ *Common code: GCMs as a family*

One final issue about GCMs concerns their relationships with each other. Because of their complexity and expense, the total number of atmospheric GCMs is not large—probably around 50 worldwide, counting only those in active use (World Climate Research Programme Working Group on Numerical Experimentation, 1999). Many of these models share a common heritage. Typically, one modelling group 'borrows' another group's model and modifies it. This avoids unnecessary replication of effort, but it also means that the 'new' models may retain problematic elements of those from which they were created. Several modellers told me that substantial segments of the computer code in modern GCMs remain unchanged from the original models of the 1960s.[6] This may be one reason for the fact that some systematic errors in GCMs are common to virtually all extant models (Boer, 1992; Bourke *et al.*, 1991; Gates, 1997).

■ VALIDATING CLIMATE MODELS: EPISTEMOLOGICAL ISSUES

The validation of climate models is another complex issue with both epistemological and empirical dimensions.

Climate scientists frequently use this term to describe the evaluation of model results against empirical observations, along with various forms of statistical testing (Gates, 1997; Kiehl, 1992; Rotmans, 1990; Schneider, 1992, 1994; Wigley *et al.*, 1992). A 'validated' model is one that reliably reproduces observed climatological patterns.

This usage of 'validation' is contested. An influential recent article by Oreskes *et al.* (1994) maintained—based on results from philosophy of science—that modellers could not properly speak of 'validation' or 'verification' of models. 'Verification', in their view, implies definitive proof of truth, but models are essentially intricate inductive arguments. Since no inductive proposition can be proven with perfect certainty, models (like most scientific theories) cannot be verified in this strict sense. The fact that a model agrees—even perfectly—with observations does not guarantee that the principles it embodies are true. The possibilities always remain either that some other model could explain the observations equally well, or that future observations will not agree with the model.

'Validation' is a somewhat less stringent standard. Strictly defined, it refers to the demonstration of internal consistency and an absence of detectable flaws. Thus a model might be valid, in this sense, without being an accurate explanation. Nevertheless, as Oreskes *et al.* pointed out, 'validation' is commonly used by scientists as a synonym for 'verification'.

Oreskes *et al.* concluded that models can at best be 'confirmed'. This term implies only that model results agree with observations. A 'confirmed' model remains within the set of viable candidates for true explanation. In other words, confirmation raises the probability that the model embodies true principles, but cannot confer absolute certainty. This view is consistent with Popper's well-known doctrine of falsificationism, which holds that scientific hypotheses (or models) can be proven false by observations, but cannot be proven true (Popper, 1959 [1934], 1962). It also accords with more general results from the philosophy and sociology of knowledge related to inductive reasoning (Hume, 1977 [1748]).

Although the Oreskes *et al.* discussion is a generalized analysis of scientific models rather than a specific critique of GCMs, it has been taken to heart by the climate science community. As a result, the IPCC now avoids the word 'validation'. In its place, the group has substituted 'evaluation', defined as assessment of 'the degree of correspondence between models and the real world they represent' (Houghton *et al.*, 1996, p. 235).

What are the current results of climate model evaluation? The most recent IPCC evaluation effort concluded that 'the large-scale features of the current climate are well simulated on average by coupled models [OAGCMs]' (p. 249). The group agreed that

> the increasing realism of simulations of current and past climate by coupled atmosphere–ocean climate models has increased our confidence in their use for projection of future climate change. Important uncertainties remain, but these have been taken into account in the full range of [IPCC] projections of global mean temperature and sea level change (p. 5).

This evaluation was based not only on comparisons of model results with observations, but also on model *inter*comparison, a relatively new methodology (Gates, 1997).

Do these epistemological issues have practical implications? Yes, and they are important. Below, I will argue that the epistemology of modelling is among the main foci of political debates about climate change. Distinguishing evaluation from validation or verification helps to clarify the proper role of models in climate change projections: rather than absolute truth claims or predictions, they provide heuristically valuable simulations or projections (Edwards, 1996).

Before turning to political issues, however, I want to focus on another epistemological issue, namely the relation of models to observational data. Here I will argue that any sharp distinction between models and observations is a misleading caricature of climate science in practice.

☐ *Data-laden models*

A 'model', in the view of Oreskes *et al.*, is a theoretical construct, deriving its results from physical laws. It uses basic information about key physical variables only as a starting point (what modellers call 'initial conditions'). This is a fine description of *one* ideal toward which modellers may strive. Yet it fails to capture the purpose of the simulation, modelling practices which now dominate many scientific fields. These 'computational' sciences attempt to model phenomena not readily subjected to standard forms of laboratory experiment. For them, the goal of modelling is less to derive a single correct explanation of some natural phenomenon (the purpose of pure theory), than to convincingly *reproduce* the phenomenon in question. As Eric Winsberg has argued, simulation modelling of this sort is really an application, rather than a test, of theory. As such, it involves 'a complex chain of inferences that serve to transform theoretical structures into specific concrete knowledge of physical systems. ... [The] epistemology [of simulation modeling] is unfamiliar to most philosophy of science, which has traditionally concerned itself with the justification of theories, not with their application' (Winsberg, forthcoming).

True analogue models, in which one physical system is used to model another (by 'analogy'), may be largely non-theoretical. Early experiments in climate modelling, for example, sometimes used analogue models, such as dishpans (representing the Earth) filled with fluids (representing the atmosphere) rotating above a heat source (representing the sun); scientists observed the flow patterns of the fluids (Hide, 1952). Today, true analogue models have virtually disappeared from the field. Nevertheless, one might argue that the extensive climate studies of Mars and Venus serve a similar function. Indeed, the digital simulations which replaced early analogue models serve an identical purpose. Like laboratory experiments, which operate by tightly constraining situations which would be far more complex in reality, both digital simulations and analogue models function as analogues whose activity is assumed to parallel real-world phenomena (Norton and Suppe, forthcoming).

As my discussion of parameterization has shown, the reality of climate modelling practice corresponds poorly to the ideal of pure-theory modelling.

Many of the basic physical laws governing atmospheric behaviour are well understood and relatively uncontroversial. Modellers call these the 'primitive equations'. But the huge range of spatial and temporal scales involved—from the molecular to the global, from milliseconds to millennia—makes it impossible to build models from these principles alone. Schneider notes that

> even our most sophisticated 'first principles' models contain 'empirical statistical' elements within the model structure. ... We can describe the known physical laws mathematically, at least in principle. In practice, however, solving these equations in full, explicit detail is impossible. First, the possible scales of motion in the atmospheric and oceanic components range from the submolecular to the global. Second are the interactions of energy transfers among the different scales of motion. Finally, many scales of disturbance are inherently unstable; small disturbances, for example, grow rapidly in size if conditions are favorable (Schneider, 1992).

