MIT Joint Program on the Science and Policy of Global Change

GLOBAL
SCIENCE
POLICY
CHANGE
MIT

Constraining Uncertainties in Climate Models Using Climate Change Detection Techniques

Chris E. Forest, Myles R. Allen, Peter H. Stone and Andrei P. Sokolov

The MIT Joint Program on the Science and Policy of Global Change is an organization for research, independent policy analysis, and public education in global environmental change. It seeks to provide leadership in understanding scientific, economic, and ecological aspects of this difficult issue, and combining them into policy assessments that serve the needs of ongoing national and international discussions. To this end, the Program brings together an interdisciplinary group from two established research centers at MIT: the Center for Global Change Science (CGCS) and the Center for Energy and Environmental Policy Research (CEEPR). These two centers bridge many key areas of the needed intellectual work, and additional essential areas are covered by other MIT departments, by collaboration with the Ecosystems Center of the Marine Biology Laboratory (MBL) at Woods Hole, and by short- and long-term visitors to the Program. The Program involves sponsorship and active participation by industry, government, and non-profit organizations.

To inform processes of policy development and implementation, climate change research needs to focus on improving the prediction of those variables that are most relevant to economic, social, and environmental effects. In turn, the greenhouse gas and atmospheric aerosol assumptions underlying climate analysis need to be related to the economic, technological, and political forces that drive emissions, and to the results of international agreements and mitigation. Further, assessments of possible societal and ecosystem impacts, and analysis of mitigation strategies, need to be based on realistic evaluation of the uncertainties of climate science.

This report is one of a series intended to communicate research results and improve public understanding of climate issues, thereby contributing to informed debate about the climate issue, the uncertainties, and the economic and social implications of policy alternatives. Titles in the Report Series to date are listed on the inside back cover.

Henry D. Jacoby and Ronald G. Prinn,
Program Co-Directors

For more information, contact the Program office:

MIT Joint Program on the Science and Policy of Global Change

Postal Address: 77 Massachusetts Avenue
MIT E40-271
Cambridge, MA 02139-4307 (USA)

Location: One Amherst Street, Cambridge
Building E40, Room 271
Massachusetts Institute of Technology

Access: Telephone: (617) 253-7492
Fax: (617) 253-9845
E-mail: *globalchange@mit.edu*
Web site: *http://web.mit.edu/globalchange/www/*

Printed on recycled paper

Constraining Uncertainties in Climate Models Using Climate Change Detection Techniques

Chris E. Forest[1][†], Myles R. Allen[2][†], Peter H. Stone[1] and Andrei P. Sokolov[1]

Abstract

Different atmosphere-ocean general circulation models produce significantly different projections of climate change in response to increases in greenhouse gases and aerosol concentrations in the atmosphere. The main reasons for this disagreement are differences in the sensitivities of the models to external radiative forcing and differences in their rates of heat uptake by the deep ocean. In this study, these properties are constrained by comparing radiosonde-based observations of temperature trends in the free troposphere and lower stratosphere with corresponding simulations of a fast, flexible climate model, using techniques based on optimal fingerprinting. Parameter choices corresponding either to low sensitivity, or to high sensitivity combined with slow oceanic heat uptake are rejected. Nevertheless, a broad range of acceptable model characteristics remains, such that climate change projections from any single model should be treated as only one of a range of possibilities.

Climate change detection studies are conducted by running a climate model under a set of prescribed forcings and then using statistical techniques to determine whether the simulated pattern of climate change is distinguishable in the observations at some level of confidence [*e.g.*, references 1, 2, 3, 4, 5]. To date, such studies have not attempted to quantify how a positive detection result relates to our level of confidence in projections of future climate. In part, this is because the detection diagnostics are not directly linked to the properties of the climate system known to be important for climate-change projections. The two most important of these properties are the equilibrium climate sensitivity to changes in radiative forcings [6] and the rate of heat uptake by the deep ocean [7, 8, 9]. Together, these properties determine the decadal-to-century timescale response to a change in forcing.

This paper presents a method for constraining these properties by comparing the modeled response to anthropogenic forcings over the 1961–1995 period against climate observations of this period. One key aspect of our study is that we can vary both our model's climate sensitivity (by varying the cloud feedback) and its rate of mixing heat into the deep ocean, and thus determine how the comparison with observations depends on these factors. This would not be possible with the most sophisticated models available, the coupled atmosphere-ocean general circulation models (AOGCMs), owing to their computational requirements and fixed representation of physical processes. A second key aspect is the use of an objective approach to model-data comparison, based on the "optimal fingerprinting" algorithm used in the detection of climate change, to quantify likely ranges for these properties in probabilistic terms. Other estimates of such constraints have

[1] Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Corresponding author: ceforest@mit.edu.
[2] Space Science Department, Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK; Department of Physics, University of Oxford, UK.

been subjective or qualitative [*e.g.*, see 10, 11, 12]. For example, the Intergovernmental Panel on Climate Change (IPCC) estimates that the climate sensitivity, *S*, lies in the range of 1.5°C to 4.5°C, but the IPCC does not specify the probability that the sensitivity lies in this range [11].

