A Multimodel Study of Parametric Uncertainty in Predictions of Climate Response to Rising Greenhouse Gas Concentrations

BENJAMIN M. SANDERSON
NCAR, Boulder, Colorado

(Manuscript received 3 November 2009, in final form 4 August 2010)

ABSTRACT

One tool for studying uncertainties in simulations of future climate is to consider ensembles of general circulation models where parameterizations have been sampled within their physical range of plausibility. This study is about simulations from two such ensembles: a subset of the climateprediction.net ensemble using the Met Office Hadley Centre Atmosphere Model, version 3.0 and the new “CAMcube” ensemble using the Community Atmosphere Model, version 3.5. The study determines that the distribution of climate sensitivity in the two ensembles is very different: the climateprediction.net ensemble subset range is 1.7–9.9 K, while the CAMcube ensemble range is 2.2–3.2 K. On a regional level, however, both ensembles show a similarly diverse range in their mean climatology. Model radiative flux changes suggest that the major difference between the ranges of climate sensitivity in the two ensembles lies in their clear-sky longwave responses. Large clear-sky feedbacks present only in the climateprediction.net ensemble are found to be proportional to significant biases in upper-tropospheric water vapor concentrations, which are not observed in the CAMcube ensemble. Both ensembles have a similar range of shortwave cloud feedback, making it unlikely that they are causing the larger climate sensitivities in climateprediction.net. In both cases, increased negative shortwave cloud feedbacks at high latitudes are generally compensated by increased positive feedbacks at lower latitudes.

1. Introduction

To better understand the relationship between general circulation model (GCM) parameterizations and the response to future increases in greenhouse gases, we have created a grand ensemble of climate simulations using the Community Atmosphere Model, version 3.5 (CAM) to complement several large-scale experiments conducted in the past few years using versions of the Met Office Hadley Centre model (Stainforth et al. 2005; Murphy et al. 2004). Perturbed physics ensembles (PPEs) are one tool for exploring the uncertainties inherent in simulations of future climate change when the values of model parameters themselves are subject to uncertainty. However, a PPE alone cannot provide any information about systematic differences between different GCMs, or common errors, which may be present in all of the current generation of climate models. Numerous methodologies exist for making probabilistic statements about future climate change based on PPE simulations, but each of these requires some assumptions to be made about systematic model errors. A purely Bayesian treatment like that proposed in Murphy et al. (2007) can incorporate systematic uncertainty by considering errors in prediction in perfect model experiments using members of the Coupled Model Intercomparison Project phase 3 (CMIP3) ensemble to represent truth. A complete treatment, however, would require a description of how this systematic error varied in the PPE parameter space and additional terms accounting for common errors between models. Similarly, regression-based approaches like Piani et al. (2005) introduce terms for systematic discrepancy based upon inflated variance in predictors; however, the degree of inflation is based upon biases between the PPE mean state and the mean state of other GCMs. Unfortunately, this term has the undesired property of making the result dependent upon the sampling strategy of the PPE itself.

Even if an analysis considers the full range of CMIP3 models in addition to the PPE, there remains the issue that the parameter uncertainty has been sampled in only one model, although other models may be more or less sensitive to similar perturbations. It is thus of some interest to consider in parallel PPEs from a range of GCMs. A recent study (Yokohata et al. 2010) has already
examined some physical differences in PPEs using the Met Office Hadley Centre Quantifying Uncertainty in Model Predictions (QUOMP) ensemble and an ensemble using Model for Interdisciplinary Research on Climate 3.2 medium-resolution version [MIROC3.2(medres)], finding considerable differences in the sensitivity of low-level clouds to parameter perturbations. As a result, the authors found a much larger range of net shortwave (SW) feedbacks in the Met Office Hadley Centre model than is possible in MIROC.

The following study adds a third model to this comparison process, introducing a new ensemble that uses the Community Atmosphere Model, version 3.5 to produce an ensemble with a similar experimental design to the existing climateprediction.net ensemble. Considerable differences in the parameter sensitivity of the two models could necessitate a reconsideration of methodologies where parameter uncertainty and systematic uncertainty are considered to be independent. Our analysis requires several steps, initially examining the global-scale behavior of models in each ensemble, both in terms of the preindustrial climate and the response to greenhouse gas forcing. The analysis that follows attempts to explain the range of global mean response to greenhouse gas forcing in each ensemble and to find physical explanations for any differences.

Section 2 describes the experimental design of the new CAM ensemble and how it relates to the existing ensembles using the Met Office Hadley Centre models. Section 3 details the results from the two experiments. Section 3a explores the climate sensitivity distributions in each ensemble, while section 3b examines how the climate sensitivity relates to parameter changes made in each case. Section 3c breaks this response down into global mean feedbacks for different components of the radiative budget, and section 3d introduces the concept of “partial sensitivity,” which allows the climate sensitivity of a model to be broken down linearly into different components. Section 3e explains the large-scale features of the perturbed climates that are well correlated with the different components of climate sensitivity and proposes physical mechanisms to explain these connections and thus explain the mechanisms leading to the variation of climate sensitivity within the two ensembles. Finally, section 4 discusses the ways in which these results should be interpreted and possible future directions for multiensemble analysis.

