EVALUATING THE EFFICACY OF MODELS IN PREDICTING CLOUD FEEDBACK PATTERNS

A Thesis

by

JACOB CHARLES MULLER

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Andrew Dessler
Committee Members,	Ping Yang
	Ping Chang
Head of Department,	R. Saravanan

December 2020

Major Subject: Atmospheric Science

ABSTRACT

Historical estimates of equilibrium climate sensitivity (ECS) potentially underestimate the true value of our climate system's ECS due to the fact that the warming pattern experienced over the last century is different from the equilibrium pattern of warming. In quantifying this so-called pattern effect, cloud feedbacks were identified as an important contributor. In particular, the ability of models to accurately resolve cloud feedback patterns and their relationship with surface temperature patterns is paramount to the process of properly estimating ECS. To determine the ability of models to predict how cloud feedbacks respond to these patterns, cloud feedbacks were calculated from three model ensembles and compared to observed cloud feedbacks. Similar relationships between cloud feedbacks and surface temperature were found in the models and observations, lending credence to the theory that understanding this pattern effect could likely contribute to reducing the uncertainty of ECS calculations.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Andrew Dessler, and my my committee members, Dr. Ping Yang and Dr. Ping Chang, for their guidance and support throughout the course of this research.

Thanks also go to my colleague Li Wei Chao for her assistance in obtaining and processing the data used during this study.

Finally, thank you to my family for their encouragement and support throughout my academic career.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professors Andrew Dessler and Ping Yang of the Department of Atmospheric Science and Professor Ping Chang of the Department of Oceanography.

The data analyzed in this study was provided by Professor Andrew Dessler and Li Wei Chao.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

This work was also made possible in part by the National Science Foundation under Grant Number AGS-1661861. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

INTRODUCTION

Climate change is, loosely defined, the long-term change in the average conditions of Earth's atmosphere, most commonly expressed in terms of temperature. Significant variations in Earth's climate are often used as important markers in Earth's history, with phenomena like the Paleocene-Eocene Thermal Maximum and various glacial maxima and minima marking important trends in paleoclimate. As such, understanding both the tendencies of the climate system and the forces driving its changes is vital to predicting how the climate will behave in the future.

In its Fifth Assessment Report (AR5), the Intergovernmental Panel on Climate Change (IPCC) reported that the three previous decades were warmer at Earth's surface than any decade since 1850. Furthermore, the 1983-2012 period was very likely the warmest 30-year period of the last 800 years in the Northern Hemisphere. Such drastic records indicate a significant shift in the climate system as a whole, as global average temperature is not a rapidly changing value. For example, the IPCC AR5 reports a global warming of 0.85°C [90% confidence interval: 0.65°C – 1.06°C] between 1880 and 2012, while it is estimated that the difference between modern July surface temperatures and ice-age July surface temperatures of 18,000 years ago is 4.9°C (Gates, 1976). In context, warming nearing 1°C over the course of a century is much more daunting when considering the much longer time scale at which the planet warmed from ice-age temperatures to current temperatures.

As temperature changes, other things will also change. In the AR5, the IPCC reports high confidence that rates of sea level rise since the mid-19th century have been greater than that of the previous two millennia, citing a global mean sea level rise of 0.19 m [90% confidence

interval: 0.17 m – 0.21 m] over the 1901 – 2010 period. While this trend is regionally dependent – the rate of rise is roughly thrice the mean in the Western Pacific but near zero in the Eastern Pacific – this leaves many low-lying coastal areas vulnerable to flooding and eventual submersion if sea levels continue to rise at their current rate.

Climate change can be forced by natural or anthropogenic changes, or it can be the result of internal, unforced variability. According to the AR5, anthropogenic forcing includes increases in atmospheric concentrations of greenhouse gases like CO_2, CH_4, and N_2O, which trap outgoing radiation within the atmosphere and lead to overall warming. The most notable of these gases is CO_2, which increased at the fastest observed decadal rate of change on record over the 2002 – 2011 period. The total anthropogenic radiative forcing has increased more rapidly since 1970 than during prior decades, lending credence to the theory that the current warming of Earth's climate is beyond the scope of internal variability.