To determine constraints, we use the MIT 2D statistical-dynamical climate model [8, 9] in conjunction with estimates of climate variability from the control integration of the Hadley Centre's second-generation coupled AOGCM, HadCM2 [13]. The MIT 2D model consists of a zonally-averaged version of the Goddard Institute for Space Studies (GISS) Model II atmospheric general circulation model [14] coupled to a Q-flux mixed-layer diffusive ocean model. The current version of the model has 24 latitude bands ($\Delta\phi = 7.826°$) and 11 vertical layers with 4 layers above the tropopause so that the response to changes in stratospheric ozone concentrations can be modeled to a first approximation. The model uses the GISS radiative transfer code, which accounts for all radiatively important trace gases as well as aerosols. The surface area of each latitude band is divided into a percentage of land, ocean, land ice, and sea ice with the surface fluxes computed separately for each surface type. This allows for appropriate treatment of radiative forcings dependent on underlying surface type such as anthropogenic aerosols. Similar to AOGCMs but in contrast with energy balance models, the model contains full representations of the atmosphere's moisture, heat, and momentum cycles, and thus includes the most important of the nonlinear dynamical interactions between components of the atmospheric system [8].

The climate sensitivity, *S*, of the MIT model can be varied by varying the cloud feedback [8]. Differences in cloud feedback appear to be the primary cause of the different climate sensitivities of AGCMs [15, 16]. The global-mean vertical thermal diffusivity below the ocean's mixed layer, K_v, is an external parameter in the MIT model, and thus the rate at which the model mixes heat into the deep ocean can be varied by changing K_v. The global response of a given 3D coupled ocean-atmosphere GCM for different forcings can be reproduced accurately by the MIT 2D model with an appropriate and unique choice of these two parameters (Fig. 1). Published values of 3D GCM model sensitivities range from 1.9° to 5.2°C [11, 16]. Comparisons between time-series of transient climate changes calculated with the MIT model and with 3D AOGCMs show that the AOGCM's equivalent vertical diffusivities range from 0.0 to 25.0 cm²/s (Fig. 1).

The MIT model is forced with changes in greenhouse gas, anthropogenic aerosol, and stratospheric ozone concentrations to produce a simulated climate record for the 20th century. The greenhouse gas and sulfate aerosol forcings are the same as those used in the HadCM2 "GSO" experiments [5]. The direct impact of sulfate aerosols is accounted for by increasing the surface albedo [17]. This increase corresponds to a globally averaged forcing of -0.5 W/m² in 1986 as compared to pre-industrial times. The version of the MIT 2D model with fully interactive chemistry [18] can handle aerosols in a more realistic way: we used this simple albedo-based parameterization for compatibility with previous studies. This also reduces the computational cost of the large number of runs that had to be performed because the climate-chemistry version of the MIT model is much more computationally demanding than the one used in this study. The changes in ozone concentrations are prescribed according to the GISS ozone concentration data set [1]. Not considered are the effects of changes in tropospheric ozone, which could partially offset aerosol effects. We also do not explicitly consider the effects from natural forcings, but the diagnostic used was specifically tailored by references [5, 19], to limit their impact.
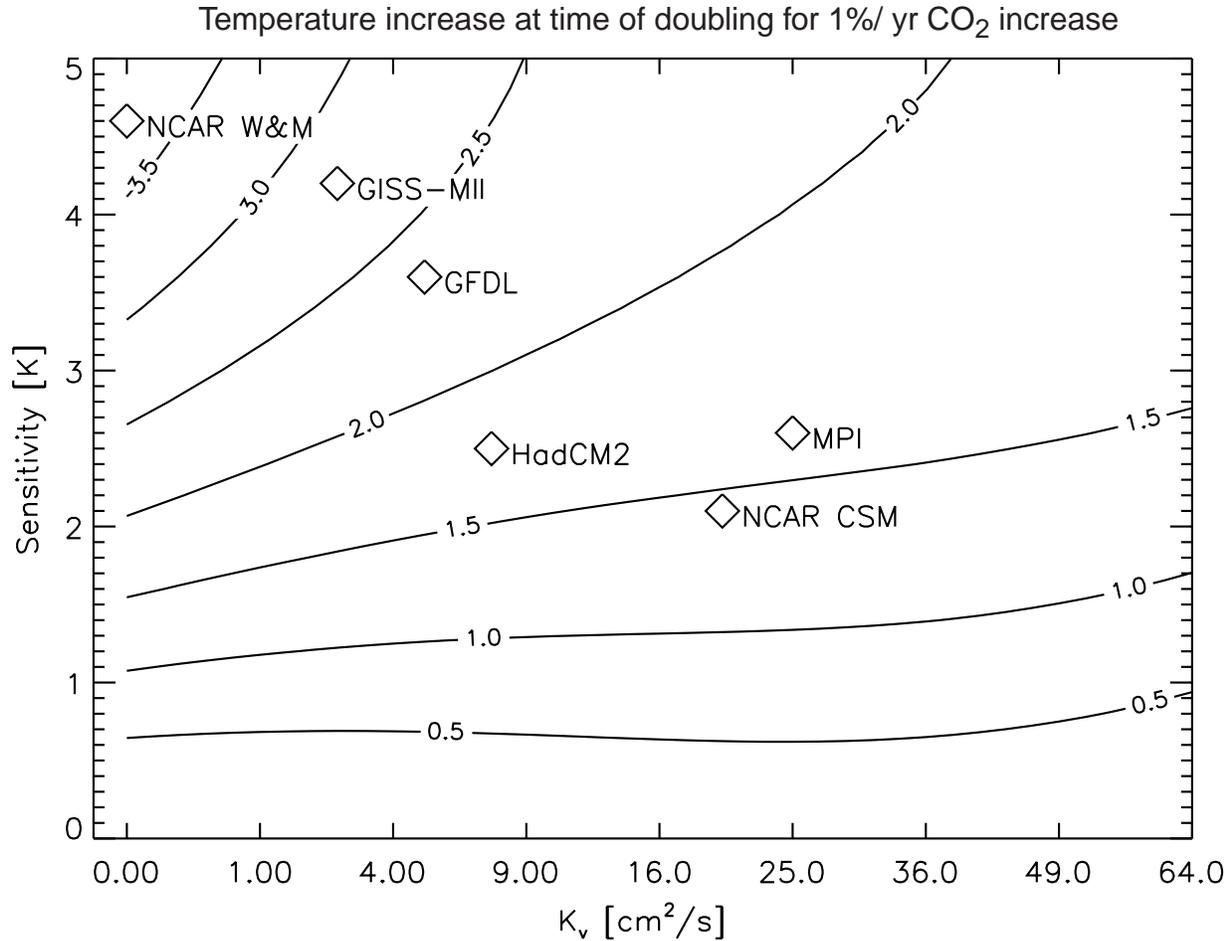
Figure 1. Response in the global mean surface temperature at the time of doubling of CO_2 for simulations with 1% per year increase in CO_2 concentration (contours). The S and K_v values for six AOGCMs are also shown (diamonds). For a given AOGCM, the model sensitivity was typically obtained from the literature and a set of MIT 2D model runs were performed to estimate the value of K_v that gave a transient response matching published AOGCM experiments with prescribed forcing, usually a 1% per year increase in CO_2 concentration. The results for the GFDL, MPI, and NCAR W&M models are described in ref. 8. For the HadCM2 model, the sensitivity and transient run data were obtained through personal communications [6]. The sensitivity for the NCAR CSM model was obtained through personal communication and the transient run data were taken from ref. 29. The parent GISS model is also included, although its ocean component is a Q-flux mixed-layer diffusive model similar to the ocean component of the MIT model.

