Uncertainty, Complexity and Concepts of Good Science in Climate Change Modeling: Are GCMs the Best Tools?

ARTICLE *in* CLIMATIC CHANGE · JANUARY 1998

Impact Factor: 3.43 · DOI: 10.1023/A:1005310109968

4 AUTHORS, INCLUDING:

Simon Shackley
The University of Edinburgh

Peter C. Young
Lancaster University UK and Australian Nat…

Brian Wynne
Lancaster University

UNCERTAINTY, COMPLEXITY AND CONCEPTS OF GOOD SCIENCE IN CLIMATE CHANGE MODELLING: ARE GCMs THE BEST TOOLS?

SIMON SHACKLEY[1], PETER YOUNG[2], STUART PARKINSON[2] and BRIAN WYNNE[1]

[1]*Centre for the Study of Environmental Change (CSEC) Bowland Tower East, Lancaster University, Lancaster, LA1 4YN, U.K.*

[2]*Centre for Research into Environmental System (CRES), Institute of Environmental and Biological Sciences, Lancaster University, LAI 4YQ, U.K.*

Abstract. In this paper we explore the dominant position of a particular style of scientific modelling in the provision of policy-relevant scientific knowledge on future climate change. We describe how the apical position of General Circulation Models (GCMs) appears to follow 'logically' both from conventional understandings of scientific representation and the use of knowledge, so acquired, in decision-making. We argue, however, that both of these particular understandings are contestable. In addition to questioning their current policy-usefulness, we draw upon existing analyses of GCMs which discuss model trade-offs, errors, and the effects of parameterisations, to raise questions about the validity of the conception of complexity in conventional accounts. An alternative approach to modelling, incorporating concepts of uncertainty, is discussed, and an illustrative example given for the case of the global carbon cycle.

In then addressing the question of how GCMs have come to occupy their dominant position, we argue that the development of global climate change science *and* global environmental 'management' frameworks occurs concurrently and in a mutually supportive fashion, so uniting GCMs and environmental policy developments in certain industrialised nations and international organisations. The more basic questions about what kinds of commitments to theories of knowledge underpin different models of 'complexity' as a normative principle of 'good science' are concealed in this mutual reinforcement. Additionally, a rather technocratic policy orientation to climate change may be supported by such science, even though it involves political choices which deserve to be more widely debated.

> 'The multiplicity of models is imposed by the contradictory demands of a complex, heterogeneous nature and a mind that can only cope with few variables at a time; by the contradictory desiderata of generality, realism, and precision; by the need to understand and also to control; even by the opposing esthetic standards which emphasize the stark simplicity and power of a general theorem as against the richness and the diversity of living nature. These conflicts are irreconcilable' (Richard Levins) [1].

1. Introduction

The information available on climate change for policy making purposes is plagued by large inherent uncertainties. This includes uncertainties in the climate change models, as well as in the models of climate change impacts, economic costs and policy responses. It also includes those uncertainties which affect the social and

Climatic Change **38:** 159–205, 1998.

policy contexts into which such knowledge is intended to be used. In such situations of endemic social and political uncertainty, policy analysts have frequently recommended a strategy of *pluralism*. They have reasoned that a diversity of policy approaches should be adopted to increase the likelihood that at least some of these will prove useful and successful despite the uncertainty that characterises both the present and future in this most challenging area of study [2]. This argument is, however, rarely extended to the provision of scientific knowledge itself. That is because objective criteria for defining ‘good science’ are widely held to be available which can be used to ‘sort out’ competing approaches to the production of new scientific knowledge, such that only the most successful and promising are supported. Thus consolidation around General Circulation Models (GCMs) in much contemporary climate change research is explained by the higher rating that the climate science community gives to GCMs in comparison with alternative approaches.

In this paper, we present a different, less widely known, viewpoint according to which the criteria defining good science for policy are not solely derived from science, with its currently cherished paradigms, but also incorporate social and policy judgements. Little has been written about the role of such non-scientific judgements in the provision of climate science for policy [3]. Past studies of environmental policy have suggested that the successful reception and implementation of environmental policy is dependent upon a coalition of assorted actors (such as scientists, policy makers, environmentalists and industrialists), often with divergent expectations and agendas, yet who come together for the specific support of certain policy measures (albeit it sometimes for different reasons) [4]. Whilst shared scientific knowledge and understanding is frequently important, such studies suggest also the major role of shared social commitments about the issue and its solution. This insight is important in two respects: it disputes the impression, still found amongst some, that science by itself is sufficient to determine environmental policy; and it suggests that, for policy determination purposes, a comparison of competing interpretations of knowledge might come to include more than purely scientific considerations.

We suggest that although GCMs are widely considered to be the ‘best science’ for the study of future climate change, this judgement is significantly influenced by factors which are not purely scientific. For the following reasons, reflection on these non-scientific factors which, we argue, also support their dominant position, is important to the overall resilience of the scientific research base for global environmental change policy making.

- Firstly, it is easier to argue a case, say, for research funding, if all the reasons in support of that case are fully appreciated.
- Secondly, the validity of non-scientific elements, especially when used in contexts of application different from those in which the criteria were originally developed, can be better appreciated and evaluated.

- Thirdly, research policy can be made more resilient in the face of changes in the social and policy worlds if it takes into account the role of social and policy factors in contributing to notions of ‘good science’.
- Fourthly, without the critical scrutiny of non-scientific elements, there is a danger that climate change science will unconsciously come to support a technocratic approach to policy which always puts its main emphasis on an instrumental management and control of the environment and society. While this eventuality may be sensible, depending upon the circumstances, it always needs to be considered carefully, and more widely debated, in terms of its possibility, negative consequences and desirability, as a way of relating to the environment and to each other.

But the importance of non-scientific elements is not the only factor to be considered when questioning the role of GCMs in the study of future climate change. Another important dimension of our argument is that there are other scientific approaches to the study of climate change which are more robust and promising than their current (low) standing at the science-policy interface suggests. We argue that the relative lack of attention to alternatives is caused, in part, through their apparent weaker connection to policy and social goals and agendas (than GCMs). But these social and policy goals are not necessarily sovereign or unchanging. Hence, analysis of the non-scientific elements involved in defining ‘good science’ enables a better appreciation of alternative policy *and* scientific approaches and argues for more plurality in science. Our position is that, while GCMs clearly have an important role to play, other scientific alternatives should be further developed for possible use at the science-policy interface so as to support and complement the information derived from GCMs.

In Section 2, we support our claim that GCMs are presently dominant in producing knowledge of possible future climate change. Sections 3 and 4 both critically examine the conventional set of reasons in support of GCMs as being dominant, and find them questionable. In Section 5 we provide a brief illustrative example of an alternative modelling strategy based on different epistemic and methodological commitments to representing complexity and uncertainty. This serves to reinforce the point that more discerning scientific and other criteria should be involved in defining ‘good science’ in the climate change domain, and these are analysed more systematically in Section 6.

2. The Climate Modelling Pyramid

This section briefly considers the current and emerging future role of GCMs in climate change research, using the common metaphor of the climate modelling pyramid, and clarifies the criteria by which models are ranked by climate scientists.

2.1. THE ROLE OF GCMS IN CLIMATE RESEARCH

One of the most prominent climate change research activities has been the development and application of General Circulation Models (GCMs) of the global climate. GCMs have been used to simulate the consequences of an increase in atmospheric CO_2 on the mean global climate, both in instantaneous doubling of CO_2 equilibrium simulations, and time-dependent simulations in which the CO_2 concentration is increased incrementally over a number of model years. Coupling of atmospheric GCMs (AGCMs) with 3-D dynamical ocean models (OGCMs), together with more or less complex models of the land surface vegetation and sea-ice, represents the present 'state of the art' in such scientific research. Because of their low resolution and large errors in the simulation of regional feedbacks, however, GCMs cannot presently provide detailed *regional* climate simulations or estimates of future regional climate change. Nevertheless, it is assumed by many GCM modellers ('GCMers') that the regional errors do not compromise the validity of the global response to CO_2 forcing. This is because, provided the major, large-scale features are simulated and energy at the top of the atmosphere (TOA) is balanced, then the model will produce a new global equilibrium position which is largely independent of what occurs regionally.

Support for this latter view arises from the reasonable level of agreement between the majority of GCMs in terms of climate sensitivity, despite the wide divergences in their different simulations of regional climate change [5]. Many climatic feedbacks, especially ones related to atmospheric chemistry and biological processes, are not currently included in GCMs and much current work at the interface of climatology, geophysics, environmental physics and ecology is aimed at understanding such feedbacks and addressing how to include them in GCMs [6].

The development and application of GCMs is not only a major component of the climate change research field in financial terms; there is also a powerful logic in their role as the 'lynch-pin' of such a growing area of research. This arises from the scientific and policy consensus that GCMs provide more reliable scientific inference of future climate change than any other model or scientific methodology currently available. And, given that computer models have been highly significant in providing 'evidence' for the enhanced greenhouse effect, this puts GCMs at the crux of the current scientific mission to provide reliable scenarios of future climate change. 'Reliable' in this context means two things: firstly, that the proponents of GCMs consider that they represent the current climate system more realistically than other models; and secondly, because of this, that GCMs are more likely than other models to incorporate accurately those climatic processes which will be important in simulating the climatic changes ensuing from the atmospheric accumulation of greenhouse gases (as well as from other changes in radiative forcing).

2.2. THE RELATIVE RANKING OF CLIMATE MODELS

GCMs are widely considered by climate scientists to represent the climate system more realistically than other models, such as the 0, 1 or 2 dimensional energy balance models, radiative-convective models and statistical-dynamical models. They do so, according to Shine and Henderson-Sellers, because of 'the relative merit and complexity with which these four [basic climate] elements [of dynamics, radiation, surface processes and resolution] are included' [7]. These authors consider that different climate models can be positioned in a pyramidal space, as shown in Figure 1, according to the completeness with which the basic elements are included. GCMs emerge as the most integrated of models, even if: 'they do not ... incorporate all aspects of the climate system and are, therefore, not at the absolute apex of the pyramid' [8]. Greater realism is defined by Shine and Henderson-Sellers, therefore, in terms of the complexity of the *interaction* of the primary processes at the highest possible resolution, rather than the complexity *per se* with which the primary processes are represented.

Thus one-dimensional radiative-convective models (RCMs), which treat the atmosphere as a one-dimensional column, sometimes have a more sophisticated handling of radiation and other physical parameterisations than do GCMs. And Wigley and Raper's one-dimensional STUGE energy balance model included the effects of CO_2 fertilisation, feedback from stratospheric ozone depletion and the radiative effects of sulphate aerosols; whereas contemporaneous GCMs did not include these key climatic processes [9].

Shine and Henderson-Sellers remark that: 'The apex [of the modelling pyramid] suggests that when all facets are correctly and adequately incorporated at a high enough resolution a model, *presumably identical to the real climate*, results' [10]. Hence, the complexity of interaction of the dynamics and physical parameterisations, properly formulated, is held to generate realism (and these criteria can be used to sort out the adequacy of different models).

The hierarchy of the climate modelling pyramid is legion. In the volume *Climate System Modeling*, for example, Schneider notes that: 'Many scientists believe that the ultimate goal of climate modeling should be fully comprehensive, three-dimensional models of all elements of the climate system including very high resolution and as much detail as possible' [11]. In the same volume, Kiehl organises a discussion of AGCMs around the model hierarchy, defining model superiority in terms of the extent to which models avoid spatial averaging of the basic hydro- and thermodynamic equations for the atmosphere: more averaging is clearly equivalent to greater approximation and less realism [12].

GCMers readily admit to the omissions of some climate-related feedbacks in their current models and, for this and other reasons, opinion is divided on how distant GCMs are from the pyramid's apex. What is widely agreed upon, however, is that the key, first-order climate feedbacks *are* incorporated, and that feedbacks can: 'be predicted credibly only by physically-based models that include the essential

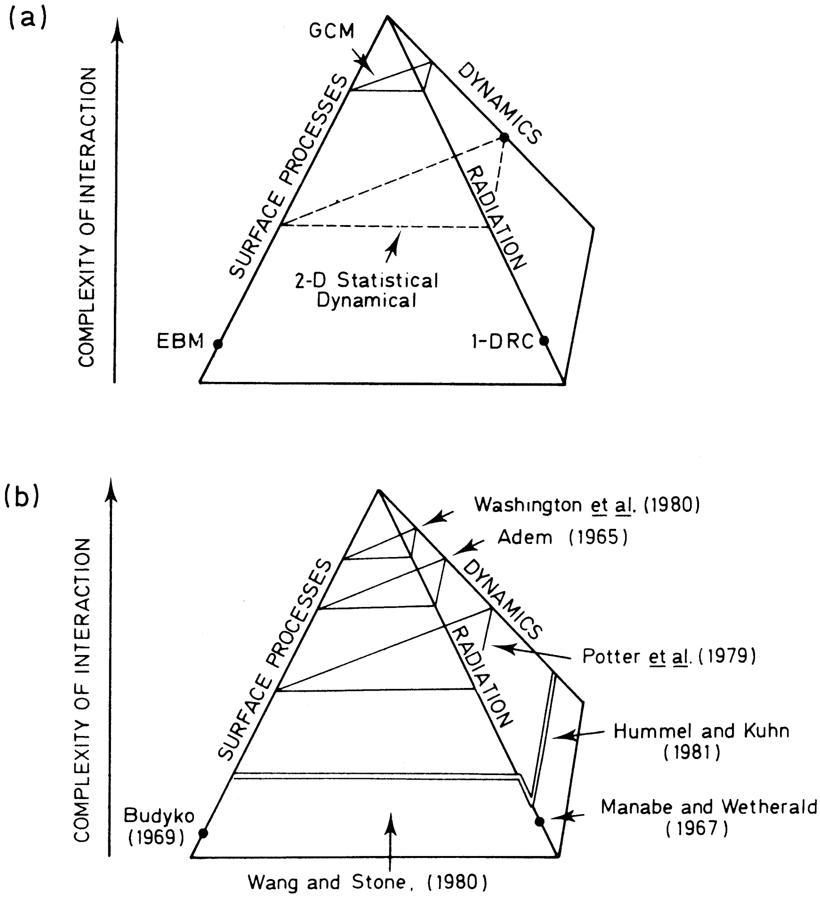

Figure 1. The climate-modelling pyramid. The position of a model on the pyramid indicates the complexity with which the three primary processes interact. The base of the pyramid can be considered hollow since there is essentially no interaction between the primary processes. Progression up the pyramid leads to greater interaction between each primary process. The vertical axis is not intended to be quantitative. (a) The positions of the four basic model types. (b) Particular climate models are seen to be based upon different methods of incorporating the primary processes and the level of complexity of the interactions (after Shine and Henderson-Sellers, 1983). (From Henderson-Sellers and McGuffie, 1987 [8]; Figure 2.1, page 37.)

dynamics and thermodynamics of the feedback processes. Such physically-based models are the general circulation models' [13]. In semi-public contexts, the physical basis of GCMs, compared to other methods, is a frequently heard claim for their superiority (though a number of physical laws, such as conservation of energy, are also incorporated into most simple climate models). When modellers are communicating to non-scientists – for example to journalists, during television broadcasts, and in front of Congressional Committees – it is not uncommon for them to respond to scepticism by pointing to the fundamental physical laws underpinning GCMs [14].

Our interviews with climate modellers provide further support for the perceived superiority of GCMs compared to other climate models. One modeller discussed the findings of the first transient, coupled AOGCM runs in the early 1990s compared to earlier, simpler models: ‘the box-diffusion model [used in the Intergovernmental Panel on Climate Change 1990 assessment] we couldn’t take too seriously. There are things you could anticipate but couldn’t prove without the coupled AOGCM model’ [15]. Hence AOGCMs provided *proof* of climatic change, whereas the one-dimensional model only indicated a possible effect. For another GCM modeller:

> ‘Simpler models are useful for isolating and illustrating points in an easily understandable manner. . . . They are essential for understanding but useless for prediction’ [16].