Hence the necessity of parameterization, much of which can be described as the integration of observationally-derived approximations or heuristics into the model core. Schneider sometimes refers to parameters as 'semi-empirical', an apt description that highlights their fuzzy relationship with observational data.

For the foreseeable future, all GCMs will contain many of these 'semi-empirical' values and equations. Thus we might say that GCMs are *data-laden*.

I use this phrase symmetrically with the well-known observation that data are 'theory-laden' (Hanson, 1958; Popper, 1959 [1934]). In one sense there is nothing odd about this, since theory in the physical sciences always includes constants (such as the gravitational constant or the sun's energy output) derived from empirical measurements. However, physical-science practice normally attempts to explain large-scale phenomena as an outcome of smaller-scale processes. The 'data-ladenness' I describe here refers to the inclusion of large-scale, empirical statistical data in models, which necessarily goes against the reductionist imperative of the physical sciences.

☐ *Model-filtered data*

Global climatological data sets are also problematic.

Some of the reasons are obvious. Many kinds of measurements, from many different instruments, are necessary to make up a data set that covers the entire global atmosphere in three dimensions and over many years. These measurements are taken under a vast variety of conditions, which differ for reasons that are not only physical (e.g. Antarctic vs. temperate zones), but social (differing levels of understanding, technical skill, and experience in different countries) and historical (changes in techniques, instrumentation, etc. over time).

Fairly good records of land and sea surface meteorology exist for the last 100 years, but changes over time in instrument quality, location, number, and measurement techniques create many uncertainties. For example, most thermometers are located on land and clustered in urban regions, where 'heat island' effects raise local temperatures above the regional average. Meteorological records at sea tend to be drawn from shipping lanes, ignoring the globe's less travelled areas. For the last several decades, records from the atmosphere above the surface have been drawn from increasingly extensive commercial aircraft, radiosonde (weather balloon), and rawinsonde (radar-tracked radiosonde) networks, but these too are concentrated in particular areas. Coverage in the tropics and in the southern hemisphere is particularly poor. Heroic efforts continue to purify these data, by estimating and correcting for systematic errors (Houghton *et al.*, 1996, chapter 3). For example, satellite data are being used to estimate the effects of urban heat island bias on global surface temperature data (Johnson *et al.*, 1994); historical sea surface temperatures have been corrected for the effects of different kinds of buckets used to draw water samples (Folland and Parker, 1995); and problems with rawinsonde data are being addressed by comparisons with satellite data and corrections for various sampling errors (Parker and Cox, 1995).

Among the chief tools of this data filtering process are what we might call 'intermediate models'. These include models of instrument behaviour, interpolation techniques (for converting actual observations into gridded data), techniques for automatic rejection of anomalous data points, and many other methods (Christy *et al.*, 1995; Hurrell and Trenberth, 1997; Karl *et al.*, 1994). NWP models themselves are also used to filter and process data, a technique known as four-dimensional data assimilation. Recently, the problem of historical change in data assimilation systems—i.e. the fact that the models used to produce uniform data sets have changed over time, leading to inhomogeneities in data characteristics—is being addressed by a number of major data 'reanalysis' projects. These projects take raw historical data and reprocess them using state-of-the-art, GCM-based data assimilation systems in order to create a more homogeneous data set (European Center for Medium

Range Weather Forecasts, 1999; National Oceanic and Atmospheric Administration, 1999).

☐ *Satellites and global data sets*

In the last two decades, satellite observations of atmospheric radiative structure, cloud cover, and land surface characteristics have attained excellent quality. Unlike all other data sources, these have the signal advantage of being genuinely global in scope. Weather satellites overfly the entire globe at least twice every day. This total coverage makes satellite data extremely attractive to climate modellers. 'We don't care about a beautiful data set from just one point', one modeller told me. 'It's not much use to us. We have one person whose almost entire job is taking satellite data sets and putting them into files that it's easy for us to compare our stuff to.'

Yet satellite data are also problematic. Satellites provide only proxy measurements of temperature at low altitudes, which may be distorted by optical effects. In addition, their lifespans are short (2–5 years) and their instruments may drift out of calibration over time(Christy *et al.*, 1995). A number of scientists, including one responsible for satellite data analysis at a major climate modelling group, told me that the quality of these data was not very good. One said that their main practical value has been for television weather images. Nevertheless, the satellite data are generally regarded as the most reliable global observational record.

The solution, again, is a suite of intermediate models. Statistical models filter out 'signals' from noise; models of atmospheric structure and chemistry are used to disaggregate radiances detected at the top of the atmosphere into their sources in the various atmospheric layers and chemical constituents below. In addition, models are used to 'grid' the data and to combine them with other data sources. Among the most important data sources are the twice-daily atmospheric analyses of the US National Meteorological Center and the European Centre for Medium-Range Weather Forecasting. 'These analyses incorporate observational data from both the surface and from satellites into a 4-D data assimilation system *that uses a numerical weather prediction model* to carry forward information from previous analyses, giving global uniformly gridded data' (Kiehl, 1992, pp. 367–8, emphasis added). Thus the twice-daily periods of actual observation are transformed into 24-h data sets *by NWP general circulation models.* These model-dependent data—integrated with other data sets that have also been processed, by models, to correct for various errors—are then used to validate (or 'evaluate') general circulation models.[7] This is exactly what it sounds like; one data analyst described the relationship between GCMs and 'data products', as they are known in the atmospheric sciences, as 'highly incestuous'.

The point here is that despite their global coverage, satellite data are no more immune than others to the need for model processing. Modelling is *required* for them to support any projections about global climate change.

What sort of behaviour, for example, should we expect in the near future based on observations of the recent past? In a highly complex system with multiple feedbacks, such as climate, there is no *a priori* reason to suppose that historical trends will continue in a linear progression. Such an assumption suffers profoundly from the *same* inductive fallacy discussed above. In effect, it too is a model of atmospheric behaviour, but without any basis in physical theory. The point here is that without *some* model of atmospheric behaviour—even this primitive and almost certainly false one—the exact shape of the curve of global climate change cannot be projected at all.

Finally, modelling is necessary to separate human from natural contributions to climate change. For example, major volcanic eruptions, such as those of El Chichón (1982) and Mount Pinatubo (1991), can inject vast quantities of particulate aerosols into the stratosphere, causing cooling near the surface and warming in the stratosphere that can last for several years. A much more difficult problem of the same type is that climate varies naturally over multiple time scales; trends over shorter periods (e.g. decades) may even be of opposite sign from long-term trends. If we are to understand the human role in global climate change, the effects of major natural events must be extracted from the global data and natural climate variability must be separated from anthropogenic (human-caused) trends. This can be done only through modelling.

What is 'global' about global climate data?

These problems suggest another important epistemological issue about the relationship between climate models and observations, namely the question of exactly what is 'global' about our knowledge of global climate.

