A multi-model study of parametric uncertainty in response to rising greenhouse gases concentrations

Benjamin M. Sanderson¹

National Center for Atmospheric Research, Boulder, Colorado, USA

DRAFT

21 October 2010

Corresponding author address: Ben Sanderson, National Center for Atmospheric Research, 1850 Table Mesa Dr, Boulder, CO, 80305, USA.

E-mail: bsander@ucar.edu
ABSTRACT

When considering uncertainty in the earth system response to future greenhouse gas emissions, General Circulation Models (GCMs) of the atmosphere, ocean, land and biosphere are essential tools for predicting the future evolution of the climate. A proper understanding of the uncertainties associated with model simulations is essential, and one tool for studying the range of possible response is to consider ensembles of model versions where parametrizations have been sampled within their physical range of plausibility. Here, we study 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 CAMcube ensemble using the Community Atmosphere Model version 3.5. We find the two ensembles produce very different distributions for climate sensitivity - with the climateprediction.net ensemble subset ranging from 1.7K to 9.9K while the CAMcube ensemble range is 2.2K to 3.2K. On a regional level, however, both ensembles show a similarly diverse range in their mean climatology. A study of model radiative fluxes in the carbon dioxide doubling experiment suggests that the major difference between the range of response of the two ensembles lies in the treatment of upper tropospheric water vapor. Large clear-sky feedbacks present only in the climateprediction.net ensemble are found to be proportional to significant increases in upper tropospheric water vapor, which are not observed in the CAMcube ensemble. Some models in the climateprediction.net ensemble which have no clearly defined inter-tropical convergence zone (ITCZ) are also shown to exhibit unrealistically strong water vapor feedbacks when an ITCZ forms in the double CO₂ simulation. Both ensembles show a range of shortwave cloud feedbacks, but increased negative feedbacks at high latitudes are generally compensated by increased positive feedbacks at lower latitudes. Using large scale observations to constrain all of these processes results in a weak constraint on sensitivity due to large systematic model errors, but models with climate sensitivities of greater than 5K can be shown to be inconsistent with observations.
1. Introduction

In order to better understand the relationship between model parametrization 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) in contrast to 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)).

In order to make this comparison as direct as possible, we have followed the experimental design for the [climateprediction.net](http://climateprediction.net) project, outlined in Stainforth et al. (2005), and applied the methodology to CAM in order to make a new ensemble, CAMcube. This involves choosing a selection of model parametrizations whose values are uncertain from fundamental principles and choosing one of three possible values for each where the extreme values represent reasonable limits of uncertainty. Multiple combinations of parameter perturbations result in an ensemble of model versions, which might show a range of responses to changing greenhouse gas forcing.

The large size of the [climateprediction.net](http://climateprediction.net) ensemble was made possible by using a distributed computing architecture described in Stainforth et al. (2005), whereas the simulations which make up CAMcube were conducted on parallel supercomputing facilities which places a limit on the potential ensemble size. However, with the benefit of previous studies (Sanderson et al. (2008b), Sanderson et al. (2008a)) we isolate the dominant parameters influencing climate sensitivity in the [climateprediction.net](http://climateprediction.net) ensemble to consider a 4 parameter subspace, sampled at 3 points in each dimension giving 81 unique parameter combinations. We find parameters most similar to these in the CAM and produce a similar 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.

The model used is the Community Atmosphere Model model version 3.5, coupled to a thermodynamic slab ocean for reasons of consistency and speed of simulation. The atmospheric model shares many components with the CAM 3.0 model used in the IPCC AR-4, with additional improvements to the dynamics and the convection scheme. The
Figure 1: Global annual mean temperature plotted for each member considered in both the climateprediction.net ensemble and the CAM ensemble. Three 15 year simulations are shown, a calibration, control and double carbon dioxide simulation.

experiments conducted are identical to those in climateprediction.net: three 15 year simulations for each model version: (1) a calibration stage to determine the ocean-atmosphere heat flux necessary to maintain observed sea surface temperatures, (2) a control stage to test the stability of the model with constant pre-industrial greenhouse gas forcing and (3) a double-co2 run in which carbon dioxide concentrations are doubled at the start of the simulation so that climate sensitivity may be inferred by assuming the global mean temperature follows an exponential decay towards a new equilibrium temperature.

2. Climate Sensitivity Distributions

Figure 1 shows model development of the simulations in the CAM ensemble as compared to a 2,000 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 double-co2 stage, and a significant number of drifting simulations in the control stage - neither of these features is clearly apparent in the CAM ensemble. Instead, most simulations show qualitatively little variation in sensitivity to external forcing, and none of the simulations display a major drift in the control stage.

Figure 2(a) demonstrates that the distribution of climate sensitivity within the CAM ensemble is noticeably narrower than is the case for climateprediction.net. The CAM
Figure 2: Distributions for climate sensitivity created from the CAM and climateprediction.net ensembles. (a) raw histograms for each of the ensembles, using 1K bins of climate sensitivity and with values normalized by the total number of simulations in the ensemble. Both a 2000 member and an 81 member subset of the climateprediction.net ensemble are shown. (b) Distributions following a methodology similar to Murphy et al. (2004), for the CAMcube (blue) and climateprediction.net (red) ensembles. For both plots, box and whisker plots show the 5th, 10th, 50th, 90th and 95th percentiles of each distribution.

ensemble values of climate sensitivity range from 2.2 to 3.2K, whereas the 2,000 model subset from the climateprediction.net ensemble range from 0.2K to 11.8K (excluding those models, as in Stainforth et al. (2005), with a significant negative drift in the control simulation). Reducing the climateprediction.net ensemble to a 4 parameter subspace containing 81 models has little effect on the distribution of sensitivities in the ensemble, with upper and lower bounds of 1.7K and 9.9K.

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) and Piani et al. (2005), have proposed methods of overcoming the effects of sampling on the resulting distribution. These two studies both 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 2(b) 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 $10^4$ simulations in each case. Frame et al. (2005) argued that such an interpolation should not, however, be viewed as a Probability Density Function (PDF) because the prior sampling distribution for the distribution is uniform within the parameter space, and thus the posterior distribution is fundamentally dependent upon arbitrary choices made during the experimental design process.