Yet another modeller, who has worked extensively with a range of climate models including GCMs, described how objections to the use of simple models emerged in the peer review process, both at the proposal and paper submission stages [17].

2.3. THE PERCEIVED ROLE OF SIMPLE MODELS

The existence of a model hierarchy does not imply that simple models are not valued within the GCM community. On the contrary, they are typically regarded as extremely useful tools in model development and testing [18]. Simpler models are portrayed as useful *heuristics* to generate insights about the processes involved in climate change, and to aid further understanding for teaching or policy purposes. But their scientific sophistication is inevitably compromised by the requirements imposed by such teaching, policy or open-ended exploration requirements, so that they are not seen as credible simulations of reality as GCMs are claimed, or aim, to be. In the dominant view, completeness and accuracy can only be traded-off against tractability and economy in the *learning stages* of the model development process, not when realism is sought. Often, the same judgement is made by many users and advocates of the simpler models, who routinely calibrate the sensitivity of simpler models to CO_2-doubling by using the simulations of GCMs. As one modeller, who normally uses simpler climate models, put it:

> ‘. . . ultimately one would have to say that [GCM development] is perhaps the only way, in the end, that we’re going to generate better answers ultimately because, in principle, the way they [GCMs] are constructed and conceived enables them to take on board all components of the climate system in a dynamic way; in a physically realistic way and with a capability for looking at feedbacks. The fact that they haven’t done so to date doesn’t mean that they are incapable of doing so. In terms of intellectual progress, then ultimately that is going to be probably the route that is most successful’ [19].

The use of the words ‘ultimately’ and ‘in principle’, and the general tone of the above passage, suggests that GCMs are assessed, to an important extent, by their *future* potential to realistically represent the interaction of primary processes.

Whilst the perspective of Meehl, Henderson-Sellers, and Kiehl is reasonably typical, there are other climatologists and modellers who take a more pluralist stance, advocating the validity of different models from across the hierarchy, or a more pragmatic, instrumental position towards climate models, in which, as one climatologist put it: ‘Different models are valuable for different purposes. At the end of the day you can’t say that one is better than the other’ [20]. Certainly, it is this view that models should be ‘objective-oriented’, that is they should be developed and utilised in different ways for different purposes, that has most appeal to the present authors.

2.4. REASONS IN SUPPORT OF THE CLIMATE MODELLING PYRAMID

The climate system is widely perceived as being complex, not only in terms of the characteristics of the processes involved (such as their non-linearity and the level of interconnection between sub-systems) but also with respect to their sheer range: i.e., the broad scope of interactive physical, chemical and biological processes in earth systems which influence climate. It is also widely accepted that this level of complexity is far beyond the power of any present climate models.

One dominant response to these arguments is the inclusion of ever more detail so as to increase the perceived accuracy of the model; i.e., to meet complexity in nature with model complexity, as a way of improving understanding of the interactive dynamical physical processes, as well as for thorough testing of the model’s sensitivity to certain processes and possible policy scenarios. This strategy assumes that a more complex model may lay claim to be a better representation of a complex system and, hence, that it has a greater truth-content than other models. From this standpoint, redundancy in model detail is assumed to be a very long way off.

Thus, the complexity of model representation becomes a central normative principle in evaluating ‘good science’ in the climate research domain. It is argued that the predictive potential of the model is enhanced; the uncertainty reduced or better defined; and, as a further bonus, the policy-usefulness of the model output is strengthened relative to simpler models. It is this ‘central dogma’ – that greater complexity equals greater realism, equals greater policy-utility – which we set out to explore critically in this paper.

3. Challenges to the Climate Modelling Pyramid I: Applying the Scientific Criteria of Evaluation

A common criticism of pluralism is to argue that not all options are as good as one another and that there are objective scientific criteria of evaluation which sort out the adequacy of the various model options. Two such criteria advocated by many climate modellers, which emerge from the discussion in Section 2, are:

- to what extent is the model physically-based?; and,
- what is the complexity with which the physical and dynamical processes, and especially their interactions, have been modelled?

We share the general perspective that different options should be critically evaluated, but question whether the above criteria alone can be used to differentiate between climate models. For this set of evaluative scientific criteria to be credible in explaining the dominance of GCMs, they would clearly need to be demonstrably applicable and widely accepted amongst the peer-group. Otherwise who decides what the criteria should be, how these might change as model applications change, and how they are applied and interpreted? In this section, some of the practical difficulties of applying criteria for climate model evaluation are discussed, using examples from the published literature.

3.1. PARAMETERISATIONS IN GCMS

Many climate phenomena occur on smaller scales than explicitly resolved in the three-dimensional grid of a GCM (which is typically several hundred kilometres in the horizontal directions). Parameter representations, or parameterisations, are used in GCMs to represent physical processes, such as clouds, collectively at the sub-grid scale. Parameterisations often consist of a statistical relationship (based on climatological data) between variables that are resolved at the grid scale and ones that are not, and which can then be used to simulate the unresolved physical process. The modeller uses his or her judgement in developing a parameterisation, based on physical theory and understanding, empirical data, intuition, other facets of the model, and so on. Parameterisations have frequently been presented more publicly as a problem involving a lack of sufficient detail in the model; a problem which can be partially solved by increasing model resolution [21]. However, this is not the view of other modellers, especially when writing in more specialist contexts. For example, Stephen Schneider comments that:

> ‘Modelers, of course, strive to keep their parameterizations as physical and non-empirical and scale-independent as practical. Thus, the validity of parameterization and overall model performance as well, depends ultimately on empirical tests, not only on the inclusiveness of the first principles. In other words, even our most sophisticated “first principles” models contain ‘empirical statistical’ elements within the model structure’ [22].

In reality, the key parameterisations, whilst they are usually based partially on the physical understanding of the scientist, are rarely fully theoretically-based. As a result, scale-dependence and some degree of arbitrariness is the rule rather than the exception [23]. Moreover, whilst the common portrayals of GCMs in public contexts present parameterisations as introducing an uncertainty at the sub-grid scale, hence making simulation at that level alone suspect, their effect also filters through at the resolved-scale level. As one modeller put it in a letter to us: ‘Since the simulation of the whole model is determined by the fidelity of the parameterizations, a climate model as a whole is not deterministic’ [24].

There is an equivocal use of the word ‘deterministic’ here: each run of the GCM is clearly deterministic in the usual scientific meaning of the word, since the effects of uncertainty are not accounted for. If uncertainty prevails in the thinking and practice of the GCM modellers and different parameterisations, or scenarios, are utilised to reflect this uncertainty, then all well and good. But such procedures are certainly not rigorous exercises in uncertainty analysis and it is determinism that permeates the overall process of the mathematical analysis.

It would seem, therefore, that even if one accepts that the physical basis of a model is the key criterion by which to assess its adequacy, GCMs do not rate quite so highly as their frequent, semi-public representation indicates. Although they might rate more highly than other climate models in terms of the attempted physical representation of interactive processes, this does not correspond with physical realism in any straightforward way for two reasons. Firstly, as already noted, particular physical processes may be more completely physically-based in simpler models. Secondly, a reductionist approach to the representation of physical processes, as in GCMs, does not automatically guarantee physical realism [25].

3.2. PARAMETERISATIONS AND MODEL INTERACTIONS

Stone and Risbey analysed the meridional flux of heat in the atmosphere simulated by three AGCMs (versions of the GFDL, GISS and NCAR AGCMs, descriptions of which were published in the early to mid-1980s) [26]. It turned out that such fluxes were poorly related to observations (with errors of typically 50%), notwithstanding the models’ successful simulation of observed surface air temperatures (SATs). If a more ‘realistic’ meridional heat transport is used to constrain the models, however, the SATs then become very inaccurate.

This raises the question as to whether it is best to maximise the simulation of either the SATs or the meridional heat flux; or whether to find a compromise between such objectives. Some modellers have argued that, since the observational record of meridional heat fluxes is inferior to that of SATs, discrepancy of the model with the former is a less reliable measure of its accuracy. From a statistical standpoint, there is no doubt that these differences in measurement uncertainty should be taken into account when modelling on the basis of such observations. Yet Stone and Risbey rebuke modellers for always assuming that it is the observations which are at fault rather than the structure and parameterisations within the models, without providing detailed reasons for this decision.

For example, many GCMers have stated a preference for the calculated meridional heat transports of the European Centre for Medium Range Weather Forecasting’s (ECMWF’s) model over the observations of the rawinsonde network analysed by Oort; where the latter shows greater discrepancy from the GCMs than the former [27]. The ECMWF's approach is based on operational analysis, i.e., one in which the model's simulation and available observational data are merged. The implicit argument of the GCMers has been that the ECMWF's data are superior because the model generates reliable data where there are no observations. But this view can be questioned in two ways: firstly, it is effectively using a model to validate a model; secondly, in a paper published in 1978, Oort actually tested the sensitivity of the heat transport to the observational rawinsonde network by supposing that the model was, in fact, a representation of reality, interpolating what would be the energy fluxes at the specific positions of the rawinsonde network, and comparing the heat transport implied by these few points to that of the fully resolved model. He found that the current observational network was, in fact, adequate (this being because the processes by which heat is transported are large-scale).

Stone and Risbey also analysed the sub-grid scale vertical flux (associated with moist convection and boundary-layer processes) which largely determines the vertical profile of temperature. To test the interactions between the sub-grid scale parameterisations and the vertical temperature profile in the absence of observational data on vertical heat fluxes, Stone and Risbey performed a sensitivity analysis by comparing the vertical dynamical heat fluxes in the GISS Model 1 using two different convection schemes: a moist adiabatic adjustment scheme and a penetrative convection scheme. As expected, the results were completely different for the small-scale convective flux, since the latter is a function of the parameterisation. Yet the large-scale flux was *also* different for the two parameterisations (50% smaller when the penetrative convection scheme was used). According to Stone and Risbey this finding is significant since:

> 'If, then, current GCMs do have an advantage (over other models) in simulating vertical dynamical heat fluxes, it must be because they can explicitly and accurately calculate the vertical fluxes associated with large scale motions. . . . GCMs' large scale heat fluxes interact so strongly with the parameterized subgrid scale physics that their calculations of the large scale fluxes are no more accurate than their (uncertain) calculation of the subgrid scale physics'.

And they conclude:

> '. . . *we doubt that the GCMs currently in use for climate change experiments are better at simulating global scale climate feedbacks and temperature changes than simpler models that do not simulate large-scale processes from first principles'* [28].

This is a very important caveat, for it recognises *that models should be related to the scale of the processes involved*: the reductionist argument that large scale behaviour can be represented by the aggregative effects of smaller scale process has never been validated in the context of natural environmental systems and is even difficult to justify when modelling complex man-made processes in engineering [29].

3.2.1. *Response from GCMers*

One GCM modeller reacted to Stone and Risbey's result in the following way:

> 'That is something you can argue either way ... say "well if the convective scheme is parameterising what it says it is simulating, the precipitation for different schemes should be the same" ... but I don't think we would necessarily expect two convection schemes to produce the same amount because they inevitably interact with large-scale processes ... What you're trying to do is to get the net effect to be right' [30].

In other words, for this modeller, the convection schemes should be evaluated by assessing how well they simulate global patterns of precipitation: this being the net result of many processes (both sub-grid *and* large scale processes) acting in different directions. The above modeller went on (though perhaps somewhat tongue-in-cheek) to explain how he thought another, world-famous, GCM modeller might react to Stone and Risbey's finding:

> 'He would take the view that the model knows what it wants to do and whether it does it [distribution of water vapour] through the convection scheme or through the large-scale rain, for example, is not so critical. The main thing is that you're getting the patterns right and the right overall sensitivity' (ibid.).

This is effectively stating that it is satisfactory to achieve the right result for the wrong reason, a position which would be challenged by many modellers. Nevertheless, these views create a dilemma for assessing the complexity of interaction upon which the modelling pyramid is commonly based. If the analytical reduction of the model to its constituent parts compromises the special interactive properties of the model, how can the complexity and accuracy of the interaction between the primary processes be assessed properly? Clearly, in delegating the issue of what is the appropriate interaction of dynamics and parameterisations to the model 'itself', an a priori commitment to the validity of GCMs (and their dominant position on the modelling pyramid) is entertained.

According to other modellers, including Stone and Risbey, this response is unsatisfactory because, if the division of flux between the sub-grid and resolved large-scale processes is changed, different dynamical processes become involved, which are also likely to have different feedbacks. Thus, the response of the model to a change in forcing, such as that from greenhouse gases, is likely to be affected by the presence of significant model errors [31]. In addition, some commentators feel that there is an unscientific circularity in some of the arguments provided by GCMers; for example, the claim that GCMs may produce a good simulation sits uneasily with the fact that important aspects of that simulation rely upon somewhat arbitrary parameterisations, tuning and prescribed boundary conditions.

3.3. FORMAL AND INFORMAL STYLES OF MODEL EVALUATION

In summary, there appear to be two different strategies in operation within the scientific community for evaluating GCMs. One involves a relatively explicit analysis of model errors, including attempts to understand the detailed nature of the model interactions. This we term a 'formal evaluation strategy'. The other approach, more common within the GCM community, is a more intuitive, model-trusting, and subjective approach, which we term an 'informal evaluation strategy' [32]. In the formal strategy, the desire is to explicitly calibrate the detail and adequacy of interaction within the model – to get 'inside' the model, so to speak. In the informal method, more trust is placed in the GCM modeller's personal *judgement* as to the complexity and adequacy of model interactions.

Our point is not to choose one strategy of evaluation over the other, since both would seem to have some value. It is rather that their existence poses a problem for the evaluation of GCMs, and their positioning vis-a-vis other climate models, according to any one set of criteria. Also the choice of the particular scientific style adopted by the modeller comes to influence the sorts of testing that can readily be performed on the model. The complexity of GCMs tends to limit the extent to which the model processes, interactions and uncertainties can be understood and analysed. This is because of the range and interdependence of such processes and the practical and analytical limitations thereby imposed upon our understanding of the model.

We have focused here upon several aspects of GCMs which are puzzling and/or troubling to the claims frequently made for such models, especially as far as their position in a modelling pyramid is concerned. There are additional features of GCMs which raise very similar questions and problems [33]. A more in-depth analysis of such additional facets of GCMs would be important in ascertaining the generality and significance of the results we have drawn upon above. We are not suggesting that other models are somehow devoid of these and other problems. We are simply pointing out an inconsistency in the arguments of many climate modellers. They frequently consider the output from simple models to be unreliable because they are 'over-parameterised', in the sense of over-relying on parameterisations to simulate climate processes [34]. GCMs by contrast are held to yield quantitatively meaningful projections of future climate change. Yet, this reasoning sits uncomfortably with the crucial dependency of GCMs on their *own* parameterisations [35].

4. Challenges to the Climate Modelling Pyramid II: Complexity and Model Construction

This section advances the challenge to the climate modelling pyramid by questioning whether the evaluative criteria which inform the climate modelling pyramid are themselves the only appropriate set of criteria.

4.1. ARGUMENTS FOR SIMPLICITY

The arguments about the strengths and limitations of GCMs tend to centre on whether or not one believes that the apparently complex global climate system can only be represented adequately by a complex simulation model. And yet one dominant theme in scientific thought has been to espouse a preference for simplicity in scientific explanation (sometimes called the principle of 'Occam's Razor') [36]. In theoretical physics, for example, there has been a strong feeling that simplicity underlies seemingly complex behaviour. As Einstein once stated it: 'In a reasonable theory, there are no (dimensionless) numbers whose values are only empirically determinable . . . this world is not such as to make an 'ugly' construction necessary for its theoretical comprehension' [37]. The underlying principle of parsimony maintains that, all else being equal (e.g., assuming it could explain or correctly predict as much behaviour), the simpler representation of a system's behaviour is preferable to unnecessary complexity. Popper, in like-mind, wrote that:

> 'The method of science depends on our attempts to describe the world with simple theories. Theories that are complex may become untestable, even if they happen to be true. Science may be described as the art of systematic over-simplification: the art of discerning what we may with advantage omit' [38].