We compare the height-latitude pattern zonal-mean temperature changes in the troposphere and lower stratosphere as simulated by the MIT 2D model with similarly-estimated trends from the UKMO radiosonde dataset [20]. The radiosonde data were used to calculate annual mean, zonal mean temperature fields on a latitude-height grid (Fig. 2). Because data were unavailable for certain regions and times, a missing-value mask was obtained for both space and time. The MIT model results were regridded to the observational grid, and then sampled using the observational data mask to compute the 1961–1980 and 1986–1995 period means. Thus, the fingerprint pattern is the latitude-height pattern of temperature change between the two period means. This fingerprint matches that used in the Hadley Centre study [5] so results can be compared. The averaging

periods were chosen to accommodate the 1958–1995 record of radiosonde data. Additionally, the averaging periods exclude the years following major volcanic eruptions, 1963–4 and 1992, to reduce their effect on the computed averages [see 19].

We produce a large number of simulations spanning the two-dimensional parameter space (S and K_v), and evaluate the degree of correspondence between individual simulations and the observational record using a goodness-of-fit criterion (noise-weighted residual sum of squares) identical to that used in optimal fingerprinting [2, 19, 21]. In this manner, a location in the model's parameter space is identified that best fits the observational record. More importantly, by varying parameters about this best fit, we can determine which parameter-combinations lead to simulations that are inconsistent with observations at a given level of confidence, and thereby assign a confidence region to the parameters that determine the model's climate properties.

The "noise" in this problem is the natural unforced fluctuations of the climate system. Internal variability in the MIT 2D model is somewhat lower than typical of AOGCMs (hence our use of an AOGCM for uncertainty analysis), but to further reduce noise in the predicted response to the climate forcings we generated four-member ensembles. The model runs were started in 1860 with the prescribed forcing and run through 1995. At the beginning of 1940, the model was stopped and its state perturbed by adding deviations from a control run to create the additional members of the ensemble. The model perturbations were created from the model state output every 10 years from a control run with present-day fixed conditions. Three perturbation runs were started in 1940 for a total of four runs with the given forcing and parameter settings. Our results are based on 61 different pairs of parameter choices, and thus required a total of 18,544 years of simulated climate changes: hence, the need for an efficient model. The characteristics of climate noise are estimated from the control run of the HadCM2 model, one of the few AOGCMs for which extended control integrations are currently available [22]. Although there is no data available for checking the adequacy of this model for simulating the variability of our chosen fingerprint, its variability on shorter time scales compares well with observed variability [22, 23].

