Clim Dyn (2010) 34:735–745
DOI 10.1007/s00382-009-0556-1

Sensitivity of climate change projections to uncertainties in the estimates of observed changes in deep-ocean heat content

A. P. Sokolov · C. E. Forest · P. H. Stone

Received: 24 September 2008 / Accepted: 17 March 2009 / Published online: 31 March 2009

Abstract The MIT 2D climate model is used to make probabilistic projections for changes in global mean surface temperature and for thermosteric sea level rise under a variety of forcing scenarios. The uncertainties in climate sensitivity and rate of heat uptake by the deep ocean are quantified by using the probability distributions derived from observed twentieth century temperature changes. The impact on climate change projections of using the smallest and largest estimates of twentieth century deep ocean warming is explored. The impact is large in the case of global mean thermosteric sea level rise. In the MIT reference ("business as usual") scenario the median rise by 2100 is 27 and 43 cm in the respective cases. The impact on increases in global mean surface air temperature is more modest, 4.9 and 3.9 C in the two respective cases, because of the correlation between climate sensitivity and ocean heat uptake required by twentieth century surface and upper air temperature changes. The results are also compared with the projections made by the IPCC AR4's multi-model ensemble for several of the SRES scenarios. The multi-model projections are more consistent with the MIT projections based on the largest estimate of ocean warming. However, the range for the rate of heat uptake by the ocean suggested by the lowest estimate of ocean warming is more consistent with the range suggested by the twentieth century changes in surface and upper air temperatures, combined with the expert prior for climate sensitivity.

Keywords Climate response · Probabilistic forecast · Uncertainty · Climate change

A. P. Sokolov (✉) · C. E. Forest · P. H. Stone
Joint Program on the Science and Policy of Global Change, Center for Global Change Science, Massachusetts Institute of Technology, 77 Mass. Ave., Cambridge, MA 02139, USA
e-mail: Sokolov@mit.edu

Present Address:
C. E. Forest
Department of Meteorology, Pennsylvania State University, 77 Mass. Ave., University Park, PA 16802, USA

1 Introduction

There are significant uncertainties in the characteristics of the climate system which define its response to external forcing, such as climate sensitivity, strength of aerosol forcing and the rate of deep ocean heat uptake (e.g., Andronova and Schlesinger 2001; Frame et al. 2005; Forest et al. 2006; Knutti et al. 2006). Consequently anthropogenic climate change can be described only in probabilistic terms, even when changes in the concentrations of greenhouse gases (GHGs) and aerosols are prescribed (e.g., Knutti et al. 2003). To account for such uncertainties, projections of future changes in global mean surface air temperature (SAT) presented in the IPCC AR4 (Meehl et al. 2007b) are based in part on a multi-model ensemble of simulations with coupled atmosphere–ocean general circulation models (AOGCMs).

There are, however, well known problems with the use of so-called "ensembles of opportunity" (Tebaldi and Knutti 2007). Among them are difficulties with defining the relative weights of different models and the fact that existing AOGCMs do not cover the full range of uncertainty in climate sensitivity and the rate of oceanic heat uptake. By contrast, earth system models of intermediate complexity can be used effectively for producing probabilistic predictions of future climate changes, due to their computational efficiency and ability to vary the above mentioned characteristics over wide ranges.

Here we use the climate component of the MIT Integrated Global System Model (Sokolov et al. 2005) to evaluate the climate response and its associated uncertainty when we prescribe the forcing from four different scenarios. The probability distributions for the uncertain input parameters were obtained by comparing twentieth Century temperature changes as simulated by the MIT model with available observations (Forest et al. 2006). In particular we use the probability distributions obtained this way when they are combined with an expert prior on climate sensitivity (Forest et al. 2006).

Distributions presented by Forest et al. (2006) are based on data for the trend in the ocean temperature averaged over the 0–3 km layer from Levitus et al. (2005). There are, however, significant differences between estimates of changes in deep ocean temperature obtained in different studies (Gouretski and Koltermann 2007; Domingues et al. 2008). Moreover estimates given by Levitus et al. (2005) and Domingues et al. (2008) exclude each other; each of them does not fall into the range of uncertainty for the other.

To evaluate the impact of these differences on climate projections, we constructed two additional sets of input parameter distributions and carried out additional ensembles of twenty-first century climate simulations for two forcing scenarios. We also compare these projections with projections produced by the ensemble of AR4 AOGCMs.

2 Model description

The climate model used in this study, as well as in Forest et al. (2006), is a modified version of the model described by Sokolov and Stone (1998). It consists of a 2-dimensional (zonally averaged) statistical–dynamical atmospheric model coupled to an ocean mixed layer model with temperature anomalies diffused below the mixed layer. The atmospheric model is derived from the Goddard Institute for Space Studies (GISS) Model II general circulation model (GCM) (Hansen et al. 1983) and uses parameterizations of the eddy transports of momentum, heat and moisture by baroclinic eddies (Stone and Yao 1987, 1990). The model uses the GISS radiative transfer code which contains all radiatively important trace gases as well as aerosols. The surface area of each latitude band is divided into fractions of land, ocean, land-ice and sea-ice, with the surface fluxes and surface temperature computed separately for each surface type. The version used here has 4° latitudinal resolution and 11 layers in the vertical. The Q-flux ocean mixed layer model and the thermodynamic sea-ice model have 4° × 5° latitude–longitude resolution and are described by Hansen et al. (1984).

The climate sensitivity of the atmospheric model (S) can be changed by varying the strength of the cloud feedback (Sokolov and Stone 1998), while the rate of deep oceanic heat uptake can be changed by varying the value of the global mean diffusion coefficient (Kv) used to mix ocean temperature anomalies below the mixed layer. This model, while exaggerating thermosteric sea level rise on longer time scales, has been shown to reproduce accurately changes in global mean climate variables simulated by the version of the MIT model which uses 3D ocean GCM through the end of twenty-first century (Sokolov et al. 2007).