Popper's emphasis on the testability of models reflects his central idea that *falsification* of knowledge-claims is the only secure route to improving the knowledge-base. Because simpler models are easier to falsify they are preferable from a Popperian standpoint.

Apparently simple models can, however, display highly complex behaviour (as in chaotic models, such as the classic Lorenzian attractor model). Indeed, falsification of an apparently simple model may be an artefact of the domain in which the model resides under strict test conditions. It has been suggested that 'falsification' may, in any case, be a problematic concept for two reasons: first, scientists normally adjust their theories in small ways to avoid discrepancy with observations; and second, falsification relies upon faith in the reproducibility of the falsification of a low-level empirical hypothesis. Some detailed studies of scientific practice have shown how theory and model adjustments are routine and reproducibility is highly indeterminate [39]. There is always, therefore, more than logic and methodology involved in deciding whether the falsification of a hypothesis or model has been achieved or reproduced.

Hence, whilst intuitively plausible, it remains an open question whether model complexity makes falsification any more problematic. But there is, possibly, a sociological and even political sense in which Popper might often be right. In the case of a simple model, there is often greater transparency to the outside observer in the assumptions imported into the model, including greater openness in the relations between different scientific disciplines involved in its production, than is the case for a complex model [40]. Arguably, such transparency improves understanding and the possibility of model analysis and testing, as well as increasing the modellers' self-awareness of the commitments entered into and their conditionalities, all of which would improve the opportunities for falsification (or confirmation).

4.2. COMPLEXITY AS A MORE HOLISTIC SCIENCE?

Those who advocate complex models respond to the above point by arguing that the traditional physical approach, and its demonstrated use of parsimony in scientific explanation, is not applicable to open systems. The real world is being observed passively and planned experiments, of the kind conducted in the laboratory, are the exception rather than the rule. Indeed, the apparent beauty and simplicity of physical theory is sometimes even regarded as an artefact of the controlled space and conditions of a laboratory, or of the epistemic commitments of scientists [41]. Against this, scientific representations of open-ended environmental systems may well appear 'ugly'; but surely only if they are judged against received and inappropriate normative criteria.

Some advocates of GCMs point to such models as representing a *more* scientific approach to a new kind of scientific problem, namely the modelling of multivariable, *open* environmental systems; in contrast to modelling the kind of closed systems that are associated with planned experimentation in the laboratory sciences [42]. They argue that it is only through complex models, which are able to simulate the subtle, often non-linear, interactions between dynamics and thermodynamics, that counter-intuitive, unexpected, and abrupt processes of change (such as those which litter the palaeoclimatological literature) will be successfully simulated [43].

An illustrative argument of this type in favour of GCMs can be put in the following way (after Palmer [44]). The non-linearity of the climate system means that the sensitivity pattern of the global atmosphere to perturbations is not coincident with its response. The consequences of increasing greenhouse gases will be transmitted through changes in the frequency and stability of different regime structures in the atmosphere. It is, therefore, imperative to understand and document the various possible spatio-temporal regime structures, so demanding a 3-D spatio-temporal representation which is capable of adequately simulating natural modes of variability, stability or instability, and multimodality. As Palmer puts it: 'the nonlinear perspective suggests that there is no simple short cut to the comprehensive modelling studies currently under way, if one requires a quantitative knowledge of the effects of enhanced CO_2' [45]. Taken to its extreme, the implication of these arguments is that even the *global* response of simple models to a change in forcing is unreliable [46].

Even if spatio-temporally distributed systems, such as the global climate, do require a modelling approach which is substantially GCMs-based, it is unclear whether current GCMs and coupled AOGCMs are adequate to the task anticipated by Palmer (a point he also makes). For them to be so, we must assume that climate GCMs can and do reliably reproduce the complex, non-linear dynamics and thermodynamics of the real climate system. Yet this is, to some extent, unknown and has to be taken on trust, or in the faith that future model development will deliver such an improved model [47].

Model approximations and tuning are routinely performed in GCM work which invest heavily in pre-suppositions about the character of the response. An example is the adoption of the assumption that the *difference* between the doubling of CO_2 perturbation and the control runs is the same as that between the ‘real’ climatology (the mean value of climate variables over a specified time-period, usually 30 years) and the ‘real’ perturbed climate. Therefore, the errors which result in the difference between the control run and climatology are not believed to move the climate feedbacks out of a linear regime. Hence, it is questionable whether the current generation of GCMs, as used in climate change experiments, can provide what Palmer outlines as necessary in order to properly simulate the spatio-temporal response to a change in forcing. This limitation of GCMs is not meant as a criticism, however; it is more an acknowledgement of the difficulty of the scientific task and is one, therefore, also shared across similar scientific fields and modelling.

The more general issue here is that the propensity of complex models to generate surprises is queried when the necessary judgement of the modeller in producing and running a model which is *plausible* (according to the prevailing understanding of the GCM community) is acknowledged [48]. Hence, some of the force of arguments in favour of complexity, such as those of Palmer, refer not to the current generation of GCMs, but rather to the anticipated, *potential* achievements of future, more complex, GCMs. To observe that this is a further judgement by GCMers is not to claim that it is right or wrong, but rather to open the way for other scientific judgements on such issues to be heard.

5. Challenges to the Climate Modelling Pyramid III: The Need for Objective-Oriented, Stochastic Models

In this paper, the model building process used to formulate and construct the GCM is considered as a prime example of ‘deterministic reductionism’. By reductionism, we mean here the process of ‘reducing’ a complex system to the sum of its *perceived* component parts (or sub-systems) and then constructing a model based on the *assumed* interconnection of the sub-models for these many parts. This is not, of course, a process which necessarily reduces the model in size at all: on the contrary, it normally leads to more complex models, like the GCM, because most scientists feel that the *apparent* complexity that they see in the natural world should be reflected in a complex model: namely a myriad of ‘physically meaningful’ and interconnected sub-systems, each governed by the ‘laws of nature’, applied at the *micro*-scale but allowed to define the dynamic behaviour at the *macro*-scale, in a manner almost totally specified by the scientist’s (and usually his/her peer group’s) perception of the system. This reductionist philosophy of the GCM model is ‘deterministic' because the models are constructed on purely deterministic principles. The scientist may accept that the model is a representation of an uncertain reality but this is not reflected at all in the model equations: the GCM is the numerical (always approximate) solution of a complex but purely deterministic set of non-linear partial differential equations over a defined spatio-temporal grid, and no attempt is made to introduce any quantification of uncertainty into its construction. The admission of uncertainty often only enters the scene when the model is later utilised in a predictive sense, normally by considering a range of different, but again deterministic, simulation 'scenarios' that reflect the model-user's lack of precise knowledge about the model parameters and its inputs. However, the propagation of uncertainty within the non-linear GCM, and its effect on the model predictions, is not quantified by any systematic, statistical analysis of the kind considered later in this section.

Attempts to quantify model uncertainty have been made in relation to other global models; e.g., those which attempt to represent global carbon cycle dynamics [49]. But the consequences of such uncertainty are then only evaluated using fairly rudimentary methods of deterministic sensitivity analysis. In particular, there appears to have been little research on the effect that stochastic influences, such as those arising from uncertain inputs and model parameters, might have on the model outputs and predictions. One notable exception is the work of Gardner and Trabalka (1985) who used *Monte Carlo Simulation* (MCS) analysis to investigate the implications of such stochastic effects on the Oak Ridge National Laboratory's World Carbon Cycle Model [50].

Deterministic reductionism, as discussed above, is by no means the norm in mathematical modelling. In many areas of environmental science, engineering, economics and statistics, a mathematical model is considered only as an approximate representation of reality, so that the quantification of uncertainty and its effects on model predictions is considered *sine qua non* for successful modelling. Also, whilst there is usually little argument over the fundamental physical laws which underlie the formulation of the model, the way in which these laws are utilised in model construction can differ widely, depending upon the perception of the model builder *and the use for which the model is intended*. Under this more pluralistic and objective-oriented modelling philosophy, the concept of a completely general, all-encompassing model, which is accepted by all scientists and satisfies all requirements, is considered an idealistic fantasy. Rather, in any particular practical application, a variety of models are likely to be required, each of which is tailored to a different, fairly specific objective and defined at an appropriate scale of behaviour and measurement.

This pragmatic, objective-oriented view of model building sits uncomfortably within the concept of a modelling pyramid: the idea that one model in the 'family' of plausible models has some special characteristics which make it universally 'better' than others is difficult to sustain in rational terms (though at any one time a specific model might be the best for a particular purpose). At this point in time, the GCM might well be considered as the ‘best’, albeit far from ideal, currently available model for *certain* objectives, such as the exploration of different hypotheses about the detailed, regional nature of the global climate dynamics, and the causes of observed patterns of climate change, or the generation of speculative regional forecasts that cannot be provided by other currently available and more aggregated models. But, as a vehicle for making probabilistic forecasts about the global mean temperature variations into the next millennium, it clearly has certain severe disadvantages, not the least being its totally deterministic nature and the complexity of its solution for such long period forecasts. It should be recalled that, even if the theoretical equations of the GCM were able to provide *exact* descriptions of the global climate, which surely most scientists would consider impossible, the numerical solution of these equations in the computer would still be approximate since an accurate, analytic solution of the equations is not possible. Moreover, the nature of the error propagation in this approximate solution is not clear, so that systematic errors become increasingly likely as the predictive interval is increased. And such systematic errors could well be similar to, and therefore not identifiable from, other errors that are endemic in the coupled AOGCM and are currently handled by devices of questionable validity such as ‘flux adjustment’ [51].

Perhaps the most publicised use of the GCM is for predicting the future levels of climate variables, such as global and regional mean temperatures with an assumed increase in atmospheric CO_2 concentration. There is a large technical literature on probabilistic forecasting and attention to this must raise some doubts about the manner in which the GCM is currently used for forecasting purposes. It would certainly suggest that the predictions obtained from the GCM should be treated with more circumspection than *seems* to be the case at the moment. It is simply a matter of ‘horses for courses’ and the highly bred, expensive GCM filly is not necessarily the best horse for the uncertain and hard ‘going’ that characterises the global climate racecourse into the next century. This is not, we hasten to add, meant as a criticism of the GCM or GCMers in total terms; merely we wish to emphasise that the supreme position of the GCM in global climate research, particularly in relation to the forecasting of global mean temperature variations into the far distant future, is not fully justified. What is justified, we believe, is a more balanced use of models in global climate research, with the GCM playing a very important but not so dominant role in the future.

Bearing the above comments in mind, this section explores the need for plurality in modelling the global climate by considering briefly a specific global climate modelling example: namely, an investigation of the Enting-Lassey (EL) (1993) model of the global carbon cycle [52]. We would have liked to utilise a GCM in these studies but, as pointed out previously, the facilities required for conducting such research are expensive and were not available to us. Moreover, the complexity of the GCM model would have militated against the completion of the research within a reasonable Ph.D. time-frame. Nevertheless, while the (EL) model is nowhere near the size and complexity of the GCM and takes the form of twenty three, interconnected, non-linear ordinary differential equations, rather than the partial differential equations of the GCM, it is still quite complex by more ordinary modelling standards. And it has played an important part in the studies carried out for the IPCC, so that its critical evaluation does raise important issues that relate also to the GCM. For detailed description of the EL model see the references in the endnote [53].

The EL model is formulated in deterministic terms and so, to start with, we consider how it may be converted into a probabilistic form which recognises *explicitly* the uncertainty that must be inherent in any scientific evaluation of the global carbon cycle. The implication is, of course, that this modified form of the model is more appropriate for predictive purposes, since it can consider the propagation of predictive uncertainty arising from the lack of accurate knowledge about the parameters and inputs of the model.

After this stochastic study, another mathematical technique, Dominant Mode Analysis (DMA), is utilised to explore whether the relative complexity of the EL model is really required for predicting global mean atmospheric CO_2 variations; or whether the apparent complexity of the model is belied by the simplicity of its dynamic behaviour in this regard. It should be noted that, within the context of our discussion on GCMs, the main purpose of introducing this example is to emphasise that, if models (criticised as being overly-simplistic by some GCMers) present certain problems when considered in terms of their treatment of uncertainty and their complexity, then there can be no doubt that the GCMs themselves present even more difficult problems in this regard; problems that we believe demand urgent attention.

5.1. QUANTIFYING THE CONSEQUENCES OF UNCERTAINTY: MONTE CARLO SIMULATION (MCS) ANALYSIS

While it is a technique which is rarely used in climate research, possibly due to the reductionist view that it is 'second best' [54], stochastic modelling, in its various forms, is a powerful tool for investigating poorly defined systems with limited observational data, such as the global climate. In this era of the powerful digital computer, probably the most useful, general approach to the stochastic modelling of non-linear systems is *Monte Carlo Simulation (MCS*) analysis [55] . It is an approach that, rather conveniently in the present context, can be applied to physically-based models that have, in the first instance, been formulated in deterministic reductionist terms [56].

In this kind of MCS analysis, the physically-based, deterministic model is converted into stochastic form by assuming that its parameters and inputs are inherently uncertain: the parameters being characterised by probability density functions (pdfs) and each input being expressed in terms of a stochastic time series model [57]. The subsequent analysis then exploits MCS; i.e., the repeated simulation of the model, with the parameter and input values for each simulation run (or 'stochastic realisation') randomly drawn from the defined pdfs. Typically, MCS involves several hundred, and sometimes several thousand, simulation runs: the exact number being dependent on the required accuracy of output statistical measures, such as cumulative distribution functions for each output variable [58].

Adopting a Bayesian approach, the pdfs are defined by reference to the current literature on the parameters values and their associated uncertainty. Since knowledge is often lacking in this area, however, a certain amount of subjectivity will naturally enter into the analysis and this must be taken into account when evaluating the results. On the other hand, if this aspect of the research is carried out carefully, the measures of uncertainty will reflect the prevailing opinions of the scientific community and it is important, from a Bayesian standpoint, that they should be considered as prior information to be introduced systematically into the modelling process. *In the absence of any evidence to the contrary,* the pdfs for the parameters are assumed to be Gaussian in nature and independent of each other. Of course, if the parameters are, in fact, not statistically independent, then this assumption *may* lead to higher estimates of uncertainty in the model responses [59]. Again, however, this reflects the uncertainty that characterises the scientific community's knowledge of the climate system and it is as well that the consequences of such limitations in knowledge are fully realised.

In the case of the EL model, the uncertainties in the model parameters and inputs required for the MCS analysis were obtained by reference to the latest literature on the subject and from information supplied by scientists working on global climate change. Here we present only the most pertinent results of the detailed research programme that followed [60].

The uncertainty associated with the ensemble of atmospheric CO_2 realisations generated by the MCS analysis was always much greater (by up to twenty times) than that in the observational record. Even taking into account the caveats mentioned above regarding the specification of the pdfs, this is a large discrepancy, which suggests that more research is required both to reduce the levels of uncertainty associated with the model parameters, and to consider whether the assumed values of the parameters are entirely appropriate to the scale of the model: it is not obvious that parameters inferred at the micro-scale are applicable at the macro-scale, particularly when the macro-scale, as in this case, is the entire globe.

Continuing with the MCS analysis, a special form of generalised sensitivity analysis reveals that the EL model is very dependent on the values of a small number of important parameters, especially the pre-industrial level of atmospheric CO_2. The high level of variance of the MCS ensemble of atmospheric CO_2 realisations is due, in part, to such high sensitivities. However, it could also arise because the model is more complex than is justified by the nature and paucity of the observations. This possibility is given further credence by the Dominant Mode Analysis (DMA) discussed in the next sub-section [61].