We start by defining the estimated residual, $\tilde{\mathbf{u}}$, for a particular choice of parameter values, S and K_v, as follows:

$$\tilde{\mathbf{u}}(S, K_v) = \mathbf{y} - \tilde{\mathbf{x}}(S, K_v) \tag{1}$$

where $\mathbf{y}$ is the vector of observations over a given period and $\tilde{\mathbf{x}}(S, K_v)$ is the ensemble-mean simulation of the MIT 2D model corresponding to those parameter values. Following standard practice in climate change detection, we assume the dominant contribution to this residual is internal climate variability. Those studies that have considered the impact of observational error on these scales find its effect to be relatively minor [24], and the diagnostic used was specifically chosen to minimize the impact of other external climate forcings such as solar variability and volcanic eruptions. Because variability simulated by the HadCM2 model is demonstrably too low on small spatial scales [23, 25], both $\mathbf{y}$ and $\mathbf{x}$ are projected onto the first κ empirical orthogonal functions (EOFs) of the control prior to computing $\tilde{\mathbf{u}}$. We now define the noise-weighted residual sum-of-squares, $\tilde{r}^2$ by:

$$\tilde{r}^2 = \tilde{\mathbf{u}}^T \hat{C}_N^{-1} \tilde{\mathbf{u}} \tag{2}$$

where $\hat{C}_N^{-1}$ is the estimated inverse of the noise covariance matrix obtained from successive

segments of "pseudo-observations" extracted from the control run for the climate model [19]. The EOFs were estimated from a separate portion of the control run, which is essential to avoid "artificial skill."

If there are m free parameters to be adjusted, we expect $\tilde{r}^2$ at the unknown location of the true parameter values to exceed $\tilde{r}^2$ at the observed minimum according to

$$\Delta \tilde{r}^2 \sim m \, F_{m,\nu}, \tag{3}$$

where ν is the degrees of freedom of the estimate $\hat{C}_N$, discussed at length in [19]. Mapping $\tilde{r}^2$ as a function of S and K_v, and comparing $\Delta \tilde{r}^2$ with the standard F-distribution [26] thus provides a means of assigning objective confidence intervals to these parameters.

We carried out simulations for 1860 to 1995, varying S from 0.4 to 11.0 K and K_v from 0.0 to 160.0 cm^2/s. We then calculated the fingerprint diagnostic and $\tilde{r}^2$ statistics by comparison with the radiosonde record. Simulated patterns of climate change are shown (Fig. 2) for three versions of the MIT model that correspond to relatively high, medium, and low responses to a given forcing. The high-response case corresponds to high S and low K_v, and *vice versa* for the low response. The response shows cooling in the stratosphere associated mainly with a prescribed decrease in stratospheric ozone concentrations as well as upper-tropospheric tropical warming and high-latitude warming associated with increased greenhouse gas concentrations. The response also shows reduced warming in the Northern Hemisphere associated with cooling from aerosol changes as compared to the response with no aerosol changes (not shown). In all, the response is qualitatively similar to zonally averaged results from AOGCMs [5]. As further evidence that the zonally averaged model provides a reliable simulation, we note the similarity between the 2D model's fingerprint and that of the Hadley Centre's coupled AOGCM [see Fig. 2 in ref. 5]. When the MIT model uses the ozone data from the HadCM2-GSO experiment and the appropriate S and K_v to match HadCM2, the pattern correlation and congruence [5] between the MIT and HadCM2-GSO fingerprints is 0.8.