2. Experimental design

To make this comparison as direct as possible, we have followed the experimental design for the climateprediction.net project, outlined in Stainforth et al. (2005), and applied the methodology to CAM to make a new ensemble, “CAMcube.” This involves choosing a selection of model parameterizations whose values are uncertain from fundamental principles and choosing one of three possible values for each where the extreme values represent reasonable and arbitrary limits of uncertainty (limits are determined by expert solicitation with model developers). The large size of the climateprediction.net ensemble was made possible by using a distributed computing architecture described in Stainforth et al. (2005), whereas the simulations that make up CAMcube were conducted on parallel supercomputing facilities that place a limit on the potential ensemble size. However, with the benefit of previous studies (Sanderson et al. 2008a,b, 2009), we isolate the dominant parameters influencing climate sensitivity in the climateprediction.net ensemble to consider a four-parameter subspace, sampled at three points in each dimension, giving 81 unique parameter combinations. We find parameters similar to these in CAM if they exist (using other parameters related to the same processes if not) and produce another ensemble in the hope that a similar range of behavior might be observed in this alternative model framework. Details of the perturbed parameters are given in appendix A.

CAMcube is based on the Community Atmosphere Model, version 3.5 (Gent et al. 2009). The atmospheric model shares many components with CAM3 (Collins et al. 2006), with a finite volume dynamical core (Lin 2004) and improvements (Neale et al. 2008; Richter and Rasch 2008) to the existing convection scheme of Zhang and McFarlane (1995). The model resolution is used at a resolution of 1.9° latitude by 2.5° longitude, with the finite-volume dynamical core based upon Lin (2004) and 26 model levels. The land component used is Community Land Model, version 3.5 (CLM3.5; Oleson et al. 2008; Stöckli et al. 2008). The experiments use a thermodynamic slab ocean for reasons of consistency and speed of simulation. The slab ocean model consists of a thermodynamic mixed layer grideded ocean layer and thermodynamic prognostic sea ice cover. The ocean mixed layer contains an internal heat source $Q$ that is determined by running a calibration simulation with predefined sea surface temperatures (SSTs). The experiments comprise three 15-yr simulations for each model version: (i) a calibration stage to determine the $Q$ flux, (ii) a control stage (CTRL), and (iii) an instantaneous double CO$_2$ run (X2CO2). Climate sensitivity is inferred by assuming the global mean temperature follows an exponential decay toward a new equilibrium temperature.

3. Results

a. Climate sensitivity distributions

Figure 1 shows the model development of the simulations in the CAM ensemble as compared to a
2000-member subset of climateprediction.net. Differences between the ensembles are strikingly apparent: whereas the climateprediction.net ensemble shows a large range of behavior in the X2CO2 stage and a significant number of drifting simulations in the control stage, neither of these features are clearly apparent in the CAM ensemble. Instead, most simulations show qualitatively little variation in sensitivity to external forcing, and none of the simulations displays a major drift in the control stage.

The annual global mean temperature in the control simulations in the CAMcube ensemble shows a clear bimodal distribution. The two groups are separated by their values for the parameter controlling the rate of change of convective parcel momentum with height (Raymond and Blyth 1992; Neale et al. 2008). Those models with the parameter set to zero effectively have convective momentum transport (CMT) switched off completely, and these models tend to show greater surface temperatures in the tropics. It has been shown previously (Wu et al. 2007) that the introduction of CMT in CAM weakens equatorial convergence, which tends to reduce convective heating around the equator.

Figure 2a demonstrates that the distribution of climate sensitivity within the CAM ensemble is noticeably narrower than is the case for climateprediction.net. The CAM ensemble values of climate sensitivity range from 2.2 to 3.2 K, whereas the 2000-model subset from the climateprediction.net ensemble ranges from 0.2 to 11.8 K (excluding those unstable models with a climatological cooling of greater than $-0.04$ K yr$^{-1}$ in the last 8 yr of the control stage). We reduce the climateprediction.net ensemble to a four-parameter subspace containing 81 models, of which 61 are stable (these stable models are highlighted throughout). The 61 remaining models have upper and lower bounds of climate sensitivity of 1.7 and 9.9 K, respectively.

The distribution of each ensemble is clearly arbitrary and dependent on the parameter choices made in the experimental design, and various studies, notably Murphy et al. (2004), Knutti et al. (2006), and Piani et al. (2005), have proposed methods of overcoming the effects of sampling on the resulting distribution. All these studies used the Met Office Hadley Centre model; Murphy et al. (2004) sampled the parameter space in a Monte Carlo fashion and then linearly interpolated between the known points in the space to produce an estimate for sensitivity at each point in the model’s parameter space, which was also shown weighted by a measure of model likelihood. Figure 2b shows a sensitivity distribution interpolated over parameter space in both CAMcube and climateprediction.net using a cubic spline, with the interpolated space sampled evenly at 10 values in each of the four parameter dimensions, giving an interpolated ensemble size of $10^5$ simulations.

The narrow distribution of sensitivity in the CAM ensemble is unexceptional in itself because it could easily be engineered by perturbing null parameters within the model. However, on a regional scale, there are large variations, both in control climatology and in the local response to greenhouse gas forcing. Figure 3 shows both preindustrial regional mean temperatures and the regional warming on CO2 doubling, and it shows that the CAM ensemble produces a range of preindustrial climates that is not significantly less diverse than for climateprediction.net, implying that we have not simply perturbed “null” parameters in the CAM ensemble (a similar diversity can be seen in distributions of precipitation in both ensembles; see the appendixes).

b. Parameter dependency

Stainforth et al. (2005) and Sanderson et al. (2008a) claimed that the sensitivity in the climateprediction.net parameter space is not well represented as a linear function. To demonstrate this, Figs. 4a and 4b show the behavior of climate sensitivity in the model parameter space of each ensemble. The significant curvature of the contours of climate sensitivity in the parameter space of both ensembles shows a strongly nonlinear parameter dependency in each case.