The second study, Piani et al. (2005), should not be so directly dependent on the sampling of sensitivity in the ensemble as the methodology relies on finding correlations between climate sensitivity and observable quantities in the pre-industrial simulation of the model. Once the correlations are found within the ensemble, additional error is added to the result to account for the prediction error (this error is derived from the ensemble itself). An estimate for climate sensitivity was obtained by using the observed correlations together with observations to predict the ‘real-world’ sensitivity. There remains an unknown ‘systematic error’, which describes how accurate the predictors might be when applied to a different climate model, or to the observations. This error was estimated by adding a bias to the model derived noise estimate, which widens the resulting distribution (although the extent of the widening is likely to be dependent on the choice of bias). Although we do not, in this work, follow the methodology of Piani et al. (2005), our approach is more akin to the basic principal of finding correlations between observable quantities and unknown response metrics and using observations to constrain the real-world response.

2.1. Regional Responses

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, it can easily be demonstrated that on a regional scale, there are large variations both in control climatology and in the local response to greenhouse gas forcing. Figure 3
Figure 3: Scatter-plots 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 pre-industrial annual mean temperature for the region, while the vertical axis shows the regional equilibrium temperature response to carbon dioxide doubling. Histograms on each axis represent the distribution of that quantity in the ensemble.
Figure 4: As for figure 3 but for annual total precipitation.
shows both pre-industrial regional mean temperatures and the regional warming on CO₂ doubling, and shows that the CAM ensemble produces a range of pre-industrial climates which is not noticeably less diverse than climateprediction.net. Many regions appear to show some correlation in one of the two ensembles between pre-industrial temperature and regional warming. However, in nearly all cases - these correlations do not appear to be significant in the other ensemble implying that they are highly dependent on the underlying model and the perturbations made. A similar plot is shown in Figure 4 showing annual mean precipitation in both ensembles, again - it is clear that both the pre-industrial precipitation and the precipitation response to CO₂ doubling varies considerably in both ensembles, implying that we have not simply perturbed ‘null’ parameters in the CAM ensemble.

It remains a question to understand why there is such a marked difference between the range of climate sensitivity in one ensemble, and not in the other. To begin to address this, we first investigate the effects of the perturbed parameters on climate sensitivity in each of the ensembles. It has been demonstrated before (Stainforth et al. 2005), (Sanderson et al. 2008a) that the sensitivity in the climateprediction.net parameter space is not well represented as a linear function. To demonstrate this, we show in Figures 5(a) and 5(b) a visualization of the behavior of climate sensitivity in the model parameter space of each ensemble. It is instantly apparent that both ensembles show a strongly non-linear parameter dependency for climate sensitivity.

In the case of the CAM ensemble shown in Figure 5(a), there is not a large variation in climate sensitivity, but the largest sensitivities of 3.4K are achieved with a specific combination of parameters - including the switching off of entrainment of mass into a convective plume (leaving the convecting parcels undiluted) (Raymond and Blyth 1992). Similar distributions for climateprediction.net are shown in Figure 5(b), where it is clear that the largest climate sensitivities require multiple perturbations from the default model. Once again a perturbation to decrease entrainment into convecting plumes is required for the model to show a climate sensitivity of greater than 6K, and as in the CAM case, this perturbation produces consistently poorer scoring models in the multi-variable metric.
Figure 5: Three dimensional contour plots representing surface contours of climate sensitivity in the model parameter space of the CAMcube ensemble. Each axis represents one of three parameter dimensions, while the three plots represent slices through a fourth parameter dimension. Values of climate sensitivity are interpolated using a cubic spline approximation between the 81 sampled points in each hypercube.
However, sensitivities of up to 6K can be achieved in the [climateprediction.net](http://climateprediction.net) without any perturbation of the convection scheme through combined perturbations of the other three parameters - and these models exhibit model scores similar to that of the unperturbed simulation, which has a climate sensitivity of 3.1K.

3. Global feedbacks

3.1. Top of atmosphere feedbacks

Given the results of Section 2, it is of some interest to know why perturbed experiments using the Hadley Centre model display so much more variation in global mean response to greenhouse gas forcing than those using the Community Atmosphere Model. In order to make this more apparent, it is useful to break down the distributions of global response into different parts of the radiative budget. Figure 6 shows histograms of the top of atmosphere response to surface warming in both ensembles. This metric shows the change in clear-sky flux and cloud radiative forcing in both the longwave and shortwave top of atmosphere budgets for a 1 Kelvin increase in global mean temperature (detailed in Section B).

In the case of the ensembles under consideration here, one would gain the impression by studying the distributions of global mean feedback shown in Figure 6 that the most dramatic differences between model response in the [climateprediction.net](http://climateprediction.net) ensemble occur as a result of vastly different shortwave cloud forcing responses to surface warming. This is evident simply by comparing the width of the feedback distribution in Figure 6(b) with the width of the other distributions, and has been suggested as the major component of top of atmosphere feedback response determining global mean response to greenhouse gas forcing in both perturbed physics and multi-model ensembles (Webb et al. (2001), Sanderson et al. (2008b), Trenberth and Fasullo (2009)).

However, this explanation fails to explain why high sensitivities are not found in the CAM ensemble. In the convention used here, more negative values of climate feedback cause the model sensitivity to become higher, and yet only three models in the subset
Figure 6: Histograms of top of atmosphere response to one degree kelvin surface warming, as calculated from the double CO2 simulation of each model. Histograms show the response of top of atmosphere cloud radiative forcing in longwave (a) and shortwave (b), and clear-sky response in longwave (c) and shortwave (d). The vertical axis shows the number of simulations in each 0.1 Wm$^{-2}$K$^{-1}$ bin in the climateprediction.net (blue) and CAMcube (red) ensembles.
of \url{climateprediction.net} considered show cloud feedback more negative than any model found in the CAM ensemble. It is clear that both ensembles display a reasonable range of global mean cloud feedback and other processes are playing a fundamental part in causing the high sensitivity in \url{climateprediction.net}.