Of course, not all atmospheric forcings are anthropogenic. Most natural forcings are caused by changes in volcanic activity and incoming solar radiation, both of which are beyond the scope of human interference. However, changes in solar irradiance only accounted for 2% of total radiative forcing in 2011 relative to 1750, and the impacts of volcanic aerosols tend to be mitigated naturally within a few years of a major eruption.

In order to predict future changes in climate, the IPCC formulated four Representative Concentration Pathways (RCPs) that correspond to different responses to climate change in the form of limiting or not limiting current emissions of greenhouse gases and other air pollutants. RCP2.6 is the most optimistic of these scenarios and involves aggressive and stringent emissions mitigation, RCP8.5 is a more dire estimate of climate change resulting from increasing emissions, while RCP4.5 and RCP6.0 are intermediate scenarios. Most estimates of a "business

as usual" approach to emissions with no significant efforts to reduce emissions tend to fall between RCP6.0 and RCP8.5 in terms of long-term effects on Earth's climate.

Table 1, taken from the Future Climate Changes, Risk and Impact section of IPCC's AR5, shows the projected changes in global mean temperature according to these four RCPs. The optimistic result of RCP2.6 shows global temperatures likely staying less than 2°C greater than pre-industrial values, while the worst-case RCP8.5 scenario could lead to warming over a 250-year period of the same magnitude as that of the 18,000-year period since the last ice age. Such a drastic increase in temperature would have irreversible effects on countless aspects of Earth's climate system.

Table 1: Predicted increase in global mean surface temperature change relative to the 1985 – 2005 period according to the four RCP scenarios of the IPCC Fifth Assessment Report. Reprinted from IPCC 2014.

		2046 – 2065		**2081 – 2100**	
	Scenario	**Mean**	**5 to 95% confidence interval**	**Mean**	**5 to 95% confidence interval**
Global Mean Surface Temperature Change (°C)	RCP2.6	1.0	0.4 to 1.6	1.0	0.3 to 1.7
	RCP4.5	1.4	0.9 to 2.0	1.8	1.1 to 2.6
	RCP6.0	1.3	0.8 to 1.8	2.2	1.4 to 3.1
	RCP8.5	2.0	1.4 to 2.6	3.7	2.6 to 4.8

The response of the climate systems to changes in forcing is the combination of a direct warming and a subsequent response to that initial temperature change known as climate feedbacks; these feedbacks serve to either amplify (positive feedback) or dampen (negative feedback) the direct warming from the forcing. Most forcings are positive and amplify an initial temperature perturbation. Examples of atmospheric feedbacks are the lapse rate, water vapor, albedo, and cloud feedbacks.

Equilibrium climate sensitivity (ECS) has long been a metric used to determine the extent of anthropogenic climate change given the direct relationship it posits between increasing concentrations of CO_2 and increasing temperatures. Specifically, the following equations have been employed in many studies (for example, Gregory et al., 2002; Otto et al., 2013, Lewis and Curry, 2015; Forster, 2016) to estimate the feedback parameter λ and ECS:

$$\lambda = \Delta(R - F)/\Delta T_s \tag{1}$$

$$ECS = -\frac{F_{2\times CO_2}}{\lambda} \tag{2}$$

where Δ indicates the change between over some time period of interest, R is net top-of-atmosphere (TOA) flux, F is the forcing, an imposed energy imbalance, T_s is global average surface temperature, λ is the feedback parameter, which relates temperature to a change in net TOA flux, and $F_{2\times CO_2}$ is the radiative forcing resulting from a doubling of CO_2 concentration. Note that λ can be decomposed into the sum of individual λ values for each feedback.

Many analyses have used Equations 1 and 2 to estimate ECS using data obtained between the mid to late 19th century and present day. However, it has been speculated that these estimates may underestimate the true value of ECS (Gregory and Andrews, 2016; Andrews et al., 2018; Marvel et al., 2018; Dessler et al., 2018) because the pattern of warming we have experienced is different from the equilibrium pattern of warming.