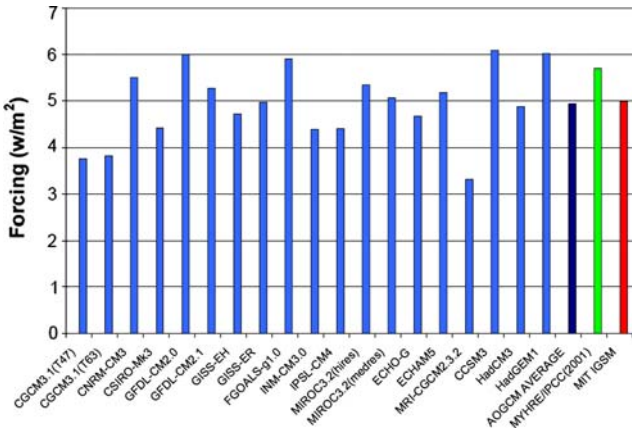

Fig. 1 Climate forcing for SRES A1B in the last decade of the twenty-first century relative to present (1980–1999) for AR4 AOGCMs, MIT IGSM and from Myhre et al. (1998)

While parameterizations of physical processes used in present atmospheric GCMs are much more sophisticated than the parameterizations which can be used in zonally-averaged models, global mean responses of different AOGCMs to an external forcing are determined by the integral properties of these models, namely effective climate sensitivity[1] and the rate of oceanic heat uptake (Raper et al. 2002; Sokolov et al. 2003). As a result, despite its simplicity the MIT 2D model, with the appropriate choice of the model's parameters (S, Kv), reproduces the behavior of different CMIP3 (Meehl et al. 2007a) AOGCMs in simulations with increasing atmospheric CO_2 concentrations. Reproducing the AOGCM results for other forcing scenarios, such as the SRES scenarios, is more complicated. In some cases different sets of forcing agents were taken into account in the simulations of climate for twentieth and 21st centuries with different AOGCMs. In addition, the strengths of aerosol forcing and the efficacies of different forcings vary among models.

As a result differences in the "climate forcing"[2] among AOGCMs (Fig. 1), are significantly larger than differences

[1] Effective climate sensitivity is defined as an equilibrium climate sensitivity corresponding to the strength of feedbacks at the time of CO_2 doubling in simulations with 1% per year increase in atmospheric CO_2 concentration.

[2] "Climate forcing" for a given scenario is calculated from the changes in SAT and heat flux at the top of the atmosphere using the feedback parameter estimated from the simulation with 1% per year CO_2 increase (Forster and Taylor 2006).

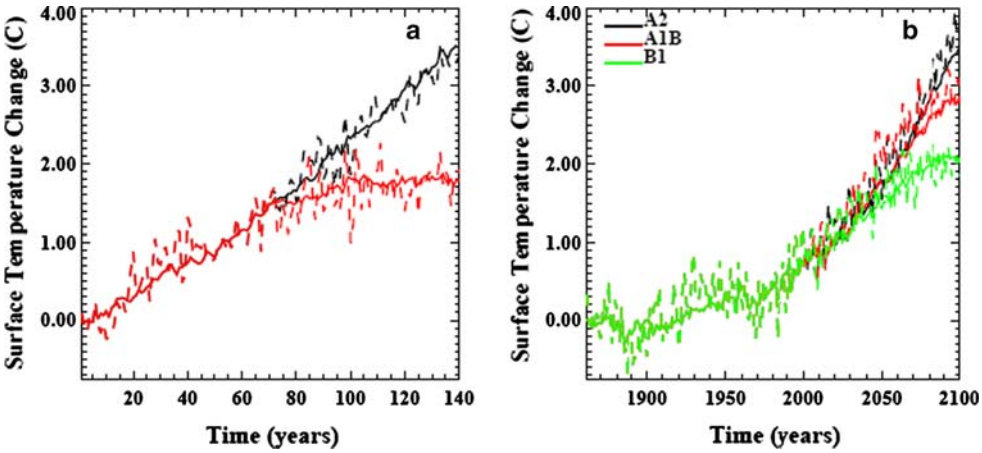

Fig. 2 Change in global mean annual mean surface air temperature simulated by the GFDL 2.1 model (*dashed lines*) and the corresponding version of the MIT model (*solid lines*) in the simulations **a** with 1% per year CO_2 increase for 70 years and fixed after that (*red line*) and for 140 years (*black line*) and **b** with SRES A2, A1B and B2 scenarios

in CO_2 forcing (Forster and Taylor (2006)). However, for some AOGCMs (e.g., GFDL CM2.1) the "climate forcing" for the A1B SRES scenario is similar to the "climate forcing" simulated by the MIT 2D model. As seen in Fig. 2, the version of the MIT model, which reproduces the behavior of the GFDL CM2.1 model for changes in CO_2 only, also reproduces its behavior for the SRES scenarios.

3 Probability distributions of input climate parameters

Performing probabilistic climate projections requires knowledge of probability distributions for climate system parameters determining model response to an external radiative forcing. Forest et al. (2002) proposed a method for obtaining such distributions based on comparison of the results of twentieth century climate simulations with observational records of surface, upper air and deep ocean temperature changes (see Appendix). Distributions presented by Forest et al. (2002, 2006, 2008) were obtained using estimates of changes in deep ocean heat content provided by Levitus et al. (2005). However, several other estimates have been published in the last few years (e.g., Gouretski and Koltermann 2007; Ishii et al. 2006; Carton and Santorelli 2008; Domingues et al. 2008). The methods used in different studies vary in their treatment of subsurface temperature measurements from different types of instruments (MBT, XBT and so on) as well as in the methods used to estimate changes in data-sparse regions.

While analyses of surface air temperature changes carried out by different groups give very similar results, published estimates of changes in deep ocean heat content are dramatically different. Only three papers, Levitus et al. (2005), Gouretski and Koltermann (2007) and Domingues et al. (2008) (LAB05, GK07 and DOM08 hereafter) provide observationally based estimates of changes in the ocean heat content for the 0–3,000 m layer (Fig. 3). All the others give results for the upper 700 m only. When we

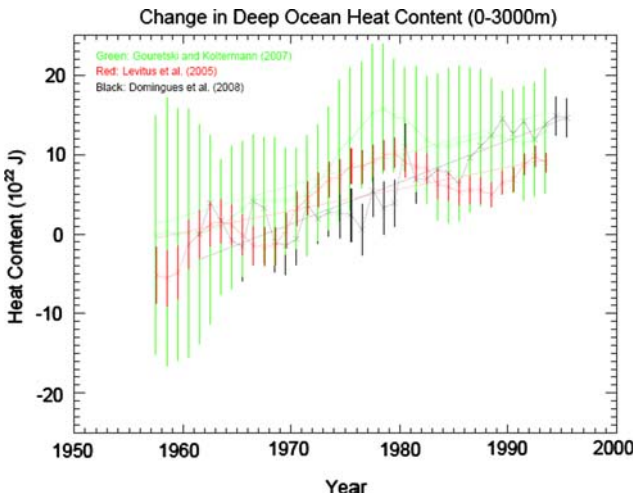

Fig. 3 Changes in ocean heat content from LAB05 (*red*), GK07 (*green*) and DOM08 (*black*) for the 0–3,000 m layer as estimated by each group. Five year running means were applied to the GK07 data and the errors combined appropriately. The *error bars* are nominally 1-sigma standard errors but represent different sources of uncertainty in each case

re-did our analyses using trends based on just the upper 700 m we found that they did not introduce any constraints on the climate parameters beyond those resulting from using just the upper air and surface temperature data. This lack of impact is due to the strong correlation between upper-ocean and sea surface temperatures and the large natural variability of the upper ocean layers.