While it is certainly true that total top of atmosphere feedback strength is inversely related to climate sensitivity, this statement ceases to be true when considering each of the component parts of the total feedback - making their interpretation impossible without knowing all the other feedbacks in the system. Two assumptions are implicitly made when one uses the inverse of the sum of top of atmosphere feedbacks to estimate climate sensitivity: that radiative response is a linear function of surface warming and that greenhouse gas forcing can be considered to be constant after a step change in CO$_2$ concentrations. However, neither of these assumptions is necessarily true. The linearity of the radiative response to changing surface temperatures is not true of the clear-sky longwave feedbacks due to the fundamentally non-linear nature of the water vapor feedback (Held and Soden 2000). The initial forcing due to CO$_2$ is also demonstrably not constant within the ensemble (Figure 15(c)). In order to provide to find a more meaningful metric for breaking down the climate sensitivity into components, we consider components of the temperature response itself, rather than its inverse.

3.2. 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 clear-sky and cloudy sky components. To achieve this exactly is impossible, because feedbacks are coupled - but a result can be approximated by considering perturbations from in the atmospheric transmissivity and albedo from the mean state and ignoring any products of perturbation terms. The method is described in Section C and the distributions of the longwave and shortwave partial sensitivity for clear-sky and cloud is shown in Figure 7.

The distributions in Figure 7 are approximate separations of the global mean surface temperature response into components, but a validation against the true model climate
Figure 7: Histograms of the partial surface temperature response due to changes in (a) longwave cloud transmittance, (b) shortwave cloud reflectivity, (c) longwave clear-sky atmospheric transmittance and (d) shortwave clear-sky reflectivity. The partial temperature responses for each model are calculated by comparing equilibrium fluxes from both the double CO$_2$ and control experiments, using Equation 4. In each case, the solid line represents the histogram of the partial temperature response in the climateprediction.net (blue) and CAMcube (red) ensembles.
sensitivity indicates that the approximations are justified (Figure 15). In the case of the partial surface temperature response to changes in cloud forcing, the results are unsurprising given the distribution of model feedbacks. Longwave partial cloud sensitivities (Figure 7(a)) are between -0.6K and +0.1K in the climateprediction.net ensemble, with a narrower distribution of -0.4K to -0.2K in the CAMcube ensemble (where lower and upper limits are the 5th and 95th percentiles of partial warming in each ensemble). Partial shortwave cloud sensitivities range from -0.9K to +0.2K in climateprediction.net and -0.5K to +0.2K in CAMcube. So, although there is clearly a greater range of cloud feedback response in climateprediction.net than in CAMcube - it is not enough to account for the large (>6K) climate sensitivities in the climateprediction.net ensemble.

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 ranges from +2.2K to +4.9K while in CAMcube, the range is +2.2K to +2.4K. When considering this values, it is useful to remember what they represent: the surface temperature change between the single and double carbon dioxide simulations which may be attributed directly to a change in longwave atmospheric clear-sky transmittance (Section C). Hence 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. What is clear from this distribution is that up to 5K of warming is due to clear-sky forcing and feedbacks alone in climateprediction.net, but in the CAM ensemble there is almost no variation. Thus it becomes of paramount importance to understand the physical difference between the clear-sky response in each ensemble.

4. Mechanisms leading to high climate sensitivity

In section 2 it is shown that there is a marked difference in the distribution of model climate sensitivity in the climateprediction.net and CAMcube ensembles, with the former showing a much larger range of global mean temperature response to greenhouse gas forcing. Explaining this systematic difference requires consideration of different feedback mechanisms which might amplify the radiative forcing of the climate system in a warming
Table 1: Parameter values in the “default” HadAM3 simulation and in a “HiSens” HadAM3 model which exhibits a large clear-sky sensitivity.

4.1. Longwave Temperature Response

Section 3.2 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, whereas little such variation is seen in the CAMcube ensemble. Explaining this difference in model response would thus show to a large extent why the model sensitivity distributions are so different. We begin by examining two models from the climateprediction.net ensemble - the default HadAM3 simulation where no parameters have been perturbed and a high-sensitivity model with a large partial clear-sky longwave sensitivity. The model perturbed parameters and longwave clear-sky response are shown in Table 1.

Figures 8(a) and (b) respectively show the annual tropical mean temperature and relative humidity distributions in the last 8 years of both the single and double CO2 simulations, as compared to the “default” HadAM3 single CO2 simulation. We see that the control “HiSens” simulation has a warmer troposphere and a cooler stratosphere than the default simulation (although the sea surface temperatures are constrained to observed climatology by the 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 significantly more tropospheric warming and
stratospheric cooling (as compared to “default”), which would be expected given an enhanced water vapor feedback which would be 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).

The other major difference between the two models lies in the tropical lower and middle troposphere, where the “HiSens” model shows significant drying in the control simulation (as compared to “default”). This drying, however, is significantly reduced in the double CO$_2$ simulation. This can be shown to be a largely dynamical effect. Figure [8(c)] shows the Northern Hemisphere summer meridional wind profile for the two models. In the control simulation, the “HiSens” simulation forms a clear double convergence zone in the tropics, however, in the double CO$_2$ simulation, the model forms a standard single ITCZ. Associated with this instantaneous formation of an ITCZ in the double CO$_2$ simulation - the “HiSens” model exhibits a large increase in humidity in the mid troposphere. This effect can easily be seen in the zonal mean 850mb relative humidity profile (shown in Figure [8(d)]) and is not true of the “default” simulation which has a well-defined single ITCZ in both single and double CO$_2$ conditions.

After establishing the mechanisms leading to a large partial clear sky sensitivity in these two models, we can examine how these processes relate to the longwave clear-sky sensitivity in both the [climateprediction.net](http://climateprediction.net) and CAMcube ensembles. Concentrating first on UTLS humidity, Figure [9(a)] shows the clear-sky longwave partial sensitivity of models in both the CAMcube and [climateprediction.net](http://climateprediction.net) ensembles as a function of 150mb specific humidity in the pre-industrial control simulation. In the [climateprediction.net](http://climateprediction.net) ensemble - these two quantities have a statistically significant correlation of 0.74. The quantities are not significantly correlated in CAMcube, mainly because the standard deviation of the clear-sky longwave sensitivity is only 0.06K, compared to the [climateprediction.net](http://climateprediction.net) standard deviation of 1.02K.