There are two sources to this so-called pattern effect: one due to internal variability within the Earth's climate system and one resulting from a difference in average transient pattern of warming and the pattern of warming at equilibrium after a doubling of CO_2 (Dessler, 2020).

While the warming pattern resulting from a doubling of CO_2 is forced, the internal variability within the Earth's climate system is unforced and can be connected to variability in sea ice loss and tropospheric warming. This paper will focus specifically on observationally quantifying the unforced pattern effect in models and comparing that to observations.

In quantifying the pattern effect, we will focus on one component of the total λ: the cloud feedback. Known to be one of the biggest contributors to uncertainty in ECS measurements (Soden et al., 2004; Dufresne and Bony, 2008), it is also one of the big contributors to the unforced pattern effect (Zhou et al., 2016; Zhou et al., 2017; Dessler, 2020). As such, understanding the spatial patterns of cloud feedbacks and their relationship to the spatial patterns of warming is the primary goal of this study. Specifically, we wanted to evaluate the ability of models to predict cloud feedback patterns and their relationship with surface temperature patterns. To do this, three model ensembles were used to calculate three separate sets of cloud feedbacks, and these data were compared to feedbacks calculated from observation and reanalysis data.

DATA

For observations, monthly data were collected from three sources in order to calculate cloud feedbacks. Observed top-of-atmosphere (TOA) fluxes were obtained from NASA's Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) top-of-atmosphere (TOA) edition 4.1 satellite data set (Loeb et al., 2018; Kato et al., 2018). European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5th Generation (ERA5) provided relative humidity, 2m temperature, and albedo data (Hersbach et al., 2020). Forcing was taken from Dessler and Forster (2018), which is an updated version of the Intergovernmental Panel on Climate Change (IPCC) AR5 forcing. Our analysis covers March 2000 to October 2017, a time period determined by the availability of the forcing data.

Three model ensembles were examined over the course of this study. The first ensemble is the Max Planck Institute (MPI) Grand Ensemble (Maher et al., 2019). This single-model ensemble uses the MPI Earth System Model version 1.1 (MPI-ESM1.1), a couple model consisting of the ECHAM6.3 atmosphere and land model and the MPI-OM ocean model (Maher et al., 2019). The data used in this study are the 100-member ensemble of historical runs, spanning 1850 – 2005 and differing between ensemble members only in initial conditions.

The final two ensembles are the pre-industrial control (piControl) runs of Phases 5 and 6 of the Coupled Model Intercomparison Project (CMIP), CMIP5 (Taylor et al., 2012) and CMIP6 (Eyring et al., 2016). These multi-model ensembles contained 26 and 27 models for CMIP5 and CMIP6, respectively. Prior to any calculations, the spatial resolution of MPI, CMIP5, and CMIP6 data was resampled to 45 latitude points by 90 longitude points.

METHODS

Feedback calculation

The first step is to estimate the cloud feedback λ_{cloud} in observations and model runs. To do this, we first estimate ΔR_{cloud}, the flux anomaly due to changes in clouds, using the adjusted-CRF method described by Soden et al. (2008) and Shell et al. (2008) and radiative kernels developed by Huang et al. (2017).

For MPI data, λ_{cloud} is calculated for each of the 100 ensemble members using monthly data from the 1996 – 2005 period. For each ensemble member, λ_{cloud} is estimated as the slope of the regression ΔR_{cloud} vs ΔT_s. For CMIP5 and CMIP6, each control run was split into 10-year periods and λ_{cloud} over each period was also estimated as the slope of the regression ΔR_{cloud} vs ΔT_s. For the observations, it was assumed that all changes in forcing over this period were in the longwave, so forcing was omitted from the shortwave cloud feedback calculation. As with the models, the feedback is estimated as the slope of the regression ΔR_{cloud} vs ΔT_s. In addition to the total cloud feedback, we have also calculated shortwave and longwave components.

