Geophysical Research Letters

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

Supporting Information:
- Supporting Information S1

Correspondence to:
C. Wang,
c.wang@princeton.edu

Compensation Between Cloud Feedback and Aerosol-Cloud Interaction in CMIP6 Models

Chenggong Wang[1] , Brian J. Soden[2] , Wenchang Yang[3] , and Gabriel A. Vecchi[1,3,4]

[1]Program in Atmospheric and Oceanic Sciences, 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]High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA

Abstract The most recent generation of climate models (the 6th Phase of the Coupled Model Intercomparison Project) 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 more positive or neutral cloud feedback models to reproduce the observed global-mean temperature change. Since anthropogenic aerosols primarily concentrate in the Northern Hemisphere, strong aerosol-cloud interaction models produce an interhemispheric asymmetric warming. We show that the observed 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 CO_2, the historical record of global-mean temperature is not a strong constraint in distinguishing these models. However, the interhemispheric difference in temperature over the 20th century provides a constraint that distinguishes the models that have a large or small sensitivity to increasing CO_2. Over the 20th century, changes in anthropogenic aerosols were mostly concentrated in the Northern Hemisphere. Consequently, models with strong or weak aerosol-cloud interactions produce different warming asymmetry over the historical period, and the observed warming asymmetry is more consistent with the models that have weak aerosol-cloud interactions (and less positive cloud feedback). This study can help us 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 GHGs 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 GHGs 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 to 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 involve changes in aerosols. Interactions between clouds and aerosols are complex and also influence the radiation budget (Penner et al., 1992). Aerosols affect the radiation directly by scattering and absorbing incoming sunlight. Additionally, aerosols can act as cloud condensation nuclei, change the cloud droplet size and alter cloud albedo, and cloud lifetime, modulating the radiation budget (Rotstayn & Penner, 2001; Twomey, 1977). The 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 increasing GHGs also tend to have a stronger cooling effect from aerosol-cloud interactions (ACI). These two effects offset each other during much of the 20th century, when both anthropogenic aerosols and GHGs 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 CO_2 and other GHGs continue to increase. We will show that the interhemispheric temperature contrast over the historical period provides a way to distinguish between low and high ECS models. Also, we find that models with a lower ECS (and weaker ACI) are more consistent with the observed interhemispheric asymmetric warming pattern during the 20th century.

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 the variables necessary to compute cloud feedback parameters in four experiment, and a piControl experiment longer than 450 years. This leaves 30 models as listed in Table S1. All anomalies in this study are reference to the monthly climatology of piControl experiment. The abrupt-4xCO2 experiment is used to estimate climate sensitivity and cloud feedback. The 1pctCO2 experiment 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 ECS by regressing the global-mean top of the atmosphere (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 CO_2. One thing to note is that this method tends to underestimate the equilibrium climate sensitivity (Armour et al., 2013; Winton et al., 2020).

To calculate the cloud feedback strength, we use the radiative kernels from the GFDL model (Soden et al., 2008) to decompose the radiative response at the TOA into the components due to changes in temperature, water vapor, surface albedo, and clouds. The cloud feedback is defined as the regression slope 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 ACI. Following Soden and Chung (2017), we decompose the total cloud radiative response (ΔR_c^{tot}) in the historical experiment into two parts: the part due to global-mean surface temperature change and the part due to ACI (the aerosol-mediated cloud radiative response, ΔR_c^{aer}). The aerosol-mediated cloud radiative response includes both the aerosol indirect effect and nonlocal changes in clouds that result from aerosol-induced changes in the large-scale circulation (Soden & Chung, 2017). The surface temperature driven part can be estimated by multiplying the global-mean temperature anomaly and the normalized cloud radiative response parameter α obtained from the corresponding 1pctCO2 experiment for each model. Therefore, the aerosol-mediated cloud radiative response can be expressed as:

$$\Delta R_c^{\text{aer}} = \Delta R_c^{\text{tot}} - \alpha_{\text{1pctCO2}} \cdot \overline{\Delta T_s}$$