All [climateprediction.net](http://climateprediction.net) models with UTLS specific humidities of greater than 20ppmv display a partial longwave clear-sky sensitivity of greater than 4K (compared to the default configuration value of 2.5K). Also shown on the plot is the humidity values taken from
Figure 8: Properties of two models in the ensemble - a high sensitivity model “HiSens” (Table I) and the “default” HadAM3 configuration, each plot is time mean over the last 8 years of the control and X2CO2 simulations respectively. (a) shows the tropical mean temperature profile, (b) is the tropical mean relative humidity profile, (c) is the zonal mean surface meridional wind strength in June, July and August (JJA) and (d) is the zonal mean relative humidity at 850mb in JJA.
Figure 9: (a) clear-sky and (b) cloudy-sky partial model climate sensitivity as a function of 150mb global mean specific humidity in the pre-industrial control simulation. Points in blue and red represent models from the climateprediction.net and CAMcube ensembles respectively. The center of the vertical gray bars lines show the mean humidity values inferred from AIRS data Gettelman and Fu (2008), while the width of the bar represents the 5-95 percentiles of the distribution about the mean (estimated using the distribution of 64x15 year periods about the mean in a CAM 3.0 control simulation). Vertical green lines show humidity values in CMIP-3 models with suitable data available from a pre-industrial control simulation.

Both AIRS observations and the CMIP-3 pre-industrial control simulation. The AIRS data is inconsistent with those climateprediction.net models which exhibit a partial clear-sky longwave sensitivity of greater than 4K. Only one model in the CMIP-3 ensemble (CSIRO Mk 3.0), exhibits a 150mb specific humidity of greater than 20 ppmv.

Also shown is the partial longwave cloudy sensitivity as a function of UTLS specific humidity (Figure 9(b)) - which is also significantly correlated in the climateprediction.net ensemble with a correlation of 0.74 (again, there is no significant correlation in CAMcube and the spread of response is much smaller).

The second clear-sky longwave mechanism causing large climate sensitivities in climateprediction.net related to a double ITCZ in the control simulation and the formation
of a single ITCZ in the double CO$_2$ simulation. This leads to a dramatic middle troposphere increase in humidity but is clearly an unphysical feedback. In order to provide a simple control simulation diagnostic for this behavior, we use the June, July, August (JJA) zonal mean meridional surface wind profile. We take principal components in the latitudinal/ensemble domain such that the first EOF has dimensions $n_{lat}$ (the number of latitudinal domains, which is 72 in this case) and is the mode which describes the most inter-model ensemble variance in the 81 member climateprediction.net subset. After removal of the mean meridional wind profile, the leading EOF clearly represents a dual ITCZ pattern (Figure 10). Hence, the coefficients of this pattern projected onto the meridional wind profile of the control simulation from any model provide a reasonable, simple metric detailing the extent of this behavior in that model. We demonstrate this in Figure 10(b), which plots the pattern coefficients for each model in the CAMcube and climateprediction.net ensembles against the tropical 850mb relative humidity increase between the control and double CO$_2$ simulations. It is notable from this plot that a double ITCZ pattern is a necessary but not sufficient condition for a significant increase in lower tropospheric tropical relative humidity. Some models displaying relatively large values of the pattern coefficient do not show an RH increase in the X2CO2 simulation. In these models, both the control simulation and the X2CO2 simulation have a double ITCZ behavior - a single ITCZ does not form upon CO$_2$ doubling.

We have isolated two mechanisms leading to an increased partial longwave clear-sky response with corresponding observable metrics in the control simulation (the two metrics being control values of global mean 150mb specific humidity and the projection of the double ITCZ mode onto the model’s control zonal mean JJA meridional wind profile).

In Figure 11 shows the partial clear-sky longwave sensitivity as a function of these metrics. In the climateprediction.net ensemble, the two metrics are not significantly correlated, but together explain 85% of the total variance in clear-sky longwave sensitivity. In CAMcube, there is no significant correlation between the metrics and the response - but as discussed earlier, the variance of the global mean clear-sky longwave response in CAMcube is 2 orders of magnitude smaller than for climateprediction.net and is not a
Figure 10: (a) The first principal component of annual, zonal-mean meridional surface wind in the climateprediction.net ensemble (mode is normalized thus the vertical axis is unitless). (b) Tropical mean change in 850mb Relative Humidity plotted as a function of coefficients of the EOF pattern projected onto each member of the climateprediction.net (blue) and CAMcube (red) ensembles. Also shown for reference are the coefficients of the pattern projected onto members of the CMIP-3 archive (green). The value from the NCEP reanalysis is shown by the center of the vertical gray bars, while the width of the bar represents the 5-95 percentiles of the distribution about the mean (estimated using the distribution of 64x15 year periods about the mean in a CAM 3.0 control simulation).
Figure 11: Partial clear-sky longwave sensitivity plotted as a function of two observable control model metrics. The horizontal axis shows global mean 150mb control state specific humidity in parts per million volume, while the vertical axis shows the coefficients when the zonal-mean meridional wind profile projected onto the EOF pattern in figure 10(a). Colored squares and diamonds are models in the climateprediction.net and CAMcube experiments respectively, and the larger squares and diamonds are the default, unperturbed model configurations - the coloration depicts the model partial clear-sky longwave sensitivity (as explained in Appendix C). Black points are models in the CMIP-3 ensemble. The horizontal line is the coefficient of the NCEP wind profile projected onto the EOF, while the vertical line is the 150mb specific humidity inferred from AIRS data (Gettelman and Fu 2008). The oval at the cross-point of the two observational lines shows the 5-95% natural variability, by removing the mean from a 500 year CAM 3.0 control simulation, and expressing the variation about the observed mean.

significant factor in determining climate sensitivity in the CAMcube ensemble.