RESULTS

Figure 1 shows maps of λ_{cloud} from the observations as well as the ensemble average from the MPI grand ensemble over the years 1996 – 2005 and averages of the 10-year periods of the CMIP5 and CMIP6 control runs. There is good correspondence between the fields, particularly for the CMIP6 models, suggesting the models have some skill in reproducing the cloud feedback in response to unforced variability. Differences do exist, however: all of the models predict strong positive cloud feedback in the Eastern Pacific. Dessler (2013) also identified this disagreement between models and observations, although the disagreement was even larger in that paper.

It is important to realize that this is not an apples-to-apples comparison. The observations cover a single 17-year period, while the models are the average of many individual periods. Because of this, internal variability has been averaged out in the models, while it still exists in the observations. To show this more clearly, Figure 2 shows a histogram of the global average λ_{cloud} from the MPI ensemble, CMIP5, and CMIP6 from the individual 10-year segments. Also shown is the 17-year global average λ_{cloud} from observed data, shown as a solid vertical line.

All three ensembles show that internal variability can lead to significant variations in λ_{cloud} between various periods, owing to different sea surface temperature (SST) patterns. The average and standard deviations of the MPI, CMIP5, and CMIP6 and distributions are 0.17±0.33 W/m^2/K, 0.56±0.58 W/m^2/K, and 0.38±0.63 W/m^2/K, respectively. The observed cloud feedback from the full period, 0.70±0.45 W/m^2/K, falls within each ensemble's distribution. This spread in the model ensembles is a measure of the unforced pattern effect and previous work has identified

processes such as differential warming of the East and West Pacific (Zhou et al., 2016; Andrews and Webb, 2018) that can drive this variability.