In the case of the CAM ensemble shown in Fig. 4a, there is not a large variation in climate sensitivity; however, the largest sensitivities of 3.4 K are achieved by setting convective mass transport to zero, leaving the convecting parcels undiluted (Raymond and Blyth 1992). Similar distributions for climateprediction.net are shown in Fig. 4b, where it is clear that the largest climate sensitivities require multiple perturbations from the default model. As was shown in Sanderson et al. (2008a, 2009), a perturbation to decrease entrainment into convecting
plumes is required for the model to show a climate sensitivity of greater than 6 K. However, sensitivities of up to 6 K can be achieved in the climate prediction.net without any perturbation of the convection scheme through combined perturbations of the other three parameters.

c. Global feedbacks

Given the increased range of sensitivity in climateprediction.net over CAMcube, it is useful to break down the distributions of global response into different parts of the radiative budget. Figure 5 shows histograms of the top-of-atmosphere response to surface warming in both ensembles. This metric shows the change in clear-sky (CS) flux and cloud radiative forcing (CRF) in both the longwave (LW) and shortwave top-of-atmosphere budgets for a 1-K increase in global mean surface temperature (detailed in appendix B).

In both ensembles, the shortwave cloud component is the largest single component of the variance in top-of-atmosphere feedback (Fig. 5b). This has been noticed before, in both perturbed physics and multimodel ensembles (Webb et al. 2001; Sanderson et al. 2008b; Trenberth and Fasullo 2009). However, this explanation fails to explain the high climate sensitivity in some climateprediction.net models, because the most positive shortwave cloud feedbacks have a similar magnitude in both ensembles.

Climate sensitivity is not linearly related to the individual feedback components in Fig. 5, which complicates their interpretation (partial derivatives of climate sensitivity with respect to individual feedback components are implicitly dependent on the magnitude of the other components).

d. Partial surface temperature response

We express the climate sensitivity to a doubling of carbon dioxide as a sum of partial sensitivities (which can be positive or negative), relating to shortwave and longwave components for both clear and cloudy skies. For example, the clear-sky longwave partial sensitivity is a linear estimate of the change in global mean surface temperature, which is attributable to changes in the global mean longwave clear-sky transmittance of the atmosphere. By considering the surface and top-of-atmosphere longwave clear-sky fluxes in the equilibrium states before and after a doubling of carbon dioxide, one can derive the change in atmospheric transmissivity. It is then possible to calculate the necessary change in surface temperature that is required to maintain top-of-atmosphere energy balance, given the change in atmospheric clear-sky transmissivity. The method is described in full in appendix C, and the distributions of the longwave and shortwave partial sensitivity for clear-sky and cloud are shown in Fig. 6.

The distributions in Fig. 6 are approximate separations of the global mean surface temperature response into components, but we find an error of less than a 5% when the sum of partial sensitivities is compared to the total climate sensitivity of models in both climateprediction.net and CAMcube (see appendixes). In the case of the partial surface temperature response to changes in cloud forcing,
FIG. 3. Scatterplots showing the joint distribution of regional temperature mean state and response to CO₂ doubling in both climateprediction.net (blue) and CAMcube (red). In each case, the horizontal axis represents the preindustrial annual mean temperature for the region, while the vertical axis shows the regional equilibrium temperature response to CO₂ doubling. Histograms on each axis represent the distribution of that quantity in the ensemble. Circled points show the unperturbed, default configuration of each model.
the results are unsurprising given the distribution of model feedbacks. Figure 6a shows that there is clearly a greater range of shortwave cloud feedback response in climateprediction.net than in CAMcube. In the most extreme positive cases in both ensembles, the shortwave cloud partial sensitivity shows up to 1 K of global mean warming in climateprediction.net, whereas all models in the CAMcube ensemble show no net warming because of shortwave cloud forcing changes. Both ensembles show a relatively small range of response to global mean longwave cloud forcing.

By far the most significant difference between the ensembles in overall climate sensitivity comes from the partial clear-sky longwave component. In climateprediction.net, this difference ranges from +2.2 to +6.3 K; in CAMcube, the range is +2.2 to +2.4 K. This includes both the greenhouse gas forcing (which is not necessarily a constant for models with dramatically different atmospheric humidity) and any clear-sky longwave feedbacks (appendix C). Up to 5 K of warming is due to clear-sky forcing and feedbacks alone in climateprediction.net; however, in CAMcube, there is almost no variation.

**e. Mechanisms leading to high climate sensitivity**

In the following section, we determine the properties of the control climatology that are well correlated with
the different components of the global temperature response determined in section 3a and that might provide both with information on the nature of the feedbacks.