Figure 4 shows the marginal posterior probability density functions for the S-Sqrt(Kv) parameter space obtained using the LAB05 and DOM08 estimates for changes in deep ocean temperature and their respective estimates of the uncertainty in the trends[3]. These estimates are

[3] The LAB05 analysis has recently been updated to take into account the systematic errors in the XBT data noted by GK07. The result, which is available on the NOAA website, is not appreciably different from the LAB05 result.

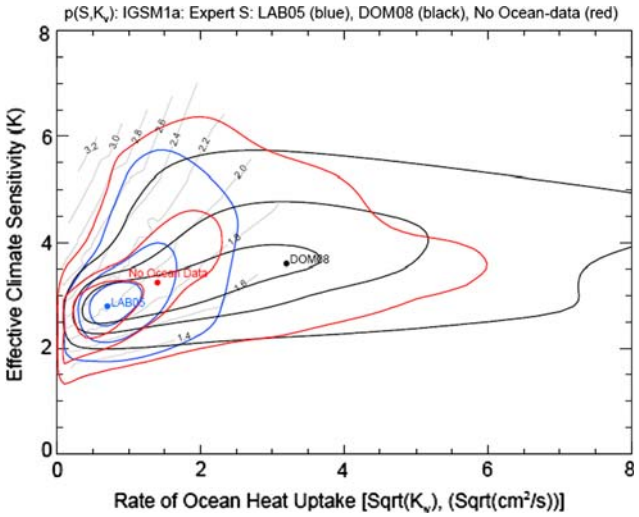

Fig. 4 The marginal posterior probability density function for S-Sqrt(Kv) parameter space obtained using data for surface, upper-air and deep-ocean temperatures by LAB05 (*blue*) andDOM08 (*black*) and using surface and upper-air temperatures only (No Ocean data) (*red*). Thick contours denote rejection regions for significance levels of 90, 50 and 10%, respectively. Colored dots indicate the median values from respective 1D marginal distribution. Thin contours show values of TCR

respectively the smallest and largest estimates of the 0–3,000 m trend. We note that each of these trends lie outside the uncertainty range cited for the other trend, and thus the published uncertainty estimates are almost certainly underestimates. As mentioned above, the probability distributions for climate parameters obtained using upper-ocean (0–700 m) data are almost identical to the distribution constructed without the use of any ocean data (i.e., No Ocean data, hereafter, NO). To evaluate what climate projections we would obtain in the absence of deep-ocean data, we also carried out simulations with the NO distribution. It is worth pointing out, that if we assume larger uncertainty in deep-ocean temperature data (e.g., larger enough to include an alternative data), then corresponding parameter distributions will also be similar to the NO distribution due to small weight of deep-ocean data.

The use of the LAB05 and DOM08 analyses strongly affects the range of acceptable values of the rate of oceanic heat uptake (Table 1). However, due to the correlation between climate sensitivity and the rate of oceanic heat uptake imposed by the data on surface air temperature, changes in SAT in response to a 1% per year increase in CO_2 concentration in simulations with the median values of S and Kv from the different parameter distributions are not very different (see Fig. 4).

Since all distributions are constrained by the same data for surface air temperature, use of different data for the oceanic heat content has changed the distribution for climate sensitivity in such a way that all model ensembles

Table 1 Percentiles of marginal distributions of Sqrt(Kv) obtained under different assumptions on deep ocean temperature changes

	2.5%	5.0%	25.0%	50.0%	75.0%	95.0%	97.5%
LAB05	0.12	0.20	0.46	0.73	1.12	1.87	2.16
NO	0.14	0.24	0.75	1.47	2.45	5.39	6.44
DOM08	0.44	0.65	1.87	3.16	4.74	7.00	7.46

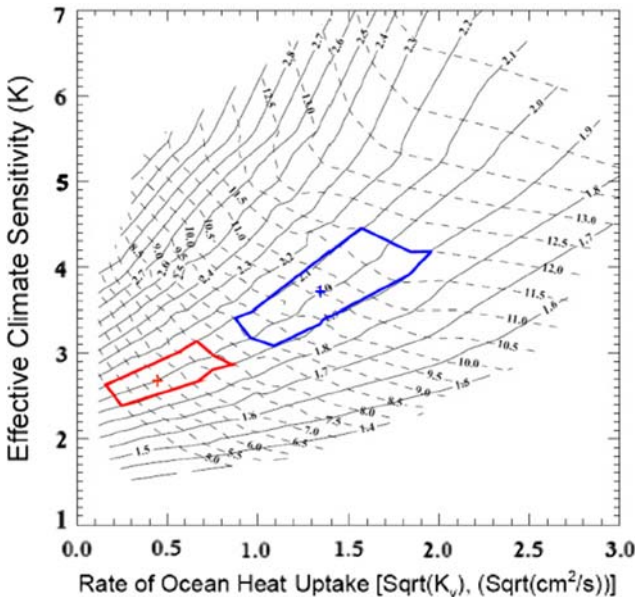

Fig. 5 Changes in surface air temperature and thermosteric sea level rise at the time of CO_2 doubling in the simulations with 1% per year increase in CO_2 concentration as a function of effective climate sensitivity and the rate of heat uptake by the ocean (see text for details)

simulate similar surface warming over the twentieth century. Different sets of models, producing similar changes in surface air temperature under one forcing scenario, are likely to simulate similar warming for other scenarios (at least for monotonically increasing forcing on comparable time scales). Figure 5 shows changes in surface air temperature and thermosteric sea level rise at the time of CO_2 doubling in simulations with 1% per year increase in CO_2 concentration as a function of effective climate sensitivity and the rate of heat uptake by the deep ocean. Two regions are indicated that contain versions of the model simulating similar surface warming (1.9–2.1C) but different sea level rise (5–8 cm and 9–12 cm for red and blue regions, respectively). Table 2 shows averaged changes in SAT and sea level simulated by the models falling in the two regions under different forcing scenarios. As can be seen, in spite of rather significant differences in the rates of the oceanic heat uptake between models in different regions, they produce very similar surface warming under all scenarios considered. These results suggest that future climate

Table 2 Change in the SAT and thermosteric sea level rise simulated by the version of the MIT climate model from two regions shown in Fig. 5

	BLUE REGION		RED REGION	
	dSAT (C)	SLR (cm)	dSAT (C)	SLR (cm)
1% CO_2 per year at the time of CO_2 doubling	2.00	7	2.00	10
B1	2.06	11	2.14	18
A1B	3.12	16	3.17	24
A2	4.06	19	4.02	27
REF	4.85	24	4.78	36
1% CO_2 per year at the time of CO_2 quadrupling	5.18	24	5.19	35

projections will be rather insensitive to differences in ocean data used to determine distributions of input parameters. Knutti and Tomassini (2008) also came to the conclusion that constraints imposed by data on surface warming on models' parameters will reduce sensitivity of projected warming on ocean heat uptake.