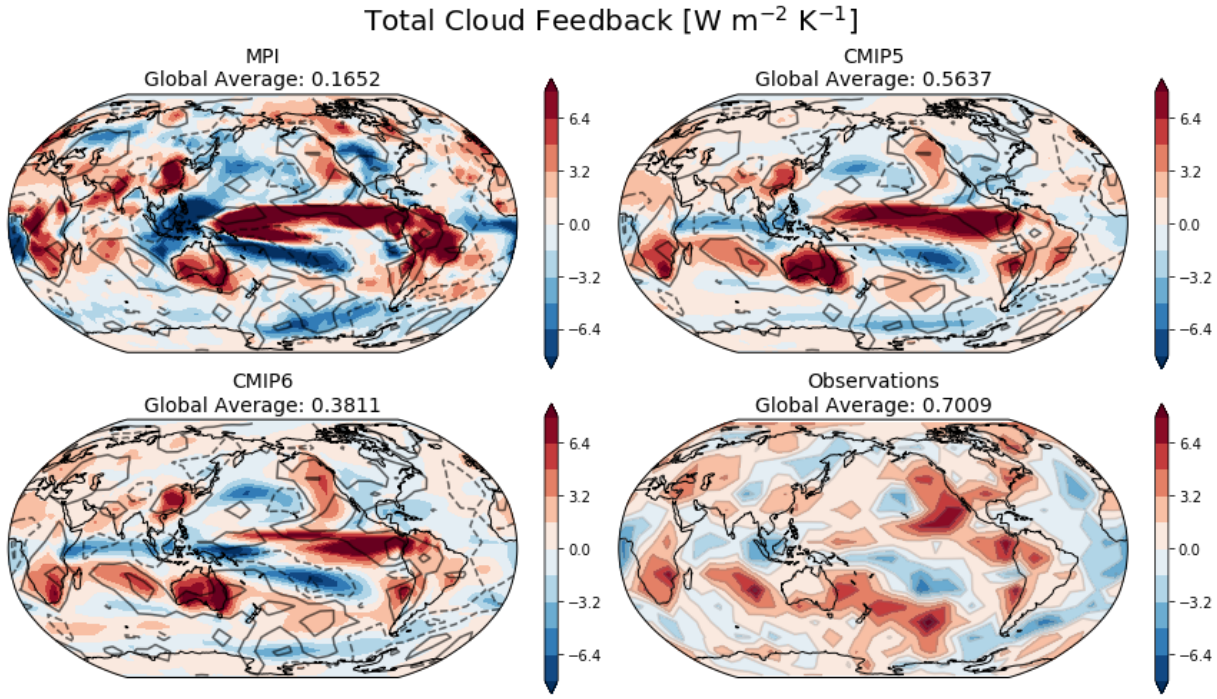

Figure 1: Total cloud feedback over the entire period for MPI, CMIP5 piControl, CMIP6 piControl, and observed data. MPI feedback is the average across all ensemble members from 1996 to 2005. CMIP5 and CMIP6 feedbacks are the average of all 10-year periods and all models. Observations are a combination of CERES data, ERA5 reanalysis, and IPCC forcings from March 2000 to October 2017. Dark contours over model plots show the feedback estimated from observations in the bottom right panel.

Differences between periods with high and low cloud feedback parameters

To better analyze the spread in λ_{cloud}, we analyze the unforced pattern effect in λ_{cloud} for the observed data. The March 2000 – October 2017 period was split into two distinct periods and λ_{cloud} was calculated for each: March 2000 – March 2010 and April 2010 – October 2017. The date at which the period was split was chosen such that the difference between the feedback parameter in each period was greatest while keeping the length of the time periods similar. This approach of splitting the CERES into two periods in order to understand internal variability was also done by Loeb et al. (2020), although they used different periods.

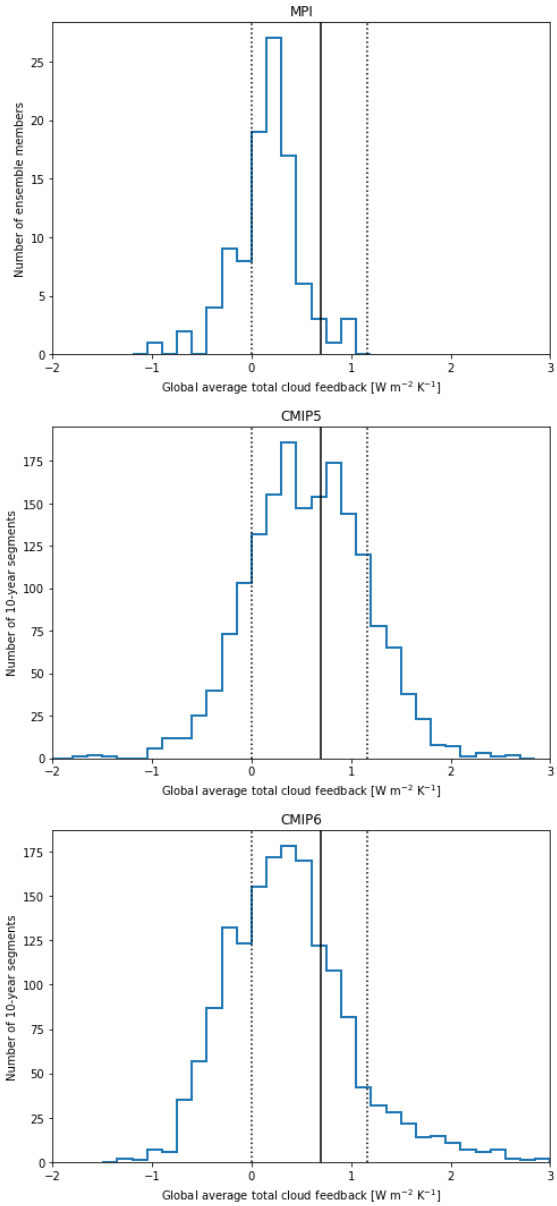

Figure 2: Histograms of global average total cloud feedback for MPI, CMIP5 piControl, and CMIP6 piControl. The MPI histogram represents all ensemble members across the 1996 – 2005 period. The CMIP5 and CMIP6 histograms represent all 10-year segments within either ensemble. The solid vertical line is the observed total cloud feedback across the entire March 2000 – October 2017 period, while the dotted vertical lines are the observed global average total cloud feedbacks for the split periods; the left-most dotted vertical line represents March 2000 – March 2010, and the right-most dotted vertical line represents April 2010 – October 2017.