1) LONGWAVE EFFECTS

Section 3d shows that the major part of the variance in climate sensitivity in the climateprediction.net ensemble is explained by the partial clear-sky longwave component, while little variation is seen in this component in the CAMcube ensemble. We begin by examining two models from the climateprediction.net ensemble—the default third Hadley Centre Atmosphere Model (HadAM3) simulation, where no parameters have been perturbed; and a high-sensitivity (HiSens) model with large, partial clear-sky longwave sensitivity. The model-perturbed parameters and longwave clear-sky response are shown in Table 1:

Figures 7a and 7b show the annual tropical mean temperature and relative humidity distributions averaged over the last 8 yr of both the single and double CO2 simulations, respectively. The control HiSens simulation has a warmer troposphere and a cooler stratosphere than the default simulation. The surface temperatures are constrained to be identical through $Q$-flux corrections. This increased temperature differential between the troposphere and stratosphere is explained by the increased upper-tropospheric, lower stratospheric (UTLS) relative humidity, which leads to an increased water vapor greenhouse effect and increased radiative cooling in the stratosphere. In the double CO2 simulation, the HiSens simulation shows more tropospheric warming and stratospheric cooling than the “default” model, which would be expected given an enhanced water vapor feedback that is associated with the increased UTLS relative humidity. This difference in behavior is not limited to the tropics, and the HiSens model exhibits increased upper-tropospheric humidity at all latitudes (not shown).

Figure 8a shows the clear-sky longwave partial sensitivity of models in both the CAMcube and climateprediction.net ensembles as a function of 150-mb specific humidity in the preindustrial control simulation. In the climateprediction.net ensemble, these quantities show a 0.86 correlation for the stable simulations. The quantities are not significantly correlated in CAMcube, which is unsurprising given the standard deviation of the clear-sky longwave sensitivity is only 0.06 K, compared to the climateprediction.net standard deviation of 1.02 K.

All climateprediction.net models with UTLS specific humidity of greater than 20 ppmv display a partial longwave clear-sky sensitivity of greater than 4 K (compared to the default configuration value of 2.5 K). Also shown on the plot is the humidity values taken from both
Atmospheric Infrared Sounder (AIRS) observations and the CMIP3 preindustrial control simulation. The AIRS data are inconsistent with those climate prediction models that exhibit a partial clear-sky longwave sensitivity of greater than 4 K. Only one model in the CMIP3 ensemble [Commonwealth Scientific and Industrial Research Organisation Mark version 3.0 (CSIRO Mk3.0)], exhibits a 150-mb specific humidity of greater than 20 ppmv.

Also shown is the partial longwave cloudy sensitivity as a function of UTLS specific humidity (Fig. 8b), which is also significantly correlated in the climate prediction.net ensemble with a correlation of 0.85 (again, there is no significant correlation in CAMcube and the spread of response is much smaller).

Some unstable models in the ensemble exhibit an additional, unphysical longwave clear-sky feedback, where changes in large-scale circulation result in large increases in tropospheric equatorial humidity. However, we can dismiss these simulations because they are not energetically stable in the control simulation (see Figs. 1 and 8).

2) SHORTWAVE EFFECTS

After the longwave clear-sky response has been accounted for, the majority of the remaining variance in climate sensitivity in both climateprediction.net and CAMcube ensembles is explained by differences in shortwave cloud response (see Fig. 6). In CAMcube, the partial shortwave cloudy sensitivity accounts for 72% of the variance in total climate sensitivity and the standard deviation of the shortwave partial cloudy sensitivity is 0.14 K. In climateprediction.net, the partial shortwave cloudy sensitivity has a larger standard deviation of 0.52 K; however, because of the dominance of the LW clear-sky response in this ensemble, the shortwave partial cloudy sensitivity accounts for only 16% of the total variance in climate sensitivity.

Previous studies have found that low cloud shortwave feedback is an important factor in determining the climate sensitivity of a parameter-perturbed GCM (Webb et al. 2006). Figure 9 shows the relationship between

![Fig. 6. Histograms of the partial surface temperature response due to changes in (a) LW cloud transmittance (LW Cld), (b) SW cloud reflectivity (SW Cld), (c) LW CS atmospheric transmittance and (d) SW CS reflectivity at the surface (Sfc.). The partial temperature responses for each model are calculated by comparing equilibrium fluxes from both the double-CO2 and control experiments, using Eq. (4). In each case, the solid line represents the histogram of the partial temperature response in the climateprediction.net (blue) and CAMcube (red) ensembles.](image)

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
<th>Default</th>
<th>HiSens</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENTCOEF</td>
<td>Entrainment coefficient</td>
<td>3.0</td>
<td>0.6</td>
</tr>
<tr>
<td>RHCRIT</td>
<td>Critical relative humidity (sfc)</td>
<td>0.7</td>
<td>0.9</td>
</tr>
<tr>
<td>VF1</td>
<td>Ice fall speed (m s⁻¹)</td>
<td>1.0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>LW CS partial sensitivity (K)</td>
<td>6.3</td>
<td>2.4</td>
</tr>
</tbody>
</table>
shortwave cloud partial temperature response and control stage shortwave cloud forcing in different regions. Figure 9a shows the relationship between control high northern latitude upgoing shortwave radiation and the partial cloudy-sky shortwave local equilibrium temperature response for regions north of 60°N. First, it is notable that in all of the simulations in both climateprediction.net and CAMcube ensembles, there is a net decrease in temperature due to a negative shortwave cloud feedback at high latitudes. What is also clear is the relationship between the control state upgoing shortwave flux and the SW cloud warming in the double CO2 simulation. (The climateprediction.net correlation is significantly −0.65, while the CAMcube ensemble shows no significant correlation, which is again expected given the small spread in both response and mean state.)