Table I compares the uncertainty measures associated with the EL model predictions of atmospheric CO_2 in the years 2050 and 2100, as generated by the MCS

Table I

Comparison of the uncertainty in future CO_2 levels due to IPCC scenario IS92a between three deterministic modelling exercises (Wigley and Raper, 1992; Enting and Lassey, 1993; and Schimel et al., 1995) and the stochastic methods used in this paper

Year	Uncertainty in atmospheric CO_2 (ppmv)	Range	Source
2050	494 to 510	16	Schimel et al. (1995)[a]
	520 to 550	30	Wigley and Raper (1992)
	491 ± 86	172	This paper, original simulation
	494 ± 22	44	This paper, modified simulation
2100	667 to 719	52	Schimel et al. (1995)[a]
	740 to 800	60	Wigley and Raper (1992)
	615 to 683	68	Enting and Lassey (1993)
	664 ± 144	288	This paper, original simulation
	669 ± 48	96	This paper, modified simulation

[a] The results presented in this paper are derived from Enting et al. (1994).

analysis in the case of the IPCC scenario IS92a; and compares them with results obtained in deterministic studies, including both Enting and Lassey, and Wigley and Raper, the latter using their STUGE model [62].

These results show that the stochastic uncertainty is larger than the range of variability found in the other studies. However, two criticisms could be levelled at the way the MCS was carried out: (i) no account was taken of recent atmospheric CO_2 observations, which are more accurate than pre-industrial estimates used to constrain the model; and (ii) as discussed before, it is possible that our estimates of the uncertainties in the parameters and inputs have been exaggerated by the assumption of independence in the pdfs. To take some account of these shortcomings, then, the MCS analysis was repeated with each realisation constrained to the uncertainty in the observed CO_2 level in 1990 (353 ± 2 ppmv), and with parametric and input uncertainties reduced by an arbitrary factor of two [63]. The results of this revised simulation are also shown in Table I. As we can see, the results of the stochastic simulation are much reduced, but they remain 50% wider than the widest of the deterministic results. This is a significant amount: moreover, it would probably be larger still if the stochastic methodology could be applied to a complete set of global carbon cycle models, as in the IPCC's comparison of deterministic simulation studies [64].

By modifying the deterministic EL model into a stochastic form and applying MCS analysis based on parametric uncertainty estimates supplied by the research community, we have produced estimates of uncertainty which are significantly greater than those gained by deterministic model studies alone. It is recognised that this kind of stochastic analysis is still in its early stages and more research, especially on parametric uncertainty, is required. Nevertheless, the present results suggest quite strongly that significant sources of uncertainty are not being considered at present by many climate researchers. This would not be surprising in the case of GCMs since their huge computational demands make such uncertainty analysis difficult, limiting any objective assessment of GCM results in relation to uncertainty. But it suggests that uncertainty analysis of this, or other, types is overdue in relation to GCMs and that, in predictive applications, GCMers need to address more rigorously and quantitatively than heretofore, the consequences of predictive uncertainty on the GCM forecasts.

5.2. SIMPLICITY OUT OF COMPLEXITY: DOMINANT MODE ANALYSIS (DMA)

Dominant Mode Analysis (DMA) seeks to analyse a given, physically-based, deterministic model by identifying objectively the small number of dynamic modes which appear to dominate the model's response to perturbations in the input variables. In contrast to the traditional reductionist modelling, this normally results in a considerable simplification of the model, which is simultaneously both reduced in order and linearised by the analysis.

The DMA methodology involves perturbing the complex and usually non-linear, physically-based model about some defined operating point, using a *sufficiently exciting* signal, i.e., one that will unambiguously reveal all the dominant modes of behaviour [65]. A low order, linear model, in the form of a transfer function, is then fitted to the resulting set of simulated input-output data, using special methods of statistical estimation that are particularly effective in this role [66]. As might be expected from dynamic systems theory, a low order linear model obtained in this manner reproduces the quasi-linear behaviour of the original non-linear model about the operating point almost exactly for small perturbations. Perhaps more surprisingly, as revealed in the present example of the EL model, the reduced order model can sometimes also mimic the large perturbational response of its much higher order progenitor.

To carry out the DMA, the full, non-linear EL model is initially set to an equilibrium condition with the (pre-industrial) atmospheric CO_2 concentration set at 275 ppmv. Then a small perturbation in the fossil fuel input is applied to the model and the resulting response is monitored over a period of 3000 years. The statistical identification and estimation analysis yields the 4th order, linear transfer function model, one interpretation of which is as a parallel connection of three, first order systems and an integrator (the latter is required because of the mass conservation assumption in the EL model) [67]. This simple dynamic model explains almost all (99.98%) of the simulated model response.

More significantly, Figure 2 compares the response of this 4th order model with that of the full, non-linear model over the entire industrial period [68]. As can be

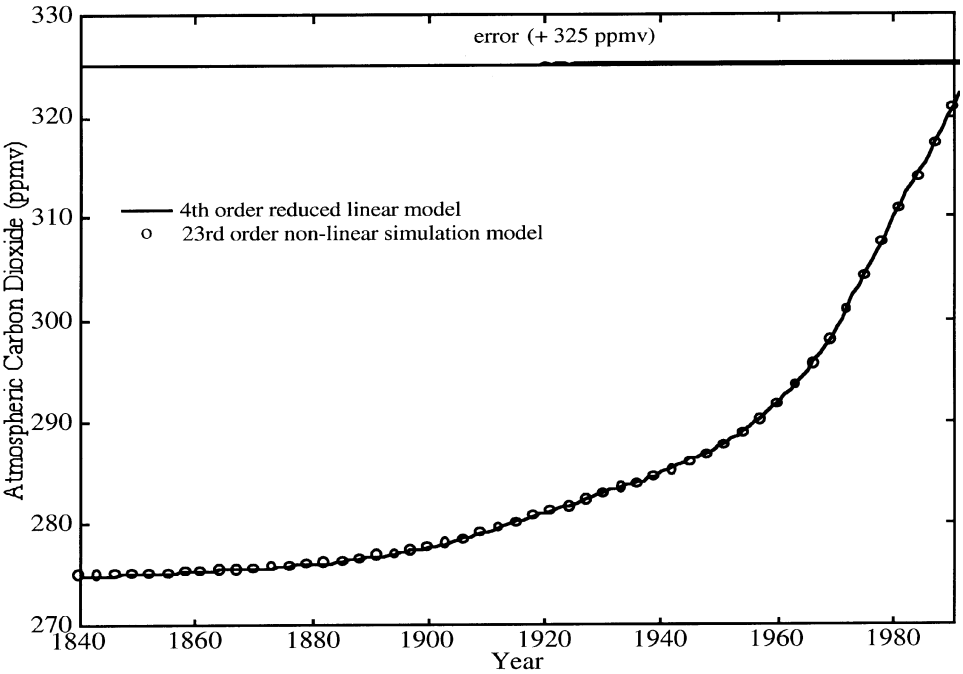

Figure 2. Comparison between the responses of the reduced fourth order, linear model and the 23rd order non-linear simulation model (Enting-Lassey Model) to fossil fuel input (error + 325 ppmv shown above).

seen, the error is very small: never greater than 0.5 ppmv, even though the model has moved 45 ppmv above the operating point at which it was estimated. These rather surprising results demonstrate the robustness of the low order, dominant mode model and suggest that the non-linearities in the original model are hardly being excited by these fairly substantial perturbations over the industrial period. Moreover, further DMA results at a range of different operating points reveal that the objectively identified, 4th order model structure does not change at all, and the behaviour of the non-linear model can be reproduced very closely with only small variations in the parameter values. This is quite a dramatic result which demonstrates the dominance of a small number of modes of behaviour in the system, as defined by the eigenvalues of the reduced order, linear model at any defined operating point.

An interesting aspect of these results is the size of the three time constants (residence times), T_i, (where $i = 1, 2, 3$) in relation to the original physically based model: $T_1(= 4.3y)$ is not dissimilar to the stratospheric turnover time of $6 \pm 2y$; whilst $T_2(= 19.1y)$ is fairly close to the air-sea exchange time of $11 \pm 4y$; and, finally, $T_3(= 467y)$ is of similar size to the deep ocean turnover time, of $900 \pm 250y$ [69]. Naturally, the complex interconnected nature of the simulation model, with its inherent feedback processes, modifies the effective 'closed loop' values of these time constants when the model is integrated and so the identified dominant mode time constants seem very reasonable [70].

What can we learn from the above DMA results? It is clear that the variations in atmospheric CO_2 produced by the high order, non-linear EL model over the whole historic industrial period can be reproduced very adequately by a 4th order linear model whose modal structure has reasonable physical significance. As regards the prediction of atmospheric CO_2, therefore, this must call into question the need for such a complex representation of carbon balance in the EL model and suggests that a much simpler representation with fewer, or at least lower order, sub-systems (e.g., the EL model has 18 compartmental levels in the ocean sub-system alone) could have produced very similar results. Moreover, such a reduced order model would be more appropriate to the amount of observational data available in this example which, in itself, makes the assumption of a high order model rather questionable on statistical grounds. Indeed, there is evidence in the literature that the EL model can only be fitted to the available data with constraints applied to many of the parameters, a common indicator of severe over-parameterisation [71].

It should be noted, however, that we have only carried out DMA for a single input and a single output of the EL model. Using more inputs and/or outputs may well produce a more complex reduced order model. The number of inputs and outputs in any DMA analysis should relate to the goals of the analysis. In this particular case, if it were felt that the use of CO_2 input from fossil fuel emissions and atmospheric CO_2 concentration output were adequate, then our example would suffice; if not, then our study would need to be extended. This is straightforward in methodological terms.

5.3. IMPLICATIONS FOR GCM DEVELOPMENT AND USE

Within the more general context of the present paper, the results outlined briefly in the previous two sub-sections suggest the need for more checks and balances in the deterministic reductionist approach to modelling which gives rise to such complex representations of decidedly uncertain physical systems. Clearly, however, the alternative modelling methodologies used in the above analysis are still being developed and themselves have shortcomings: e.g., the stochastic simulation modelling is limited by the difficulty of quantifying fully the parametric and input uncertainties; and by the fact that it has not been used to estimate output uncertainty due to uncertainty in the model structure (though it could perhaps be developed to do so). In addition, such alternative methods have as yet to be applied to spatio-temporal models described by partial differential equations, as in the case of GCMs.

Nevertheless, all of these methodologies (amongst others) offer useful additions to conventional deterministic reductionist approaches, with considerable potential for revealing new information and insight within the area of climate change, as they have in other fields. Such methods could be exploited, for example, to develop an interactive, multi-disciplinary modelling strategy in which they help to raise meaningful questions about the quantification of uncertainty and the relevance of processes or feedbacks in more traditional reductionist models; questions which could then be addressed and answered, on a continuing basis, by the reductionist modeller. This would help to rationalise the large models by avoiding excessive complexity and over-parameterisation (in the sense given in note 34) and it would ensure that sufficient attention is given to the consequences of uncertainty.

These alternative modelling methods also seem particularly attractive when one considers the comparatively low demand on funding and computational facilities made by them in comparison to those required for GCMs. In addition, the new analyses of uncertainties are likely to be important inputs to policy debate which, given the potential importance of global climate change, must surely take cognisance of any available quantifiable measure of uncertainty when evolving risk-averse strategies. For these reasons, the development of such techniques should, we believe, be placed on a more equal footing with the more commonly accepted reductionist modelling philosophy that underlies GCMs.

Of course, in highlighting alternative techniques we are not in any sense denying the relative importance of GCMs. It is simply our objective to point out that they score well with respect to certain criteria, but not to others. For example, if the goal is defined as providing long-term climate predictions which are based on current scientific perception of the physical mechanisms involved, then GCMs are certainly a strong candidate. But while GCMs are, therefore, capable of exploring some key features of climate change, including variable time-horizons and, *potentially*, regional scales and variability, they are much less suitable for integrating with other physical and socio-economic models, or for performing uncertainty analysis and associated stochastic prediction. Moreover, they are resource-intensive and rather intractable. There is a wide-spread experience with them although, as we have discussed, confirmation of their reliability, especially for the purpose of making projections, remains a difficult issue. Finally, they are not very accessible or transparent and feedback from other scientific communities as to their validity is not readily achieved.

Our main argument is that the role of GCMs can only emerge through considering fully the other possible methods that are available. An impediment to this is that, at any one time, research tends to coalesce around certain methods as if they are the only ones applicable for the given purpose, even though judgements might change radically with the benefit of hindsight. Several alternative approaches to studying future climate change that have been proposed and, in part, developed over the past several decades, now seem to have 'gone out of fashion' or disappeared; or else they are only recognised in small research enclaves (and seem to be especially marginalised when research results are presented to policy makers). Alternatively, they might be ascribed a role as 'second-rate' to GCMs or as 'merely' a heuristic or educational tool.

In the following section we go on to discuss the wider policy and social dimenions of GCMs which we earlier defined as also critical to an understanding of the relative role of different methods.

6. Why are GCMs Perceived as the Best Climate Models?

To account fully for the dominance of deterministic reductionism in climate modelling would require a detailed study of such a scientific style in its rich historical, institutional and intellectual context. Our purpose here is less ambitious, namely to consider the role of social, institutional and policy relationships in helping secure the position of GCMs. We explore whether GCMs provide unique and critical information for policy making (and hence are seen as 'better' models for this reason), as well as a more sociological interpretation which emphasises the role of GCMs in linking-up different scientists and policy makers.

6.1. ARE GCMS ESPECIALLY POLICY USEFUL?

At first sight, GCMs appear to be at a distinct advantage for policy purposes: relative to simpler climate models, they provide regional climate change scenarios which can be used in assessing climate impacts and, hence, in devising suitable policy responses. But, as already noted, since GCMs do not currently produce robust regional simulations, this apparent advantage must be scheduled into, and stored-up for, the future – at which point such simulations will become robust. Or so it is anticipated. The promise of reasonably robust regional climate simulations can, nonetheless, strengthen the case for current continued development of GCMs as the only models that are *potentially* capable of producing regional simulations.

Yet, GCMs would also seem to be at a disadvantage in policy-making because of their complexity which: (a) demands massive computing and personnel resources, so restricting the number of GCMs and modelling centres; (b) prevents the rapid inclusion of new processes into the GCMs as new scientific insights, understanding or hypotheses emerge; (c) precludes the opportunity of performing extensive sensitivity and uncertainty analysis; and (d) prevents the use of GCMs in a 'responsive mode' to rapidly explore the implications of different model assumptions and policy scenarios [72]. Evidence that these disadvantages have been experienced emerges from the demonstrated policy-usefulness of simple models in the IPCC process: because of their greater flexibility and ability to address specific policy questions, simple models have a clear advantage in this role when compared to GCMs [73]. As one government policy maker put it: 'He [a scientist using simple models] could come up with the goods very quickly with his simple box model' [74].

If models are sufficiently user-friendly and simple, advisory scientists and even policy makers can run their own simulations, sensitivity and scenario analyses.

By contrast, GCMs create a dependency on just a few centres and their resident experts. It is possible that this dependency is not welcomed by many policy makers (though it may be accepted as inevitable), since they (or their scientific advisors) are less able to gauge the merits of the model outputs through, for example, not having direct access to the modellers' advice and interpretation [75].

On the other hand, such dependency might be welcomed by scientists, given that non-specialist users of models are more liable to run the model uncritically or misinterpret model output through not appreciating its framing assumptions, uncertainties and limitations. And so the policy maker can be misled into assuming that the *apparent* simplicity of the model implies a simple interpretation of the implications of the model's output for the purposes of 'reality' and hence policy decisions [76]. Use of GCMs, by contrast, are more likely to involve the judgement of the modeller as an integral component of the scientific advice [77]. But let there be no doubt that we feel such involvement is essential if models are to be used for policy purposes: regardless of their complexity, all scientific models should be used, and their results interpreted, carefully by scientists who are able to fully understand their nuances and, if possible, have been involved in their synthesis.