λ_{cloud} values for these sub-periods are very different: 0.0042±1.00 W/m^2/K for Period 1 and 1.16±1.16 W/m^2/K (also plotted in Figure 2). This confirms the results of the model ensembles that internal variability can drive large changes in the cloud feedback.

We now estimate the pattern effect in the ensembles. To estimate pattern effect in the MPI ensemble, ensemble members with the 10 highest and 10 lowest λ_{cloud} during the 1996 – 2005 period were identified. The low λ_{cloud} members were averaged and then subtracted from the

average of the high λ_{cloud} members. We refer to this difference as $\Delta\lambda_{cloud}$, and it is one estimate of the pattern effect in this ensemble.

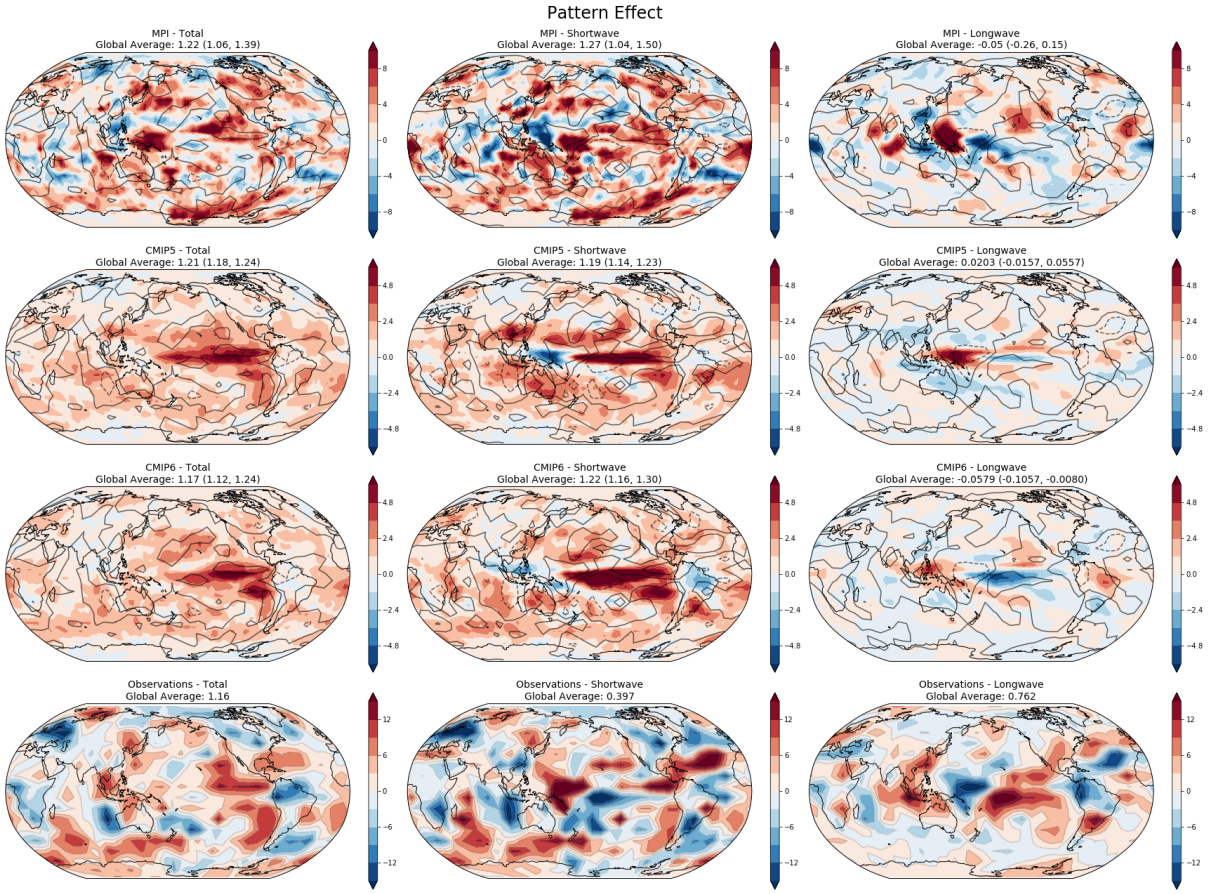

Figure 3: Difference between total, shortwave, and longwave cloud feedbacks for high and low segments within MPI, CMIP5 piControl, CMIP6 piControl, and observed data. MPI differences were found by taking the difference between the average of the ensemble members with the 10 highest and 10 lowest global average total cloud feedbacks across the 1996 – 2005 period. CMIP5 and CMIP6 differences were found by taking the difference, for each model, between the average of the 10-year periods with the 5 highest and 5 lowest global average total feedbacks from 1850 to 2017, then averaging across all models. Differences in observations were found by splitting the full period (March 2000 – October 2017) into two distinct periods (March 2000 – March 2010, April 2010 – October 2017) and taking the difference between the two periods. The date at which to split the two periods was chosen by finding the split date that led to the greatest difference between the two periods while maintaining similar time period lengths. The numbers over each panel represent the global average and, in parentheses, 5th and 95th percentiles calculated via bootstrap resampling. Dark contours over model plots represent observed data.

For each model within CMIP5 and CMIP6, the 10-year segments with the 10 highest and 10 lowest global average λ_{cloud} were identified. These high and low λ_{cloud} values were then

averaged and differenced, yielding an estimate of $\Delta\lambda_{cloud}$ for each model, and then the $\Delta\lambda_{cloud}$ for each model was averaged to obtain an ensemble average of $\Delta\lambda_{cloud}$.