The range of control simulation high-latitude upgoing shortwave radiation in climateprediction.net is extensive— a spread of more than 50 W m$^{-2}$—with the models, with the largest upgoing flux showing the largest decrease in temperature upon CO2 doubling. The range of control state upgoing radiation in CAMcube is too small to test this conjecture, but the ensemble results are consistent with the climateprediction.net-derived relationship. The results imply that a model that produces a larger area of low-lying, high-latitude cloud in the control simulation will also produce a larger increase in low-lying cloud in the double CO2 simulation. Interestingly, the CMIP3 models are equally widely distributed, implying that high-latitude upgoing radiation is not in itself a well-tuned quantity among current GCMs. The Clouds and the Earth’s Radiant Energy System, Edition-2 (CERES-2) observed value, however, is more consistent with a smaller upgoing flux; thus, if the climateprediction.net-derived relationship were robust, then this would imply relatively limited high-latitude negative feedback to surface warming.

Figure 9b shows the same relationship as Fig. 9a, but for the tropics and the midlatitudes. Here, the situation is very different: most models show a net positive shortwave cloud feedback in these regions, with a net increase in shortwave radiation hitting the surface. Once again, there is a large range upgoing radiation in the climateprediction.net control models (of almost 100 W m$^{-2}$, but only 20 W m$^{-2}$ in CAMcube), and there is a relationship between the control case upgoing radiation and the double CO2 response (a significant correlation of 0.50 in climateprediction.net and no significant correlation in CAMcube).
Models with an increased upgoing SW flux in the control simulation show a greater temperature increase in the double CO$_2$ simulation. Again, the change in cloud forcing upon surface warming is proportional to the original forcing. A major difference from the high-latitude case is that there is relatively more consensus among the CMIP3 models on the strength of the control state upgoing radiation (only a 12 W m$^{-2}$ range). Using Fig. 9b as a transfer function, all the CMIP3 models and the CERES-2 observations are consistent with a shortwave-induced temperature change of between $-0.5$ and $1.5$ K averaged over the tropics and midlatitudes in the double CO$_2$ climate.

Figure 10 shows the global mean shortwave cloud partial sensitivity as a function of both midlatitude and high-latitude cloud cover. This two-parameter space explains more than 74% of the total variance in cloudy-sky shortwave partial sensitivity in climateprediction.net and 42% of the variance in CAMcube. But, noticeably, the shortwave cloud partial sensitivities in CAMcube are consistent with climateprediction.net, given their low- and high-latitude fluxes.

Models with a large temperature increase due to shortwave cloud feedback have initially a large initial shortwave cloud forcing in the midlatitudes and a small forcing at high latitudes. In the double CO$_2$ experiment in the default Met Office Hadley Centre model configuration, shortwave cloud forcing tends to increase at high latitudes and decrease at low latitudes. We show here that these changes scale with the amount of cloud forcing initially present. Using the ensemble-derived regression, the CERES-2 observational estimates are actually consistent with a temperature increase due to shortwave cloud feedback, which is greater than either the default HadAM3 or the CAM3.5, suggesting a slightly positive net global partial shortwave cloud temperature response of 0.3 K with 95% confidence intervals of $-3.6$ and $+4.1$ K.

In climateprediction.net, the standard deviation of the shortwave clear-sky partial sensitivity is 0.2 K, accounting for 1% of the overall variance in climate sensitivity. In CAMcube, the standard deviation of the shortwave clear-sky response is 0.1 K, accounting for 10% of the variance in climate sensitivity. We do observe some correlation in climateprediction.net between
the global clear-sky shortwave partial sensitivity and the control state high-latitude shortwave top-of-atmosphere flux (shown in Fig. 9d). Models with a small amount of high-latitude cloud cover in the control simulation show a greater sensitivity to land albedo changes because the land surface is not masked in the shortwave by clouds overhead.

4. Discussions and conclusions

We have studied perturbed physics ensembles created by perturbing the atmospheric models of two atmospheric general circulation models—the Community Atmosphere Model version 3.5 and the Met Office Hadley Centre Atmosphere Model version 3.0. We have used a subset of 81 models from the preexisting climateprediction.net ensemble (Stainforth et al. 2005) and produced an identical-size ensemble using the CAM model; however, we find dramatically different distributions of global climate sensitivity in these two ensembles: the climate sensitivity ranges from 1.7 to 9.9 K in the climateprediction.net ensemble, whereas the range is only 2.2–3.2 K in the CAM ensemble. This in itself is not significant (in both ensembles, the parameter perturbations themselves are arbitrary and thus so are the distributions of climate sensitivity). However, it is at least an academic curiosity to determine why these two models respond so differently to similar parameter perturbations. Both ensembles show a diverse range of behavior when models are examined at a regional level, where both precipitation and temperature diagnostics show differences in both the preindustrial simulations and in the
response to greenhouse gas forcing to be comparable in both ensembles.