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

this is included in the α parameter estimated from 1pctCO2 experiment, but not in the cloud feedback estimated from abrupt-4xCO2 experiment. As shown by Soden and 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. As a further test of the method we use to estimate ACI, our results are highly correlated to the ACI 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 nine most positive cloud feedback models (the "top nine" or T9) and the nine weakest cloud feedback models (the "bottom nine" or B9). In the main text, we compute the differences in cloud feedback, radiation and temperature between the multimodel 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 multimodel ensemble mean (using a two-sided *t*-test with a *p*-value<0.05). As shown in the supplementary material, the conclusions of this study 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 GISTEMP 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 CO_2, 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, $r^2 = 0.69$) (Meehl et al., 2020; Zelinka et al., 2020). This strong ECS-cloud feedback relationship is consistent with 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 (Figures 2a–2c) differs considerably between the models with the most positive cloud feedback (the "top nine" or T9) and those with the least positive cloud feedback (the "bottom nine" or B9). The more positive global-mean cloud feedback in the T9 models arises principally from a substantially more positive cloud feedback in the Southern Hemisphere. The differences are statistically significant in the southeast regions of the Pacific, Atlantic and Indian Ocean as well as in the Southern Hemisphere midlatitude (Figure 2c). The more positive cloud feedback, mostly due to more positive shortwave low cloud feedback in the Southern Hemisphere midlatitude, 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 investigate CMIP6 models forced by 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.
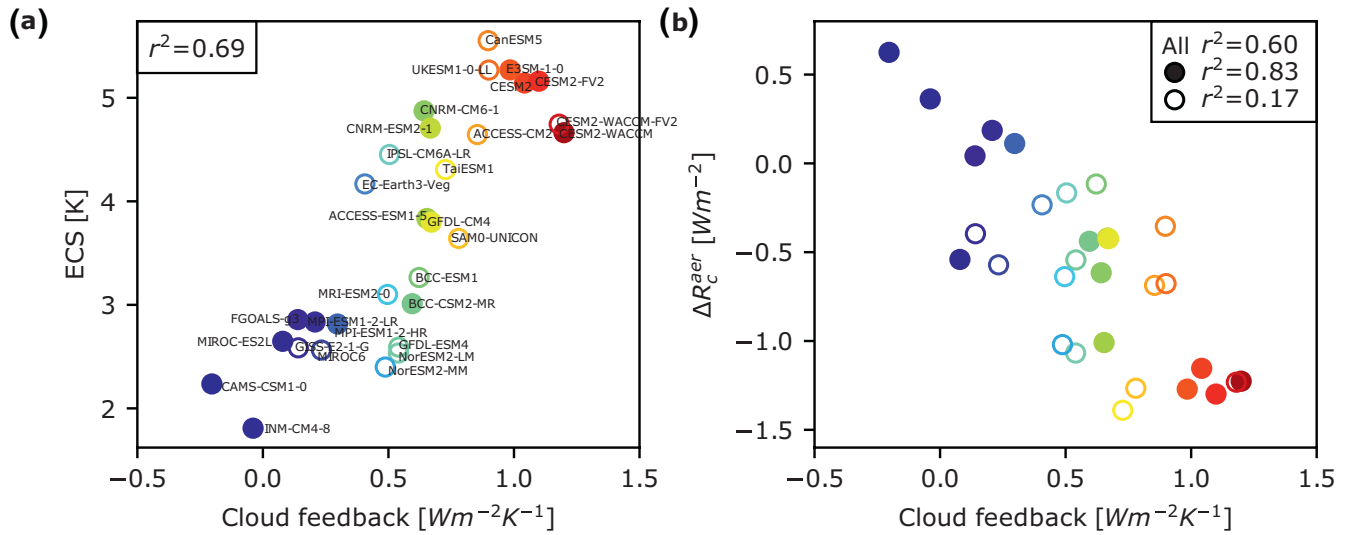
Figure 1. Cloud feedbacks, ECS, and aerosol-mediated cloud radiative responses (ΔR_c^{aer}) in the CMIP6 models. (a) Scatter plot of ECS and cloud feedback parameter. (b) Intermodel relationship between cloud feedback and ΔR_c^{aer}. The cloud feedback and ECS are computed from the response to 4xCO$_2$ forcing and the ΔR_c^{aer} 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 are more consistent with the observation on historical warming (1990–2014 mean minus 1880–1909 mean) than the other 15 models (open circles). CMIP6, the 6th Phase of the Coupled Model Intercomparison Project; ECS, effective climate sensitivity.

