Compensation between cloud feedback and aerosol-cloud interaction in CMIP6 models

Chenggong Wang¹, Brian Soden², Wenchang Yang³, Gabriel A. Vecchi¹,³,⁴

1. Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA
2. Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL, USA
3. Department of Geosciences, Princeton University, Princeton, NJ, USA
4. Princeton Environmental Institute, Princeton University, Princeton, NJ, USA

To be submitted to GRL
September 2020

*Corresponding author address: Chenggong Wang, c.wang@princeton.edu
Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA 08544

Key Points:
1. Models with strong positive cloud feedback tend to have strong aerosol-cloud interaction
2. This compensation relationship enables the models to match the historical warming even with a large spread in climate sensitivity
3. Historical interhemispheric warming indicates the high climate sensitivity models overestimate the aerosol-cloud interaction
Abstract

The new generation of climate models yields estimates of effective climate sensitivity (ECS) that are much higher than past generations due to a stronger amplification from cloud feedback. If plausible, these models require substantially larger greenhouse gas reductions to meet global warming targets. We show that models with a more positive cloud feedback also have a stronger cooling effect from aerosol-cloud interactions. These two effects offset each other during the historical period when both aerosols and greenhouse gases increase, allowing either strong or weak cloud feedback models to reproduce the observed global-mean temperature change. Since anthropogenic aerosols primarily occur in the Northern Hemisphere, strong aerosol-cloud interaction models produce a distinct hemispheric asymmetry in warming. We show that the observed interhemispheric warming asymmetry during the mid to late 20th century is more consistent with low ECS (weak aerosol indirect effect) models.

Plain Language Summary

The response of clouds to surface temperature change can amplify or dampen the greenhouse gas induced warming, also known as cloud feedback. We find that in the latest generation of climate models, those models with a more positive cloud feedback tend to have a stronger cooling effect from aerosol-cloud interaction. The compensation between cloud feedback and aerosol-cloud interaction enables models to reproduce the historical global mean temperature change. In spite of having significantly different surface temperature sensitivity to increasing CO₂, the historical record of globally-averaged temperature is not a strong constraint in distinguishing these models. However, the inter-hemispheric difference in temperature over the 20th century provides a
constraint that distinguishes the models that have a large or small sensitivity to increasing CO₂.

Historical aerosol emission from human activities has largely been in the Northern Hemisphere.

Consequently, models with strong or weak aerosol-cloud interaction produce distinctly different warming patterns over the historical record. The observed warming asymmetry is inconsistent with the strong cloud feedback models, and is more consistent with the weak cloud feedback models. This study can help us to better understand, and reduce, the uncertainty in the projected future warming.

1. Introduction

Clouds exert a profound influence on global climate by modulating the flow of energy through the atmosphere. The radiative effects of clouds are complex: clouds can both cool climate by reflecting incoming sunlight, and warm it by absorbing and reemitting thermal radiation (Ramanathan et al., 1989). The net impact of these competing effects depends on the distribution, macrophysical and microphysical properties of clouds (Hartmann et al., 1992). As the planet warms from increasing greenhouse gases (GHGs), it is not yet clear whether changes in cloud properties will further amplify or dampen the GHG induced warming, or by how much. Uncertainties in predicting this radiative feedback from clouds are the largest cause of spread in model predictions of future global warming (Boucher et al., 2013; Ceppi et al., 2017; Zelinka et al., 2020).

Current estimates of cloud feedback range from effectively neutral to substantially positive in response to GHG forcing (Chung & Soden, 2015; Vial et al., 2013; Zelinka et al., 2013, 2016). The latest climate models from the Sixth Phase of the Coupled Model Intercomparison Project (CMIP6) has produced a number of models with significantly higher effective climate sensitivity
(ECS) compared to previous generations (Zelinka et al., 2020). This higher ECS has been shown to result primarily from a more positive cloud feedback in models. The ECS ranges from 1.8-5.6 K in the CMIP6 models, with seven of them having an ECS greater than 4.7 K, the upper bound of ECS in CMIP5 (Andrews et al., 2012; Flato et al., 2014).