The distribution of $\tilde{r}^2$ (Fig. 3) indicates that we can discriminate between regions of parameter space using this diagnostic, but the constraint is not strong enough to place clear bounds on S and K_v in all directions. The shaded regions represent the 80th, 95th, and 99th percentile confidence regions (corresponding to $\Delta \tilde{r}^2 = 3.7$, 7.8, and 13.9, respectively) and indicate that two regions can be rejected with 95% confidence. The first is the high-response region at the upper left, which corresponds to both high S and low K_v. The second region is for S less than 0.6 K for all values of K_v. We examined the dependency of this distribution on κ, to remove possible effects from small-scale features of variability in higher order EOFs. After examining the EOF structures, we truncated at $\kappa = 6$; this truncation includes the patterns of variability that contain the largest spatial scales and are most likely to be sampled well by the HadCM2 control run. This change in κ does not affect the low-response rejection region but does reduce somewhat the area of rejection for high-response cases (Fig. 3B). We also estimated the $\tilde{r}^2$ distribution by using only lower tropospheric temperature changes (lowest 3 layers) and found little difference in the results (not shown). Thus, the results do not depend strongly on reproducing the observed cooling in the stratosphere. This is reassuring because, on physical grounds, we would expect this stratospheric cooling to be almost independent of S and K_v (being primarily a radiative effect). Moreover, we have grounds to distrust the simulation of stratospheric variability in HadCM2 [23].
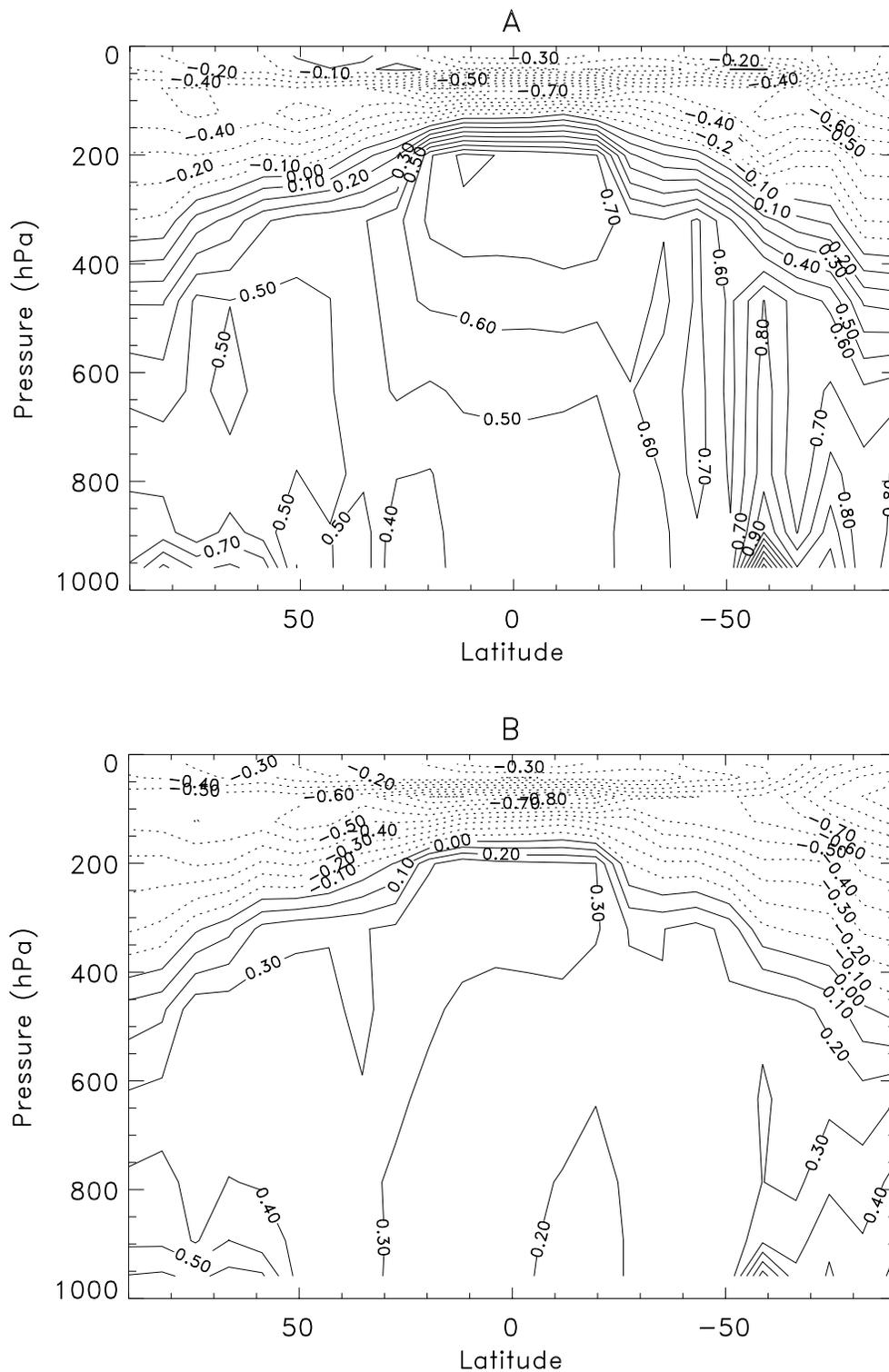
Figure 2. Vertical patterns of temperature change from 1961–1980 to 1986–1995 simulated by three versions of the MIT model and from radiosonde observations (D). The three versions correspond to relatively high ($S = 4.5$ K, $K_v = 0.16$ cm^2/s) (A), medium ($S = 3.0$ K, $K_v = 2.5$ cm^2/s) (B), and low ($S = 1.6$ K, $K_v = 40.0$ cm^2/s) (C) responses of the MIT model. Contour interval is 0.1 K and negative contours are dashed.
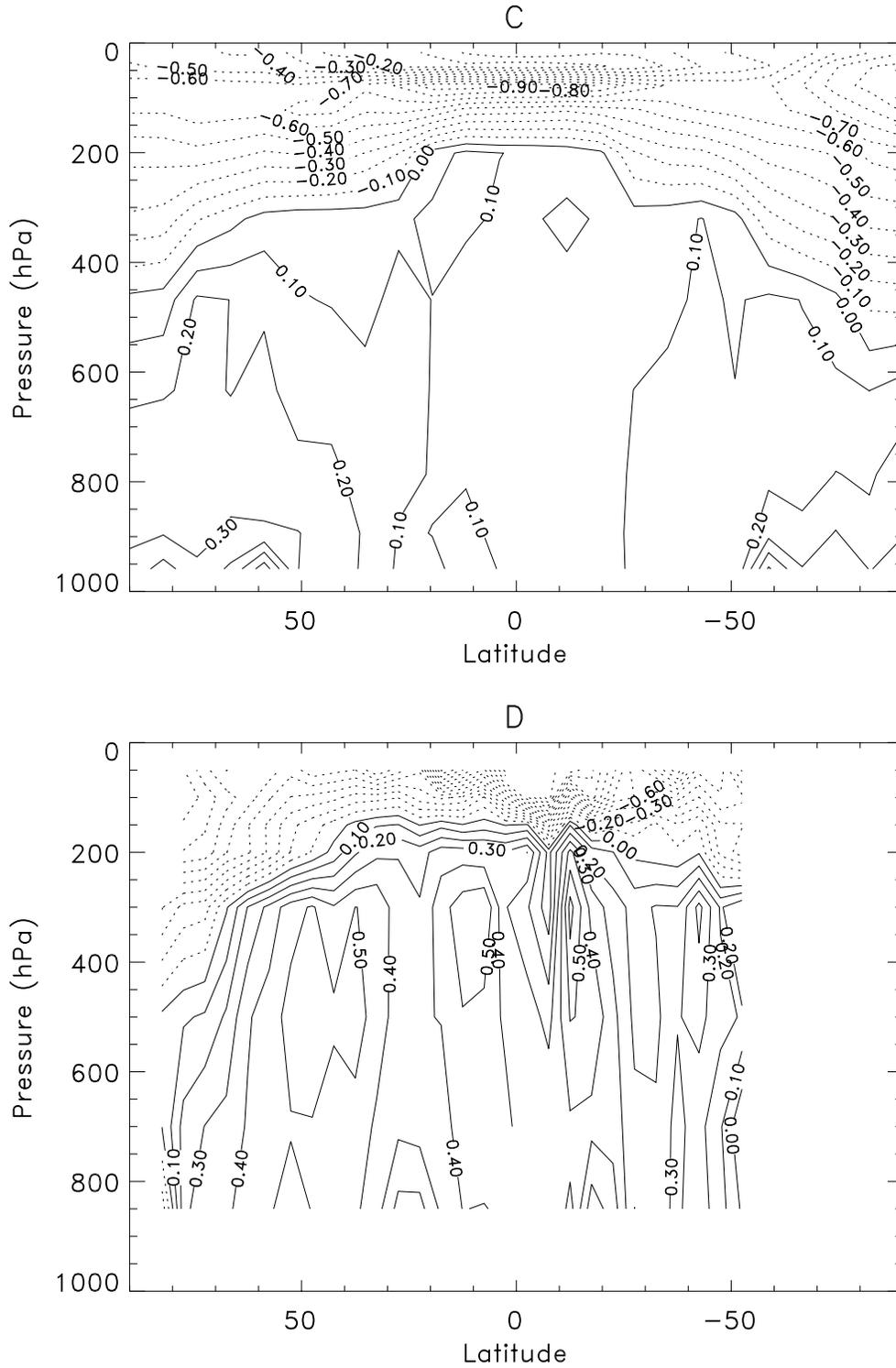
Figure 2 (continued). Vertical patterns of temperature change from 1961–1980 to 1986–1995 simulated by the MIT model (C) and from radiosonde observations (D). The above simulation corresponds to relatively low (S = 1.6 K, K_v = 40.0 cm^2/s) responses of the MIT model. Contour interval is 0.1 K and negative contours are dashed.