Also shown in Figure 11 are the predictive metrics diagnosed from control pre-industrial simulations using models in the CMIP-3 archive, as well as ‘observed’ values of the metrics. For upper tropospheric humidity, we use the humidity profiles calculated from AIRS satellite data (Gettelman and Fu 2008), while for the zonal mean meridional wind profile, we use a 40 year mean of NCEP reanalysis output. We estimate natural variability around this observed mean by using a 500 year CAM 3.0 control simulation, removing the mean value and expressing the variability around the observed value. In both CAMcube and climateprediction.net only models with longwave clear-sky sensitivities between 2.5K and 2.75K fall within the 5-95% range of observed the ‘observed’ diagnostics. All models with a partial clear sky longwave sensitivity of greater than 4K show either large upper
troposphere humidity, a dual ITCZ in the control simulation or both.

Some of the CMIP-3 models, however, do display properties consistent with higher longwave clear-sky sensitivities in the \texttt{climateprediction.net} ensemble. The CSIRO Mk 3.0 model has a climate sensitivity of 3.1K, but has upper tropospheric humidities of 23.8 ppmv (which is comparable to the \texttt{climateprediction.net} maximum of 25.7 ppmv, a model which has a climate sensitivity of 7.5K). Hence, clearly this relationship is influenced by systematic differences between the models. Similarly, the NCAR PCM-1 model projects strongly onto the double ITCZ pattern, but has a relatively low climate sensitivity of 2.1K.

4.2. Shortwave Temperature Response

After the longwave clear-sky response has been accounted for, the majority of the remaining variance in climate sensitivity in both \texttt{climateprediction.net} and CAMcube ensembles, is explained by differences in shortwave cloud response. 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 0.14K. In \texttt{climateprediction.net}, the partial shortwave cloudy sensitivity has a larger standard deviation of 0.52K, but 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.

Studies have found previously that low cloud shortwave feedback is an important factor in determining the climate sensitivity of a parameter perturbed GCM ([Webb et al. 2006]). However, given the large range of control-state cloud distribution in the perturbed model versions of \texttt{climateprediction.net} and CAMcube, one might hope that the control state cloud distribution might provide some constraint on cloud forcing changes in the double CO$_2$ stage. [Sanderson et al. 2008a] and [Sanderson et al. 2009] proposed some relationships between modes of variability in the control state cloud cover and in modes of cloud response. However, interpreting these modes is troublesome and in this work we are seeking large scale observables, hopefully less dependent on any single model architecture.
Figure 12: (a) Shortwave local partial cloudy-sky equilibrium temperature change plotted as a function of control annual mean upgoing shortwave for latitudes northwards of 60°N, for both the climateprediction.net (blue) and CAMcube (red) ensembles. Vertical lines show values diagnosed from CMIP-3 models (green) and CERES-2 observations (pink bar) (b) as for (a) but for latitudes 60°S-60°N. (c) Tropical (23.4°S-23.4°N) local partial shortwave cloudy-sky equilibrium temperature response plotted as a function of the annual, zonal mean meridional wind projected onto the EOF pattern shown in Figure 10(a). Projection of the NCEP 40 year mean onto the EOF is depicted by the light blue bar. For NCEP and CERES-2, the width of the bar represents 5-95th percentiles of the distribution of 15 year mean values in a 500 year control simulation, with the mean subtracted and the result added to the observed value.
and representing some physical principles.

We show in Figure 12 the relationship between shortwave cloud partial temperature response and large scale observables in three zonal bands. Figure 12(a) 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. Firstly, 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 (climateprediction.net correlation is significantly -0.65, while the CAMcube ensemble shows no significant correlation - which is 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, with a spread of over 50 Wm⁻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 which 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 CMIP-3 models are equally widely distributed, implying that high latitude upgoing radiation is not in itself a well-tuned quantity. The CERES-2 observed value, however, is more consistent with a smaller upgoing flux and thus if the climateprediction.net derived relationship were robust, this would imply relatively limited high latitude negative feedback to surface warming.

Figure 12(b) shows the same relationship as 12(a), but for the tropics and the mid-latitudes. 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 100Wm⁻², but only 20Wm⁻² in CAMcube), and there is a
relationship between the control case upgoing radiation and the double CO\textsubscript{2} response (a significant correlation of 0.50 in climate\textunderscore prediction\textunderscore 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\textsubscript{2} simulation. In some ways, this is consistent with the high latitude effect - the change in cloud forcing upon surface warming is proportional to the original forcing. Another major difference from the high latitude case is that there is more consensus amongst the CMIP-3 models on the strength of the control state upgoing radiation (only a 12Wm\textsuperscript{-2} range). Using Figure 12(b) as a transfer function, all the CMIP-3 models and the CERES-2 observations are consistent with a shortwave induced temperature change of between -0.5K and 1.5K averaged over the tropics and mid-latitudes in the double CO\textsubscript{2} climate.

For tropical cloud feedback, it is also worth noting in Figure 12(c), that the double ITCZ pattern identified in Figure 10 also has some influence over the extent of SW cloud forcing changes in the tropics. Those models which develop an ITCZ during the course of the double-CO\textsubscript{2} simulation tend to experience a net surface heating in tropical regions, presumably as regions of subsistence become more extensive and shortwave forcing is reduced, allowing more radiation to reach the surface. Once again, these feedbacks are unphysical because the control state of the model does not have a functioning ITCZ.

We can use the relationships shown in Figure 12(a) and (b) to show a simple observable space to predict global mean shortwave cloud temperature response in the double CO\textsubscript{2} simulation. Figure 13 shows the global mean shortwave cloud partial sensitivity as a function of both mid-latitude and high-latitude cloud cover. This two parameter space explains over 74\% of the total variance in cloudy sky shortwave partial sensitivity in climate\textunderscore prediction\textunderscore net, and 42\% of the variance in CAMcube. But, noticeably the shortwave cloud partial sensitivities in CAMcube are consistent with climate\textunderscore prediction\textunderscore net, given their low and high latitude fluxes.

We find that generally, the models with a large temperature increase due to shortwave cloud feedback have initially a large initial shortwave cloud forcing in the mid-latitudes.
Figure 13: As for Figure 11, except showing global cloudy-sky shortwave sensitivity as a function of high latitude and mid-latitude control stage upgoing shortwave radiation. Observational mean is taken from CERES-2 for the corresponding region, and the variance about the mean is estimated from a 500-year CAM control simulation.

and a small forcing at high latitudes. In the double CO₂ experiment in the default 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 CAM model, suggesting a slightly positive net global shortwave cloud feedback of 0.26 with 95% confidence intervals of -3.6 and +4.13K.