In addition to changes in GHGs, climate forcing over both the historical era and projected future scenarios will involve changes in anthropogenic tropospheric aerosols. Interactions between clouds and aerosols are complex and also influence the radiation budget (Penner et al., 1992). Aerosols affect the flow of radiation directly by scattering and absorbing incoming sunlight. Additionally, aerosols act as cloud condensation nuclei and can alter cloud albedo and cloud life time which, in turn, further modulate the radiation budget (Rotstayn & Penner, 2001; Twomey, 1977). For example, increasing aerosols reduces the cloud droplet size and, for a given liquid water path, increases cloud albedo resulting in a cooling effect. These indirect effects are both highly uncertain and often larger than the direct radiative impact of aerosols (Lohmann et al., 2010; Myhre et al., 2013; Smith et al., 2020; Zelinka et al., 2014).

In this study, we show that models with a more positive cloud feedback in response to greenhouse gases also tend to have a stronger cooling effect from aerosol-cloud interactions. These two effects offset each other during much of the 20th century, when both anthropogenic aerosols and GHG emissions increased. Thus, both models with low and high ECS are able to reproduce the observed changes in global-mean temperature. However, this compensation does not occur in future emission scenarios where aerosols are projected to decrease as CO2 and other greenhouse gases continue to increase. We will show that the inter-hemispheric temperature contrast over the
historical period provides a way to distinguish between low and high ECS models. We find that models with a lower ECS (and smaller aerosol-cloud interactions) are more consistent with the observed interhemispheric asymmetric warming pattern during the 20th century than are high ECS models.

2. Data and Methods

We use monthly model data from historical, piControl, abrupt-4xCO2, and 1pctCO2 experiments in CMIP6 (Eyring et al., 2016). We limit our analysis to models that have 1) all four experiments available, 2) the variables necessary to compute cloud feedback parameters are archived, and 3) a piControl experiment of at least 450 years length available. This leaves 30 models as listed in Table S1. For each model, monthly climatology (450 years mean) of piControl experiment provides a pre-industrial reference that is used to compute the anomalies in the other experiments. The abrupt-4xCO2 experiment instantaneously increases the CO2 concentration to 4 times the pre-industrial level and is used to estimate climate sensitivity. The 1pctCO2 experiment increases the CO2 concentration by 1% per year and is used to quantify the cloud radiative response to surface warming under a transient emission scenario. The global analysis of surface temperature observations is obtained from the GISS Surface Temperature Analysis (GISTEMP v4) (Lenssen et al., 2019).

Following Gregory et al. (2004), we calculate the effective climate sensitivity (ECS) by regressing the global-mean TOA radiation anomaly on the global-mean surface temperature anomaly of the first 150 years in abrupt-4xCO2 experiment. Half of the x-intercept of the regression is considered as ECS, which is defined in terms of the doubling CO2. Although this method tends to
underestimate the equilibrium climate sensitivity (Armour et al., 2013; Winton et al., 2020), the equilibrated change in global-mean surface temperature in response to a doubling of CO$_2$, it is widely used due to simplicity and practicality.

The radiative kernels from the GFDL model are used to calculate the cloud radiative response (detailed in Soden et al. (2008)). With the radiative kernels, we decompose the radiative response at the top of the atmosphere (TOA) that results from changes in temperature, water vapor, surface albedo, and clouds. The cloud feedback is defined as the slope of the regression of the cloud radiative response on global-mean temperature anomaly in the abrupt-4xCO2 experiment.

The radiative response due to changes in clouds in the historical simulation arises from both surface warming induced changes (i.e., cloud feedback) and aerosol-cloud interactions. Following Soden & Chung (2017), we decompose the total cloud radiative response ($\Delta R_c^{tot}$) in the historical experiment into two parts: the part due to global-mean surface temperature change and the part due to aerosol-cloud interaction (the aerosol-mediated cloud radiative response, $\Delta R_c^{aer}$). The aerosol-mediated cloud response includes both the aerosol indirect effect and non-local changes in clouds that result from aerosol-induced changes in the large-scale circulation (Soden & Chung, 2017). The first part can be estimated by multiplying the global-mean temperature anomaly and the normalized cloud radiative response parameter $\alpha$ obtained from the corresponding 1pctCO2 experiment for each model. Therefore, the aerosol-mediated cloud radiative response can be expressed as:

$$\Delta R_c^{aer} = \Delta R_c^{tot} - \alpha_{1pctCO2} \cdot \Delta T_s$$
We note that clouds also exhibit a fast response to CO$_2$ forcing (Andrews & Forster, 2008; Chung & Soden, 2015; Colman & McAvaney, 2011; Gregory & Webb, 2008; Vial et al., 2013; Zelinka et al., 2013) and that this is implicitly included with the cloud feedback in computing $\alpha$ in 1pctCO2 experiment, but not in the cloud feedback estimated in abrupt-4xCO2 experiment. As shown by Soden & Chung (2017), this approach successfully reproduces the estimates of aerosol induced cloud radiative response calculated using single forcing (i.e., aerosol-only) experiments with fixed SSTs to suppress the surface temperature driven cloud feedbacks. These estimates are highly correlated to the aerosol-cloud interaction cooling effect estimated by the approximate partial radiative perturbation method in Smith et al. (2020) (Figure S11).

The independent two-sample $t$-test is applied to distinguish the statistically significant features between the strongest nine cloud feedback models (the “top nine” or T9) and the weakest nine cloud feedback models (the “bottom nine” or B9). In the main text, we compute the differences in cloud feedback, radiation and temperature between the multi-model ensemble mean of T9 and B9 models. All the plots only show the differences that reject the null hypothesis that the T9 and B9 models have the same multi-model ensemble mean (using a two-sided $t$-test with a $p$-value<0.05). As shown in the supplementary material, the conclusions of this paper are not sensitive to the number of models chosen to represent the top or bottom range of the intermodel spread in cloud feedback.

We evaluate how well the models simulate the global-mean historical warming by the GOOD HIST index: the absolute difference in historical warming between CMIP6 models and GISSE TEMP data. The historical warming is defined as the averaged surface temperature in 1990-
2014 minus that in 1880-1909. So, the models that are good at simulating the historical warming have a small GOOD HIST index (see values in Table S1).

3. Results

3.1. Cloud Feedback and ECS

In response to increasing CO₂, models show warming and substantial climate changes that feed back onto the warming, including changes in the amount and distribution of clouds (Wetherald & Manabe, 1988). The part of cloud radiative response (units of W m⁻²) due to a change in global-mean surface temperature (units of K) is defined as the cloud feedback (W m⁻² K⁻¹). In CMIP6, the cloud feedback tends to be positive and there is a strong relationship between cloud feedback and ECS: models with more positive cloud feedback show higher ECS (Figure 1a, r²=0.69) (Meehl et al., 2020; Zelinka et al., 2020). This strong ECS-cloud feedback relationship is consistent with the previous studies showing that cloud feedback is the dominant source of the uncertainty of climate sensitivity (Cess et al., 1990; Colman, 2003; Dufresne & Bony, 2008; Soden & Held, 2006; Webb et al., 2013; Zelinka et al., 2020).

The spatial pattern of cloud feedback (Figure 2a-c) differs considerably between the models with the strongest cloud feedback (the “top nine” or T9) and those with the weakest cloud feedback (the “bottom nine” or B9). The strong global-mean enhancement in cloud feedback in the T9 models arises principally from a substantially higher cloud feedback in the Southern Hemisphere. The differences are significant in the southeast of the Pacific, Atlantic, the Indian Ocean as well as the Southern Ocean (Figure 2c). The larger cloud feedback is the primary cause of the substantially
higher ECS in CMIP6 compared to previous coupled model ensembles (Meehl et al., 2020; Zelinka et al., 2020)

### 3.2. Aerosol-Cloud Interaction in the Historical Period

To better understand the cloud radiative response in the historical period, we explore this same model ensemble subject to the historical radiative forcing over 1850-2014. The historical experiments allow us to: i) examine the behaviors of clouds in response to more complex emission scenarios that involve both aerosols and GHGs; and ii) ascertain the extent to which observations can constrain the range of cloud feedbacks and/or ECS.