However, climate changes observed over twentieth century are rather small relative to natural variability and observational errors. As a result constraining climate parameters from twentieth century climate simulations requires using a prior for climate sensitivity. In addition, none of the data sets on changes in oceanic heat content rules out very small rates of heat uptake, which in all distributions are bounded by zero. These two factors affect shapes of input distributions and, as will be shown later, make projected surface warming more sensitive to the choice of input parameter distribution than would be expected from the results shown above. The relatively large sensitivity of projected SAT increase to input climate parameters also may be associated with differences in temporal patterns of twentieth and tweny-first century forcings, primarily due to taking into account volcanic forcing in twentieth century climate simulations.

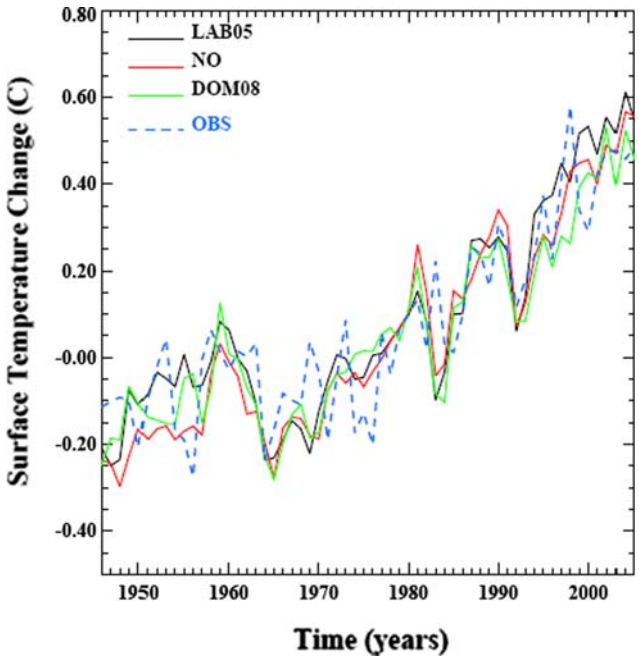

Fig. 6 Changes in global-mean and annual-mean surface air temperature in simulations with median values of climate parameter from different distributions. Observations are from Jones and Moberg (2003)

4 Simulations

First we carried out three 250-member ensembles of simulations from year 1860 to year 2000, using different values of the climate sensitivity, the rate of the oceanic heat uptake and the strength of the aerosol forcing for each of the distributions based on the LAB05, NO, and DOM08 results. Different combinations of model parameters controlling these characteristics were chosen using the Latin Hypercube Sampling algorithm (Iman and Helton 1988) from the probability density functions described in the previous section. The distribution of each model parameter was divided into 250 segments of equal probability and sampling without replacement was performed, so that every segment was used once. Details on the sampling procedure can be found in Webster et al. (2003).

In each simulation, the MIT 2D climate model was forced by the observed changes in GHGs, stratospheric aerosols from volcanic eruptions, tropospheric and stratospheric ozone, solar irradiance and changes in vegetation due to land use change. Details of the forcings are given in Forest et al. (2006). As can be seen from Fig. 6, changes in SAT in simulations with median values of climate parameters from all three distributions agree well with observations. Then we carried out a number of 250-member ensembles of simulations for several forcing scenarios using different distributions of the climate parameters (Table 3). Each simulation was started from the end of the corresponding twentieth Century simulation and simulated years 2001–2100. Data for GHG concentrations for the SRES scenarios B1, A1B and A2 scenarios were taken from the web site of the Goddard Institute for Space Studies

Table 3 List of simulations with different input parameters and different forcing scenarios

Input parameters distributions	Forcing scenarios			
	SRES B1	SRES A1B	SRES A2	MIT REF
LAB05	X	X	X	X
NO	–	X	–	X
DOM08	–	X	–	X

(http://data.giss.nasa.gov/modelforce/ghgases/GCM_2004.html). In addition to changes in CO_2 [ISAM reference (Houghton et al. 2001)], N_2O, CH_4, CFC-11 and CFC-12, these data include changes in additional trace gases (Hansen and Sato 2004). Changes in the loading of sulfate aerosols were prescribed according to the IPCC TAR's description (Houghton et al. 2001). Changes in black and organic carbon aerosols were not included.

In the simulations with the MIT REF scenario we used GHG and sulfate aerosol concentrations obtained in a reference simulation with the full version of the MIT IGSM2.2 with "business as usual" emissions. Detailed information on this scenario can be found in Prinn et al. (2008). It is worth noting that the total radiative forcing and SAT changes in simulations with the MIT REF scenario are close to those in simulations with the SRES A1FI scenario. The chosen set of simulations allows us to explore the dependency of the projected climate changes on both the forcing scenarios and the choice of ocean data.

5 Results and discussion

The use of different pdfs for input climate parameters has a smaller effect on the distributions of projected surface warming than on the distribution of projected sea level rise (Fig. 7, Tables 4, 5) because a correlation between the input parameters is imposed by the twentieth century surface air temperature data. For example, the mean values of sea level rise in the ensembles with NO and DOM08 input distributions exceed the mean value for LAB05 distribution by 40 and 51%, respectively, while changes in the mean values of the increase in SAT are only 6 and 13%. The choice of input distributions also significantly changes the shape of the probability distribution for projected thermosteric sea level rise.