In contrast to the GHGs-only forcing experiment, the total cloud radiative response (ΔR_c^{tot}) to the more complex historical forcings involves both surface temperature driven and aerosol-mediated changes in clouds.
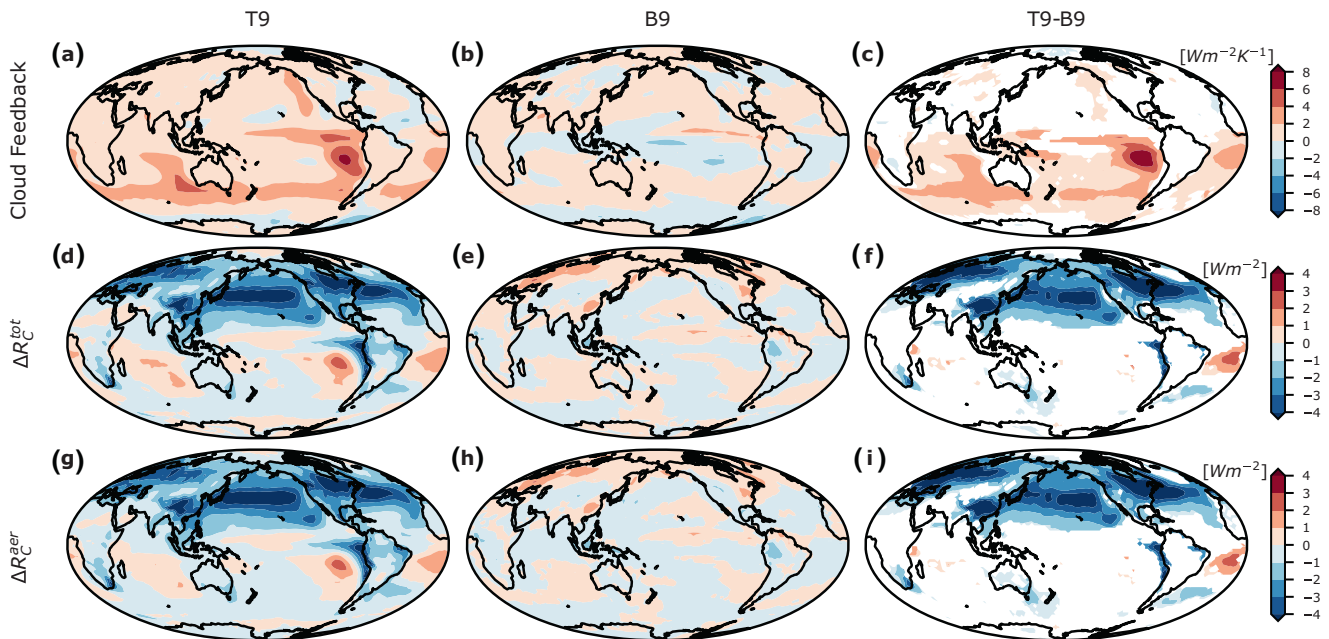

Figure 2. Cloud feedback (top row), historical cloud radiative response (1950–2000 mean of ΔR_c^{tot}, mid row) and aerosol-mediated cloud radiative response (1950–2000 mean of ΔR_c^{aer}, bottom row) in the CMIP6 ensemble. (a, d, and g): Ensemble mean of T9 models. (b, e, and h): Ensemble mean of B9 models. (c, f, and i): The differences between T9 and B9 models. Only the regions that the difference passed t-test ($p < 0.05$) are shown. CMIP6, the 6th Phase of the Coupled Model Intercomparison Project.
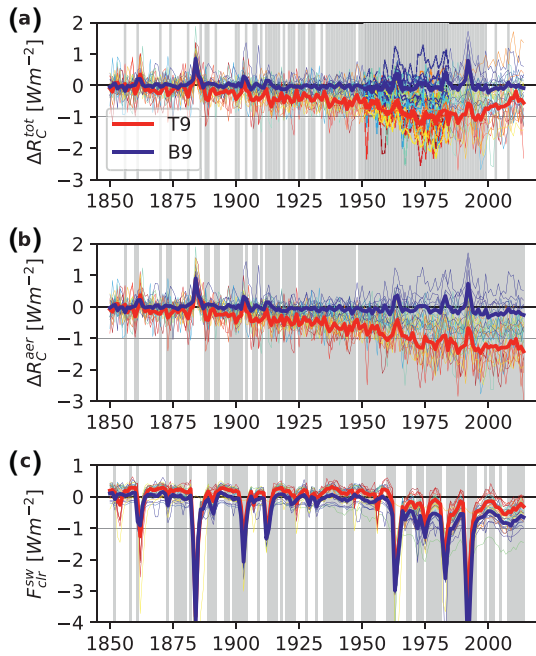
(a)

(b)

(c)

Figure 3. Radiative response in the historical experiments from the CMIP6 ensemble. (a) The total cloud radiative response (ΔR_c^{tot}). (b) The aerosol-mediated cloud radiative response (ΔR_c^{aer}). (c) TOA clear-sky shortwave forcing $\left(F_{\text{clr}}^{\text{sw}}\right)$. 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, $p < 0.05$). CMIP6, the 6th Phase of the Coupled Model Intercomparison Project; TOA, top of the atmosphere.