In contrast to the GHG-only forcing experiment, the total cloud radiative response ($\Delta R^{\text{tot}}_c$) to the more complex historical forcings involves both surface temperature driven and aerosol-mediated changes in clouds. Figures 3a and 3b compare the anomalies of the total cloud radiative response ($\Delta R^{\text{tot}}_c$) and aerosol-mediated cloud response ($\Delta R^{\text{aer}}_c$). In models with a more positive cloud feedback (T9; red thick line in Figure 3a), $\Delta R^{\text{tot}}_c$ actually exhibits a cooling effect in the historical simulations; i.e, $\Delta R^{\text{tot}}_c < 0$. Even more surprising, the models with the strongest positive cloud feedback (T9; thick red line) have a larger cloud cooling effect than the models with a weaker cloud feedback (B9; thick blue line Figure 3a). That is, the models with a more positive cloud feedback in response to CO2 doubling show a more negative cloud response due to historical radiative forcing. This is particularly evident after 1950 (gray shading in Figure).

The negative $\Delta R^{\text{tot}}_c$ in T9 models arises almost entirely from a negative aerosol-mediated cloud response $\Delta R^{\text{aer}}_c$ (Figure 3b). In contrast, the B9 models (with weak cloud feedback) have a very
small aerosol-mediated cloud response. Thus, models with a more positive cloud feedback (T9) tend to have a larger negative aerosol-mediated response compared to those with a weaker cloud feedback (B9). This occurs despite both T9 and B9 models having similar clear-sky aerosol forcing, as evident in the reflected clear-sky shortwave TOA fluxes (Figure 3c). This indicates that, at least for the shortwave, the intermodel difference in the total aerosol forcing is dominated by aerosol-cloud interactions not the direct aerosol forcing. The aerosol emissions, especially sulfur dioxide, are relatively constant after 1980, while the GHG emissions continue to increase (Hoesly et al., 2017). Thus, the total cloud radiative response increases with increasing global mean temperature in T9 models (red line in Figure 3a) after 2000, while the aerosol-mediated cloud response remains nearly constant, reflecting a more dominant role of cloud feedback in determining the total cloud radiative response.

In the T9 models, there are four models developed at the same modeling center, which raises the possibility that lack of model independence and a single family of models may bias the composite model results discussed above. To evaluate this possibility, we repeat the analysis using two other groupings. First, selecting only one model per modeling center in our composites and consider the top and bottom six models (T6 and B6). Second, selecting a broader range of models (i.e., further from the extremes) while also restricting the analysis to one model per center (the top and bottom eight models, T8 and B8). These alternative composite analyses (see Supplemental Materials) reproduce the main aspects of the T9/B9 composite results, indicating robustness relative to the details of model selection within this ensemble.
The spatial pattern of the cloud radiative response in the historical experiment also differs from that obtained in the abrupt-4xCO2 experiment. In Figure 2d, the cloud radiative response (1950-2000 mean) is negative in the Northern Hemisphere in high cloud feedback models (T9). The cooling effect of clouds is as large as -4 W m^-2 over many of the northern extratropics and subtropics in the T9 models, while the B9 models have little change in the total cloud radiative response compared to the pre-industrial period (Figure 2e). The aerosol-mediated cloud response in the T9 models is responsible for the negative total cloud radiative response in the northern midlatitude during 1950-2000 (Figure 2g). Since the $\Delta R_{c}^{tot}$ and $\Delta R_{c}^{aer}$ of B9 models are small, the differences in the total cloud radiative response between the T9 and B9 models are almost identical to the differences in the aerosol-mediated cloud radiative response (Figure 2f and Figure 2i).

These results imply a compensation between the cloud feedback from CO2-induced surface warming and the aerosol-mediated cloud response. This anti-correlation is more clearly shown in Figure 1b, which compares the global-mean cloud feedback for each model from the abrupt-4xCO2 simulations with the corresponding aerosol-mediated cloud response from the historical simulations. Models with a more positive cloud feedback tend to have a larger negative aerosol-mediated response ($r^2$=0.60). This helps to explain why models with a higher ECS also tend to have a larger net aerosol radiative effect (Meehl et al., 2020). Dividing the 30 models into two groups based upon how well they simulate the observed global-mean historical warming (defined in Data and Methods) illustrates the importance of this compensation. The correlation between the warming effects of cloud feedback and the cooling effect of aerosol-cloud interactions is higher for the 15 models that better reproduce the observed warming ($r^2$=0.83 filled circles in Figure 1b) compared to those that don’t ($r^2$=0.17, open circles in Figure 1b).
3.3. Interhemispheric Warming Asymmetry