Using the full latitude-height pattern of temperature changes, only the NCAR W&M model is rejected with 95% confidence. All other 3D AOGCMs (Fig. 1) are within the region of plausible parameters. Previous detection studies [4, 5] have indicated that a positive detection implies a non-zero climate sensitivity. Our result goes a step further by rejecting the possibility of negative feedbacks strong enough to cause S to be less than 1.2 K (at the 80% confidence level) or 0.6 K (at the 95% confidence level). The utility of these constraints is illustrated further when we consider the uncertainty of the climate system response to future forcings. Projections of future climate corresponding to the acceptable parameter values still vary considerably, as illustrated in Fig. 1. For example, for the scenario in which CO_2 increases 1% per year, comparison of Figs. 1 and 3A indicates an 80% probability that global warming, at the time of doubling, would range from 1.1° to 2.8°, and a 95% probability that it would range from 0.5° to 3.3°.

We note that uncertainty in the tropospheric aerosol forcing remains substantial (0.0 to -2.5 W/m^2 when possible indirect effects are included [11]) and was not addressed here. We have also not explicitly considered the possible influence of natural external forcings, although the diagnostic used was designed to minimize their impact by focusing on decadal means and excluding years contaminated by volcanic eruptions. In using $\Delta \tilde{r}^2$ statistics to assign confidence intervals to model parameters, we are assuming that the statistical properties of climate variability simulated by the HadCM2 model are both accurate [23] and independent of the parameters being estimated. This is clearly a simplification because we have reason to believe on physical grounds that a high-S, low-K_v climate would display greater variability than a low-S, high-K_v climate. Given, however, the uncertainties inherent in estimating climate variability from an AOGCM, quantifying the dependence of this variability on parameters is currently unfeasible.

These caveats indicate that our estimate of the uncertainty in S and K_v may be too low. Conversely, by confining attention to the 35-year period for which atmospheric observations are available, and not making use of the longer surface temperature record, we may be giving a somewhat inflated estimate of these uncertainties. In a complementary study by Allen *et al.* [27], surface observations were directly compared with GCM simulations and the estimates of uncertainty were interpreted in terms of simple models, rather than by explicitly varying the simulation model, as here. As a result, that study does not draw conclusions concerning climate properties but focuses on the implications of the observed signal for the transient response under a specific forcing scenario for the next 50 years.