As we saw in Figure 7(d), the shortwave clear-sky response accounts for a very small percentage of the variance in overall climate sensitivity (in either ensemble). In climateprediction.net, the standard deviation of the shortwave clear-sky partial sensitivity is 0.17K, accounting for 1% of the overall variance in climate sensitivity. In CAMcube, the standard deviation of the shortwave clear-sky response is 0.06K, accounting for 10% of the variance in climate sensitivity. We do observe some correlation in climateprediction.net between the control state high-latitude shortwave cloud forcing (shown in figure 12(d)). 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
Table 2: Observable quantities used in regression shown in Figure 14. The observables are zonal mean Top Of Atmosphere Shortwave Flux (TOA SW), the first EOF of Zonal Mean Meridional Wind (ZM Mer. Wnd.) derived within the climateprediction.net ensemble and globally averaged Specific Humidity at 150mb (S.Hum)

shortwave by clouds overhead.

5. Predictions of likely Sensitivity

In Section 4, we established several potential mechanisms which allow for large climate sensitivities in the climateprediction.net ensemble and not in the CAMcube ensemble. For each of these, we have found aspects of the control climatology which are to some extent correlated with the amplitude of the response with at least some plausible physical mechanism as to why they might be correlated. In this section, we intend to take these small number of large scale observable metrics and establish a basic predictor of climate sensitivity. The observable metrics used are shown in Table 2 and are further detailed in Section 4.

We cannot assume that these observable quantities are statistically independent within the ensemble (indeed, they are not), and so a conventional multi-linear regression approach would not be valid. In order to relate the observables to model climate sensitivity, we use the partial least-squares regression algorithm of de Jong (1993), which does not assume independence of the predictors. The algorithm is used with the data from the climateprediction.net ensemble - because the range of climate sensitivity in CAMcube is too small to obtain robust statistics or correlations between observable quantities and climate sensitivity. The CAMcube ensemble can however be used verify the regression to estimate the systematic error associated with any predictions made using real-world observations.
Figure 14: A plot in the fashion of Piani et al. (2005), where true model climate sensitivity is shown as a function of predicted model climate sensitivity. Points represent a single model in the climateprediction.net, CAMcube or CMIP-3 ensemble. The curve on the horizontal axis shows the ‘best-guess’ distribution of predictions for climate sensitivity obtained by applying the climateprediction.net derived regression coefficients to real-world mean observations assuming variability about the mean from a 500 year CAM control simulation. Curves on the vertical axis use the CMIP-3, CAMcube or climateprediction.net ensembles as a transfer function to estimate the probability distribution after predictive error has been considered. Box and whisker plots show the 90th and 95th percentiles of each distribution.
Figure 14 shows the distribution of predicted climate sensitivity, based upon partial least squares regression within the climateprediction.net ensemble using the four predictors in Table 2. Within climateprediction.net, the regression explains 70% of the variance in total climate sensitivity. When the same predictors were applied to the CAMcube ensemble, they explain 49% of the variance, but only 17% in CMIP-3 due to 4 significant outliers.

Once the regression coefficients have been established, they can be used to predict the sensitivity from observed values. We use observational estimates of real-world values to produce a ‘best-guess’ of real world climate sensitivity. We account for variability in these observed values by taking a 500 year pre-industrial control simulation of CAM 3.0, and separating into 66, 15 year periods, subtracting the mean and then adding the residuals to the ‘observed’ or reanalysis values detailed in Figure 2 to create 66 independent estimates of climate sensitivity. Thus, excluding any potential error in prediction, this gives 95th percentiles of 2.44 and 3.54K for climate sensitivity.

We estimate prediction error using three different ensembles, climateprediction.net itself, CAMcube and the CMIP-3 ensemble. In each case, the prediction error is taken as the distribution of residuals for estimated sensitivities in that ensemble. We measure 4 moments of that distribution - mean, variance, skewness and kurtosis, once these have been measured, we simulate a 1000 member random distribution with the same moments in order to produce the PDF. Roe and Baker (2007) show that the probability distribution is likely to be a Normal Inverse Gaussian distribution, but in this case we make as few assumptions about the error distribution as possible, rather assuming that each ensemble represents a sample from a skewed normal distribution with unknown kurtosis. The total error is then estimated by expressing the prediction-error PDF around each of the 66 observational predictions for climate sensitivity.

Using climateprediction.net to estimate the prediction error, we obtain an approximately Normal Inverse Gaussian PDF. The 90th percentiles lie at 1.6 and 4.5K, while the 95th percentiles are 1.3 and 5.2K. Using CAMcube to estimate the prediction error actually results in a smaller error estimate of 2.4 and 3.5K for the 90th percentiles and
2.2 and 3.6K for the 95th percentiles. Using CMIP-3 to estimate the error gives a less skewed distribution with 90th percentiles of 1.0 and 4.8K and 95th percentiles of 0.3 and 5.3K.

6. Discussion

We have studied perturbed physics ensembles created by perturbing the atmospheric models of two of the world’s major General Circulation Models - the Community Atmosphere Model version 3.5 and the Hadley Centre Model Atmosphere Model version 3.0. We have used a subset of 81 models from the pre-existing [climateprediction.net](http://climateprediction.net) ensemble ([Stainforth et al. 2005](http://climateprediction.net)) and produced an identical size ensemble using the CAM model, but we find dramatically different distributions of global climate sensitivity in these two ensembles: in the [climateprediction.net](http://climateprediction.net) ensemble the climate sensitivity ranges from 1.7K to 9.9K, whereas in the CAM ensemble the range is only 2.2 to 3.2K. 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, and in some cases identical, parameter perturbations.