Due to the larger cooling effect of the aerosol-cloud interactions, T9 models simulate slightly colder surface temperature anomalies during the mid to late 20th century compared to the B9 models (Figure 4a), even though the T9 models have a more positive cloud feedback and a higher effective climate sensitivity. While this difference between the B9 and T9 models’ surface temperature anomaly is small when globally averaged (and only few scattered years are significantly different – indicated by the gray shading), the hemispheric asymmetry of the historical aerosol forcing induces substantial differences in the interhemispheric warming asymmetry (Figure 4b). Here, we use the surface temperature change in Northern Hemisphere minus that in the Southern Hemisphere to evaluate the interhemispheric warming asymmetry. The meridional asymmetry in the temperature evolution over the late 20th century distinguishes the T9 and B9 models: the T9 models warm more in the SH than the NH during the last century, and the differences in the interhemispheric warming asymmetry between the T9 and B9 models are significant for the entire period 1950-2000 (gray shading in Figure 4b).

Observed interhemispheric warming asymmetry over the 20th century is more consistent with the models with weaker cloud feedback and aerosol indirect effect (B9) than with those with the more positive cloud feedback and aerosol indirect effect (T9). The observed global- and annual-mean temperature anomalies are broadly consistent with both sets of models (Figure 4a). However, when the interhemispheric difference in surface temperature is considered (Figure 4b), the model ensemble mean of B9 more closely reproduces the observed hemispheric contrast in warming over most of the historical period. The ensemble of B9 models provides a more calibrated estimate of
the probability distribution of the observed interhemispheric temperature evolution than does the ensemble of T9 models, with the observed temperature anomaly rank being approximately uniformly distributed within the B9 ensemble, but concentrated at the upper end of the T9 ensemble (Figure S1).

Discussion

The seeming consistency of global-mean temperature evolution between strong cloud feedback (high ECS) models and observations requires a strong aerosol indirect effect that leads to an interhemispheric temperature evolution that is inconsistent with observations. Because of the strong negative correlation between a model’s cloud feedback in response to CO₂ (and its CO₂-induced ECS) and its indirect effect (Figure 1b), the global-mean temperature evolution in strong and weak cloud feedback models is not well separated over the historical period (Figure 4a) when increases in CO₂ are accompanied by strong aerosol forcing. Both strong (high ECS) and weak (low ECS) cloud feedback models are able to simulate the observed global-mean temperature record, but T9 models do it through a combination of strong warming from GHGs and strong cooling from aerosols, while B9 models do it with moderate warming from GHGs and modest cooling from aerosols. Because historical aerosol forcing has been larger in the Northern Hemisphere, the strong aerosol response in T9 models produces a distinctive historical interhemispheric surface temperature evolution (red line in Figure 4b), which is inconsistent with that in observations over 1950-2000 (black line in Figure 4b). These results mirror the recent findings that CMIP6 models that more faithfully capture the observed evolution of surface anomalies across
a range of quantities over 1980-2014 tend to have lower 21st century projected warming (Brunner et al., 2020).

Successfully reproducing the observed changes in global-mean temperature over the 20th century is an important test of climate models. Given the importance of clouds in determining ECS, it seems unlikely that a model with a strong cloud feedback and a weak aerosol-cloud interaction, or vice-versa, could achieve this important benchmark. Thus, the compensation could result from implicit or explicit efforts to tune the representation of clouds in models to reproduce the observed global-mean temperature record when forced with historical emissions (Mauritsen & Roeckner, 2020; Schmidt et al., 2017). Indeed, the compensation is more evident within the models that simulate the historical warming well ($r^2=0.83$) and not significant for models that do not reproduce the historical warming ($r^2=0.17$). This helps to reconcile why Meehl et al. (2020) found a significant positive correlation between the total aerosol forcing and ECS, while Smith et al. (2020) did not. In previous generations of models, which largely did not include the aerosol indirect effect, a significant correlation was found between the aerosol direct forcing and climate sensitivity (Kiehl, 2007; Knutti, 2008). Although across all CMIP5 models, Forster et al. (2013) found no significant intermodel correlation between the present-day forcing and ECS, they found a significant correlation for the models that simulate the historical warming well.