6.1.1. *Sununu and Simple Models*

A pertinent example of how simple models may be employed in policy terms is the use made by John Sununu, President Bush's Chief of Staff, of a simple climate model provided to him in the late 1980s by Warren Washington of the National Center for Atmospheric Research (NCAR) [78]. Sununu explored the effects upon surface temperature of changing the depth of the model ocean's mixed layer; and, in tandem with the results of Washington's early AOGCM transient runs, concluded that any warming would be taken-up by the ocean. His exposure to climate models apparently led him to characterise them as 'bad models' – a charge which would presumably be refuted by climate scientists on the basis that Sununu was guilty of having over-interpreted or misinterpreted the output of a simple climate model [79].

Yet, the problem of interpretation faced by Sununu was not exclusively related to simple climate models, since Washington also exposed Sununu to GCMs. As Washington now tells the story, the early transient run of the coupled AOGCM which he showed to Sununu in 1989 did not show a slowing down of the thermohaline circulation (THC) with an increased atmospheric concentration of CO_2 with consequent effects on the take-up of heat into the deep ocean. The slowing-down of the THC only occurred subsequently in the apparently improved transient runs [80].

We cannot conclude much about the use of simple or complex models in policy from just one example. The limited disputes over the use of knowledge from climate models in policy have not, to date, revolved around the simple/complex distinction, but more about conflicting interpretations of the *same* simple and complex models, and their relationship with other sources of knowledge, such as instrumental temperature records [81]. Simple and complex models may, nevertheless, serve different purposes in the climate change policy domain. GCMs, perceived as the most authoritative, 'realistic' models sustain the legitimacy of climate change as a central issue for international environmental policy. Simple models, on the other hand, serve a more instrumental role in addressing policy-driven questions; for example, concerning the effects of specific scenarios and projections. In conclusion, the climate modelling pyramid cannot be explained by the special policy-role of GCMs (even if policy usefulness were to be admitted as a valid criterion in this context).

6.2. TOWARDS A SOCIOLOGICAL UNDERSTANDING OF THE POSITION OF GCMs

A more sociological understanding of the role of GCMs, is to point out the way in which they link-up a diverse range of sciences and policies, both now and potentially in the future. The suggestion here is that the more extensive are the links between GCMs and other sciences and policy issues, then the more support is implicitly being provided to GCMs. Their position as the 'best science' is buttressed by the implicit or explicit recognition and uptake of GCMs by others, both scientists and policy actors.

We argue that this occurs through a number of processes *of mutual exchange and reinforcement* of the rationale and validity of a scientific or policy enterprise. The most important of these are summarised in Figure 3, to which the reader should refer throughout this section. Three sets of relations are stressed:

- those between GCMers and the policy community, both of which share a commitment to the task of long-term climate projections at global and regional scales;
- those between GCMers and the climate impacts community, both of which share a commitment to producing regional climate change simulations;
- those between GCMers and surrounding sciences, both of which share a commitment to including more processes and detail into GCMs.

The mutual reinforcement which emerges out of these, partially over-lapping, shared commitments is an important element in understanding the role and position of GCMs in climate change research.

The end-result may be that GCMs become enmeshed in a network which constitutes a powerful and resilient set of relationships: it is perhaps this resilience which gives the impression that GCMs are 'naturally' at the top of a climate modelling pyramid. Another way of putting this is that GCMs (*contra* other models or methods) come to act as a sort of *common currency* between groups of scientists and policy makers – each considers they have something to gain in intellectual, scientific, funding and social terms – from being involved in their development and use – and that this commonality serves as a way of linking-up such groups into

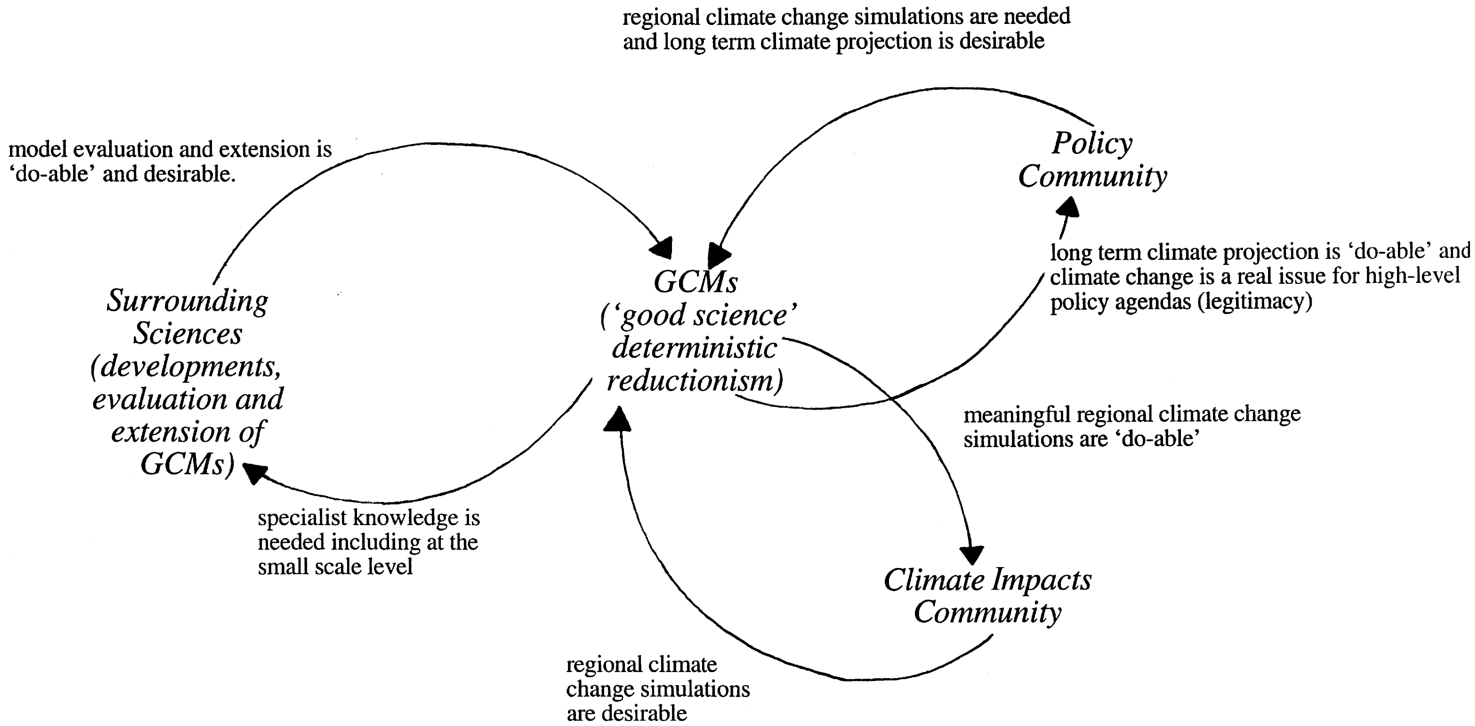

Figure 3. The interdependencies between climate modellers, surrounding disciplines and the policy community. The direction of the arrow on the lines indicates the message received by one group from the other, e.g., the polilcy community receive from the GCM modellers the key messages that long-term climate projection is 'do-able' and that anthropogenic climate change is a 'real' issue.

loose coalitions. Hence, GCMs come to have a wider symbolic significance than implied by their scientific credentials alone.

A necessary part of our argument is that these mutually reinforcing commitments emerge in part from the lack of knowledge and understanding of each group when dealing with the other. For example, GCMers typically (and understandably) operate with a somewhat narrow conception of what policy ‘needs’ from science and this comes to limit the sorts of research they promote in their dealings with policy makers. The subsequent response of the policy maker in ‘requesting’ particular sorts of science is already partially conditioned by the assumptions of the scientist. As far as the scientist is concerned, however, this merely reinforces his or her beliefs about what scientific advice and research is needed by policy! However, as apparent from Section 5, it is also part of our argument that most scientists themselves are frequently (and again understandably) limited in their acknowledgement of the possible range of scientific approaches. The existence of particular scientific commitments (say to deterministic reductionism) also comes to limit what the scientists communicate to each other and to policy makers.

Exactly the same argument applies to policy makers; i.e., they too presuppose the role and capabilities of science as well as routinely operating with a limited appreciation of the possible range of policy approaches. The existence of convergences between these limited accounts of science and policy, which tend to be assumed and not subjected to critical evaluation, helps explain their reinforcement of each other, as if they are the only possible versions of science and policy. It also has a major influence upon how GCMs are evaluated relative to other climate models [82].

6.2.1. *Involvement of Other Scientists*

GCMs appear to have great, but only partly realised, potential to interconnect diverse areas of environmental science. For example, Figure 4 is an ironic version of these interrelationships from the perspective of the research co-ordinator at one GCM centre. No other model promises this degree of integration or collaboration between scientists. A modeller using a simple model who wants to include the carbon fertilisation feedback or the effect of aerosols in the model, for example, does not require an extended collaboration with biologists or atmospheric chemists. What is needed are a few ‘best estimates’ from the world’s experts on these subjects, probably accessible from the published literature.

In GCM ‘extension work’, by contrast, ecologists and hydrologists are needed to model, for example, the intricacies of the movement of water from the soil through vegetation to the boundary layer, to extend micro-level models of catchments so they can be applied at the resolution of GCM grid-squares; and so on through a myriad of other possible examples. And again, without the help of atmospheric chemists, GCMers would be unable to attempt the integration of the dynamics and chemistry of different greenhouse gases into the three-dimensional structure of GCMs.

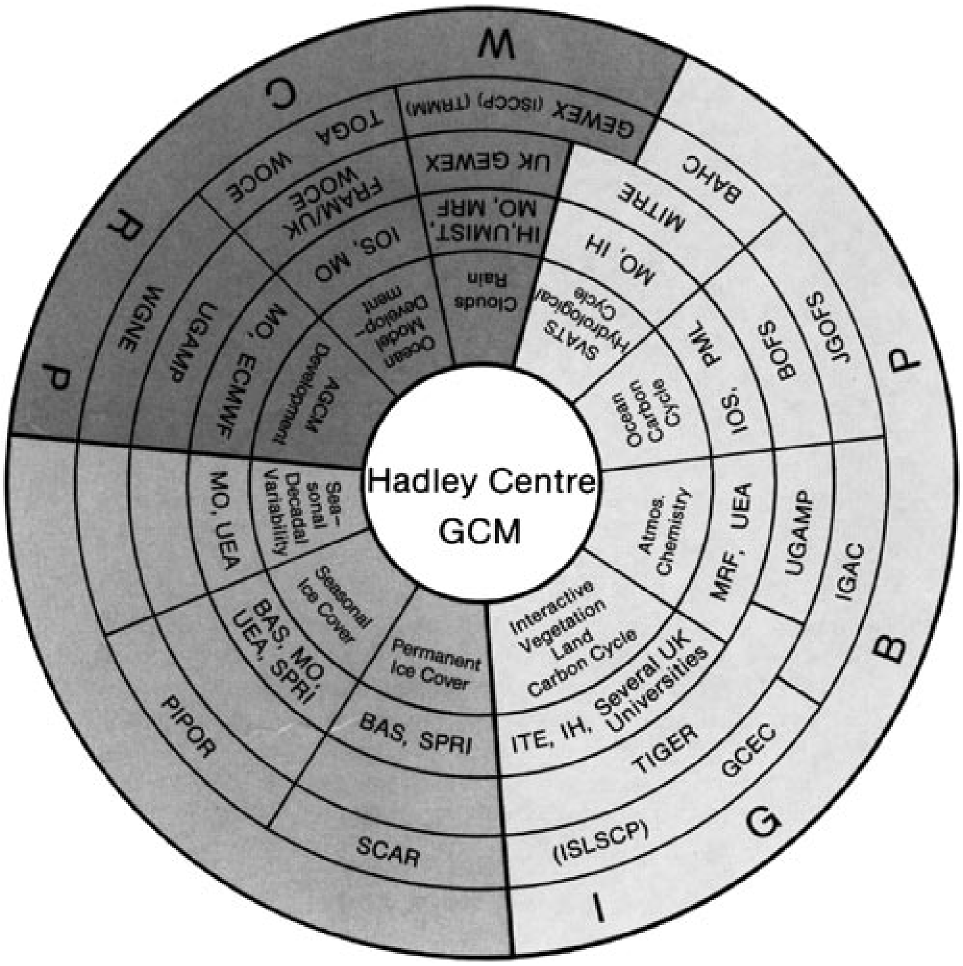

Figure 4. The Hadley Centre perspective of the science base for climate prediction research (after Copernicus). The outer ring contains the international framework programmes with some of their constituent projects listed in the next ring. Inside these there are main U.K. programmes followed by the U.K. institutions with which the Hadley Centre is most closely associated. The inner ring gives a brief description of the products of highest priority to the Hadley Centre programme. (From *A Review of the Science Base Underpinning Climate Prediction Research*, Hadley Centre, U.K. Meteorological Office 1992, page 1.

There is a clear role, then, for many scientists, other than GCMers, in the continuing development of GCMs such that they are seen to be ever more inclusive, realistic and comprehensive. By perceiving the advantages and opportunities of collaboration with GCMers in their own research work, these other enrolled scientists are effectively endorsing and advocating GCMs. Their own work practices and disciplinary pre-occupations then change in the service of the greater model (and, for this reason, they can be considered as mavericks or treated with suspicion from the perspectives of their native disciplines) [83]. This, then, is a vision of GCMs as a sort of 'collective good' amongst environmental and earth system scientists of many disciplinary persuasions [84]. They act as integrators of diverse empirical and theoretical pursuits [85]. Hence the set of links between GCMs and other fields should be especially dense.

The extended GCM community includes also those meteorologists and climatologists who specialise in handling observational data-bases, and who increasingly argue that their data-sets are essential to the proper validation of GCMs. This includes the satellite agencies, who have made persuasive arguments about the value of global remote sensing and have secured a large proportion of the total funds for global environmental change research, as indicated in Table II. Some observational scientists have gone so far as to pin the full rationale for data collection upon the need for improved GCMs, sometimes to the discomfort of GCMers [86] Other observational scientists have become linked with GCM groups, identifying the validation of GCMs as a new role for themselves, the variable spatial and temporal scales of GCMs allowing flexibility in what analysis is actually done; a role which supports the relevance of data collection and collation and provides a respectable home for their data-bases. It is true that simpler models also need data for validation; but they need much more highly aggregated data organised into a few global variables. By contrast, because GCMs are simulating climate at many thousands of grid points and over many time-scales, there is a much greater demand for disaggregated, comprehensive regional data, whose interpretation and extrapolation to grid points (to provide the model's boundary-conditions or for validation purposes) requires detailed interaction with observational scientists. [87].

In addition to satellites, remote sensing and other forms of instrumentation, GCMs and GCMers are connected to the technological enterprises of supercomputing, networking and software development, including massively parallel processing. By contrast, most other climate models can now be run on a desktop computer; and even computationally intensive procedures such as Monte Carlo Simulations, which are virtually impossible for GCMs, can be applied to quite complicated (non-GCM) models by linking several such simple processors in parallel [88].

6.2.2. *Scientific 'Users' of GCMs*

Further participants in the construction of the GCM-network are the various scientific communities using the output from GCMs. For example, the 'climate impacts community' explores the effects of climate change on agriculture by deriving data from individual grid-points of GCMs and feeding it into crop productivity models [89]. Such impacts models are reliant upon the regional output of GCMs, which varies a great deal between different models, and is dependent upon the specific formulation, such as the parameterisation of land-surface processes. Impact modellers become, to an extent, dependent upon the judgements of the climate modellers

Table II
Approximate expenditure on global change (U.S.) and global climate change (U.K.) research by subject area

Broad area	Approximate expenditure in millions of U.S. dollars for 1993–1994	
	U.K. (1993/94)	U.S. (1993)
Observations		
Satellites	100	560
Ground-based observations	50	25
Informational management	25	275
Sub-total	175	860
Detection	1	Not known
Understanding global change	65	455
Predicting global change	20	50
Impacts/evaluating the consequences	20	40
Response strategies	40	Not known
Assessment tools	Not known	15
TOTAL	320	1420

These figures are highly approximate and rounded off to the nearest 5 million dollars. They should be regarded as no more than a rough guide: for example other estimates for the U.K. differ by approximately 15% for reasons which are not known. The source for the U.S. figures is Global Change Research Program: 1995, *Our Changing Planet*, U.S. Government Printing Office, Washington, DC. For the U.K., estimates have been made from information assembled by the U.K. Department of the Environment; a conversion ratio of $2 to £1 has been used in the table.

as to which GCM model runs to use and how to construct suitable future climate scenarios.