Figure 3 shows plots of $\Delta\lambda_{cloud}$ for the three ensembles, including the shortwave (SW) and longwave (LW) components. The total $\Delta\lambda_{cloud}$ for each ensemble shows a general spatial pattern similar to that of the observations, specifically including large values of $\Delta\lambda_{cloud}$ over the equatorial Eastern Pacific. This shows that the models agree, generally at least, with the pattern of observed $\Delta\lambda_{cloud}$.

Shortwave $\Delta\lambda_{cloud}$ spatial patterns appear to largely resemble those of total $\Delta\lambda_{cloud}$ spatial patterns in all datasets, showing that the unforced pattern effect is mainly a shortwave phenomenon.

The areas with the strongest longwave $\Delta\lambda_{cloud}$ tend to be co-located with areas of shortwave $\Delta\lambda_{cloud}$ of similar magnitude but opposite sign. This is a manifestation of the well-known cancellation between LW and SW effects of high clouds (Kiehl, 1994; Zhou et al., 2013). This results in a total feedback of low magnitude in these regions of cancellation and is found in similar regions across all three model ensembles, such as the area near the northeast coast of Australia and equatorial Western Pacific.

Examining differences between periods with high and low cloud feedback parameters that correlate highly with observations

While the comparison in Figure 3 shows that the ensembles produce results with aspects similar to the observations, there are also clear differences. Loeb et al. (2020) showed that driving the models the observed SST pattern produces good agreement between CERES and a set of models, suggesting that differences in Figure 3 between the observations and ensemble

estimates of $\Delta\lambda_{cloud}$ are due to the coupled models not producing the exact difference in SST patterns in the observations.

Such a difference could arise for two reasons. First, it could be that the coupled models do not accurately simulate unforced variability in SST. Second, it could be that the periods covered by the observations are statistically unlikely, and therefore ensembles we have are sampling a bigger phase space of variability. Given that, we now perform other tests to ascertain how accurately the models reproduce the observations of $\Delta\lambda_{cloud}$.