Figures 3a and 3b compare the anomalies of the ΔR_c^{tot} and aerosol-mediated cloud radiative response (ΔR_c^{aer}). In models with a more positive cloud feedback (T9; thick red line in Figure 3a), ΔR_c^{tot} actually exhibits a cooling effect in the historical simulations; that is, $\Delta R_c^{\text{tot}}<0$. Even more surprising, the models with the most positive cloud feedback (T9; thick red line) have a larger cloud-induced cooling effect than the models with a weaker cloud feedback (B9; thick blue line in Figure 3a). That is, the models with a more *positive* cloud feedback in response to CO_2 show a more *negative* cloud radiative response in historical simulations. This is particularly evident after 1950 (gray shading in Figure 3).

The negative ΔR_c^{tot} in T9 models arises almost entirely from a negative aerosol-mediated cloud radiative response (ΔR_c^{aer}; Figure 3b). In contrast, the B9 models (with weak cloud feedback) have a very small aerosol-mediated cloud radiative response. In other word, models with a more positive cloud feedback (T9) tend to have a more negative ΔR_c^{aer} compared to those with a weaker cloud feedback (B9). This occurs despite the B9 models having a more negative value in the TOA clear-sky shortwave forcing ($F_{\text{clr}}^{\text{sw}}$; Figure 3c), representing the shortwave clear-sky aerosol direct forcing (considering the GHGs have tiny effect on shortwave radiation and the solar irradiation change is small during the historical period). Since the difference between T9 and B9 models in aerosol direct forcing is much smaller than that in ACI, it indicates that, at least for the shortwave, the intermodel difference in the total aerosol forcing is dominated by ACI, not the aerosol direct forcing. The aerosol emissions, especially sulfur dioxide, are almost fixed after 1980, while the GHGs emissions continue to increase (Hoesly et al., 2018). Thus, ΔR_c^{tot} increases with increasing global-mean temperature in T9 models (red line in Figure 3a) after 2000, while ΔR_c^{aer} 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 Supporting Information) reproduce the main aspects of the T9/B9 composite results, indicating robustness relative to the details of model selection within this ensemble.

Figure 3b imply a compensation between the cloud feedback from CO_2-induced surface warming and the aerosol-mediated cloud radiative response. This anticorrelation 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 ΔR_c^{aer} from the historical simulations. Models with a more positive cloud feedback tend to have a more negative ΔR_c^{aer} ($r^2 = 0.60$). This helps to explain why models with a higher ECS also tend to have a larger net aerosol radiative cooling effect (Meehl et al., 2020).

The spatial pattern of the cloud radiative response in the historical experiment also differs from that obtained in the abrupt-4xCO2 experiment because of the regional imbalance in the aerosol emission during the historical period. 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 ΔR_c^{tot} compared to the preindustrial period (Figure 2e). Due to smaller global warming before
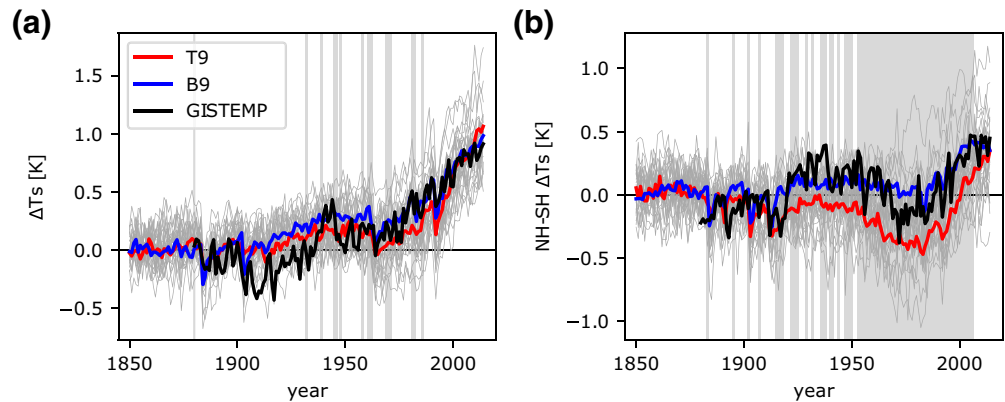
Figure 4. Modeled and observed surface temperature change. Annual time-series of (a) the global-mean surface temperature anomaly and (b) the interhemispheric contrast of surface temperature anomaly. The black line is from the GISTEMP and is rebased to match the model ensemble mean of the 1951–1980 period 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, $p < 0.05$).