Finally, Fig. 3 illustrates that our model's response has distinct regions in which the statistics vary smoothly, albeit with considerable noise (*i.e.*, multiple minima within the acceptable region). However, the sharp transition from the medium- to high-response cases suggests that the linearity assumption behind standard statistical models, including ours, will not be valid if the "true" climate system properties lie near such a border. Indeed, owing to an amplification by the ice-albedo feedback [28], the model response shows a marked change in southern mid-latitudes that alters the fingerprint and consequently increases $\tilde{r}^2$ values. Only if the "true" climate system behavior is far from such a "border" in Fig. 3, will the assumption of Gaussian statistics be appropriate. Also, the position of this "border" is dependent on the assumed aerosol forcing, uncertainty in which was not considered here. Our results clearly emphasize the need to estimate the location of the "true" climate system in this two-dimensional space.
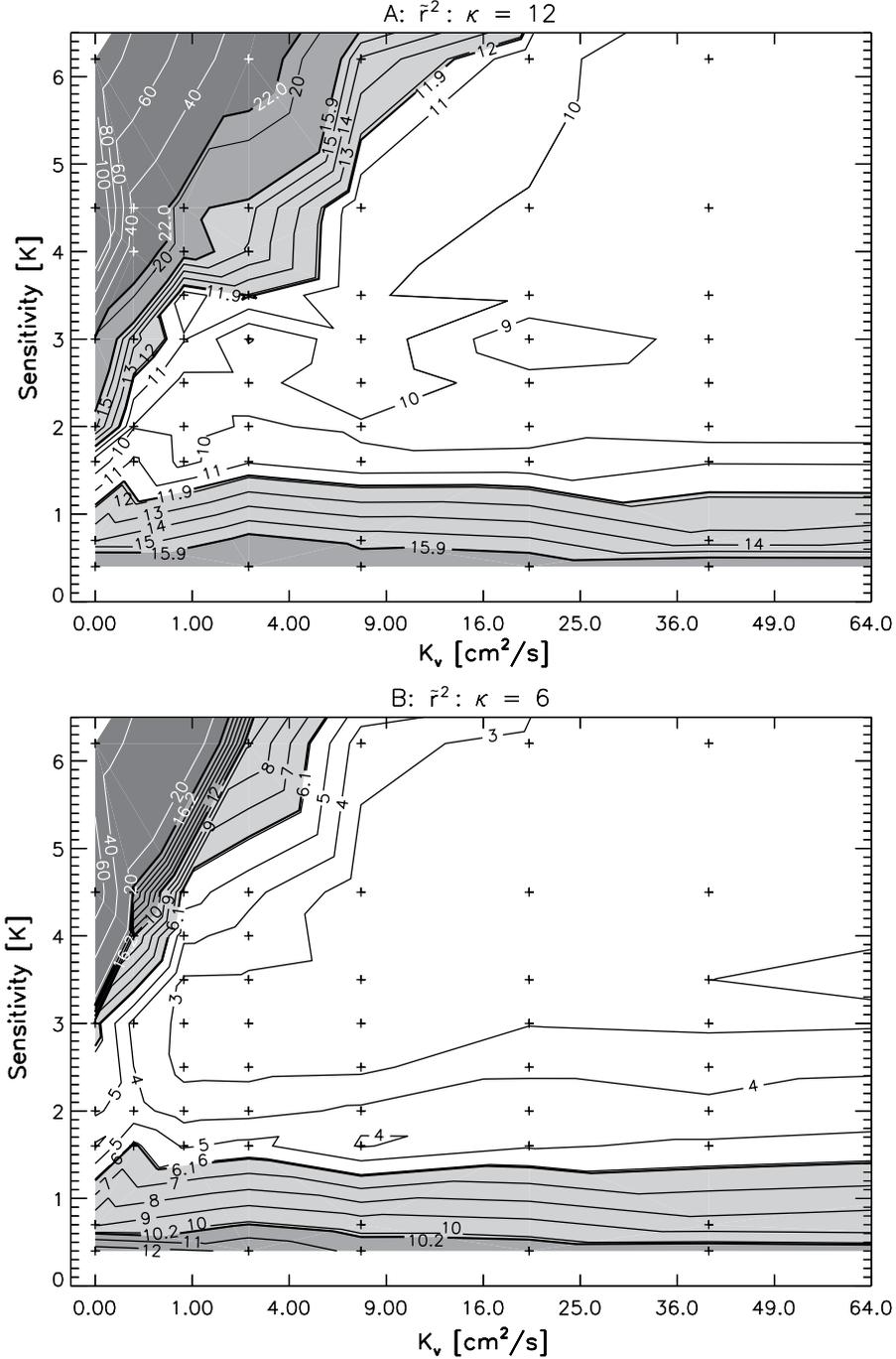
Figure 3: Dependence of $\tilde{r}^2$ (contours) on climate sensitivity (S) and oceanic heat uptake (K_v) for the model response to forcing as compared to observations for $\kappa = 12$ (A) and 6 (B). Model runs were performed at parameter locations shown by the + symbol. Contours extending to the boundary are based on model runs outside this region. Contour interval below 15 is 1; above 20 it is 20. The shaded regions designate the boundaries of the 80th, 95th, and 99th percentile regions corresponding to $\Delta \tilde{r}^2$ values of 3.7, 7.8, and 13.9 respectively. This difference is with respect to the minimum $\tilde{r}^2$ value of 8.2 at $K_v = 20.0$ cm^2/s and $S = 3.0$ K. In each figure, all estimates of $\tilde{r}^2$ use the same number of EOFs in the comparisons with observations. The truncation at 12 EOFs was chosen to match ref. 5. We also show the results with $\kappa = 6$ that indicates the shape of the rejection region does not strongly depend on the truncation number. Both observations and model data are weighted by the mass of the respective layers, which reduces the effect of stratospheric variability on the EOF structure [19].