The most natural suspicion for the different behavior of global sensitivity in the two ensembles would be that effective ‘null’ parameters have been perturbed in CAM. However, we suspect this is not the case for two reasons: firstly, the ensemble shows a diverse range of behavior when models are examined at a regional level - where we examine both precipitation and temperature diagnostics and find differences in both the pre-industrial simulations and in the response to greenhouse gas forcing comparable to those seen in the [climateprediction.net](http://climateprediction.net) ensemble. Secondly, of the four parameters perturbed - two of them, the critical relative humidity and the ice fall speed, are similarly defined in both models - so that almost identical parameter changes may be applied. A third parameter relating to convective entrainment is also perturbed in both models, and although the parameters are slightly differently defined, in both ensembles the convective perturbation causes significant alterations to tropical mean climatology which makes us confident that
our results are not based on the perturbation of null parameters.

Previous studies (Webb et al. (2006), Sanderson et al. (2008b)) of climateprediction.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 - then we obtain similar results in this study, that 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 which 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 which 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 results of this breakdown tell a different story to that which is inferred from considering the various top-of atmosphere feedback components. By far the most significant difference between the climateprediction.net and CAM ensembles is seen in the clear-sky longwave component of the surface temperature change. This is not the inverse of the top of atmosphere clear-sky longwave feedback, rather it shows the component of surface temperature change which is due to decreasing clear-sky transmittance. The partial clear-sky longwave sensitivity as defined here is a function of the initial transmittance of the atmosphere (which is the net transmittance, including the effect of existing clouds), but is proportional to perturbations in the clear-sky transmittance only.

The longwave clear-sky component includes both the final forcing due to the doubled CO₂ concentrations, together with any water vapor feedback which might further decrease the atmospheric transmittance. This is important because the initial CO₂ forcing in the models in either ensemble is not a constant (Figure 15(c)), and the final forcing at the end of the double CO₂ simulation is not knowable without conducting a Partial Radiative
Perturbation type experiment ([Wetherald and Manabe 1988](#)), substituting pre-industrial carbon dioxide concentrations into an equilibrated double CO$_2$ simulation. Only with this information could the longwave clear-sky term could be further separated into a forcing and a feedback.

Given that the major difference between the partial temperature responses between the two ensembles appears to lie in the longwave clear-sky component - where the variance is an order of magnitude greater in climateprediction.net than for the CAM ensemble, we search for aspects of the humidity profiles in the model which are strongly correlated with the temperature response. This process suggests that upper tropospheric (UT) water vapor may be playing two critical roles in the processes leading to high sensitivity, firstly elevated upper-tropospheric humidities in the pre-industrial simulation decrease the atmospheric transmittance and amplify any forcing applied to the system. In addition, those models with increased UT water vapor in the pre-industrial simulation tend also to show large increases in UT water vapor during the double 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 response of the same magnitude.

This hypothesis could be tested in further work by artificially increasing UT water vapor in simulations using a different GCM to attempt to produce models of high climate sensitivity. But if we accept for a moment that the mechanism is a plausible explanation for the very large sensitivities in climateprediction.net (and their absence in the CAM ensemble) - then several questions remain for future study. Firstly, what aspect of the HadAM3 model (or the perturbations made within it) make this high-level water vapor feedback possible, and why are these processes not apparent within the CAM ensemble? Secondly, is there any statement of likelihood for ‘real world’ climate sensitivity which might be made from an understanding of the mechanisms present in the ensemble?

There are other important processes in determining the clear-sky response of the model. Some models in the climateprediction.net ensemble can be shown to lack a clearly defined single ITCZ in the control simulation, a property which can be seen easily in the meridional wind profiles and which results in an overly dry tropical mid-troposphere.
These models require a large input of heat to the slab ocean in order to remain stable during the control simulation. However, in the double CO₂ simulation, some (but not all) of these models develop an ITCZ and mid-tropospheric humidity increases dramatically, causing a large and unphysical water vapor feedback. We found that a good proxy for this behavior in the control simulation was to use the first ensemble-derived EOF of annual mean meridional wind, zonally averaged in the control simulation of the climateprediction.net models.

Taking the two observed metrics together, we have shown that to achieve partial clear-sky longwave sensitivities of greater than 4K, the model must display either high upper tropospheric humidity or be lacking in an ITCZ in the control simulation. None of the CAMcube simulations display these characteristics, and we do not see a large range in clear-sky sensitivity. This evidence is not conclusive, but suggests that the dominant cause of the differing sensitivity distributions in the two ensembles lies in the longwave clear-sky response.

Of course, 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 mid-latitudes 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 (polewards of 60°) exhibit a more negative high-latitude cloud feedback than the average model. Similarly, models with a large control mid-latitude 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 mid-latitude cloud forcing provides a reasonable metric for the shortwave cloud response in the double CO₂ simulation. We do not find any significant correlation between longwave cloud forcing in the control and in the double CO₂ simulation, but as in the clear-sky longwave case, we find that the longwave cloudy-sky
sensitivity is strongly affected by upper tropospheric humidities. Finally, the variation in shortwave clear-sky sensitivity is sufficiently negligible in both ensembles that finding observable correlated quantities is impossible.

Combining the above observables, we can perform a multi-linear regression based approach to predict real-world values of climate sensitivity. The resulting probability distribution suggests best-guess 95th percentiles for real-world climate sensitivity of 2.44 and 3.54K, before we account for any prediction error. The prediction error can be estimated by considering the residual distribution of sensitivity predictions within each of the three ensembles. Using climateprediction.net itself to estimate the prediction error gives a skewed probability distribution such that it remains difficult to eliminate the small possibility of sensitivities greater than 5K.

The predictors do show some skill in predicting the model sensitivities in CAMcube and using CAMcube to estimate prediction error results in a smaller width of distribution, but the small range of sensitivity present in the ensemble makes the ensemble a poor sample to estimate the prediction error. With CMIP-3, the predictors have some skill for 15 of the 19 models (with prediction errors of less than 1K), but there are 4 significant outliers where sensitivity is under or over predicted by a large amount. As a result, these outliers are strongly influential in determining the width of the PDF when CMIP-3 is used to estimate the prediction error, resulting in an estimate for climate sensitivity with 95th percentiles at 0.3 and 5.3K.