1980s (Figure 4a), the surface temperature driven cloud radiative response is also small and ΔR_c^{aer} in the T9 models is responsible for the negative ΔR_c^{tot} in the northern midlatitude during 1950–2000 (Figure 2g). Also, because of the small ΔR_c^{tot} and ΔR_c^{aer} in B9 models, the differences in ΔR_c^{tot} between the T9 and B9 models are almost identical to the differences in ΔR_c^{aer} (Figures 2f and 2i). The intermodel differences in the spatial pattern of the ΔR_c^{aer} lead to a distinct warming pattern and can be useful to constrain the aerosol indirect effect with observation temperature.

3.3. Interhemispheric Warming Asymmetry

Due to the larger cooling effect of the ACI, 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 ECS. 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 during 1950–2000 (gray shading in Figure 4b).

The observed interhemispheric warming asymmetry over the 20th century is more consistent with the models with weaker cloud feedback and aerosol indirect effect (B9) than those with more positive cloud feedback and aerosol indirect effect (T9). Although the observed global- and annual-mean temperature anomalies are broadly consistent with both sets of models (Figure 4a), the B9 model ensemble mean of B9 more closely reproduces the observed hemispheric contrast in warming over most of the historical period (Figure 4b). The rank of the observed NH-SH temperature anomaly pooled from the B9 model ensemble produces an approximately uniformly distribution, but pooling from T9 model ensemble produces a skewed distribution, indicating that the B9 model ensemble is a more reliable representation (Figure S1).

4. Discussion

The seeming consistency of global-mean temperature evolution between more positive cloud feedback (high ECS) models and observations requires a strong aerosol indirect cooling effect that leads to an inter-hemispheric temperature evolution that is inconsistent with observations. Because of the strong negative correlation between a model's cloud feedback in response to CO_2 (and its CO_2-induced ECS) and its aerosol indirect effect (Figure 1b), the global-mean temperature evolutions in more positive and less positive cloud feedback models are not well separated over the historical period (Figure 4a) as both CO_2 and aerosol increase. Both more positive (high ECS) and less positive (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 ACI cooling effect 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 support the recent findings that the CMIP6 models 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).

Reproducing the observed global-mean temperature evolution over the 20th century is an important test for climate models. It seems unlikely that a model with a more positive cloud feedback and a weak ACI, 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). Dividing the 30 models into two groups based upon how well they simulate the observed global-mean historical warming (defined in Data and Methods) implies the potential of model tuning based on observed global-mean surface temperature changes in the process of model development as a cause of this compensation relationship. The correlation between the warming effects of cloud feedback and the cooling effect of ACI 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 do not ($r^2 = 0.17$, open circles in Figure 1b). 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), in which many of the used models are less consistent with the observed global-mean surface warming, 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). In CMIP5 models, Forster et al. (2013) also found a significant intermodel correlation between the total aerosol forcing and ECS for the models that simulate the historical warming well.

An alternative interpretation of the intermodel correlation between ECS and ACI, is that there could be a physical process that is intrinsic to models that links cloud feedback and ACI. For example, models with a more positive cloud feedback produce a very different temperature change pattern than those 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; Wang et al., 2016), which might lead to differences in aerosol transport and lifetime, or to different changes in cloud regime. Additionally, both ACI and cloud feedbacks can be affected by the mean cloud field, and it is possible that a mean state that is advantageous to larger positive cloud feedbacks also makes large ACI more likely—so that mean cloud biases could impact ACI and cloud feedback, and lead to the intermodel correlation. However, we have not been able to find evidence for this interpretation in our analyses.

The differences in the spatial pattern of warming induced by strong ACI 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 (Deser et al., 2020; 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; Villarini & Vecchi, 2012, 2013; 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 ACI (Stevens, 2015) as well as shifts in climatic quantities of more direct societal relevance, such as tropical cyclones and rainfall.