Climate impacts work is not the only sub-discipline to take the output of GCMs ‘on trust’. Economists who study the costs of climate change, for example, also require information on factors such as: radiative forcing factors, greenhouse gas life-times, damage functions, the rate of climate change, and the likely time-scale for the reduction of critical scientific uncertainties, some of which information is derived from GCMs [90]. As already noted, those who use simple climate models also rely upon GCMs for their calibration; and this includes some economists and engineers developing integrated assessment models. By using GCM data, both impacts scientists and economists are also engaged in constructing and maintaining the network, and the pivotal position of GCMs. At the same time, GCMers are lending their implicit support to the impacts and economics work by providing seemingly reliable scenarios of ‘realistic’ climate change.

6.2.3. *Policy Makers*

The rationale for climate impacts and economic studies is, ostensibly, to contribute to better climate change policy-making. Hence, the GCM network becomes central to knowledge-based policy-making aspirations. Yet, can it be claimed that government or other policy makers attach themselves particularly to GCMs in preference to other modelling or climate research activities? Funding reveals some preferences (see Table II) although a direct numerical comparison is not especially meaningful given the different costs of different sorts of research (impacts and integrated assessment research are much cheaper than GCM modelling, which again is cheaper than satellite development and application). Nevertheless, it is noticeable that, for example, all of the climate modelling funded by the U.K.'s Department of the Environment is that involving GCMs, even though it could support much simpler model development for relatively little money.

Why, then, should the U.K. DoE attach itself to GCMs rather than simpler modelling? Clearly part of the reason derives from the scientific community's evaluation of GCM work *vis-à-vis* simpler modelling. The argument that scientific credibility (and preferably certainty) is necessary for the accomplishment of political consensus appears to hold powerful sway in many political and policy cultures and, if GCMs are held out as the most reliable and robust of climate models, it follows that they will enjoy an elevated status in policy circles. The wish to maintain a specific national influence in international policy negotiations is also seen by some policy makers to depend upon the national government being amongst the global scientific leaders which, assuming the existence of the modelling-pyramid, means supporting GCMs.

Given the above situation, part of the scientific evaluation of the GCM science as 'the best' climate science derives from this policy status and influence [91]. Furthermore, of all climate models, only GCMs seem to offer the promise of the detailed, regional, 'hard' scientific knowledge necessary to implement proposed local policies for sustainable development (as specified by Agenda 21 of the Rio Earth Summit), despite the far-from-achieved efficacy of such regional modelling. As a result, the understandable criteria amongst scientists for the evaluation of different model types, in terms of *future* potential simulations, may have been extended to an evaluation of the *current* policy usefulness of different models.

A further possible and implicit advantage for policy makers of using GCMs may be their sheer complexity and impenetrability. As such, they cannot easily be challenged by critics, whether environmentalists or industrialists, most of whom cannot possibly comprehend the complexities of such a large computer model, let alone articulate its deficiencies. There is no conscious deception here since, as far as the policy makers are concerned, the models are impenetrable *because* they represent the most sophisticated science.

Certain models or areas of science, other than GCMs, could even be a *threat* to the vision of using predictive knowledge of natural (and social) systems for the purposes of global environmental risk management. For example, stochastic models might be seen to fit into this category because they take the uncertainty of our knowledge as their starting point, so problematically questioning the attempted deterministic closure of networks around the current science-policy nodes. Chaos theory and some palaeoclimatology research also seem to qualify as problematic scientific partners to a global environmental management policy because they question the very predictability of climate. Not only do these other areas of scientific research not 'gel' with the global environmental change project in its policy-oriented formulation of the last decade or so, to an extent they have to be actively resisted or ignored by it [92].

6.3. IMPLICATIONS

We have tried to show that the emergence of consensus around particular concepts of 'good science' are neither predetermined by nature, nor are they without larger political significance. This is because they inadvertently assist the development of political consensus as to the appropriate policy response. We have argued that it is partly these political implications of the global scientific culture which place GCMs at the top of a 'natural' hierarchy of 'good science' and valid knowledge for international policy actors [93].

Sciences which promise *integration* may be particularly appealing in the context of increasing senses of policy and political flux and insecurity globally. From this perspective, GCMs can be seen as critical instruments by which policy agendas aimed at the control and management of global climate change are shaped and sustained. They are perceived to lend much credibility to the prospect and legitimacy of uniting and globalising diverse activities, including national and regional policy making. They achieve this through the application of standardised techniques, inputs and assumptions, as well as through promoting and legitimising the development, in other related domains, of such standard methods and techniques. So, for example, GCMs provide justification for the creation of greenhouse gas emission inventories, as well as for changing social practices in a way which reduces such emissions, whether through energy efficiency programmes, or a change in agricultural practices. If they are believed, deterministic models such as GCMs and their applications could become enormously powerful, both in technical and social terms, because they stand to acquire the power of all those diverse constitutive activities carried out over decades in different scientific sub-fields.

But is such power deserved? And could it not be dangerous? After all the efficacy or otherwise of the GCM in predictive terms cannot be tested more rigorously in scientific terms until sometime into the next millennium. A large part of the scientific community believes that climate change is unlikely to be predictable in a precise fashion. Hence, climate change management strategies based on the availability of precise-knowledge of short- and medium-term events are unlikely to succeed [94].

The impression that climate change can be so predicted and managed is not only misleading, but it could also have negative repercussions should policy makers act on this assumption. Such negative consequences might occur through the subsequent failure to develop policies which encourage societal *resilience* to unpredictable change. Additionally, there is a danger that policy authorities will appear to others to have 'solved the problem', whereas they will have done so only in the unlikely event that the predictions or projections made are correct. The societal perception that the 'climate change problem' is being adequately handled could inhibit the emergence of, and support for, creative social, policy and economic responses to the challenge of coping with a possibly inherently unpredictable system such as climate.

7. Conclusions

We have argued that the common relative ordering of different climate models – which champions GCMs as the 'ultimate' climate model – is not fully justified by scientific arguments alone, at least not for the simulation and projection of future climate change. The criteria by which GCMs are conventionally assessed as being superior to other climate models are at least questionable from without (and to some extent from within) the GCM community. In assessing GCMs one is drawn into an open-ended discussion concerning:

- how the inclusion of perceived complexity should be traded-off against a better understanding and analysis of dominant model processes, interactions, feedbacks and uncertainties;
- what likely scientific and technological advances can be anticipated;
- what complexity is and how it can or should be modelled and measured; and,
- the impossibility of validating such large, complex models.

The existing dominant notion of how to model complexity in the climate change field is accepted by climate modellers as the most adequate. However, complexity and uncertainty are explored using different modelling philosophies and styles in other areas of environmental science. These alternative approaches are taken as defining good science in those fields. Yet such methods are not being sufficiently assimilated into 'state of the art' global climate modelling.

In a situation where the system is nominally very complex and the data are scarce, it often makes good sense to construct high order simulation models which represent the 'state-of-the-art' scientific understanding of the system. Indeed, large, complex simulation models play an important part in many fields of science and engineering and the second author has been involved in many studies which have exploited such models to good effect [95]. Consequently, let there be no doubt that we believe models such as the GCMs and their less complex relatives, the global carbon cycle models, are valuable elements in climate research. But, as we have tried to demonstrate, such large deterministic simulation models have their limitations and are not the *only* type of model that should be considered when the uncertainties, or even ambiguities, inherent in the climate system are taken into consideration. There are, we believe, clear advantages in a 'spectrum' of models that are designed to satisfy different objectives and which *together* provides the scientific input to policy making exercises.

Some may argue that the diversity we call for already exists in practice. That is clearly true to an extent. But only to an extent, and one which is not reflected in policy and user circles. What we are arguing, therefore, is that the choice of method should depend more explicitly on the sort of multi-dimensional criteria suggested here and that there should be a more *explicit* debate on goals, epistemology, experiences, and the policy context. Additionally, the reasons for adopting particular methods should be better appreciated and acknowledged. This, we believe, will assist in a better role of science in policy making and in public debate, and allow for science to be used more intelligently (and realistically) by the user community. For example, if we accept that GCMs are supported in part because they help to create an interactive and international community of scientists and policy makers – and not just because they are 'the best' model – then we can at least question and discuss whether that is the sort of community we want to emerge to inform our political processes.

We have suggested some possible sociological reasons why GCMs have come to occupy a predominant position, especially their unique ability to integrate and implicitly support diverse areas of scientific research and climate policy. We have hinted that the lack of recognition of such factors is especially acute for complex and comprehensive models, since the judgements and choices taken – scientific, social and policy – become more easily hidden in the complexity of the model. The range of model components and interlinkages, the craft-ladenness of their construction and use, and their history (having being produced by teams, often over decades) means that they are enormously difficult to comprehend at an overarching level, or in terms of the dominant model interactions, even for many of the modellers themselves.

Integration of sub-models within GCMs provides a pragmatic solution to the horrendous problems of delimiting and characterising the system to be modelled, which is itself ever more comprehensive and globally inclusive. It is also a way of coping with uncertainties, ambivalence, scientific judgements and disagreements, though it frequently does so largely through hiding such issues in the complexity of the model. The effect of the networking around the 'hub' of GCMs is that implicit political and value judgements, assumptions and commitments are being incorporated, though not through deliberate debate or the conscious intent of any one actor: rather, these emerge from the network. The incorporation of such implicit commitments and beliefs into models also contributes to their removal from wider discussion and debate in more political contexts (though we are not suggesting this is the intended outcome). The achievement of an apparently resilient universality for the scientific culture across all its original disparity, and despite its remaining incoherence, is nonetheless an enormous social, political and intellectual accomplishment.

At one level, the success of such model-based networks in policy may depend on their denial of the more open-ended scientific and social choices, in favour of their *presentation* as objectively given by natural science, in which politics and values play little or no part. This is largely because of the consequent exclusion of all social and meta-scientific issues from the realm of legitimate debate, as well as of the actors promoting them; a delimitation which facilitates political consensus-formation.

Yet this apparent success may be earned only at the cost of a lack of effectiveness at different, and arguably more important, levels of the policy process; in particular, at the stages of influencing change in individual and collective perceptions and behaviour to – *inter alia* – reduce greenhouse gas emissions, reduce energy use, change consumption patterns and so forth. For example, the more the scientific and social choices are effectively 'delegated' to the model [96], the more ambivalence is likely amongst those policy actors implicated. In consequence, subtle political and institutional changes could result in radical shifts in the perception of GCMs. If our analysis is at all valid, however, the credibility of GCMs would suffer, perhaps greatly, were political and policy actors to display relatively little interest in them. The cost of impenetrability is fragility of the entire edifice.

Even the conscious deliberation of all scientific social and policy judgements and choices by the present community of modellers, other research scientists, research managers and policy-makers, could not possibly illuminate adequately the validity and wisdom of many of the assumptions and judgements involved in the construction and scientific policy use of GCMs. That task requires at least an understanding of how climate and climate change is perceived by a range of policy actors and the public. Such an understanding could probably only come from a more open and reflective discussion in a newly constituted public domain [97].

Nor does robustness in the social understanding incorporated into the policy use of climate science imply correctness, since many of the pre-suppositions will be legitimate interpretations based on judgement and, frequently, more than one alternative will present itself. In the absence of a single 'correct' analysis, responsiveness to the outcome of policy based on a particular set of assumptions or judgements is important. In other words, iterative learning from trial and error should be a necessary part of the policy and research process; and such openness to learn and change is what constitutes a resilient policy and scientific approach to the climate change issue. This implies no less than the development of new public institutions and processes for the interaction of science, policy and the public. We envisage that many of these institutions will be largely exploratory and developed incrementally to better define who might be involved in exploring scientific and social assumptions and commitments, as well as critically reviewing the effects of policy based upon them.

If our analysis of the mutual reinforcement of particular scientific and policy approaches has any purchase, the above processes will also raise questions about the relative merits of alternative scientific methods and styles. Indeed, we anticipate that such a discussion will become vitally necessary to the provision of insights and understanding into the range and potentialities of intellectual tools for the support of policy options and public debate.

In this spirit, we finish by inviting climate scientists to take part in such discussions: in a sense, to (re)consider the scientific and implicit social and policy commitments they have entered into through pursuit of a particular scientific research programme. Only through such debate will we be able to more fully address the question of whether the lack of recognition of the alternative scientific, social and policy elements in the definition of ‘good science’ has held back some of the alternative approaches which could be further developed and used profitably in the widening climate science and policy domain.

Acknowledgements

This paper has benefited greatly from the help and advice of James Risbey (Carnegie Mellon University) and Peter Stone (Massachusetts Institute of Technology). Ian Enting (CSIRO, Division of Atmospheric Research) provided us with a copy of the Enting-Lassey global carbon cycle model for which we are most grateful. We are also grateful to the following for their generous assistance and readiness to respond to our questions: David Bennnetts, David Carson, Chris Folland, Klaus Hasselmann, Paul Jarvis, Geoff Jenkins, John Houghton, Mick Kelly, Adrian Kelsey, Mike Hulme, Richard Lindzen, Jerry Mahlman, Suki Manabe, Paul Mason, John Mitchell, Tim Palmer, Roger Pielke, Colin Prentice, Stefan Rahmstorf, Michael Schlesinger, Edward Schneider, Ron Stouffer, David Viner, Hans von Storch, David Warrilow, Warren Washington, Robert Watson, Ian Woodward, Tom Wigley, and Carl Wunsch. We are especially grateful to Stephen Schneider, Paul Edwards, Naomi Oreskes and two anonymous reviewers, for their further advice and guidance. It goes without saying that the opinions expressed in this paper are not attributed to any of the above. The research was funded by the Economic and Social Research Council (grant number Y302/28/3001) and the Natural Environment Research Council (grant number GT4/90/AAPS/27) both of the U.K.

Notes and References

1. Levins, R.: 1966, ‘The Strategy of Model Building in Population Biology’, *Amer. Sci.* **54** (4), 421–431.
2. Lindblom, C. and Cohen, D.: 1979, *Usable Knowledge*, Yale University Press, London, p. 129; Hogwood, B. and Gunn, L.: 1984, *Policy Analysis For The Real World*, Oxford University Press, Oxford, p. 289; Brunner, R.: 1996, ‘Policy and Global Change Research: A Modest Proposal’, *Clim. Change* **32** (2) 121–147.