None of the individual observational data sets available remotely approach what might be construed as a minimal requirement for truly global climatological data: namely, coverage of the entire Earth on (say) a 100 × 100 km grid, using standardized measuring techniques and well-calibrated instruments, with at least twice-daily sampling over a period of at least 100 years. Instead coverage is spotty, inconsistent, poorly calibrated, and temporally brief.

Rather, it is the *models* which are global. They make inaccurate, incomplete data *function as* global by correcting, interpolating, completing, and gridding them. The sheer size of the data sets makes it unlikely that much could be learned from them without the computer models that make them comprehensible. Furthermore, global uniformly gridded data would never have been generated in the first place without the models which required data in that form. The dynamics of the Earth's atmosphere could not be understood without them—at least not at a level of detail which would confer the ability to make long-term projections.

☐ *The model/data relationship as symbiosis*

The model/data relationship in climate science is thus exceptionally complex. Models contain 'semi-empirical' parameters, or heuristic principles derived from observations. Meanwhile, global data sets are derived from direct observations by modelling. Since the problems of scale that create this situation are present in all global dynamic processes, the same could be said of all world-scale models. Seen in this light, the issues of model validation, confirmation, and evaluation take on a rather different cast.

What should we make of all this? While it looks very little like our idealized image of science, in which pure theory is tested with pure data, that image was always a false one (Collins and Pinch, 1993; Galison, 1987; Kuhn, 1962; Lakatos and Musgrave, 1970; Latour and Woolgar, 1979). However, neither is it a viciously circular relationship. Instead, the interdependence is symbiotic. In other words, in computational sciences theory and data feed on each other, in a mutually beneficial but also mutually dependent relationship.

This is not (or need not be) a form of contamination. Interdependence is not identity; data sets are still derived *primarily* from observation, and models *primarily* from theory. Multiple sources of data; multiple, independently developed models; and cross-correlation and intercomparison of all of these make the increasing convergence of climate model results and climate data unlikely to be artefactual. Yet none of this eliminates the model-dependency of data or the data-ladenness of models.

I think we should, instead, view these phenomena as support for the claims of philosophers Frederick Suppe and Stephen Norton. Defining scientific methods essentially as ways of controlling for the possibility of artefactual results, they argue that the blurry model/data relationship pervades all science. Even the laboratory sciences, 'with suitable computational capabilities and sufficient instrumentation, removal of or correction for artifactual elements in the data analysis stage often provides more effective laboratory experimental control than do attempts to remove potential contaminating effects via physical control'. Therefore, 'heavy reliance on modeling in no way impugns the epistemological status of the [climate modellers'] claims' (Norton and Suppe, forthcoming). If they are right (and I believe they are), the purity of models or data (in the sense of maintaining their separate status) is not the important issue. Instead, the question is how well scientists succeed in controlling for the presence of artefactual elements in both theory *and* observation.

The model/data relationship should be viewed as symbiotic, rather than oppositional, because the purpose of simulation models is not to explain or theorize, but to forecast by creating analogues based in both theory and data. As Sergio Sismondo puts it, '[a]pplied theory isn't simply theory applied, because it instantiates theoretical frameworks using a logic that stands outside of those frameworks'. Describing Winsberg's analysis, Sismondo continues:

> simulations and their components are evaluated on a variety of fronts, revolving around *fidelity* to either theory or material; assumptions are evaluated as close enough to the truth, or unimportant enough not to mislead; approximations are judged as not introducing too much error; the computing tools are judged for their transparency; graphics systems and

> techniques are expected to show salient properties and relationships (Sismondo, forthcoming, emphasis in original).

Thus the concept of purity on which received notions of model/data relations often rely is misguided, even irrelevant, in the 'semi-empirical' world of simulation modelling.

THE POLITICS OF MODELS AND DATA

All this may seem like terminological hair-splitting, of interest only to philosophers. Yet the epistemological status of models and their relationship to observation have become significant, even central issues in global change debates. In recent climate politics, contests between what Samuel Hays has called 'frontier' and 'high-proof' scientists have taken centre stage (Hays, 1987, 1989). The two groups differ essentially in their attitude toward uncertainties and imperfections in models.

'Frontier' scientists prioritize theory. In general they believe that models, while still imperfect, are an appropriate scientific tool. Rather than wait for an unambiguous 'signal' of climate change to show up in the observational data, they are willing to accept intermediate levels of data/model agreement as preliminary, if uncertain, confirmation of model results. Increasing convergence between observations and models is enough to convince them that the models are to some extent reliable, especially when models reproduce not only global climatic averages but observed *patterns* of climatic behaviour (Santer *et al.*, 1993, 1996; Schneider, 1994; Wigley *et al.*, 1992). Other kinds of confirmation, such as close agreement among several models produced by independent groups, may also increase their faith in model results.

'High-proof' scientists, by contrast, prioritize observation. Their tolerance for uncertainty is lower; they seek high levels of empirical confirmation before accepting model results. This group tends to perceive variation among models, and systematic errors within them, as evidence of their fundamental inadequacy (Spencer *et al.*, 1997). For them, parameterization is often a particularly troubling concern. Richard Lindzen, for example, frequently refers to parameterizations of water vapour and cloud physics as a severe problem with climate GCMs (Lindzen, 1990, 1992, 1996).

Obviously, 'frontier' and 'high-proof' describe opposite ends of a spectrum of possible views. The frontier/high-proof conflict, ubiquitous in climate politics, a phenomenon common to many public scientific debates, especially in the environmental arena (Jasanoff, 1990, 1991a, b). Virtually every press report on climate change quotes modellers who express limited confidence in model results, 'balanced' by others who point to discrepancies between models and data and express scepticism about the very possibility of accurate modelling (see Schneider, 1989, chapter on 'Mediarology'). This conflict might be dismissed as an artefact of American media standards of 'objective' journalism, but it appears directly in political debate as well. I will discuss two recent examples.

☐ *The Scientific Integrity Hearings*

In 1995, the US House of Representatives Subcommittee on Energy and Environment convened a series of hearings entitled 'Scientific Integrity and the Public Trust'. Chaired by Rep. Dana Rohrabacher (R-CA), the hearings were part of a sweeping attack on established Federal environmental policymaking techniques by the 104th Congress' newly-elected Republican majority. Each of the three hearings addressed a particular environmental issue where 'abuse of science' was alleged to have occurred: climate change, ozone depletion, and dioxin regulation.

In each case, the challenge took a similar form. Scientific witnesses of the high-proof school were called (as well as others). Some of them, such as Patrick Michaels and S. Fred Singer, testified that empirical observations failed to bear out the theoretical predictions of the science 'establishment', theories embodied in computer models. These 'sceptic' scientists went on to claim that the observational data failed to confirm the models. Many, including Michaels, Singer, and Sallie Baliunas, also claimed that their interpretations of observational data, and/or their own alternative theories or models, had been systematically ignored by the science establishment (e.g. in the case of climate change, by the IPCC). This establishment's self-interest in maintaining government funding for its research was alleged to be among the corrupting influences leading to this deliberate suppression of 'sound science'.