It has been noted in several papers (e.g., Knutti et al. 2008; Meinshausen et al. 2008) that, for the given set of AOGCM models, the shapes of the distributions for SAT increase are similar under different SRES scenarios. Our simulations yielded similar results. Ratios of the different percentiles of the given distribution to the corresponding mean differ only slightly between ensembles with different forcing scenarios (see Table 6). At the same time, the ratios of the SAT changes in the simulations with different forcing scenarios for *a given version* of the MIT IGSM are not defined just by the ratios of forcings, but also depend on the values of climate sensitivity and the rate of heat uptake by the ocean.

For each of the ensembles listed in Table 3, the SAT change in a simulation using the median values of the input

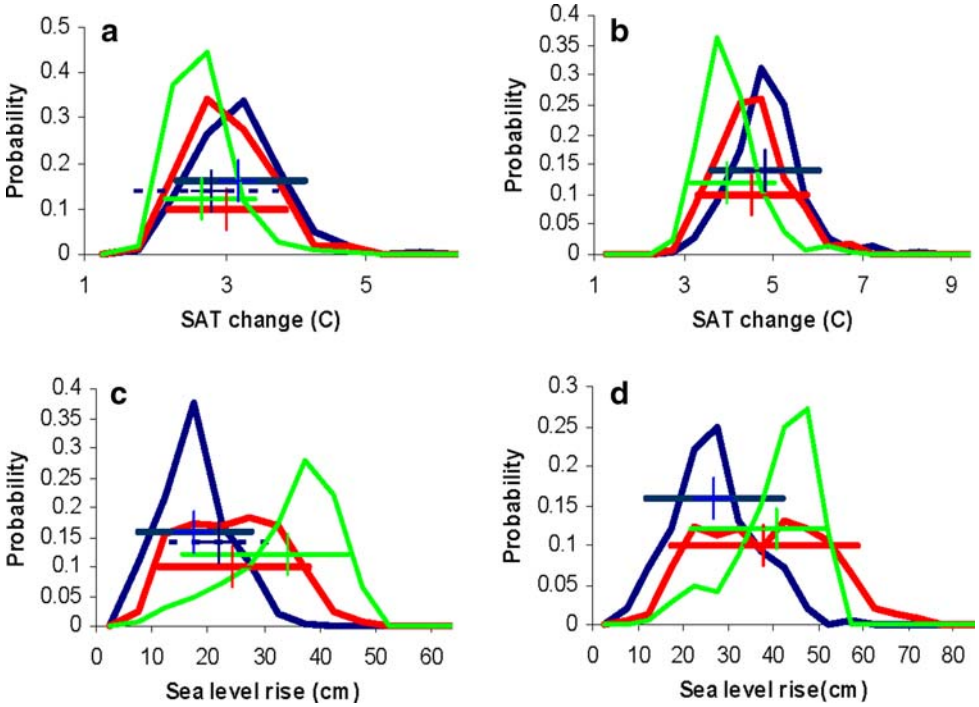

Fig. 7 Frequency distributions for SAT increase under SRES A1B (**a**) and MIT REF (**b**) scenarios and corresponding thermosteric sea level rise (**c**, **d**) in 2091–2100 relative to the 1981–2000 average in simulations with LAB05 (*blue*) DOM08 (*green*) and NO (*red*) input parameter distributions. *Solid horizontal bars* show 5–95% ranges from 250-member ensembles of simulations with the MIT model, *dashed horizontal bars* show 5–95% ranges from the multi-model IPCC AR4 AOGCM ensemble [from figure TS.27 of Solomon et al. (2007)]

Table 4 Distributions of temperature changes in the last decade of twenty-first century relative to the 1981–2000 mean

	5.0%	16.7%	50.0%	Mean	83.3%	95.0%
LAB05						
B1	1.51	1.69	2.10	2.08	2.37	2.71
A1B	2.30	2.61	3.17	3.17	3.64	4.11
A2	3.00	3.43	4.06	4.03	4.55	4.98
REF	3.62	4.11	4.86	4.83	5.41	6.02
NO						
A1B	2.20	2.45	2.94	3.01	3.55	3.86
REF	3.33	3.76	4.50	4.50	5.23	5.78
DOM08						
A1B	2.13	2.31	2.58	2.66	3.00	3.44
REF	3.16	3.43	3.91	3.99	4.50	5.06

Table 5 Distributions of thermosteric sea level rise in the last decade of twenty-first century relative year to the 1981–2000 mean

	5.0%	16.7%	50.0%	Mean	83.3%	95.0%
LAB05						
B1	5.23	8.85	13.02	12.40	17.85	21.26
A1B	7.67	12.16	17.45	16.77	23.35	27.60
A2	10.64	15.80	22.10	21.35	28.65	33.56
REF	11.80	18.91	26.95	26.11	35.73	42.11
NO						
A1B	11.27	14.65	24.40	24.54	33.32	38.07
REF	17.65	23.10	37.83	37.80	51.45	58.66
DOM08						
A1B	15.26	25.73	36.00	34.34	42.76	45.26
REF	21.50	32.55	42.71	40.82	49.15	51.47

climate parameters is close to the median value of SAT changes from the corresponding ensemble of simulations. This fact, in combination with the similarity of the statistical properties of the distributions of projected surface warming for different forcing scenarios, allows us to approximate the probability ranges for surface warming under the SRES B1 and A2 scenarios for the NO and DOM08 sets of input parameters without running the corresponding ensembles. Namely, we run simulations with the median values of climate parameters for the remaining scenarios and then used ratios from Table 6 to estimate probability ranges. Those ranges are shown in Figs. 7 and 8 by dashed lines.

The values of Sqrt(Kv) required for the MIT IGSM to reproduce the results of different AR4 AOGCMs in simulations with 1% per year CO_2 increase, range from 0.9 to 2.0 cm $s^{-1/2}$. The Sqrt(Kv) values for practically all the AR4 models fall in the upper half of the Sqrt(Kv) range suggested by the pdf using the LAB05 ocean heat content data and in the low half of the range of the Sqrt(Kv) distribution based on the DOM08 results.