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 which can be used to physically exclude some, but not all, high sensitivity models in 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.
Table 3: Definition of perturbed parameters as used in the 81 member subset of climateprediction.net experiments used in this analysis. Parameters marked † are actually defined on 19 model levels, with the mean over model levels shown here. Values column indicates the possible parameter 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>ENTCOEF</td>
<td>Entrainment coefficient</td>
<td>[0.6, 3.0, 9.0]</td>
</tr>
<tr>
<td>RHCRIT</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 speed (ms(^{-1}))</td>
<td>[0.5, 1, 2]</td>
</tr>
</tbody>
</table>

A. Choosing perturbed parameters

In order to create a manageable ensemble size from the climateprediction.net grand ensemble, we select 4 parameters which previous studies have shown to be most important in controlling model sensitivity (Murphy et al. (2004), Sanderson et al. (2008a), Sanderson et al. (2008b), 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 3.

Because the original ensemble sampled over a much larger range of parameters, it is not always possible to find a model with the correct permutation of the four parameters used while keeping all other parameters at their default value. In these cases, we consider all models with the required combination of the four parameters together with other, random perturbations - from this distribution we choose the model with the median value of climate sensitivity. This approximation makes the assumption that other perturbations in the ensemble are equally likely to increase or decrease the climate sensitivity.

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 different parametrization schemes were used in HadAM3 and CAM. Critical relative humidity is defined differently in both models - 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 which only varied significantly in the upper troposphere, so in the CAM ensemble we choose to perturb
<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMPDZ</td>
<td>Parcel fractional mass entrainment rate (m^{-1})</td>
<td>([0, -0.5, -1] \times 10^{-3})</td>
</tr>
<tr>
<td>RHMINH</td>
<td>Critical relative humidity for high cloud</td>
<td>([0.7, 0.8, 0.9])</td>
</tr>
<tr>
<td>VICE _SMALL</td>
<td>Ice fall speed (\text{ms}^{-1})</td>
<td>([0.5, 1, 2])</td>
</tr>
<tr>
<td>ICRITC</td>
<td>Threshold for auto-conversion of cold ice</td>
<td>([2, 5, 16] \times 10^{-6})</td>
</tr>
</tbody>
</table>

Table 4: As for Table 3, but for the CAM experiment.

only the “high” critical relative humidity. Ice fall speed parametrization is similar in both models - which both use Heymsfield (1977).

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 CAM 3.5, 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 auto-conversion of cloud water to liquid parameter ‘ct’ in the CAM model, so instead we perturb the threshold for the auto-conversion of cold ice ‘icritc’ - which is a resolution dependent tuning parameter in the model. Perturbed parameters and values are listed in 4.

B. Top of atmosphere feedbacks

Top of atmosphere feedbacks shown in Figure 6 are calculated using information from the double CO₂ simulation. Clearsky 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 which indicates the change in global mean flux for 1K global mean surface temperature rise. For the cloud feedbacks, the Cloud Radiative Forcing (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 1K surface temperature rise.
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 \( S_o \), we can write an equation for the equilibrium temperature \( T_g \):

\[
\sigma T_g^4 = \frac{S_o(1-\alpha)}{\gamma}.
\]

(1)

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

\[
\gamma = \gamma_{cs} + \gamma_{cld}
\]

(2)

\[
(1-\alpha) = (1-\alpha_{cs})(1-\alpha_{cld})
\]

(3)

where \( \gamma_{cs} \) and \( \gamma_{cld} \) are top of atmosphere clear-sky radiation and longwave cloud radiative forcing 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 \( \alpha_{cs} \) is the top of atmosphere upgoing all-sky and clear-sky flux 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:

\[
\Delta T_g \approx T_g \left( \frac{\Delta \gamma_{cs}}{\gamma_{cs} + \gamma_{cld}} + \frac{\Delta \gamma_{cld}}{\gamma_{cs} + \gamma_{cld}} - \frac{\Delta \alpha_{cs}}{(1-\alpha_{cs})} - \frac{\Delta \alpha_{cld}}{(1-\alpha_{cld})} \right),
\]

(4)

thus allowing the total surface temperature response to be approximately separated into four components relating to changes in longwave atmospheric opacity in the clear-sky and cloudy sky, and in shortwave reflectivity in the clear-sky and cloudy sky. We can verify that our approximations are reasonable by comparing the \( \Delta T_g \) to the true climate sensitivity of the model (Figure 15(a)). When compared to a simple inversion of the
Figure 15: (a) A plot of the approximated global mean surface temperature response as calculated by Equation 4 compared to the actual Climate Sensitivity of the model (shown on the vertical axis). (b) A plot of global mean temperature response taken by assuming a forcing due to CO$_2$ doubling of 3.7 W m$^{-2}$ (Ramaswamy et al. 2001), and dividing by the net top of atmosphere feedback to increasing warming in W m$^{-2}$ K$^{-1}$. (c) Histograms of the true initial CO$_2$ forcing in the climateprediction.net (red) and CAMcube (blue) ensembles. The forcing is inferred by extrapolating net top-of-atmosphere energy imbalance back to the start of the double-CO$_2$ simulation using the methodology of Gregory et al. (2002).

net feedback parameter (Figure 15(b)), the sum of partial temperature responses shows considerably less spread.

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

$$CRF_{LW} = \gamma_{CLD}\sigma T_g^4$$

$$CS_{LW} = \gamma_{CS}\sigma T_g^4,$$

such that the derivatives with temperature show the traditionally defined clear sky and cloudy sky feedbacks:

$$\frac{dCRF_{LW}}{dT} = \Delta\gamma_{CLD}\sigma T_g^4 + 4\gamma_{CLD}T_g^3\Delta T$$

$$\frac{dCS_{LW}}{dT} = \Delta\gamma_{CS}\sigma T_g^4 + 4\gamma_{CS}T_g^3\Delta T.$$
equations 7, the right-hand terms are dependent only on the initial cloud and clear-sky transmittance, and the surface temperature rise. Thus, the clear sky and cloudy sky feedbacks as defined in Equation 7 can be non-zero even if there is no change in $\gamma_{cld}$ or $\gamma_{cs}$.

In contrast, in Equation 4 - the cloudy 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, Equations 7 can be used to derive the longwave temperature response - identical to that of equation 4 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