'Sound science', in this context, was the phrase used by Republican representatives to promote a new set of standards in science-for-policy: near-absolute empirical confirmation before action.[8]

> Rep. Doolittle, in his prepared statement for the ozone hearing, stated that 'sound science must be the basis for all future decisions we make on this important issue.' In seeking to clarify the definition of sound science, Ms. Rivers asked '... [W]hat would you consider to be sufficient evidence for action to be taken in this area?' Mr. Doolittle responded, 'I think we need a clear scientific conclusion that there is a definite cause for the problem and that so-called problem is producing definite effects. Theories or speculation about it are not sufficient. We need science, not pseudo-science. I think we've been in an era of pseudo-science where these dire consequences are portrayed in order to achieve a certain political objective.' Similar statements were made by other Members in the global change hearing with respect to projections from computer models and in the dioxin reassessment hearing with respect to choices of models of dioxin receptor activity (Brown, 1996, section IV.D).

The slogan referred directly to high-proof standards: 'science programs must seek and be guided by empirically sound data' rather than theory or models. On one level this principle is salutary. But in the context of the Energy and Environment Subcommittee's attempt to discredit the consensus opinions of the IPCC, the 'sound science' slogan served another, rather obvious political purpose. In a report on the hearings Rep. George W. Brown Jr, the committee's ranking minority member, accused the Republican majority of a 'totally unrealistic view both of science's present capabilities and of the relationship between data and theory in the scientific method'. This approach to science, he said, 'can lead to near paralysis in policy-making' because it requires an 'impossible standard' of certainty. 'Uncertainty is not the hallmark of bad science; it is the hallmark of honest science', Brown stated (section IV.D).

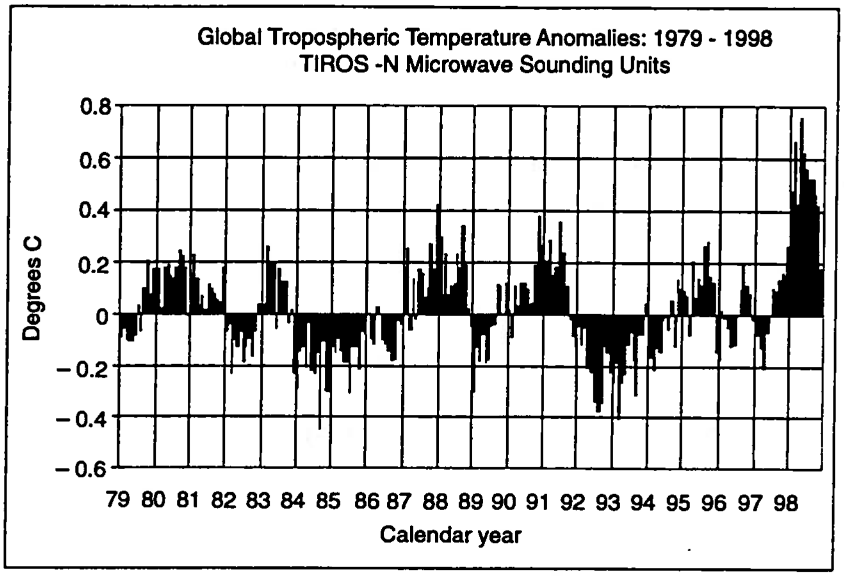
The symbiotic model/data relationship I have described shows why this 'impossible standard' would be fatally flawed even if it were not motivated primarily by anti-regulatory ideology and an unrealistic view of science. First, the model/data distinction, on which the 'sound science' standard is based, does not survive close scrutiny. *All* modern global data sets are computer-processed to some degree, even if only to reject spurious readings and interpolate grid-point values from irregularly spaced collection points. (See the discussion of satellite data sets above, and additional discussion below.) Second, the assumption that observed historical trends can be extrapolated linearly into the future *is itself a model.* In the case of a highly complex system like climate, this model is almost certainly among those least likely to be true, since it is not based on physical theory. Furthermore, it commits—in a much less defensible manner—the *same* logical fallacy of induction that makes it impossible ever to verify or validate more formal models. In other words, if observations cannot validate GCMs because other models might explain them equally well, even less can observations prove that a *linear* extrapolation of historical trends represents the most likely future outcome.

If the epistemological relationship between models and data is not one of priority, but one of symbiosis, then genuine 'sound science' must pursue them both.

Chapter 8 controversy

On 12 June 1996, just days after the formal release of the IPCC Second Assessment Report (SAR), the *Wall Street Journal* published an op-ed piece entitled 'A Major Deception on Global Warming'. The article, by the physicist Frederick Seitz, President Emeritus of Rockefeller University, accused some IPCC scientists of the most 'disturbing corruption of the peer-review process' he had ever witnessed (Seitz, 1996). Seitz's proclaimed distress stemmed from the fact that the lead authors of the SAR's Chapter 8 had altered some of its text *after* the November, 1995 plenary meeting of IPCC Working Group I (WGI), in Madrid, at which time the chapter was formally 'accepted' by the Working Group.

Chapter 8 deals directly with the question of whether models and data together can yet support conclusions about whether climate change is occurring ('detection') and how much of this change, if any, can be attributed to human activities ('attribution'). The chapter was the source of the statement quoted above that despite large remaining uncertainties, 'the balance of evidence suggests that there is a discernible human influence on global climate'. Quoting several sentences deleted from the final version of Chapter 8, Seitz argued



Microwave sounding unit (MSU) satellite data show a slight cooling of the lower atmosphere, along with disturbances such as the eruptions of El Chichon in 1983 and Mount Pinatubo in 1991. However, MSU data depend upon a model/data symbiosis.
Credit: Reproduced by permission of NASA/Marshall Space Flight Center, produced by Roy Spencer (NASA) and John Christy (University of Alabama in Huntsville), webpagehttp://wwwssl.msfc.nasa.gov/newhome/headlines/essd12jan99_1.htm.

that the changes and deletions 'remove[d] hints of the skepticism with which many scientists regard claims that human activities are having a major impact on climate in general and on global warming in particular'. According to Seitz, since the scientists and national governments who accepted Chapter 8 were never given the chance to review the truly final version, these changes amounted to deliberate fraud and 'corruption of the peer-review process'. Not only did this violate normal peer review procedure, Seitz charged; it also violated the IPCC's own procedural rules.