Table 6 Ratios of the percentiles values to the means for probability distributions shown in Table 4

	5.0%	16.7%	50.0%	83.3%	95.0%
LAB05					
B1	0.72	0.81	1.01	1.14	1.30
A1B	0.73	0.82	1.00	1.15	1.30
A2	0.74	0.85	1.01	1.13	1.23
REF	0.75	0.85	1.01	1.12	1.25
NO					
A1B	0.73	0.81	0.98	1.18	1.28
REF	0.74	0.84	1.00	1.16	1.28
DOM08					
A1B	0.80	0.87	0.97	1.13	1.29
REF	0.79	0.86	0.98	1.13	1.27

As a result, the simulations using the LAB05 input parameter distribution suggest stronger warming than the ensemble of the AR4 AOGCMs (see Figs. 7 and 8)—namely by the end of the twenty-first century (2091–2100) surface air temperature will increase above the present level (1980–1999) by 1.7C–2.4C (16.7–83.3 percentiles) for B1, 2.6C–3.6C for A1B and 3.4C–4.6C for A2. The corresponding increases in the mean are 2.1C, 3.2C and 4.0C, respectively. From the AR4 AOGCM ensemble (also shown in Figs. 7, 8), the mean increases are 1.8C, 2.8C and 3.4C for the B1, A1B and A2 scenarios, respectively. The corresponding ranges for sea level rise due to thermal expansion of sea-water in LAB05 simulations are: 9–18 cm for B1, 12–23 cm for A1B and 16–29 cm for A2 (Table 5).

The upper bounds of the distributions for SAT increases obtained in the LAB05 ensemble also significantly exceed the upper bounds simulated by the AR4 AOGCMs (Fig. 5a). For example, the probability of surface warming exceeding 4.1 C by the end of twenty-first century under the A2 scenario is 5% according to the AR4 AOGCM ensemble [Figure TS.27 of Solomon et al. (2007)], but 46% according to our results. The SAT changes simulated by the AR4 AOGCMs almost completely lie below the median suggested by our LAB05 projections for the high emissions scenario (A2) (Fig. 8 b) and below 83.3 percentile for the other two scenarios (Figs. 7a, 8a). The larger difference for A2, rather than for the other two scenarios is explained, in part, by the fact that the AR4 simulations for different SRES scenarios were carried out with different sets of AOGCMs. Thus the MIROC3.2 (hires) AOGCM, which produces the highest warming in the simulations for the A1B and B1 scenarios, was not used in the simulations with SRES A2.

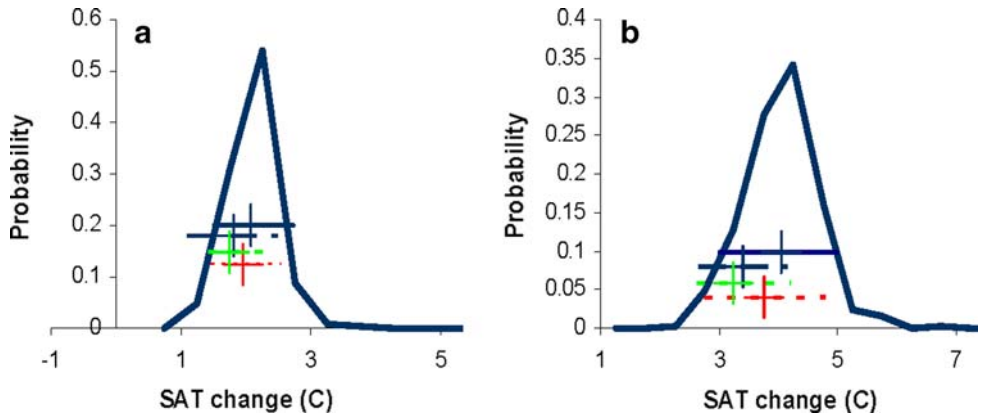

Fig. 8 Frequency distributions for SAT increase in 2091–2100 relative to the 1981–2000 average for the B1 (**a**) and A2 (**b**) SRES scenarios in simulations with LAB05 input parameter distribution. Solid horizontal bars show 5–95% ranges from 250-member ensembles of simulations with the MIT IGSM; *green and red dashed horizontal bars* show 5–95% ranges for DOM08 and NO approximated by scaling (see text for details); *blue dashed horizontal bars* show 5–95% ranges from the multi-model IPCC AR4 AOGCM ensemble [from figure TS.27 of Solomon et al. (2007)]

On the other hand our projections using the DOM08 analysis are more consistent with the AR4 AOGCM projections. However, our NO projections are closer to our LAB05-based projections, which indicates that, for the assumed prior for climate sensitivity, the observed changes in surface and upper air temperature are more consistent with the weaker ocean warming trend in the LAB05 analysis than with the stronger trend in the DOM08 analysis.

The differences in forcing between AOGCMs (illustrated in Fig. 1) are unlikely to affect our comparison in a significant way for two reasons. First, the climate forcing in the MIT IGSM simulations with mean values of climate sensitivity, rate of oceanic heat uptake and strength of aerosol forcing turned out to be very similar to the forcing averaged over the set of AR4 AOGCMs (see Fig. 1). Second, the correlation between TCR, which provides a good measure of model response to an external forcing on time scales of 50–100 years, and the strength of climate forcing for the AR4 AOGCMs is very weak. Such a correlation, if it existed, could result in a widening or narrowing of the range of simulated warming, depending on the sign of the correlation.

Kiehl (2007) and Knutti (2008) have shown that forcings in twentieth century simulations with different AOGCMs are correlated with the models' climate sensitivities. This correlation is explained by the differences in the strength of aerosol forcing in different AOGCMs. The lack of the correlation between TCR and forcing in simulations with the A1B SRES scenario for the AR4 AOGCMs is, likely, explained by the smaller contribution of aerosol forcing to total forcing over twenty-first century than over twentieth century.

As noted above, the use of input distributions based on different estimates of changes in the ocean heat content has a strong effect on the projected sea level rise due to thermal expansion. The range of values produced by the AR4 AOGCMs for the SRES A1B scenario falls in between ranges obtained in the different MIT ensembles (Fig. 7c).

6 Conclusions

Uncertainties in the estimates of the twentieth century changes in the deep ocean heat content have a strong effect on projections of the thermosteric sea level rise. The median value of the thermosteric sea level rise under the MIT reference forcing scenario is 41 cm if the DOM08 data is used, but 27 cm if the LAB05 data is used. The effect on SAT increases is more modest because of the correlation between climate sensitivity and ocean heat uptake required by observed twentieth century temperature changes. Nevertheless the effect is still significant. The median SAT increase at the end of the twenty-first century for the MIT reference scenario is 3.9 C if the DOM08 data are used for producing the probability distributions, but 4.9 C if the LAB05 data are used.

For the forcing scenarios used in this study, the statistical properties of the distributions of changes in surface air temperature do not depend on the choice of distribution for input climate parameters (see Table 6). Such similarity in the distributions for projected surface warming is explained by the correlation between S and Sqrt(Kv) imposed by the transient changes in SAT observed during twentieth century and by the fact that in all scenarios considered here forcing increases monotonically with time. Distributions for the equilibrium SAT changes under scenarios where the forcing is stabilized will be defined primarily by the input distributions for climate sensitivity.