3. See, however: Glantz, M.: 1979, 'A Political View of CO_2', *Nature* **280**, 189–190; Stewart, T. and Glantz, M.: 1985, 'Expert Judgement and Climate Forecasting: A Methodological Critique of "Climate Change to the Year 2000" ', *Clim. Change* **7**, 159–183; Schneider, S.: 1985, 'Science by Consensus: The Case of The National Defense University Study "Climate Change to the Year 2000" – An Editorial', ibid., pp. 153–157; Edwards, P.: 1996, 'Global Comprehensive Models in Politics and Policymaking', *Clim. Change* **32** (2), 149–161; Pielke, R. Jr.: 'Usable Information for Policy: an Appraisal of the U.S. Global Change Research Program', *Policy Sci.* **28**, 39–77; Jamieson, D.: 1991, 'The Epistemology of Climate Change: Some Morals for Managers', *Soc. Natural Res.* **4**, 319–329.
4. For example, Hajer, M.: 1995, *The Politics of Environmental Discourse*, Oxford University Press, Oxford, p. 332; Haas, P.: 1992, 'Introduction: Epistemic Communities and International Policy Coordination', *Int. Organ.* **46** (1), 1–35.
5. Grotch, S. and MacCracken, M.: 1991, 'The Use of General Circulation Models to Predict Regional Climate Change', *J. Clim.* **4**, 286–303; von Storch, H. and Hasselmann, K.: 1995, *Climate Variability and Change*, Report no. 152, Max-Planck-Institut für Meteorologie, Hamburg, p. 26.
6. Natural Science and Technology Council: 1996, *Our Changing Planet: The FY 1996 U.S. Global Change Research Program*, A report by the Subcommittee on Global Change Research, Committee of Environment and Natural Resources Research, A Supplement to the President's Fiscal Year 1996 Budget, Washington, D.C., p. 152.
7. Page 82, Shine, K. and Henderson-Sellers, A.: 1983, 'Modelling Climate and the Nature of Climate Models: A Review', *J. Climatol.* **3**, 81–94.
8. Page 55, Henderson-Sellers, A. and McGuffie, K.: 1987, *A Climate Modelling Primer*, John Wiley, Chichester, p. 217.
9. Wigley, T. and Raper, S.: 1992, 'Implications for Climate and Sea Level of Revised IPCC Emissions Scenarios', *Nature* **357**, 293–300.
10. Page 84 in [7], emphasis added.
11. Page 17, Schneider, S.: 1992, 'Introduction to Climate Modeling', in Trenberth, K. (ed.), *Climate System Modeling*, Cambridge University Press, Cambridge, pp. 3–26.
12. In a further contribution to the book, Meehl extends the concept of a model hierarchy to coupled models, relating comprehensiveness to realism and reliability of model output. (Meehl, G.: 1992, 'Global Coupled Models: Atmosphere, Ocean, Sea Ice', in Trenberth, K. (ed.), *Climate System Modeling*, Cambridge University Press, Cambridge, pp. 555–581; Kiehl, J.: 1992, 'Atmospheric General Circulation Modeling', in Trenberth, K. (ed.), *Climate System Modeling*, Cambridge University Press, Cambridge, pp. 319–369.)
13. Page 760, Schlesinger, M. and Mitchell, J.: 1987, 'Climate Model Simulations of the Equilibrium Climatic Response to Increased Carbon Dioxide', *Rev. Geophys.* **25**, 760–798. Similar sentiments are expressed in the IPCC 1995 report, e.g., see Gates, L. et al.: 1996, 'Climate Models – Evaluation', in Houghton, J. et al. (eds), *Climate Change 1995: The Science of Climate Change*, Cambridge University Press: Cambridge, pp. 572.
14. In evidence before the Subcommittee on Environment of the Committee on Science, Space and Technology of the U.S. House of Representatives (in 1992) the following statements have been made by research managers: 'I think we are confident in their [GCM's] fundamental physics. They are based on fundamental physical principles that we believe in. Of course there are many uncertainties associated with the parameterisations . . . . The clouds, the oceans, the ecosystems, and simply just a matter of resolution' (oral evidence of Dr. A. Patrinos, U.S. Department of Energy, U.S. Congress, House, 1992. Committee on Science, Space and Technology. *In U.S. Global Change Research Program*, Hearing before the Subcommittee on Environment, 102nd Congress, 2nd Session, 5 May, 1992, p. 91. Washington, D.C., U.S. Government Printing Office, p. 200). And: 'Right now, those GCMs incorporate the best scientific knowledge, the best understanding of the physics and transport of the atmosphere, and some coupled models include representations of the ocean as well that the science community can give you' (oral evidence of Dr. C. Riordan, U.S. Environmental Protection Agency, ibid, p. 92). Meanwhile, in the U.K., Dr. John Houghton in a speech delivered at the opening of the Hadley Centre (at which Prime

Minister Thatcher was in attendance) stated that: ‘The numerical model is an effective way – in fact the only way we know – of dealing in a meaningul way with all these non-linearities involved. . . . these is a good expectation of being able, through the use of models, to predict at least in general terms the change in climate which is likely to occur because of man’s increased burning of fossil fuels’ (text of speech delivered 25th May 1990).

15. Interview with a GCMer by first author, 11 October 1992. In this paper we use evidence obtained from semi-structured interviews with climate scientists and policy makers. Where quoting interviewees in this paper, we have sometimes slightly edited the verbatim statements to clarify what we perceived were the intended meanings. Where the citation is to a ‘personal communication’ this indicates a written communication.
16. Personal communication to first author from a GCMer, 22 May 1995.
17. He told us: ‘I frequently find that when I submit a proposal or a paper which involves some of the simpler models I run into the criticism – what’s the point of doing this? Why didn’t you use a GCM? A GCM is obviously better!’. The same scientist told us of an application he had submitted to a funding agency which was turned down on the basis of peer-reviews, one of which assessed the proposal as using an over-simple model, whilst another thought it intended to use an over-complex model. He described a subsequent conversation with an agency official as follows: ‘We had a long conversation and it was a little frustrating because he was not willing to be too explicit. . . . He talked at great length about all the GCM efforts that he is supporting. . . . He also mentioned that he was aware of my criticisms of GCMs and I couldn’t help wondering if that had something to do with this decision’. (Interview with first author, 19 April 1994.)
18. E.g., p. 18, Schneider [11]. Many well known climate scientists have developed and used simple models including Stephen Schneider, Martin Hoffert, Danny Harvey, Tom Wigley, and Peter Stone.
19. Interview with two climatologists by first author, 5 and 6 August 1992.
20. Ibid. Also, Land, K. C. and Schneider, S. H.: 1987, ‘Forecasting in the Social and Natural Sciences: An Overview and Analysis of Isomorphisms’, *Clim. Change* **11**, 7–31. Our argument is an inevitable simplification of how GCMs are perceived and in reality perceptions of GCMs are often more finely-textured and undergoing change. For example, new relationships have been emerging between GCMers and other scientists, such as ecologists and hydrologists, in which GCMs are less predominant (e.g., Root, T. L. and Schneider, S. H.: 1995, ‘Ecology and Climate: Research Strategies and Implications’, *Science* **269**, 334–341); and in the U.S.A. there are not only more research centres than in Europe, but greater diversity in the key issues pursued. However, GCMs still dominate in the U.S.A. when it comes to providing policy-relevant knowledge.
21. For example, the Policymakers Summary of the IPCC’s 1990 report commented that: ‘These models (GCMs) are based on the laws of physics and use descriptions in simplified physical terms (called parameterisations) of the smaller-scale processes . . . ’ (IPCC, 1990: xxv). In the IPCC’s 1995 report the Technical Summary states that: ‘Many physical processes, such as those related to clouds, take place on much smaller spatial scales and therefore cannot be properly resolved and modelled explicitly, but their average effects must be included in a simple way by taking advantage of physically based relationships with the larger scale variables (a technique known as parametrization)’ (p. 31, Technical Summary, IPCC 1996, *Climate Change 1995: The Science of Climate Change*, Cambridge University Press, Cambridge).
22. P. 19 in Schneider [11]. The IPCC’s presentation is not inconsistent with Schneider, but rather incomplete and with a different emphasis, e.g., upon the physical basis of parameterisations rather than their empirical statistical dimension.
23. Kiehl, in [12]. See also Kiehl, J. and Williamson, D.: 1991, ‘Dependence of Cloud Amount on Horizontal Resolution in the National Center for Atmospheric Research Community Climate Model’, *J. Geophys. Res.* **96**, 10955–10980; Slingo, T.: 1990, ‘Sensitivity of the Earth’s Radiation Budget to Changes in Low Clouds’, *Nature* **243**, 49–51; Randall, D., Harshvardhan, Dazlich, D., and Corsetti, T.: 1989, ‘Interactions among Radiation, Convection, and Large-Scale Dynamics in a General Circulation Model’, *J. Atmos. Sci.* **46** (13), 1943–1970.
24. Personal communication to first author from a GCMer, 7 February 1995.

25. Risbey has elaborated on this point: 'GCMs aim to account for physical processes from first principles, i.e., from the known laws governing the behaviour of the relevant phenomenon. Simpler models generally do not make this attempt, accepting more aggregate and empirically based representations. In practice however this distinction is not as strong as it sounds in practice, since even GCMs are highly dependent on empirical data in their parameterizations, albeit at smaller scales. This is (one reason) why the so-called "reductionist approach" of GCMs "does not automatically guarantee physical realism" ' (personal communication, September 1996).
26. Stone, P. and Risbey, J.: 1990, 'On the Limitations of General Circulation Climate Models', *Geophys. Res. Lett.* **17** (12), 2173–2176; Stone, P. and Yao, M-S.: 1991, 'Vertical Eddy Heat Fluxes from Model Simulations', *J. Clim.* **4** (3), 304–317.
27. Interview with a climate modeller by first author, 10 April 1994. (Michaud, R. and Derome, J.: 1991, 'On the Meridional Transport of Energy in the Atmosphere and Oceans as Derived from Six Years of ECMWF Analyses', *Tellus* **43A**, 1–14; Oort, A. H.: 1978, 'On the Adequacy of the Rawinsonde Network for Global Circulation Studies Tested through Numerical Model Output', *Mon. Wea. Rev*. **106**, 174–195; Oort, A. H.: 1983, *Global Atmospheric Circulation Statistics, 1958–1973*, NOAA Professional Paper 14, U.S. Government Printing Office, Washington, D.C.; Keith, D. W.: 1995, 'Meridional Energy Transport: Uncertainty in Zonal Means', *Tellus* **47A** (1), 30–44; the on-going debate is reflected in Trenberth, K. and Solomon, A.: 1994, 'The Global Heat Balance: Heat Transports in the Atmosphere and Ocean', *Clim. Dyn*. **10**, 107–134). A major comparison of models with respect to their heat transport has now been conducted (Glecker, P. et al.: 1995, 'Cloud Radiative Effects on Implied Oceanic Energy Transports as Simulated by Atmospheric General Circulation Models', *Geophys. Res. Lett.* **22** (7), 791–794) which reveals the extent of the probable errors).
28. Pp. 2175 and 2176 (our italics) respectively, in [26].
29. Young, P., Parkinson, S., and Lees, M.: 1996, 'Simplicity out of Complexity: Occam's Razor Revisited', *J. Appl. Statis.* **23**, 165–210.
30. Interview by first author with a GCM modeller, 11 March 1993.
31. Some evidence for this view comes from research with relatively simple models (e.g., Marotkze, J. and Stone, P.: 1995, 'Atmospheric Transports, the Thermohaline Circulation, and Flux Adjustments in a Simple Coupled Model', *J. Phys. Ocean.* **25**, 1350–1364; Nakamura, M., Stone, P., and Marotzke, J.: 1994, 'Destabilisation of the Thermohaline Circulation by Atmospheric Eddy Transport', *J. Clim.* **7**, 1870–1882). Note, however, that not all GCMers would accept this evidence since simple models were used in these experiments.
32. It has been suggested to us that the evaluation of climate models has become increasingly formalised since the early 1990s, with projects such as the Atmospheric Model Intercomparison Project (AMIP) and its off-shoots (Mike Hulme, personal communication, February 1997). An interesting question is the extent to which such formalisation is related to the perceived policy requirement for validation in a context of uncertain science for policy, rather than such a process being driven solely by the climate modelling community. There are generic similarities to the strategies of molecular biologists and particle physicists (Knorr-Cetina, K.: 1991, 'Epistemic Cultures: Forms of Reason in Science', *Hist. Political Econ.* **23**, 105–122).
33. A discussion and classification of the limitations and uncertainties of climate and integrated assessment models can be found in van der Sluijs, J.: 1997, *Anchoring Amid Uncertainty*: *On the Management of Uncertainties in Risk Assessment of Anthropogenic Climate Change*, Ludy Feyen, Leiden, p. 260.
34. Note that in mathematical statistics and engineering, 'over-parameterisation' is usually taken to mean that the model has more parameters (coefficients) than can be justified by the available data on which the model is based: this is, of course, quite different from the usage here.
35. Personal communication by first author with Dr. James Risbey, 6 July 1994.
36. Hempel, C. G.: 1966, *Philosophy of Natural Science*, Prentice-Hall, Englewood Cliffs, NJ, p. 116; Hesse, M.: 1974, *The Structure of Scientific Inference*, Macmillan, London, p. 309; Popper, K.: 1959, *The Logic of Scientific Discovery*, Hutchinson, London, p. 480.
37. P. 234 in Hesse [36].

38. P. 44 in Popper, K.: 1982, *The Open Universe*, Hutchinson, London, p. 185; Tennekes, H.: 1992, 'Karl Popper and the Accountability of Numerical Weather Forecasting', *Weather* **47**, 343–346.
39. Gieryn, T.: 1995, 'Boundaries of Science', in Jasanoff, S. et al. (eds.), *Handbook of Science and Technology Studies*, Sage, London, pp. 393–443; Collins, H.: 1992 (1985), *Changing Order: Replication and Induction in Scientific Practice*, Chicago University Press, London, p. 187.
40. Shackley, S. and Wynne, B.: 1995, 'Integrating Knowledges for Climate Change: Pyramids, Nets and Uncertainties', *Global Environ. Change* **5** (2), 113–126.
41. Cartwright, N.: 1994, 'Fundamentalism vs. the Patchwork of Laws', *Proc. Aristotelian Soc.* **94**, 279–292; Hacking, I.: 1992, 'The Self-Vindication of the Laboratory Sciences', in Pickering, A. (ed.), *Science as Practice and Culture*, Chicago University Press, Chicago, p. 474.
42. Interview by first author with a GCMer, 10 September 1992.
43. We avoid use of the term 'reductionism' here because of its commonly-held definition by climate (and other environmental) modellers. At least two definitions seem to be used. The more common definition of reductionism – namely that the environmental system is analytically decomposable to the governing physical laws – is widely accepted by modellers (a viewpoint we will critically explore in the next section). At the same time, however, many climate modellers are keen to distance themselves from the 'high-physics' approach to representing systems by employing a few rather abstract physical principles, a stance which they also appear to have negatively characterised as reductionism. Historically, part of the reason for the development of this second perspective may have been to counter the improper extension of a physics-derived commitment to simplicity into the subject matter of the environmental and earth sciences. Physical complexity was at least a step away from such overt reductionism. (Personal communication to first author from Dr. Naomi Oreskes, 5 February 1996.) Discussing the implications for model validation, Oreskes et al. have noted that the more sophisticated the model becomes, the more difficult it is to provide confirmation or falsification of the model (see also, Oreskes, N., Shrader-Frechette, K., and Belitz, K.: 1994, 'Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences', *Science* **263**, 641–646).
44. Palmer, T. N.: 1993, 'A Nonlinear Dynamical Perspective on Climate Change', *Weather* **48** (10), 314–326.
45. Ibid.: 324.
46. As one modeller puts it: 'there is a view that if you are only interested in global average temperature then you don't really need GCMs – that one dimensional models will be more or less correct. And that you only need the GCM models if you are interested in homing-in on a specific region. But I think that is a very simplistic view of the world and you just cannot make that assumption in a non-linear system. That is really important' (interview by first author with a GCMer, 21 June 1993). A problem with Palmer's analysis, however, is that it is based on transient results only, whereas thermodynamic argument would suggest it applies much less forcefully for the case of an equilibrium response (personal communication by first author with a climate modeller, 6 June 1994).
47. A further example which possibly illustrates the need for GCMs is the model representation of the thermohaline circulation. Simple models tend to produce a much more unstable thermohaline than models with higher degrees of freedom, though the modelling evidence is somewhat mixed. The reliability of the representation of the thermohaline circulation in current coupled AOGCMs, and its sensitivity to changes in freshwater flux and temperature, remains questionable, however (see, e.g., Rahmstorf, S.: 1995, 'Bifurcations of the Atlantic Thermohaline Circulation in Response to Changes in the Hydrological Cycle', *Nature* **378**, 145–149).
48. Cf. Ascher, W.: 1981, 'The Forecasting Potential of Complex Models', *Policy Sci.* **13**, 247–267.
49. Enting, I. G. and Pearman, G. I.: 1987, 'Description of a One-Dimensional Carbon Cycle Model Calibrated Using Techniques of Constrained Inversion', *Tellus* **39B**, 459–476; Wigley and Raper [9]; Enting, I. and Lassey, K.: 1993, *Projections of Future* CO_2, Division Of Atmospheric Research, Technical Paper 27, CSIRO, Australia, p. 42; Schimel, D., Enting, I. G., Heimann, M., Wigley, T. M. L., Raynaud, D., Alves, D., and Siegenthaler, U.: 1995, 'CO_2 and the Carbon Cycle', in Houghton, J., Meira Filho, L., Bruce, J., Lee, H., Callander, B., Haites, E., Harris, N., and Maskell, K.: 1995, *Climate Change 1994: Radiative Forcing of Climate Change and*

an Evaluation of the IPCC IS92 Emission Scenarios, Cambridge University Press, Cambridge, p. 339.