Jackson et al. (2008) also produced perturbed versions of the Community Atmosphere Model and found a similarly small range of climate sensitivity in their ensemble. The authors had a slightly different goal of producing a small number of optimal models that could be used to produce posterior probability distributions for various quantities. In contrast, our results are based upon a Latin hypercube sampling strategy that should, in theory, sample the most extreme model behavior. The fact that both experiments produce a similar range of climate sensitivity suggests that underlying structural differences between the CAM model and the Met Office Hadley Centre model are important in understanding their different behavior in a perturbed physics experiment.

Other studies (Webb et al. 2006; Sanderson et al. 2008b) of climate prediction.net and QUMP (Murphy et al. 2004) have suggested that differences in the initial state and development of low-level boundary layer clouds provide the major component of ensemble variation in climate sensitivity. Webb et al. (2006) found that 85% of the variance in total feedback in the QUMP ensemble is attributable to cloud feedbacks, with the majority of that figure associated with variation in negative cloud feedbacks resulting from increases in low-level cloud amount. Their analysis was based on the cloud feedback classification of Cess and Potter (1988), where cloud feedback is defined as the change in cloud radiative forcing during the simulation (which was further decomposed into local shortwave and longwave components).

In this work, we find that this finding is heavily dependent on how one defines cloud feedback. If, like Webb et al. (2006) and Sanderson et al. (2008b), we take cloud feedback to be simply the change in cloud radiative forcing and then we obtain similar results in this study, then the largest portion of variance in total feedback is explained by shortwave cloud feedbacks indicative of changes in low-level cloud amounts. However, although the climateprediction.net ensemble shows a greater variance of global shortwave cloud feedback (defined as the change in shortwave cloud radiative forcing per unit surface temperature change), the strongest positive shortwave cloud feedbacks in climateprediction.net and the CAM ensemble are of a similar magnitude, eliminating this as the cause of the highly sensitive models in climateprediction.net.

The potential problem with the approach of using change in cloud radiative forcing to represent cloud feedbacks, as the authors of Webb et al. (2006) acknowledge, is the presence of so-called cloud masking effects, in which changes in clouds also change the clear-sky fluxes, thus underestimating the true radiative effect of cloud changes. A simple example of this phenomenon can be observed when an increase in cloud at a given model level increases humidity, which is counted as part of the clear-sky feedback. However, the converse is also true—the use of changes in cloud radiative forcing as a measure of cloud feedback can also lead one to believe that there has been cloud feedback without any change in cloud distribution.

Another aspect of the Webb et al. (2006) and Sanderson et al. (2008b) studies is that they both look for processes...
that explain variance in global feedback parameter (the inverse of climate sensitivity), but it is not necessarily true that the same leading-order processes explain the most variance in climate sensitivity itself. Given most model-simulated future climate change scales approximately linearly with the climate sensitivity of the system (Knutti et al. 2008), it makes sense to find processes that explain variance in climate sensitivity itself, rather than its inverse.

With this in mind, we use a methodology for approximately separating the equilibrium temperature response of each model into different components, associated with changes in longwave transmittance and shortwave reflectivity of the atmosphere and ground. If one assumes that changes in these properties are perturbations on the initial state (and ignoring products of any perturbation terms), then the equilibrium temperature change of the system can be expressed as a sum of components relating to changes in the clear-sky and cloud transmittance of the atmosphere, and the clear-sky and cloud reflectivity.

The longwave clear-sky component includes both the final forcing due to the doubled-\(\text{CO}_2\) concentrations, together with any water vapor feedback that might further decrease the atmospheric transmittance. This is important because the final forcing at the end of the double-\(\text{CO}_2\) simulation is not knowable without conducting a partial radiative perturbation-type experiment (Wetherald and Manabe 1988), substituting preindustrial carbon dioxide concentrations into an equilibrated double-\(\text{CO}_2\) simulation. Only with this information could the longwave clear-sky term could be further separated into a forcing and a feedback.

Correlations between the control state and partial clear-sky sensitivity suggest increased upper-tropospheric (UT) water vapor concentrations may be decreasing the atmospheric transmittance and amplifying initial greenhouse gas forcing. In addition, those models with increased UT water vapor in the preindustrial simulation tend also to show large increases in UT water vapor during the double-\(\text{CO}_2\) simulation, which acts as a further positive feedback in the system. None of the models in the CAM ensemble show any UT water vapor concentrations of the same magnitude. This raises several questions for future study. First, what aspect of the HadAM3 model (or the perturbations made within it) makes this high-level water vapor feedback possible, and why are these processes not apparent within the CAM ensemble? Second, is there any statement of likelihood for “real world” climate sensitivity that might be made from an understanding of the mechanisms present in the ensemble?

We have shown that to achieve partial clear-sky longwave sensitivities of greater than 4 K, the model must display high upper-tropospheric humidity in the control simulation. None of the CAMcube simulations show these characteristics, and we do not see a large range in clear-sky sensitivity. This evidence is not conclusive, but it suggests that the dominant cause of the differing sensitivity distributions in the two ensembles lies in the longwave clear-sky response. A recently accepted paper (Joshi et al. 2010) supports this conclusion when analyzing the low entrainment simulations in the QUMP ensemble, finding that the large sensitivities are largely attributable to high upper-tropospheric water vapor concentrations.