Data Availability Statement

All the CMIP6 model output is available at https://esgf-node.llnl.gov/projects/cmip6/. The GISS Surface Temperature Analysis version 4 (GISTEMP v4) is available at https://data.giss.nasa.gov/gistemp/.

References

Allen, R. J. (2015). A 21st century northward tropical precipitation shift caused by future anthropogenic aerosol reductions. *Journal of Geophysical Research: Atmospheres*, *120*(18), 9087–9102. https://doi.org/10.1002/2015JD023623

Allen, R. J., Evan, A. T., & Booth, B. B. B. (2015). Interhemispheric aerosol radiative forcing and tropical precipitation shifts during the late Twentieth Century. *Journal of Climate*, *28*(20), 8219–8246. https://doi.org/10.1175/JCLI-D-15-0148.1

Andrews, T., & Forster, P. M. (2008). CO_2 forcing induces semi-direct effects with consequences for climate feedback interpretations. *Geophysical Research Letters*, *35*(4), L04802. https://doi.org/10.1029/2007GL032273

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, *39*(9), 1–7. https://doi.org/10.1029/2012GL051607

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, *26*(13), 4518–4534. https://doi.org/10.1175/JCLI-D-12-00544.1

Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., & Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, *484*(7393), 228–232. https://doi.org/10.1038/nature10946

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., et al. (2013). Clouds and aerosols. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Doschung, et al. (Eds.), *Climate change 2013: The physical science basis*. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 571–658). Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.016

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics Discussions*, *6*, 1–23. https://doi.org/10.5194/esd-2020-23

Ceppi, P., Brient, F., Zelinka, M. D., & Hartmann, D. L. (2017). Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdisciplinary Reviews: Climate Change*, *8*(4), e465. https://doi.org/10.1002/wcc.465

Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Del Genio, A. D., Déqué, M., et al. (1990). Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *Journal of Geophysical Research*, *95*(D10), 16601. https://doi.org/10.1029/JD095iD10p16601

Chung, E. S., & Soden, B. J. (2015). An assessment of direct radiative forcing, radiative adjustments, and radiative feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, *28*(10), 4152–4170. https://doi.org/10.1175/JCLI-D-14-00436.1

Chung, E. S., & Soden, B. J. (2017). Hemispheric climate shifts driven by anthropogenic aerosol-cloud interactions. *Nature Geoscience*, *10*(8), 566–571. https://doi.org/10.1038/NGEO2988

Colman, R. (2003). A comparison of climate feedbacks in general circulation models. *Climate Dynamics*, *20*(7–8), 865–873. https://doi.org/10.1007/s00382-003-0310-z

Colman, R., & McAvaney, B. J. (2011). On tropospheric adjustment to forcing and climate feedbacks. *Climate Dynamics*, *36*(9–10), 1649–1658. https://doi.org/10.1007/s00382-011-1067-4

Deser, C., Phillips, A. S., Simpson, I. R., Rosenbloom, N., Coleman, D., Lehner, F., et al. (2020). Isolating the evolving contributions of anthropogenic aerosols and greenhouse gases: A new CESM1 large ensemble community resource. *Journal of Climate*, *33*(18), 7835–7858. https://doi.org/10.1175/JCLI-D-20-0123.1

Dufresne, J. L., & Bony, S. (2008). An assessment of the primary sources of spread of global warming estimates from coupled atmosphere-ocean models. *Journal of Climate*, *21*(19), 5135–5144. https://doi.org/10.1175/2008JCLI2239.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2014). Evaluation of climate models. In Intergovernmental Panel on Climate Change. (Ed.), *Climate change 2013: The physical science basis* (pp. 741–866). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.020

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research: Atmospheres*, *118*(3), 1139–1150. https://doi.org/10.1002/jgrd.50174

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., et al. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3), L03205. https://doi.org/10.1029/2003GL018747

Gregory, J. M., & Webb, M. (2008). Tropospheric adjustment induces a cloud component in CO2 forcing. *Journal of Climate*, *21*(1), 58–71. https://doi.org/10.1175/2007JCLI1834.1

Hartmann, D. L., Ockert-Bell, M. E., & Michelsen, M. L. (1992). The effect of cloud type on Earth's energy balance: Global analysis. *Journal of Climate*, *5*(11), 1281–1304. https://doi.org/10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., et al. (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*(1), 369–408. https://doi.org/10.5194/gmd-11-369-2018