The estimates of SAT increase produced by the multi-model ensemble of the IPCC AR4 AOGCMs slightly overestimate the results of the MIT ensemble for input distributions based on the DOM08 data, and strongly underestimate the surface warming suggested by the MIT results using the LAB05 data. For the sea level rise due to thermal expansion the situation is the opposite. As noted above, the LAB05 analysis is more consistent with the observed twentieth century surface and upper air temperature changes.

The results presented above emphasize the necessity for producing reliable estimates for changes of deep ocean heat content.

Existing AOGCMs do not cover the full ranges of uncertainties in either climate sensitivity or rate of heat uptake by the ocean. In addition, results produced by "ensembles of opportunity" are rather sensitive to the choice of the AOGCMs included in the ensemble. Most of the CMIP3 AOGCMs don't include parameterization of some physical processes affecting uncertainty in climate projections, such as the interaction between thecarbon cycle and climate (Meehl et al. 2007b; Knutti et al. 2008). As a result, Earth system models of intermediate complexity are likely to remain a primary tool for producing probabilistic climate forecasts in the near future.

Acknowledgments We would like to thank Piers Forster and Karl Taylor for providing data on climate forcing for AR4 AOGCMs and Catia Domingues for providing data on ocean heat content. We also thank two anonymous reviewers for their comments and suggestions. This work was supported in part by the Office of Science (BER), U.S. Dept. of Energy Grant No. DE-FG02-93ER61677, NSF, and by the MIT Joint Program on the Science and Policy of Global Change. We thank Marcus Sarofim for LHS samples of input parameters and Josh Willis for discussions on the ocean heat content data. 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 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

Appendix: Estimating probability density functions for input parameters

The methodology for estimating the PDFs for the climate model parameters follows the basic method first presented in Forest et al. (2002) with updates in Forest et al. (2006, 2008) (These latter papers primarily demonstrate the importance of including the full set of anthropogenic and natural forcings for simulating the twentieth century). The method can be summarized as consisting of two parts: (1) simulations of the twentieth century climate record and (2) the comparison of the simulations with observations using optimal fingerprint diagnostics. First, we require a large sample of simulated records of climate change in which climate parameters have been systematically varied over a sufficiently large portion of the parameter space. This implies that we can find regions of parameter space where the model and observations are clearly statistically inconsistent with each other (as defined by the second part). Second, we employ a method of comparing model data to observations that appropriately filters 'noise' from the pattern of climate change. The variant of optimal fingerprinting proposed by Allen and Tett (1999) provides this tool and yields detection diagnostics that are objective estimates of model-data goodness-of-fit. In effect, this amounts to a calibration of the parameter space of the model to provide simulations that are consistent with each of the multiple diagnostics used. Each diagnostic is used to estimate a likelihood function for the parameters and these are then combined using Bayes' theorem.

In this work, three temperature change diagnostics are used to calibrate the model parameter space. The diagnostics were chosen and designed to represent the large-scale response of the climate system to the anthropogenic and natural forcings since the pre-industrial era. The diagnostics are: (1) the decadal mean surface temperature changes over four equal-area latitude bands for the period 1946–1995 referenced to the 1905–1995 climatology; (2) the trend in the global mean ocean temperature (down to 3 km depth) and (3) the latitude-height pattern of the zonal mean upper air temperature difference between the 1961–1980 and 1986–1995 periods. The model diagnostics are discussed in detail in Forest et al. (2006) with appropriate references therein. As discussed in this paper, we explicitly explored the sensitivity to the choice of the ocean temperature data.

The likelihood function for each diagnostic depends on an estimate of the unforced variability (i.e., noise) for the pattern of temperature change. In the case of the upper-air and surface data, this pattern is a multivariate pattern and so an estimated covariance matrix is required. For the trend in the deep-ocean temperatures, this is a univariate diagnostic and so only an estimated variance is required. Errors in the data and their effect on the likelihood have been ignored in the surface and upper-air temperatures, because they are small compared to the natural variability. For the ocean diagnostic, the uncertainty in the estimated trend is also included. In all cases, observational data is insufficient to estimate the contribution of natural variability to the variance or covariance matrix. This implies that an alternative source for the natural variability must be used. Multi-century control simulations from AOGCMs in which the external forcings are held constant provide a data set that can be used to estimate the variability information. The control data are divided into appropriate segments corresponding to the length of the temperature records and then averaged in the same way as the observational data. Each

segment is considered to be a realization of the variability in the absence of climate forcings. A covariance matrix is estimated from the set of these segments taken from the control simulation. From this, the inverse of the covariance matrix is estimated and then used to estimate the goodness-of-fit statistic in the likelihood function.

Sansó and Forest (2009) has analyzed the sensitivity to how the covariance matrices for the surface temperature diagnostic are estimated given relatively little information from the AOGCMs. (NB: A 500-year simulation contains only ten non-overlapping 50-year segments that can be considered independent.) This has yet to be incorporated into the full posterior distribution using all three diagnostics.

References

Allen MR, Tett SFB (1999) Checking for model consistency in optimal fingerprinting. Clim Dyn 15:419–434. doi:10.1007/s003820050291

Andronova N, Schlesinger ME (2001) Objective estimation of the probability distribution for climate sensitivity. J Geophys Res 106:22605–22612. doi:10.1029/2000JD000259

Carton JA, Santorelli A (2008) Global decadal upper ocean heat content as viewed in nine analyses. J Clim 21:6015–6035. doi:10.1175/2008JCLI2489.1

Domingues CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, Barker PM, Dunn JR (2008) Improved estimates of upper-ocean warming and multi-decadal sea-level rise. Nature 453:1090–1094. doi:10.1038/nature07080

Forest CE, Stone PH, Sokolov AP, Allen MR, Webster M (2002) Quantifying uncertainties in climate system properties with the use of recent climate observations. Science 295:113–117. doi:10.1126/science.1064419

Forest CE, Stone PH, Sokolov AP (2006) Estimated PDFs of climate system properties including natural and anthropogenic forcings. Geophys Res Lett 33:L01705. doi:10.1029/2005GL023977

Forest CE, Stone PH, Sokolov AP (2008) Constraining climate model parameters from observed 20th century changes. Tellus 60A:911–920

Forster PMD, Taylor KE (2006) Climate forcing and climate sensitivities diagnosed from coupled climate model integrations. J Clim 19:6181–6194. doi:10.1175/JCLI3974.1