References

1. Hansen, J., *et al*., 1997, *J. Geophys. Res*., **102**(D22):25,679.

2. Hasselmann, K., 1993, *J. Climate*, **6**:1957; Hasselmann, K., 1997, *Clim. Dyn*., **13**:601.

3. Hegerl, G., *et al*., 1996, *J. Climate*, **9**(10):2281; Santer, B., *et al*., 1995, *Clim. Dyn*., **12**(2):77.

4. Hegerl, G., *et al*., 1997, *Clim. Dyn*. **13**(9): 613; Santer, B., *et al*., 1996, *Nature*, **382**(6586):39.

5. Tett, S., J. Mitchell, D. Parker, M. Allen, 1996, *Science*, **274**:1170.

6. Climate sensitivity is defined as the equilibrium change in global-mean surface temperature due to a doubling of CO_2 concentrations and is usually estimated from a run of an atmospheric GCM coupled to a mixed-layer ocean model. This response would not include long timescale feedbacks (*e.g*., multi-century feedbacks like those assosciated with changes in sea ice distribution or ocean circulation that are present in the true climate system. For example, a simulation with the HadCM2 model has shown that the sensitivity changes with time. Sensitivity was estimated for the coupled model to be 2.5 K for initial stages of a doubled CO_2 simulation and revised to be 3.3 K from the final stage [C. Senior, 1998, personal communication].

7. Hansen, J., *et al*., 1985, *Science*, **229**:857.

8. Sokolov, A., P. Stone, 1998, *Clim. Dyn*., **14**:291.

9. Sokolov, A., C. Wang, G. Holian, P. Stone, R. Prinn, 1998, *Geophys. Res. Let*., **25**:3603.

10. Hansen, J., *et al*., 1981, *Science*, **213**:957; Hansen, J., *et al*., 1984, in: "Climate Processes and Climate Sensitivity," J. Hansen and T. Takahashi, editors, *Geophysical Monograph*, **29**:130–163, American Geophysical Union, Washington, D.C.; Titus, J., V. Narayanan, 1996, *Climatic Change*, **33**:151; Wigley, T., P. Jones, S. Raper, 1997, *Proc. Nat. Acad. Sci*., **94**:8314.

11. Houghton, J., *et al*., editors, 1996, *Climate Change 1995, The Science of Climate Change*, Cambridge University Press, Cambridge, UK.

12. The estimates of climate sensitivity from paleoclimate data for changes from glacial to interglacial conditions correspond to responses on timescales longer than a century and may not be applicable to predictions for the 21st century [see ref. 6]. An example of such an estimate can be found in: Hansen, J., M. Sato, A. Lacis, R. Ruedy, 1997, *Phil. Trans. Roy. Soc. London*, B**352**:231.

13. Johns, T., *et al*., 1997, *Clim. Dyn*., **13**:103.

14. Hansen, J., *et al*., 1983, *Mon. Weath. Rev*., **111**:609.

15. Cess, R., *et al*., 1990, *J. Geophys. Res*., **95**:16601.

16. Senior, C., J. Mitchell, 1993, *J. Climate*, **6**:393.

17. Mitchell, J., R. Davis, W. Ingram, C. Senior, 1995, *J. Climate*, **8**:2364.

18. Wang, C., R. Prinn, A. Sokolov, 1998, *J. Geophys. Res*., **103**(D3):3399.

19. Allen, M., S. Tett, 1999, *Clim. Dyn*., in press.

20. Parker, D., *et al*., 1997, *Geophys. Res. Let*., **24**:1499, with corrections to Southern Hemisphere data as described in ref. 5.

21. Hegerl, G., G. North, 1997, *J. Climate*, **10**:1125.

22. Tett, S., T. Johns, J. Mitchell, 1997, *Clim. Dyn*., **13**:303.

23. Gillett, N., M. Allen, S. Tett, *Clim. Dyn*., in press.

24. Barnett, T., G. Hegerl, B. Santer, K. Taylor, 1998, *J. Climate*, **11**:659.

25. Stott, P., S. Tett, 1998, *J. Climate*, **11**:3282.

26. Mendenhall, W., D. Wackerly, R. Scheaffer, 1990, *Mathematical Statistics with Applications*, PWS-Kent Publishing, Boston.

27. Allen, M., P. Stott, R. Schnur, C. Forest, 1999, paper presented at the 10th Symposium on Global Change Studies, American Meteorological Society, Dallas, TX, 14 January.

28. We note that this amplification can only occur after a sufficient warming has occured at high-latitudes to melt enough ice and change the albedo. Thus, we find a threshold behavior of the climate system as noted in: Murphy, J., 1995, *J. Climate*, **8**:496, and ref. 8.

29. NCAR CSM data were obtained at http://www.cgd.ucar.edu/csm/experiments/.