The *Wall Street Journal* op-ed set off a lengthy chain of exchanges lasting several months. The main participants in the public controversy were Seitz, Chapter 8 lead author Benjamin Santer, other Chapter 8 authors, the Co-Chairmen of the IPCC (Sir John Houghton and Bert Bolin), and climate-change sceptics S. Fred Singer and Hugh Ellsaesser. In a letter to the *Wall Street Journal*, Singer wrote that Chapter 8 had been 'tampered with for political purposes'. The IPCC, he claimed, was engaged in a 'crusade to provide a scientific cover for political action' (Singer, 1996). Semi-privately, in electronic mail exchanges involving many additional participants (and widely copied to others), the debate became intense and sometimes quite bitter. Both the public and the private exchanges themselves became objects of further press reports, covered in the scientific journal *Nature* (Masood, 1996) and widely disseminated by the news wire services.

A full discussion of this debate, which highlights the unusual character of the IPCC as a hybrid scientific/political organization, is beyond the scope of this paper (see Edwards and Schneider, forthcoming). Here, it serves to illustrate my contention that the epistemology of modelling is a central focus of climate politics. As it continued, the debate spread from the initial issues about peer review and IPCC procedure, to include questions about the validity of Chapter 8's scientific conclusions. Contrarians claimed that Chapter 8 dismissed or ignored important scientific results that raised doubts about the global warming hypothesis. They argued that the allegedly illegitimate changes to Chapter 8 made this problem even more acute.

The basis of many of the contrarian claims—both here and in the scientific integrity hearings—were the microwave sounding unit (MSU) satellite data. In their 'raw' form these data, collected since 1979, show a slight cooling of the troposphere (lower atmosphere), at an apparent rate of about –0.06°C per decade. They are contradicted by radiosonde data from the same period, which show an average warming trend of about 0.1°C per decade since the 1950s.

A closer look at the MSU data sets immediately reveals the model/data symbiosis I have been discussing. Satellites cannot read lower-atmosphere temperatures directly (i.e. by actual contact with the air). Instead, tropospheric temperature readings are created from top-of-atmosphere sensing records *by means of models* of atmospheric structure and chemical constituents. Sceptical opinions based on MSU records thus rely on data that are more model-bound than those taken from the radiosonde networks; one might have expected proponents of 'empirical' science to place their greatest trust in the direct instrument readings.

Furthermore, efforts to reconcile MSU and radiosonde data sets—an enterprise in which the climate-change contrarians do not seem to be interested, since it diminishes the basis of their claims—largely succeed 'if the diverse response of the troposphere and surface to short-term events such as volcanic eruptions and El Niño are taken into account'. (Both of these events occurred several times during the period of MSU records.) 'After adjustment for these transient effects, both the tropospheric and surface data show slight warming (about 0.1°C per decade for the troposphere and nearly 0.2°C at the surface) since 1979' (Houghton *et al.*, 1996, pp. 27–28). These adjustments are carried out with models. In addition, recent research suggests that instrument differences among satellites, orbital drift, and problems in the radiosonde data (caused by such factors as changes in launch site locations during the observation period) may account for much of the discrepancy (Hurrell and Trenberth, 1997).

Finally, reconciliation of various data sources is an absolute necessity for near-term climate science. The 18-year MSU record spans too short a period to make a useful comparison set for climate. To provide global data sets covering even the most minimally acceptable timespan, it must be integrated with other data sources. This too, as we have seen, requires modelling.

I am not suggesting that models and data are the same thing. I *am* saying that the distinction has often been strongly overstated, and unfounded conclusions based upon it. In climate science, at least, models and data are symbiotic. 'Raw' data are noisy, shapeless, and uninterpretable. Models are the skeleton that give them a definite form. Like flesh and bone, neither models nor data alone can support a living understanding of physical phenomena.

■ THE USES OF UNCERTAINTY

I have made two main arguments in this essay.

One is epistemological: that in global climate science (and perhaps in every model-based science), neither pure data nor pure models exist. Not only are data 'theory-laden'; models are 'data-laden' as well. Thus models and data are symbiotic. Models function as analogues to reality. They allow experimental methods to be applied to phenomena which cannot be studied using traditional laboratory techniques. They also allow the creation of coherent global data sets which could not exist without them.

The second argument is sociological: that to a large degree, the politics of climate to date have occurred at the interface between science and policymaking, and they have been centrally *about* the relationship of models to data. 'High-proof' participants in these debates have relied upon the apparent epistemic solidity of observation to attack the apparent epistemic fragility of models. Even 'friendly' philosophical critique has relied upon a sharp distinction between models and data in defining the limits of scientific knowledge.

The distinction is an epistemological error, but it is one which translates well into a political sphere where 'sound science' is construed as the priority of empirical science over theory and models. This vision of science meshes neatly with a cultivated political image of hard-headed realism and its corollary, the rejection of 'frontier' positions as fearful and reactionary. I have counterposed a more complex picture which is difficult to translate into partisan politics, for it acknowledges fundamental limits to scientific methods at the temporal and spatial scales of global change.

☐ *Uncertainty and 'social construction'*

Perception of these issues differs dramatically, depending largely on the particular perspective of the audience.

For scientists, the model/data relation I describe falls under the well-known category of 'scientific uncertainty'. Virtually all scientists recognize this as a major, legitimate issue. Most see the elimination of uncertainty as an asymptotically approachable goal, albeit never attainable.

But uncertainty has many meanings in science, and these receive unequal amounts of attention from scientists. There are empirical uncertainties, such as those that stem from sensor calibration (in data) or from 'physically-based' parameterization (in models). Other uncertainties stem, for example, from the limits on computer power which constrain model resolution. These kinds of uncertainties are quantifiable and (in principle) reducible. Scientists are comfortable with them and prepared to work with them.

Other types of uncertainty are less tractable. The Oreskes *et al.* critique of validation illustrates how a fundamental uncertainty arises from the logical limits of inductive reasoning. The complex model/

data relationship in climate science constitutes another *unquantifiable and irreducible* form of uncertainty. Empirical observation is the ultimate control by which science eliminates artefactual elements from theory. But since global models rely on embedded data (parameters), and global data must be filtered by models, this control is fundamentally limited: there will always be some indeterminate degree of artefactuality to each.

Scientists are understandably less comfortable with these *epistemological* uncertainties, perhaps because there is little they can do about them. Most scientific discussion of uncertainty concerns the quantifiable, potentially reducible sort. To the extent that epistemological uncertainties are discussed, they are often lumped together with other types; this makes them easy to forget. 'Uncertainty' can thus become a sort of conceptual garbage can into which untreatable problems are tossed.

To its great credit, the IPCC has in general avoided this path. As it has done with other aspects of uncertainty, it has taken on board the critique of validation (Oreskes *et al.*, 1994). Although I believe the IPCC bowed too easily to this simplistic critique, in doing so it at least rendered explicit its position on this aspect of the epistemology of modelling. Similarly, in Rep. George Brown Jr's report to the Democratic Caucus of the House Science committee on the scientific integrity hearings, he attempted to move political debate towards an understanding of the model/data relationship, urging that 'both observational evidence and theoretical models are essential to constructing an *understanding* of what is being observed. Neither in isolation is sufficient nor superior from an intellectual standpoint, as suggested by the Republican vision statement. ... An empirical extrapolation alone is subject to major uncertainties and misinterpretation ...' (Brown, 1996, Appendix section B.1a, italics in original). My arguments here support these positions and, elaborate the reasons behind them.