The cloud response also plays a role in determining the sensitivities of models in both ensembles. On a large scale, we find that all models in both ensembles tend to exhibit a reduction in net shortwave cloud forcing in the midlatitudes and an increase at high latitudes. Generally, we find that these changes scale with the amount of shortwave cloud forcing present in the control simulation. Hence, models with a large control shortwave cloud forcing at high latitudes (poleward of 60°) exhibit a more negative high-latitude cloud feedback than the average model. Similarly, models with a large control midlatitude shortwave cloud forcing exhibit a more positive than average shortwave cloud feedback. We find that parameters tend to influence high- and low-latitude shortwave cloud forcing in a similar way, so globally we see a large amount of cancellation between high- and low-latitude shortwave cloud feedbacks. However, the difference between the control state high-latitude and midlatitude shortwave cloud forcing provides a reasonable metric for the shortwave cloud response in the double-\(\text{CO}_2\) simulation.

We do not find any significant correlation between longwave cloud forcing in the control and in the double-\(\text{CO}_2\) simulation; however, as in the clear-sky longwave case, we find that the longwave cloudy-sky sensitivity is strongly affected by upper-tropospheric humidity. Finally, the variation in shortwave clear-sky sensitivity is sufficiently negligible in both ensembles that finding observable correlated quantities is impossible.

This study highlights various correlations between observable quantities and model response to greenhouse gas forcing in the climatedprediction.net and CAMcube ensembles. Further study could be warranted in using these relationships to extend the analysis of Knutti et al. (2006), which used regional seasonal cycles of surface temperature to predict model climate sensitivity in climatedprediction.net. The correlations and physical mechanisms proposed here may provide additional information to the neural network prediction algorithm used by Knutti et al. (2006), especially in tropical regions where there is no significant seasonal cycle to act as a predictor.

This study serves to illustrate the extreme dependency of perturbed physics ensembles on both the underlying
climate model and on the chosen parameter sampling strategy. A process-based examination of climate feedbacks in both ensembles has revealed various possible mechanisms that can be used to physically exclude some, but not all, high-sensitivity models in the subset of the climateprediction.net ensemble considered. The importance of upper-tropospheric and lower-stratospheric water vapor in determining the longwave response of the HadAM3 model suggests that further study of upper-level humidity in both model- and observation-based studies is warranted.

Acknowledgments. This work was only made possible with the efforts of the team who developed and continue to support the climateprediction.net project, in particular, Myles Allen, David Stainforth, David Frame, Tolu Aina, Carl Christensen, and Neil Massey, who were instrumental in preparing the particular simulations used in this study, as well as the members of the public who donated their idle computing time to the project. Thanks to Gerald Meehl, Richard Neale, Charles Jackson, and William Ingram for their helpful input and discussions and for valuable input from anonymous reviewers. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP’s Working Group on Coupled Modelling (WGCM), for their roles in making available the WCRP CMIP3 multimodel dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

APPENDIX A

Choosing Perturbed Parameters

To create a manageable ensemble size from the climateprediction.net grand ensemble, we select four parameters that previous studies have shown to be most important in controlling model sensitivity (Murphy et al. 2004; Sanderson et al. 2008a,b; Knight et al. 2007). Given each parameter was sampled at three values, all possible perturbations of these four parameters yield 81 models. The parameters are listed in Table 2.

The process for the CAM ensemble was much simpler because simulations were run in house, hindered only by the fact that identical parameter perturbations were not always possible because of different parameterization schemes used in HadAM3 and CAM. Critical relative humidity is defined differently in both models; that is, in climateprediction.net, it is defined on model levels, whereas in CAM it is defined on “high” or “low” levels. In climateprediction.net, three different vertical profiles were assumed that only varied significantly in the upper troposphere; thus, in the CAM ensemble, we choose to perturb only the “high” critical relative humidity. Ice fall speed parameterization is slightly different in each model. The Met Office Hadley Centre model uses empirical data from Heymsfield (1977), where the fall velocity (VF1) parameter scales the ice fall velocity of all particles. CAM models the terminal velocity of small ice particles (~40 μm) as spheres in Stokes’s equation (Boville et al. 2006), while larger particles are modeled empirically as a linear function of particle radius, according to measurements by Locatell and Hobbs (1974). The V400 parameter accounts for uncertainty in these measurements.

The convective schemes used in each model are different. HadAM3 uses Gregory and Rowntree (1990), where convection is strongly modified by the entrainment coefficient. In CAM3.5, with no such parameter in the deep convection (Zhang and McFarlane 1995) or shallow convection (Hack 1994) schemes, we instead perturb the treatment of convective mass flux (Raymond and Blyth 1992). Finally, there is no direct analog for the cloud water autoconversion of cloud water to liquid parameter “ct” in the CAM model, so instead we perturb the threshold for the autoconversion of cold ice “iccrit,” which is a resolution-dependent tuning parameter in the model. Perturbed parameters and values are listed in Table 3.

APPENDIX B

Top-of-Atmosphere Feedbacks

Top-of-atmosphere feedbacks shown in Fig. 5 are calculated using information from the double-CO2 simulation. Clear-sky feedbacks are calculated by assuming a linear relationship between top-of-atmosphere clear-sky longwave or shortwave radiative flux, using ordinary least squares regression to calculate the gradient that

---

Table 2. Definition of perturbed parameters in the 81-member subset of climateprediction.net experiments used in this analysis. Parameters marked with a dagger are actually defined on 19 model levels, with the mean over model levels shown here. The right column indicates the possible values used in the experiment for each parameter, with units (where applicable) shown in parenthesis.