Hwang, Y. T., Frierson, D. M. W., & Kang, S. M. (2013). Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century. *Geophysical Research Letters*, *40*(11), 2845–2850. https://doi.org/10.1002/grl.50502

Jacobson, T. W. P., Yang, W., Vecchi, G. A., & Horowitz, L. W. (2020). Impact of volcanic aerosol hemispheric symmetry on Sahel rainfall. *Climate Dynamics*, *55*(7–8), 1733–1758. https://doi.org/10.1007/s00382-020-05347-7

Kang, S. M., Held, I. M., Frierson, D. M. W., & Zhao, M. (2008). The response of the ITCZ to extratropical thermal forcing: Idealized slab-ocean experiments with a GCM. *Journal of Climate*, *21*(14), 3521–3532. https://doi.org/10.1175/2007JCLI2146.1

Kang, S. M., Held, I. M., & Xie, S. P. (2014). Contrasting the tropical responses to zonally asymmetric extratropical and tropical thermal forcing. *Climate Dynamics*, *42*(7–8), 2033–2043. https://doi.org/10.1007/s00382-013-1863-0

Kiehl, J. T. (2007). Twentieth century climate model response and climate sensitivity. *Geophysical Research Letters*, *34*(22), 1–4. https://doi.org/10.1029/2007GL031383

Knutti, R. (2008). Why are climate models reproducing the observed global surface warming so well? *Geophysical Research Letters*, *35*(18), 1–5. https://doi.org/10.1029/2008GL034932

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP Uncertainty Model. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6307–6326. https://doi.org/10.1029/2018JD029522

Lohmann, U., Rotstayn, L., Storelvmo, T., Jones, A., Menon, S., Quaas, J., et al. (2010). Total aerosol effect: Radiative forcing or radiative flux perturbation? *Atmospheric Chemistry and Physics*, *10*(7), 3235–3246. https://doi.org/10.5194/acp-10-3235-2010

Mauritsen, T., & Roeckner, E. (2020). Tuning the MPI-ESM1.2 global climate model to improve the match with instrumental record warming by lowering its climate sensitivity. *Journal of Advances in Modeling Earth Systems*, *12*(5), e2019MS002037. https://doi.org/10.1029/2019MS002037

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., et al. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981. https://doi.org/10.1126/sciadv.aba1981

Merlis, T. M., Zhao, M., & Held, I. M. (2013). The sensitivity of hurricane frequency to ITCZ changes and radiatively forced warming in aquaplanet simulations. *Geophysical Research Letters*, *40*(15), 4109–4114. https://doi.org/10.1002/grl.50680

Ming, Y., & Ramaswamy, V. (2011). A model investigation of Aerosol-Induced changes in tropical circulation. *Journal of Climate*, *24*(19), 5125–5133. https://doi.org/10.1175/2011JCLI4108.1

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., et al. (2013). Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. *Atmospheric Chemistry and Physics*, *13*(4), 1853–1877. https://doi.org/10.5194/acp-13-1853-2013

Penner, J. E., Dickinson, R. E., & O'Neill, C. A. (1992). Effects of aerosol from biomass burning on the global radiation budget. *Science*, *256*(5062), 1432–1434. https://doi.org/10.1126/science.256.5062.1432

Ramanathan, A. V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., et al. (1989). Cloud-radiative forcing and climate: Results from the earth radiation budget experiment linked references are available on JSTOR for this article: Cloud-radiative forcing and climate: Results from the Earth radiation budget experiment. *Science*, *243*(4887), 57–63.

Rotstayn, L. D., & Penner, J. E. (2001). Indirect aerosol forcing, quasi forcing, and climate response. *Journal of Climate*, *14*(13), 2960–2975. https://doi.org/10.1175/1520-0442(2001)014<2960:IAFQFA>2.0.CO;2

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., et al. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, *10*(9), 3207–3223. https://doi.org/10.5194/gmd-10-3207-2017

Smith, C., Kramer, R., Myhre, G., Alterskjær, K., Collins, W., Sima, A., et al. (2020). Effective radiative forcing and adjustments in CMIP6 models. *Atmospheric Chemistry and Physics*, *20*(16), 9591–9618. https://doi.org/10.5194/acp-20-9591-2020