The perception of my analysis by social scientists is likely to be different. Many are likely to read my discussion as implying that climate science results are social constructs. Like 'uncertainty', the 'social construction' category has many meanings. At its core, the concept implies that resolutions of the model/data relationship and the frontier/high-proof debate are decided by social negotiation.

The key to this view is the undeniable fact that the physical reality of climate cannot determine a single correct understanding of these issues. Instead, social factors ranging from scientific tradition (training, laboratory culture, etc.) to the politics of funding (which shape the direction and amount of research) and the enrolment of allies (not only other scientists and scientific disciplines, but politicians, public opinion, etc.) affect which of the range of scientific opinions ends up dominant. This part of what natural scientists see as 'uncertainty' is understood by social scientists as an irreducibly social element in the generation of knowledge.

There is an important element of truth in this perspective. Social-construction theory predicts that scientists will *legitimately and necessarily* differ in their understanding of the model/data relationship and its importance, and my discussion of frontier vs high-proof opinion in climate science bears this out. It also predicts that extra-scientific groups (e.g. the House Energy and Environment Subcommittee and the Conferences of Parties to the Framework Convention on Climate Change) will play a part in negotiating this understanding, and that it can be made important not only within science, but in politics as well. This prediction too is borne out. The IPCC, for example, has bent over backwards to address the scientific concerns of the sceptics, in large part because of the political pressure brought to bear by industry lobby groups and affiliated scientists. The political controversies discussed above also illustrate this point. Finally, it predicts, in this case correctly, that such controversies can never be resolved once and for all. Under the right conditions, virtually any settled scientific controversy can be reopened later (Callon *et al.*, 1986; Collins and Pinch, 1993; Latour, 1987).

However, social and political factors do not necessarily *determine* scientific conclusions. I doubt that any serious scholar would defend this position, but the widely reported claims that some do, in the recent 'science wars' controversy, require that I make the point explicit (Gross and Leavitt, 1994; Ross, 1996).

The powerful mechanism of peer review, while far from omnipotent, vastly reduces political influence within science, in part by ensuring a constant cross-talk among scientists from different local cultures. Although it cannot prevent the shaping of research directions by political decisions about funding, peer control (e.g. at the National Science Foundation) has a strong influence even on this. Finally and most importantly, though data are model-bound, they still play by far the largest part in constraining the range of scientifically legitimate theories.

If correctly understood, the 'uncertainty' and 'social-construction' perspectives are not contradictory, but complementary. Uncertainty is inevitable in science, as in probably all aspects of human understanding of complex systems. It is a many-dimensional category which has both reducible/quantifiable (empirical, practical) and irreducible/unquantifiable (epistemological) elements.

Precisely *because* uncertainty—especially of the latter sort—cannot be eliminated, social construction *also* plays a fundamental and irreducible role in human knowledge. Reality constrains, but never fully determines, what we make of it. Scientific methods are the best ways we know to increase the degree to which the physical world constrains our understanding.

Yet uncertainty—of all sorts—makes inevitable a range of scientific opinions that can all be seen as legitimate and well-supported by current standards. Within this range, social construction processes fundamentally shape what counts as knowledge. They also shape the standards by which knowledge is judged, sometimes known as 'warranting' procedures (Shapin and Shaffer, 1985). This position has been aptly called 'realist constructivism' (Cole, 1992).

CONCLUSION

What practical conclusions can be drawn?

First, climate scientists should explicitly recognize the multiple forms of 'uncertainty', and they should avoid using the category to conceal unresolvable epistemological problems. As they have already done with the concept of validation, climate scientists should make explicit the fundamental ambiguities in the model/data relationship, rather than attempt to paper them over. At the same time, they should articulate the reasons why these ambiguities are not flaws which reliance on 'pure' observation could correct, but rather fundamental features of model-based science at the global scale.

Second, all parties engaged in climate science and politics should take careful note of how translations are made between scientific and political arenas. In politics, scientific uncertainty becomes a rhetorical resource which can and will be employed by different interests in different ways (Edwards, 1996; Shackley and Wynne, 1996). For opponents of immediate precautionary action, uncertainties in global models provide a time-worn rationale for shunting funds and attention back to basic research, or for denying any validity to climate change projections (Lempert *et al.*, 1994; Lindzen, 1990, 1992; Schlesinger and Jiang, 1991). Policymakers who want to delay precautionary action ally with high-proof scientists, holding out for very high degrees of empirical confirmation.

But this link can be a two-way street. Proponents of near-term action can *also* use uncertainties politically. They should argue that precisely because uncertainties can never be entirely eliminated, the choice of how much empirical confirmation is enough is ultimately a value choice most appropriately decided in the political arena—where decisions are always made under uncertainty. This strategy requires that models be presented not as predictions, but rather as heuristic projections or general forecasts about the likely direction and nature of global change (Edwards, 1996; Liverman, 1987; Oreskes *et al.*, 1994; Schneider, 1994).

If models are heuristic guides, then the political issue becomes what kind of bets to place. Should we centre our planning on the outcome viewed as most likely? To what degree should we plan for extreme, but relatively unlikely, predicted outcomes? These boil down to questions about how much risk a society is willing to take, and how much it is willing to pay to reduce it. This construction—rather than the caricature in which science appears as a source of final certainty—places science in its most valuable and responsible role: as a very important source of information which cannot and should not *by itself* determine policy.

Finally, non-scientist observers of climate science—myself included—should pay close attention to the role their own discussions can play in climate politics. For example, arguments of the radical social-constructivist type were employed by conservative forces in the 'scientific integrity' hearings. Both contrarian scientists and Republican politicians argued that a self-interested science 'establishment' enforces the acceptance of false theories:

> Citing historical instances where unconventional theories successfully overturned conventional wisdom, some Members

> and witnesses suggested that scientific 'truth' is usually more likely to be found at the scientific fringes than in the conventional center. As the Subcommittee chair stated, 'I am not swayed by arguments that "here's a big list of scientists that are on my side and you only have a smaller group of scientists on your side". I'm just not swayed by that at all.' A similar sentiment was echoed by the Chairman of the Science Committee: 'My experience in 18 years of watching science policy being made is it is [sic] often those small groups of scientists, though, who differ with conventional wisdom that in fact are always producing the scientific judgments of the future' (Brown, 1996, section II.B).

Galileo was cited as one example.

I believe this use of history and social theory is disingenuous. But I also believe that social scientists have sometimes been guilty of similar excesses. Hiding behind the conceit that we are merely describing what we see, without at the same time intervening, is an old ruse we need to give up (Hacking, 1983; Taylor, 1993). Instead, we need to take responsibility for evaluating, within the limits of our competence, the situations we describe.