Frame DJ, Booth BBB, Kettleborough JA, Stainforth DA, Gregory JM, Collins M, Allen MR (2005) Constraining climate forecasts: the role of prior assumptions. Geophys Res Lett 32:L09702. doi:10.1029/2004GL022241

Gouretski V, Koltermann KP (2007) How much is the ocean really warming? Geophys Res Lett 34:L01610. doi:10.1029/2006GL027834

Hansen J, Sato M (2004) Greenhouse gas growth rates. Proc Natl Acad Sci USA 101:16109–16114. doi:10.1073/pnas.0406982101

Hansen J, Russell G, Rind D, Stone P, Lacis A, Lebedeff S, Ruedy R, Travis L (1983) Efficient three-dimensional global models for climate studies: models I and II. Mon Weather Rev 111:609–662. doi:10.1175/1520-0493(1983)111<0609:ETDGMF>2.0.CO;2

Hansen J, Lacis A, Rind D, Russell G, Stone P, Fung I, Ruedy R, Lerner J (1984) Climate sensitivity: Analysis of feedback mechanisms.In: Climate processes and climate sensitivity, Geophys.Monogr., No. 29, Amer. Geophys. Union, pp 130–163

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Xiaosu D (eds) (2001) Climate change 2001: the scientific basis. Cambridge University Press, UK, p 944

Iman RL, Helton JC (1988) An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal 8:71–90. doi:10.1111/j.1539-6924.1988.tb01155.x

Ishii M, Kimoto M, Sakamoto K, Iwasaki S-I (2006) Steric sea level changes estimated from historical subsurface temperature and salinity analyses. J Oceanogr 62:155–170. doi:10.1007/s10872-006-0041-y

Jones PD, Moberg A (2003) Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. J Clim 16:206–223. doi:10.1175/1520-0442(2003)016<0206:HALSSA>2.0.CO;2

Kiehl JT (2007) Twentieth century climate model response and climate sensitivity. Geophys Res Lett 34:L22710. doi:10.1029/2007GL031283 doi:10.1029/2007GL031383

Knutti R (2008) Why are climate models reproducing the observed global surface warming so well? Geophys Res Lett 35:L18704. doi:10.1029/2008GL034932

Knutti R, Tomassini L (2008) Constraints on the transient climate response from observed global temperature and ocean heat uptake. Geophys Res Lett, vol 35. doi:10.1029/2007GL032904

Knutti R, Stocker TF, Joos F, Plattner G-K (2003) Probabilistic climate change projections using neural network. Clim Dyn 21:257–272. doi:10.1007/s00382-003-0345-1

Knutti R, Meehl GA, Allen MR, Stainforth DA (2006) Constraining climate sensitivity from the seasonal cycle in surface temperature. J Clim 19:4224–4233. doi:10.1175/JCLI3865.1

Knutti R, Allen MR, Friedlingstein P, Gregory JM, Hegerl GC, Meehl GA, Meinshausen M, Murphy JM, Plattner G-K, Raper SCB, Stocker TF, Stott PA, Teng H, Wigley TML (2008) A review of uncertainties in global temperature projections over the twenty-first century. J Clim 21:2651–2663. doi:10.1175/2007JCLI2119.1

Levitus S, Antonov J, Boyer TP (2005) Warming of the World Ocean, 1955–2003. Geophys Res Lett 32:L02604. doi:10.1029/2004GL021592

Meehl GA, Covey C, Delworth T, Latif M, McAvaney B, Mitchell JFB, Stouffer RJ, Taylor KE (2007a) The WCRP CMIP3 multimodel dataset: a new era in climate change research. Bull Am Meteorol Soc 88:1383–1394. doi:10.1175/BAMS-88-9-1383

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C (2007b) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Meinshausen M, Raper SCB, Wigley TML (2008) Emulating IPCC AR4 atmosphere–ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: MAGICC 6.0. Atmos Chem Phys Discuss 8:6153–6272. http://www.atmos-chem-phys-discuss.net/8/6153/2008/acpd-8-6153-2008.html

Myhre G, Highwood EJ, Shine KP, Stordal F (1998) New estimates of radiative forging due to well mixed greenhouse gases. Geophys Res Lett 25:2715–2718. doi:10.1029/98GL01908

Prinn R, Paltsev S, Sokolov A, Sarofim M, Reilly J, Jacoby H (2008) The influence on climate change of differing scenarios for future development analyzed using the MIT Integrated Global System Model, MIT JP Report 163. (http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt163.pdf)

Raper CSB, Gregory JM, Stouffer RJ (2002) The role of climate sensitivity and ocean heat uptake on AOGCM transient

temperature and thermal expansion response. J Clim 15:124–130. doi:10.1175/1520-0442(2002)015<0124:TROCSA>2.0.CO;2

Sansó B, Forest C (2009) Statistical calibration of climate system properties. Appl Stat (in press). http://www.ams.ucsc.edu/share/technical-reports/2008/ucsc-soe-08-22.pdf

Sokolov AP, Stone PH (1998) A flexible climate model for use in integrated assessments. Clim Dyn 14:291–303. doi:10.1007/s003820050224

Sokolov AP, Forest CE, Stone PH (2003) Comparing oceanic heat uptake in AOGCM transient climate change experiments. J Clim 16:1573–1582

Sokolov AP et al (2005) The MIT integrated global system model (IGSM) Version 2: model description and baseline evaluation, MIT JP Report 124. (http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt124.pdf)

Sokolov AP, Dutkiewicz S, Stone PH, Scott JR (2007) Evaluating the use of ocean models of different complexity in climate change studies. MIT Joint Program for the Science and Policy of Global Change Rep. 128, 23 p. (http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt128.pdf)

Solomon S et al (2007) Technical summary. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL et al (eds) Climate change (2007), the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Stone PH, Yao M-S (1987) Development of a two-dimensional zonally averaged statistical–dynamical model. Part II: The role of eddy momentum fluxes in the general circulation and their parameterization. J Atmos Sci 44:3769–3786. doi:10.1175/1520-0469(1987)044<3769:DOATDZ>2.0.CO;2

Stone PH, Yao M-S (1990) Development of a two-dimensional zonally averaged statistical–dynamical model. Part III: The parameterization of the eddy fluxes of heat and moisture. J Clim 3:726–740. doi:10.1175/1520-0442(1990)003<0726:DOATDZ>2.0.CO;2

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philos Trans R Soc A 365:2053–2075. doi:10.1098/rsta.2007.2076

Webster M et al (2003) Uncertainty analysis of climate change and policy response. Clim Change 61:295–320. doi:10.1023/B:CLIM.0000004564.09961.9f