Soden, B. J., & Chung, E. S. (2017). The large-scale dynamical response of clouds to aerosol forcing. *Journal of Climate*, *30*(21), 8783–8794. https://doi.org/10.1175/JCLI-D-17-0050.1

Soden, B. J., & Held, I. M. (2006). An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate*, *19*(14), 3354–3360. https://doi.org/10.1175/JCLI3799.1

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, *21*(14), 3504–3520. https://doi.org/10.1175/2007JCLI2110.1

Stevens, B. (2015). Rethinking the lower bound on aerosol radiative forcing. *Journal of Climate*, *28*(12), 4794–4819. https://doi.org/10.1175/jcli-d-14-00656.1

Twomey, S. (1977). The influence of pollution on the shortwave albedo of clouds. *Journal of the Atmospheric Sciences*, *34*(7), 1149–1152. https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2

Vecchi, G. A., & Soden, B. J. (2007). Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature*, *450*(7172), 1066–1070. https://doi.org/10.1038/nature06423

Vial, J., Dufresne, J. L., & Bony, S. (2013). On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dynamics*, *41*(11–12), 3339–3362. https://doi.org/10.1007/s00382-013-1725-9

Villarini, G., & Vecchi, G. A. (2012). Twenty-first-century projections of North Atlantic tropical storms from CMIP5 models. *Nature Climate Change*, *2*(8), 604–607. https://doi.org/10.1038/nclimate1530

Villarini, G., & Vecchi, G. A. (2013). Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. *Journal of Climate*, *26*(10), 3231–3240. https://doi.org/10.1175/JCLI-D-12-00441.1

Wang, H., Xie, S. P., & Liu, Q. (2016). Comparison of climate response to anthropogenic aerosol versus greenhouse gas forcing: Distinct patterns. *Journal of Climate*, *29*(14), 5175–5188. https://doi.org/10.1175/JCLI-D-16-0106.1

Webb, M. J., Lambert, F. H., & Gregory, J. M. (2013). Origins of differences in climate sensitivity, forcing and feedback in climate models. *Climate Dynamics*, *40*(3–4), 677–707. https://doi.org/10.1007/s00382-012-1336-x

Wetherald, R. T., & Manabe, S. (1988). Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, *45*(8), 1397–1416. https://doi.org/10.1175/1520-0469(1988)045<1397:CFPIAG>2.0.CO;2

Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., et al. (2020). Climate sensitivity of GFDL's CM4.0. *Journal of Advances in Modeling Earth Systems*, *12*(1), 1–17. https://doi.org/10.1029/2019MS001838

Xie, S. P., Deser, C., Vecchi, G. A., Ma, J., Teng, H., & Wittenberg, A. T. (2010). Global warming pattern formation: Sea surface temperature and rainfall. *Journal of Climate*, *23*(4), 966–986. https://doi.org/10.1175/2009JCLI3329.1

Yang, W., Vecchi, G. A., Fueglistaler, S., Horowitz, L. W., Luet, D. J., Muñoz, Á. G., et al. (2019). Climate impacts from large volcanic eruptions in a high-resolution climate model: The importance of forcing structure. *Geophysical Research Letters*, *46*(13), 7690–7699. https://doi.org/10.1029/2019GL082367

Zelinka, M. D., Andrews, T., Forster, P. M., & Taylor, K. E. (2014). Quantifying components of aerosol-cloud-radiation interactions in climate models. *Journal of Geophysical Research: Atmospheres*, *119*(12), 7599–7615. https://doi.org/10.1002/2014JD021710

Zelinka, M. D., Klein, S. A., Taylor, K. E., Andrews, T., Webb, M. J., Gregory, J. M., & Forster, P. M. (2013). Contributions of different cloud types to feedbacks and rapid adjustments in CMIP5. *Journal of Climate*, *26*(14), 5007–5027. https://doi.org/10.1175/JCLI-D-12-00555.1

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), 1–12. https://doi.org/10.1029/2019GL085782

Zelinka, M. D., Zhou, C., & Klein, S. A. (2016). Insights from a refined decomposition of cloud feedbacks. *Geophysical Research Letters*, *43*(17), 9259–9269. https://doi.org/10.1002/2016GL069917

Zhang, R., Delworth, T. L., Dixon, K. W., Held, I. M., Ming, Y., Msadek, R., et al. (2013). Have aerosols caused the observed atlantic multidecadal variability? *Journal of the Atmospheric Sciences*, *70*(4), 1135–1144. https://doi.org/10.1175/JAS-D-12-0331.1