In this essay, I have tried to demonstrate this practice by example. I did not simply describe the many problems with global climate models and leave readers to draw their own conclusions. Rather, I argued that the epistemology of modelling undermines the contrarian position that models should be subordinated to observation.

I also showed how better understandings of these problems could improve the quality of political debate, by transforming it from a battle over truth to a debate about how to act within uncertainty. This is a fundamentally more democratic way of thinking about a problem like climate change, since it demotes the scientist from seer (or charlatan) to expert advisor. Such understanding can help to build a more balanced relationship between the critical information which science *can* provide to support political decisions, and the value choices which it cannot make alone.

☐ *NOTES*

1. The oceans, too, absorb and hold heat; they play a major role in the overall climate system. Here, for the sake of brevity, I focus only on the atmosphere.

2. 'Zero-dimensional' models compute energy balances as if the Earth were a single point in space rather than a volume. One-dimensional models usually compute according to latitudinal bands (without vertical depth). Two-dimensional models add either longitude (horizontal east–west) or air pressure (vertical) dimensions to produce a grid (horizontal or vertical). Three-dimensional models extend the grid either longitudinally or vertically to produce a gridded volume—the atmosphere as a whole.
3. The designation 'three-dimensional' is slightly misleading. Most GCMs are really four-dimensional, the fourth dimension being time.
4. The most popular modern modelling technique, 'spectral' analysis, does not use grids in this simple sense. Spectral models represent the atmosphere as a series of interacting waves. They are mathematically complex and difficult to grasp intuitively. For my purposes, however, the simple description given here is adequate.
5. Since 1994 I have interviewed some 40 climate scientists, most of them modellers. The identities of the modellers quoted in this essay are confidential.
6. I am presently completing a 'family tree' of the major GCMs that traces their common heritage.
7. Long-term, contemporary data sets are not the only ones against which to test climate models. The seasonal cycle provides a well-known, reasonably well-understood benchmark. Paleoclimatic (prehistoric) data from a variety of 'proxy' sources, such as tree rings, ice cores, and fossilized pollen, are also available. Model inputs can be set to the different conditions (orbital precession, trace gas concentration, etc.) of past periods and evaluated by how well they simulate the paleoclimatic record. Naturally, the level of detail in paleoclimatic data is far lower than in contemporary instrumental observations.
8. Other groups, such as the Union of Concerned Scientists (UCS), have also adopted this phrase—with a quite different meaning—in an attempt to block Republican co-optation of the term. See the UCS Sound Science Initiative, available at http://www.ucsusa.org/.

☐ *REFERENCES*

Boer, G. J. (1992) 'Some results from and an intercomparison of the climates simulated by 14 atmospheric general circulation models', *Journal of Geophysical Research*, 97: 12771–86.

Bourke, W. *et al.* (1991) 'Systematic errors in extended range predictions', CAS/JSC Working Group on Numerical Experimentation, WCRP (WMO/TD-No. 444), World Meteorological Organization, Geneva. Report No. 16.

Brown, G. E., Rep. (1996) 'Environmental science under siege: fringe science and the 104th Congress', Report, Democratic Caucus of the Committee on Science, US House of Representatives. 23 October, Washington, DC. Available from http://www.house.gov/science_democrats/archive/envrpt96.htm.

Brunner, R. (1996) 'Policy and global change research: a modest proposal', *Climatic Change*, 32: 121–47.

Callon, M., Law, J. and Rip, A. (Eds) (1986) *Mapping the Dynamics of Science and Technology*. London: Macmillan.

Christy, J. R., Spencer, R. W. and McNider, R. T. (1995) 'Reducing noise in the MSU daily lower tropospheric global temperature data set', *Journal of Climate*, 8: 888–96.

Cole, S. (1992) *Making Science: Between Nature and Society*. Cambridge, MA: Harvard University Press.

Collins, H. and Pinch, T. (1993) *The Golem: What Everyone Should Know about Science*. Cambridge: Cambridge University Press.

Edwards, P. N. (1996) 'Global comprehensive models in politics and policymaking', *Climatic Change*, 32: 149–61.

Edwards, P. N. (in press) 'The world in a machine: origins and impacts of early computerized global systems models', in T. P. Hughes and A. C. Hughes (Eds), *Systems, Experts, and Computers*. Cambridge, MA: MIT Press.

Edwards, P. N. and Schneider, S. H. (forthcoming) 'Governance and peer review in science-for-policy: the case of the IPCC Second Assessment Report', in C. Miller and P. N. Edwards (Eds), *Changing the Atmosphere: Expert Knowledge and Global Environmental Governance*. Cambridge, MA: MIT Press.

European Center for Medium Range Weather Forecasts (1999) 'The ECMWF Re-Analysis (ERA) project', available from http://www.ecmwf.int/html/ERA/index.html.

Folland, C. K. and Parker, D. E. (1995) 'Correction of instrumental biases in historical sea surface temperature data', *Quarterly Journal of the Royal Meteorological Society*, 121: 319–67.

Galison, P. (1987) *How Experiments End*. Chicago: University of Chicago Press.

Gates, W. L. (1997) 'AMIP: the atmospheric model intercomparison project', PCMDI Report, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory. Report 7, December 1992 (updated March 1997).

Gross, P. and Leavitt, N. (1994) *Higher Superstition: The Academic Left and Its Quarrels with Science*. Baltimore: Johns Hopkins University Press.

Hacking, I. (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.

Hanson, N. R. (1958) *Patterns of Discovery*. Cambridge: Cambridge University Press.

Hays, S. P. (1987) *Beauty, Health, and Permanence: Environmental Politics in the United States, 1955–1985*. New York: Cambridge University Press.

Hays, S. P. (1989) 'Three decades of environmental politics: the historical context', in M. J. Lacey (Ed.), *Government and Environmental Politics*. Washington, DC: Woodrow Wilson Center Press.

Hide, R. (1953) 'Some experiments on thermal convection in a rotating liquid', *Quarterly Journal of the Royal Meteorological Society*, 79: 161.

Houghton, J. J. *et al.* (Eds) (1996) *Climate Change 1995: The Science of Climate Change*. 3 vols. Vol. 1. Cambridge, UK: Cambridge University Press.

Hume, D. (1977 [1748]) *An Enquiry Concerning Human Understanding*. Indianapolis, IN: Hackett Publishing Co.

Hurrell, J. W. and Trenberth, K. E. (1997) 'Spurious trends in satellite MSU temperatures from merging different satellite records', *Nature*, 386 (6621): 164–6.

Jasanoff, S. (1990) *The Fifth Branch: Science Advisors as Policymakers*. Cambridge, MA: Harvard University Press.

Jasanoff, S. (1991a) 'Acceptable evidence in a pluralistic society', in D. G. Mayo and R. D. Hollander (Eds), *Acceptable Evidence: Science and Values in Risk Management*. New York: Oxford University Press.