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENTCOEFF</td>
<td>Entrainment coefficient</td>
<td>[0.6, 3.0, 9.0]</td>
</tr>
<tr>
<td>RHCritt</td>
<td>Critical relative humidity</td>
<td>[0.65, 0.73, 0.90]†</td>
</tr>
<tr>
<td>CT</td>
<td>Accretion constant (s⁻¹)</td>
<td>[40, 10, 5] × 10⁻⁵</td>
</tr>
<tr>
<td>VF1</td>
<td>Ice fall velocity for small concentrations (cm s⁻¹)</td>
<td>[0.5, 1, 2]</td>
</tr>
</tbody>
</table>
indicates the change in global mean top-of-atmosphere cloud forcing or clear-sky radiative flux for a 1-K global mean surface temperature rise. For the cloud feedbacks, the CRF is calculated both in the longwave and shortwave components of the budget by subtracting the clear-sky flux from the total upgoing top-of-atmosphere flux. In a similar fashion, a regression is performed to calculate the CRF change for each 1-K surface temperature rise.

**APPENDIX C**

**Partial Surface Temperature Response**

We consider a simple slab atmosphere over a black-body surface with a transmittance $\tau$ and reflectivity $\alpha$. With an incoming solar flux $S_\text{s}$, we can write an equation for the equilibrium temperature $T_g$,

$$\sigma T_g^4 = \frac{S_\text{s}(1 - \alpha)}{\gamma}.$$  \hfill (C1)

Furthermore, we can separate the $\gamma$ and $\tau$ terms into clear-sky and cloudy-sky components:

$$\gamma = \gamma_{cs} + \gamma_{cld}$$ \hfill (C2)

$$(1 - \alpha) = (1 - \alpha_{cs})(1 - \alpha_{cld}),$$ \hfill (C3)

where $\gamma_{cs}$ and $\gamma_{cld}$ are top-of-atmosphere clear-sky radiation and longwave cloud radiative forcing, respectively, expressed as a fraction of upgoing surface longwave radiation. This “normalized greenhouse effect” was a concept introduced by Raval and Ramanathan (1989).

In the shortwave, $\alpha$ and $\alpha_{cs}$ are the top-of-atmosphere upgoing all-sky and clear-sky fluxes, respectively, expressed as a fraction of the downward top of atmosphere solar flux (allowing one to trivially solve for $\alpha_{cld}$).

We then introduce perturbations ($\Delta\gamma_{cs}$, $\Delta\gamma_{cld}$, $\Delta\alpha_{cs}$, $\Delta\alpha_{cld}$) on each of the terms ($\gamma_{cs}$, $\gamma_{cld}$, $\alpha_{cs}$, $\alpha_{cld}$) and solve for the temperature-resulting equilibrium temperature perturbation $\Delta T_g$, ignoring any products of perturbation terms,

$$\Delta T_g \approx \frac{T_g}{4} \left( \frac{\Delta\gamma_{cs}}{\gamma_{cs} + \gamma_{cld}} + \frac{\Delta\gamma_{cld}}{\gamma_{cs} + \gamma_{cld}} - \frac{\Delta\alpha_{cs}}{1 - \alpha_{cs}} - \frac{\Delta\alpha_{cld}}{1 - \alpha_{cld}} \right).$$  \hfill (C4)

thus allowing the total surface temperature response to be approximately separated into four components relating to changes in the following: longwave atmospheric opacity in the clear sky and cloudy sky and shortwave reflectivity in the clear sky and cloudy sky.

We can also express the longwave cloud radiative forcing and the clear-sky top-of-atmosphere flux in the terms defined above:

$$\text{CRF}_{lw} = \gamma_{cld}\sigma T_g^4$$ \hfill (C5)

$$\text{CS}_{lw} = \gamma_{cs}\sigma T_g^4$$ \hfill (C6)

such that the derivatives with temperature show the traditionally defined cloudy-sky and clear-sky feedback

$$\frac{d\text{CRF}_{lw}}{dT} = \Delta\gamma_{cld}\sigma T_g^4 + 4\gamma_{cld} T_g^3 \Delta T$$ \hfill (C7)

$$\frac{d\text{CS}_{lw}}{dT} = \Delta\gamma_{cs}\sigma T_g^4 + 4\gamma_{cs} T_g^3 \Delta T.$$ \hfill (C8)

These statements serve to illustrate why the CRF/clear-sky feedback separation partitions the response differently to the partial surface temperature response technique. In Eq. (C7), the right-hand terms are dependent only on the initial cloudy-sky and clear-sky transmittance, and the surface temperature rise. Thus, the clear-sky and cloudy-sky feedbacks as defined in Eq. (C7) can be nonzero even if there is no change in $\gamma_{cld}$ or $\gamma_{cs}$.

In contrast, in Eq. (C4), the cloudy-sky and clear-sky terms making the total longwave temperature response are proportional to $\Delta\gamma_{cld}$ and $\Delta\gamma_{cs}$, respectively, and the response is amplified by the total opacity of the atmosphere. If the greenhouse gas forcing is known, then Eq. (C7) can be used to derive the longwave temperature response—identical to that of Eq. (C4)—so neither formulation is incorrect. However, the partial temperature response approach has the advantage that the components are linearly related to the climate sensitivity, and it is easier to relate each term to a physical change in the system itself, rather than an aspect of the initial state.

**REFERENCES**