Alternatively, there could be a physical process that is intrinsic to models that links cloud feedback and aerosol-cloud interaction. Cloud feedback is known to be sensitive to the spatial pattern of surface temperature change (Andrews et al., 2015; Armour et al., 2013; Winton et al., 2010; Zhou et al., 2016). Models with a strong cloud feedback produce a very different temperature change
pattern than do models with a weak cloud-feedback (Figure S3). Such hemispheric asymmetry in warming can induce changes in the large-scale circulation (Allen, 2015; Allen et al., 2015; Chung & Soden, 2017; Hwang et al., 2013; Ming & Ramaswamy, 2011) which, in turn, lead to hemispheric differences in the pattern of cloud feedback (Soden & Chung, 2017).

The differences in the spatial pattern of warming induced by strong aerosol-cloud interactions also impact many other aspects of the simulated response to anthropogenic forcing. For example, the meridional structure of sea surface temperature and heating changes has been connected to the evolution of tropical rainfall (Jacobson et al., 2020; Kang et al., 2008, 2014; Xie et al., 2010; Yang et al., 2019) and tropical cyclone activity (Booth et al., 2012; Merlis et al., 2013; Vecchi & Soden, 2007; Yang et al., 2019). The extent to which observed changes in the meridional structure of rainfall and tropical cyclone activity can be ascribed to past radiative forcing change will also depend in part on the realism of the cloud response to historical aerosol forcing (Booth et al., 2012; Chung & Soden, 2017; Zhang et al., 2013). The hemispheric asymmetry of surface warming over the historical period can thus provide an important constraint on both aerosol-cloud interactions as well as shifts in climatic quantities of more direct societal relevance, such as tropical cyclones and rainfall.
Acknowledgments

This work supported in part by Awards NA20OAR4310393 and NA18OAR4310418 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, and the Carbon Mitigation Initiative at Princeton University. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We also thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. All the CMIP6 model output is available at https://esgf-node.llnl.gov/projects/cmip6/.

References


https://doi.org/10.1029/JD095iD10p16601


Practice and philosophy of climate model tuning across six US modeling centers.


pg. 22


Figure 1. Cloud feedbacks, effective climate sensitivity (ECS), and aerosol-mediated cloud radiative responses ($\Delta R_{c}^{\text{aer}}$) in the CMIP6 models. a) Scatter plot of ECS and cloud feedback parameter. b) Inter-model relationship between cloud feedback and aerosol-mediated cloud radiative responses. The cloud feedback and ECS are computed from the response to 4xCO2 forcing and the aerosol-mediated cloud radiative response is calculated from the historical experiments (1950-2000 mean). Each dot represents a single model. The colors from red to blue indicate high cloud feedback models to low cloud feedback models. The filled circles represent the 15 models that simulate the historical warming (1990-2014 mean minus 1880-1909 mean) closer to the observation than the other 15 models (open circles).
Figure 2. Cloud feedback (top row), historical cloud radiative response (1950-2000 mean of $\Delta R_c^{\text{tot}}$, mid row) and aerosol-mediated cloud radiative response (1950-2000 mean of $\Delta R_c^{\text{aer}}$, bottom row) in the CMIP6 ensemble. a), d), g): Ensemble mean of T9 models. b), e), h): Ensemble mean of B9 models. c), f), i): The differences between T9 and B9 models. Only the regions that the difference passed t-test (p<0.05) are shown.
Figure 3. Radiative response in the historical experiments from the CMIP6 ensemble. a) The total cloud radiative response ($\Delta R^\text{tot}_c$). b) The aerosol-mediated cloud radiative response ($\Delta R^\text{aer}_c$). c) The change of TOA clear-sky net shortwave radiation. Each thin line presents a single model, and the color has the same meaning as in Figure 1. The thick lines represent the model ensemble mean of T9 (red), B9 (blue), and the gray shadings indicate the years that the difference between T9 and B9 are significant (passed t-test, p<0.05).
Figure 4. Modeled and observed surface temperature change. Annual time-series of a) the global-mean surface temperature anomaly, and b) the interhemispheric temperature anomaly difference. The black line is from the GISTEMP and is rebased to match the model ensemble mean of the 1951-1980 period value. Each thin gray line represents a single ensemble from one model. The red and blue lines indicate the model ensemble mean of the T9 and B9 models, respectively. The gray shadings indicate the years that the difference between T9 and B9 are significant (passed t-test, p<0.05).