Jasanoff, S. (1991b) 'Science, politics, and the renegotiation of expertise at EPA', *Osiris*, 7: 195–217.

Johnson, G. L. *et al.* (1994) 'Estimating urban temperature bias using polar-orbiting satellite data', *Journal of Applied Meteorology*, 33: 358–69.

Karl, T. R., Knight, R. W. and Christy, J. R. (1994) 'Global and hemispheric temperature trends: uncertainties related to inadequate spatial sampling', *Journal of Climate* 7 (7): 1144–63.

Kerr, R. A. (1997) 'Climate change: model gets it right—without fudge factors', *Science*, 276 (5315): 1041.

Kiehl, J. T. (1992) 'Atmospheric general circulation modeling', in K. E. Trenberth (Ed.), *Climate System Modeling*. Cambridge: Cambridge University Press.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago.

Lakatos, I. and Musgrave, A. (Eds) (1970) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Latour, B. (1987) *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Latour, B. and Woolgar, S. (1979) *Laboratory Life: The Social Construction of Scientific Facts*. London: Sage.

Lempert, R., Schlesinger, M. and Hammitt, J. (1994) 'The impact of potential abrupt climate changes on near-term policy choices', *Climatic Change*, 26: 351–76.

Lindzen, R. E. (1992) 'Global warming: the origin and nature of the alleged scientific consensus', *Regulation—Cato Review of Business & Government*, Spring: 87–98.

Lindzen, R. S. (1990) 'Some coolness concerning global warming', *Bulletin of the American Meteorological Society*, 71 (3): 288–99.

Lindzen, R. S. (1996) 'Testimony before the Committee on Science, Hearing on U.S. Global Change Research Programs: Data Collection and Scientific Priorities', US House of Representatives. 6 March, Washington, DC.

Liverman, D. (1987) 'Forecasting the impact of climate on food systems: model testing and model linkage', *Climatic Change*, 11: 267–85.

Masood, E. (1996) 'Climate report "subject to scientific cleansing"', *Nature*, 381 (6583): 546.

Meehl, G. A. (1992) 'Global coupled models: atmosphere, ocean, sea ice', in K. E. Trenberth (Ed.), *Climate System Modeling*. Cambridge: Cambridge University Press.

National Oceanic and Atmospheric Administration (1999) 'NCEP/NCAR CDAS/ Reanalysis Project', available from http://wesley.wwb.noaa.gov/reanalysis.html#intro.

Norton, S. D. and Suppe, F. (forthcoming) 'Epistemology of atmospheric modeling', in C. Miller and P. N. Edwards (Eds), *Changing the Atmosphere: Expert Knowledge and Global Environmental Governance*. Cambridge, MA: MIT Press.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994) 'Verification, validation, and confirmation of numerical models in the Earth sciences', *Science*, 263: 641–6.

Parker, D. E. and Cox, D. I. (1995) 'Towards a consistent global climatological rawinsonde data-base', *International Journal of Climatology*, 15: 473–96.

Popper, S. K. R. (1959 [1934]) *The Logic of Scientific Discovery*. New York: Basic Books.

Popper, S. K. R. (1962) *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books.

Ross, A. (1996) 'Science wars', *Social Text*, 46–47: 1–252.

Rotmans, J. (1990) *IMAGE: An Integrated Model to Assess the Greenhouse Effect*. Boston: Kluwer Academic Publishers.

Santer, B. D. *et al.* (1996) 'A search for human influences on the thermal structure of the atmosphere', *Nature*, 382 (6586): 39–46.

Santer, B. D., Wigley, T. M. L. and Jones, P. D. (1993) 'Correlation methods in fingerprint detection studies', *Climate Dynamics*, 8: 265–76.

Schlesinger, M. and Jiang, X. (1991) 'A phased-in approach to greenhouse-gas-induced climatic change and climatic response to increasing greenhouse gases', *Eos*, 72 (53): 596–7.

Schneider, S. H. (1989) *Global Warming: Are We Entering the Greenhouse Century?* New York: Vintage Books.

Schneider, S. H. (1992) 'Introduction to climate modeling', in K. E. Trenberth (Ed.), *Climate System Modeling*. Cambridge: Cambridge University Press.

Schneider, S. H. (1994) 'Detecting climatic change signals: are there any "fingerprints"?', *Science*, 21 January: 341–7.

Seitz, F. (1996) 'A major deception on global warming', *Wall Street Journal*, 12 June.

Shackley, S. *et al.* (1998) 'Uncertainty, complexity and concepts of good science in climate change modelling: are GCMs the best tools?', *Climatic Change*, 38: 159–205.

Shackley, S. and Wynne, B. (1996) 'Representing uncertainty in global climate change science and policy: boundary-ordering devices and authority', *Science, Technology and Human Values*, 21 (3): 275–302.

Shapin, S. and Shaffer, S. (1985) *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton: Princeton University Press.

Singer, S. F. (1996) 'Letter to the Editor', *Wall Street Journal*, 11 July.

Sismondo, S. (forthcoming) 'Models, simulations, and their objects', *Science in Context* 12 (2).

Spencer, R., Christy, J. R. and Horack, J. (1997) 'Global climate monitoring: the accuracy of satellite data', NASA Marshall Space Sciences Laboratory. 12 March, available from http://wwwssl.msfc.nasa.gov/newhome/headlines/essd12mar97_1.htm.

Taylor, P. (1993) 'Building on construction: an exploration of heterogeneous constructionism, using an analogy from psychology and a sketch from socio-economic modeling', *Perspectives on Science*, 3 (1): 66–98.

US Global Change Research Program (1998) *Our Changing Planet: the FY 1999 US Global Change Research Program*. Washington, DC: Global Change Research Information Offices. Available from http://www.gcrio.org/ocp00/toc.html.

Wigley, T. M. L., Pearman, G. I. and Kelly, P. M. (1992) 'Indices and indicators of climate change: issues of detection, validation, and climate sensitivity', in I. M. Mintzer (Ed.), *The Science of Climate Change*. Cambridge: Cambridge University Press.

Winsberg, E. (forthcoming) 'Sanctioning models: the epistemology of simulation', *Science in Context*, 12 (2).

Wirth, T. E. (1996) 'Statement on behalf of the United States of America', Second Conference of Parties to the Framework Convention on Climate Change, 17 July. Geneva, Switzerland.

World Climate Research Programme Working Group on Numerical Experimentation (1999) 'Atmospheric model intercomparison project', Lawrence Livermore National Laboratories, Program for Climate Model Diagnosis and Intercomparison. Available from http://www-pcmdi.llnl.gov/amip/.

Yanai, M., Esbensen, S. and Chu, J.-H. (1973) 'Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets', *Journal of Atmospheric Science*, 30: 611–27.