50. Gardner, R. H. and Trabalka, J. R.: 1985, *Methods of Uncertainty Analysis for a Global Carbon Dioxide Model.* Publication DOE/OR/21400-4. U.S. Dept. of Energy, Washington D.C., p. 41.
51. Schneider, E.: 1996, 'Flux Correction and the Simulation of Changing Climate', *Annales Geophysicae* **14**, 336–341.
52. Enting and Lassey [49].
53. Further details can be found in Parkinson, S.: 1995, *The Application of Stochastic Modelling Techniques to Global Climate Change*, Ph.D. Thesis, Lancaster University, U.K., p. 263; Young et al., 1996 [29]; Parkinson, S. and Young, P.: 1997, 'Uncertainty and Sensitivity in Global Carbon Cycle Modelling', submitted manuscript.
54. Huggett, R.: 1993, *Modelling the Human Impact on Nature: Systems Analysis of Environmental Problems*, Oxford University Press, Oxford, p. 192.
55. See e.g., Scavia, D.: 1993, 'Lake Ecosystem Modelling', in Young, P. (ed.), *Concise Encyclopedia of Environmental Systems*, Pergamon Press, Oxford, pp. 318–320.
56. One GCM group recently wrote a paper entitled: 'Monte Carlo climate change forecasts with a global coupled ocean-atmosphere model' (Cubasch, U., Sausen, R., Maier-Reimer, E., and Voss, R.: 1994, 'Monte Carlo Climate Change Forecasts with a Global Coupled Ocean-Atmosphere Model', *Clim. Dyn.* **10**, 1–19). In this study they ran the AOGCM model only four times, which according to the standards of many modellers would not count as MCS at all! Nevertheless the significance of the model's initial conditions was clearly indicated, so it had some positive contribution.
57. We feel it is necessary to clarify our definitions of 'deterministic' and 'stochastic', since there appears to be some variation in their meanings. A *deterministic model* for us is one whose temporal evolution can be calculated exactly, given a set of initial conditions. A *stochastic model*, on the other hand, is one whose parameters and/or inputs are defined probabilistically, so that the model response can only be defined in probabilistic terms. Thus, a *model* of the global climate system can be deterministic or stochastic, depending upon the views of the model builder. But uncertainty is surely unavoidable in any investigation of the natural environment and it seems more appropriate that it should be considered in stochastic terms.
58. See [53].
59. It is also possible, however, that this assumption may also lead to lower estimates of uncertainty. It depends very much upon the nature of the model being considered – see discussion in Parkinson and Young (1997) [53].
60. See [53].
61. This empiricist approach to model evaluation would not be accepted by those using a more formal structural epistemology.
62. Schimel et al., Enting and Lassey [49]), Wigley and Raper [9], and Parkinson and Young [53]. The uncertainty in each of the deterministic exercises is due to the following: variation between different models; variation in one model due to different assumptions about feedbacks; and limited parametric uncertainty.
63. The uncertainty (66% confidence limits) of the pre-industrial level of CO_2, $(p_a)_{pi}$ is thus reduced from 275 ± 15 ppmv to 275 ± 7.5 ppmv which is closer to the uncertainty discussed by Schimel et al. [49] – a value published too late to be considered explicitly by this analysis.
64. Schimel et al. [49].
65. See [53].
66. Young, P.: 1984, *Recursive Estimation and Time-Series Analysis*, Springer-Verlag, Berlin, p. 300; Young, P.: 1985, 'The Instrumental Variable Method: A Practical Approach to Identification and System Parameter Estimation', in Barker, H. and Young, P. (eds.), *Identification and System Parameter Estimation: Vol. 1 and 2*, Pergamon Press, Oxford, pp. 1–16.
67. See [53]. Note that the structural decomposition of the transfer function is ambiguous: other decompositions are possible, including feedback structures that may provide a better physical interpretation than the parallel decomposition. However, such discrimination is not important to the discussion here.

68. The input due to land-use changes is set to zero in the original model in order to carry out this comparison. Hence, a lower atmospheric carbon dioxide level than expected is seen in Figure 2 during the simulation.
69. These values come from the continuous-time reduced order model reported in Young and Parkinson (1996), which explains the EL model data a little better than the discrete-time model described previously in Young et al. (1996) [29]. However, the time constants of both models are quite similar and the mechanistic interpretations are the same. (Young, P. and Parkinson, S.: 1996, *Simplicity out of Complexity in Forecasting Climate Change*, Technical Note Centre for Research on Environmental Systems and Statistics, Lancaster University, p. 37.)
70. This final stage of interpreting the reduced order model in physically meaningful terms is part of a technique which we term Data-Based Mechanistic (DBM) modelling; a technique which can be applied within DMA analysis to simulated data and, more generally, to real data (see Young, P. and Beven, K.: 1994, 'Data-Based Mechanistic Modelling and the Rainfall-Flow Nonlinearity', *Environmetrics* **5**, 335–363 (Special Issue on 'Environmental Time Series Analysis')).
71. Enting and Lassey, Enting and Pearman [49].
72. The IPCC 1990 report itself stated that: 'Because running coupled ocean-atmosphere GCMs is expensive and time-consuming, many of our conclusions about global trends in future climates are based upon simplified models' (Houghton, J., Jenkins, G., and Ephraums, J.: 1990, *Climate Change: The IPCC Scientific Assessment*, Cambridge University Press, Cambridge, p. 365). Henderson-Sellers has presented similar arguments to ours. (Henderson-Sellers, A.: 1996, 'Climate Modelling, Uncertainty and Responses to Predictions of Change', *Mitigat. Adapt. Strat. Global Change* **1**, 1–21; and 'Bridging the Climate Gap', in Giambelluca, T. and Henderson-Sellers, A. (eds.) 1996, *Climate Change: Developing Southern Hemisphere Perspectives*, John Wiley, Chichester, pp. 35–60.)
73. We have also observed the enthusiastic response at a meeting of IPCC advisory scientists and policy makers to Wigley's 1-D-model, because of its perceived policy-usefulness.
74. Interview by first author with a government official, 23 February 1993.
75. Such dependency could have serious implications *vis-à-vis* the commitment of developing countries to climate change as a policy issue. Given a context of distrust between policy makers in the North and South, the question must be asked of whether GCM climate change scenarios developed in the North will be trusted, especially given the difficulty of independently evaluating the judgement which goes into construction of those scenarios.
76. Polanyi, K.: 1958, *Personal Knowledge*, Routledge and Kegan Paul, London, p. 428.
77. Leading scientists themselves may prefer the level of control over scientific assessment offered by fewer centres. For example, a previous director of the U.K. Meteorological Office, Sir John Mason, stated before the House of Lords Select Committee on Sustainable Development in 1994 that: '... the number of people or the number of groups around the world involved in climate modelling has increased rather dramatically in the last two or three years and not wholly for the better in my view. It is easier when you have only three or four major groups, particularly if one of them is in the United Kingdom, to make judgments than when there are a lot of people in the field and perhaps introducing more turbulence into the subject' (p. 15, Minutes of Evidence taken before the Select Committee on Sustainable Development, Tuesday 3 May 1994, London, HMSO, HL Paper 66-ii).
78. Interview by first author with a GCMer, 11 April 1994. The episode is mentioned in Gore, A.: 1992, *Earth in the Balance*, Houghton Miflin Company, New York, p. 407. A similar point is made by Parsons, E.: 1995, 'Integrated Assessment and Environmental Policy Making: In Pursuit of Usefulness', *Energy Policy* **23**, 463–476.
79. Technology Review: 1992, 'The Political Pleasures of Engineering: An Interview with John Sununu', *Technol. Rev.* **95**, August/September, 22–28.
80. Research results do not clearly indicate the effects of changes in the THC on the take-up of heat, and subsequent feedbacks upon surface temperature and precipitation. See, e.g., Harvey, L. D. Danny: 1994, 'Transient Temperature and Sea Level Response of a Two-Dimensional Ocean-Climate Model to Greenhouse Gas Increases', *J. Geophys. Res.* **99** (C9), 18,447–18,466.

81. A further example is the argument by Schlesinger and Jiang that a ten year delay in greenhouse gas emission reductions would have a minimal effect on potential warming. Schlesinger and Jiang were criticised by Risbey, Handel and Stone for basing strong policy conclusions on simple climate models which failed to take account of non-linear, possibly abrupt climate change, as well as probably under-recognising the extent of regional climate change. The debate between Schlesinger and Jiang and Risbey, Handel and Stone subsequently revolved, however, around different and conflicting interpretations of the *same* GCM model runs, indicating that it was, in practice, difficult to divorce the evaluation of the simple models from more complex ones, and vice versa. A further difference between Schlesinger and Jiang and Risbey, Handel and Stone was the degree to which models, whether simple or GCMs, were held to be sufficiently robust to act as the basis of policy-decisions to delay action. Schlesinger and Jiang defended the use of both types of models for this purpose, whilst Risbey, Handel and Stone implied that neither simple models nor GCMs were currently adequate to act as the basis for such decisions. (Schlesinger, M. and Jiang, X.: 1991, 'Revised Projection of Future Greenhouse Warming', *Nature* **350**, 219; Schlesinger, M. and Jiang, X.: 1991, 'A Phased-In Approach to Greenhouse-Gas-Induced Climatic Change and Climatic Response to Increasing Greenhouse Gases', *Eos Trans. A.G.U.* **72** (53), 596–597; Risbey, J., Handel, M., and Stone, P.: 1991, 'Should We Delay Responses to the Greenhouse Issue?' and 'Do We Know What Difference a Delay Makes?', *Eos Trans. A.G.U.* **72** (53), 593.)
82. In this section we draw upon theories developed in the field of social studies of science. See, for example, Callon, M.: 1986, 'Some Elements of a Sociology of Translation: Domestication of the Scallops and of the Fisherman of St. Brieuc Bay', in Law, J. (ed.), *Power, Action and Belief*, Methuen, London, pp. 196–233; Latour, B.: 1988, *The Pasteurisation of France*, Harvard University Press, London, p. 273; Latour, B.: 1992 (1987), *Science in Action*, Harvard University Press, London, p. 274.
83. Interviews by first author with global vegetation dynamics modellers, 15 December 1993 and 16 September 1993.
84. Compare, Perry, J.: 1992, *The United States Global Change Research Program: Early Achievements and Future Directions*, National Academy Press, Washington D.C., p. 20.
85. As a senior scientist at a GCM modelling centre put it: 'This (modelling centre) of course, is in a sense the tip of an iceberg though. It depends on a community of international and national research to actually achieve its task . . . if you don't have something like (the modelling centre), you actually run the risk that you make a lot of investment but as a nation, actually, you lack perhaps the primary benefit of that integrated output. . . . And I think that's one of the frustrations, perhaps, of some of the smaller European nations, that they actually contribute quite a bit to the science base of it, but they don't get the high profile of the actual prediction activity'. (Interview by first author with a research manager, 6 May 1993.)
86. Interview by first author with a GCM research manager 5 April 1993.
87. Interview by first author with an observational climatologist, 10 February 1993.
88. This was done by Parkinson and Young [53] in generating the results quoted in Section 5.
89. Parry, M.: 1990, *Climate Change and World Agriculture*, Earthscan, London, p. 157; Carter, T., Holopainen, E., and Kanninen, M. (eds): 1993, 'Techniques for Developing Regional Climatic Scenarios for Finland', *Publications of the Academy of Finland 2/93*, Painatuskeskus, Helsinki, p. 63; Viner, D. and Hulme, M.: *Climate Change Scenarios for Impact Studies in the U.K.*, Climatic Research Unit, University of East Anglia, p. 70.
90. Admittedly, economists also make extensive use of simple models. Nordhaus, W. D.: 1994, *Managing the Global Commons: The Economics of Climate Change*, MIT Press, London, p. 213.
91. It might be objected that this is a circular argument, since the status of GCMs in policy derives itself in part from the prominence and credibility of GCMs within science. However, what we are proposing is a process of mutual reinforcement of status in which both processes occur concurrently. (Shackley, S. and Wynne, B.: 1995, 'Global Climate Change: The Mutual Construction of an Emergent Science-Policy Domain', *Sci. Publ. Pol.* **22** (4), 218–230.)
92. It could be countered that versions of the science emphasising the climate system's chaotic nature might act to strengthen the rationale for global environmental policy action, especially if

the precautionary principle is accepted. Note, however, that a dominant response to the suggestion that the climate system might face abrupt, chaotic and unexpected changes, has been model-based analyses of whether this feature affects our ability to find an optimal solution to the problem of managing climate change. Hence it attempts to re-impose a control and management ethos at a subsequent level. (e.g., see: Lempert, R., Schlesinger, M., and Hammitt, J.: 1994, 'The Impact of Potential Abrupt Climate Changes on Near-Term Policy Choices', *Clim. Change* **26**, 351–376). Nevertheless, as James Risbey points out, Palmer's analysis may point to a happy marriage between chaos and GCMs, since the latter are needed to represent regime structure, identify climate attractors and perform ensemble climate forecasts. A similar point has been made by a GCM modeller, who noted that: 'GCMs provide the most practical means of investigating instability, given that the details of the "mean" climate or attractor determine the nature of the instability. . . . to get the right answer the detailed shape of the attractor may be important' (personal communication, November 1995).

93. An alternative argument, suggested by one of our referees, is that policymakers are not so much interested in 'good science' as in 'good scientists'. Although we can think of a few scientists who do not personally use GCMs and are also considered within the IPCC to be amongst the leading scientists, they still draw extensively upon GCMs in their provision of scientific advice. Moreoever, none that we can think of departs from the consensus view that GCMs are the 'best' amongst climate models (for the reasons discussed in Section 2.4). This cursory observation does not confirm the argument either way, but raises the significant question of whether the evaluation of scientists can be divorced from the methods and tools used by those scientists.
94. Many scientists would agree with this of course and it appears to be a major rationale for the development of new integrated assessment models (IAMs) as heuristic tools. However, IAMs are still considered by many scientists to be primitive when compared to GCMs and the extent to which they should be trusted as a aid to decision-making in climate negotiations surely needs to be critically discussed. See also van der Sluijs [33].
95. Young et al. [29].
96. Van Asselt, M.: 1994, *Global Integrated Assessment Models as Policy Support Tools*, Masters Thesis, University of Twente, Netherlands, p. 94.
97. For example, there is good reason to believe that the overwhelmingly technical framing of the climate change issue structures the policy agenda in a way which fails to engage with the diverse constituencies (such as local government, industry and lay members of the public) whose commitment would be necessary for any serious policy on global climate change. (Macnaghten, P., Grove-White, R., Wynne, B., and Jacobs, M.: 1995, *Public Perceptions and Sustainability in Lancashire: Indicators, Institutions and Participation*, Lancashire County Council, Preston, p. 96; Macnaghten, P. and Jacobs, M.: 1997, 'Public Identification with Sustainable Development: Investigating Cultural Barriers to Participation', *Global Environ. Change* **7** (1), 5–24.)

(Received 17 July 1995; in revised form 30 June 1997)