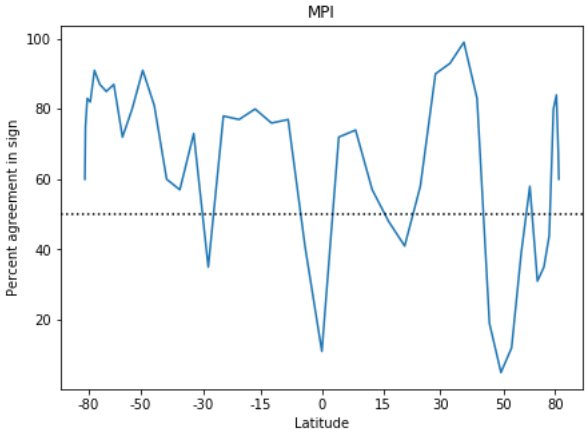

Figure 4: Percent of combinations of high and low MPI ensemble members with the same sign of zonal average total cloud feedback as in the observations at each latitude.

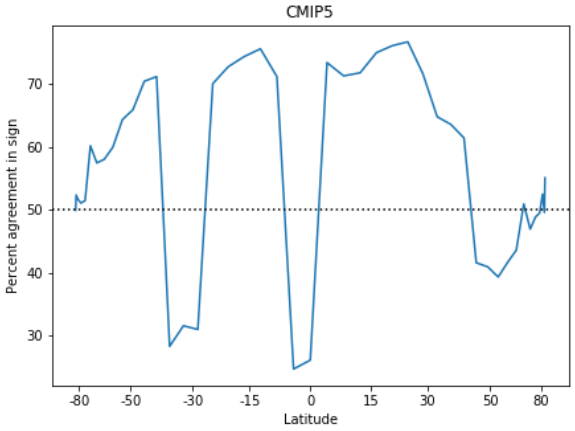

Figure 5: Percent of combinations of high and low 10-year segments within CMIP5 piControl with the same sign of zonal average total cloud feedback as in the observations at each latitude, averaged across all models.

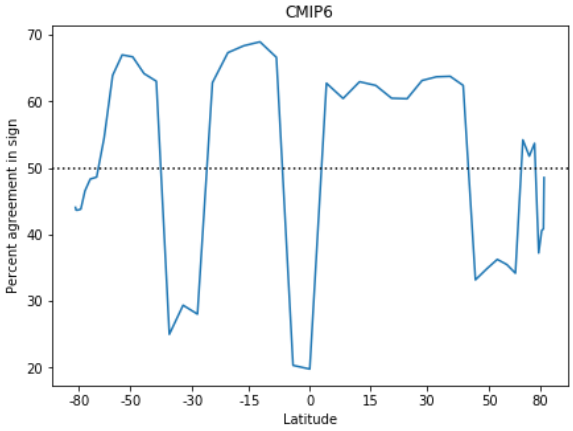

Figure 6: Percent of combinations of high and low 10-year segments within CMIP6 piControl with the same sign of zonal average total cloud feedback as in the observations at each latitude, averaged across all models.

For MPI, we can create 100 possible pairs of the ensemble members from the 10 highest and 10 lowest global average cloud feedback parameters across the 1996 – 2005 period. For each pair, we then subtract the low member from the high member. Figure 4 plots the fraction of these 100 differences that has the same *sign* of $\Delta\lambda_{cloud}$ at each latitude. At most latitudes, 70-80% of the pairs have the same sign. This result suggests that ensemble pairs with the largest $\Delta\lambda_{cloud}$ have some similarity to the observed difference.

Similar plots for CMIP5 and CMIP6 are shown in Figures 5 and 6, respectively. Much like MPI, there are multiple regions where the pairs of 10-year segments with the largest $\Delta\lambda_{cloud}$ tend to agree in sign with observed $\Delta\lambda_{cloud}$ more often than not. Of note is that these regions appear to overlap almost perfectly between CMIP5 and CMIP6 with some deviation in MPI. This implies that there may be confounding processes occurring in the regions where the signs of predicted $\Delta\lambda_{cloud}$ and observed $\Delta\lambda_{cloud}$ disagree that cause this method of evaluation to be inaccurate.

For another test, the zonal average of each of the 100 difference pairs from the MPI ensemble was regressed against the zonal average of the observed $\Delta\lambda_{cloud}$, and the 10 difference pairs with the highest correlation were averaged. Figure 7 shows the average of these 10 pairs.

This tests whether the model is capable of producing the observed pattern of $\Delta\lambda_{cloud}$, and the answer is that it can.

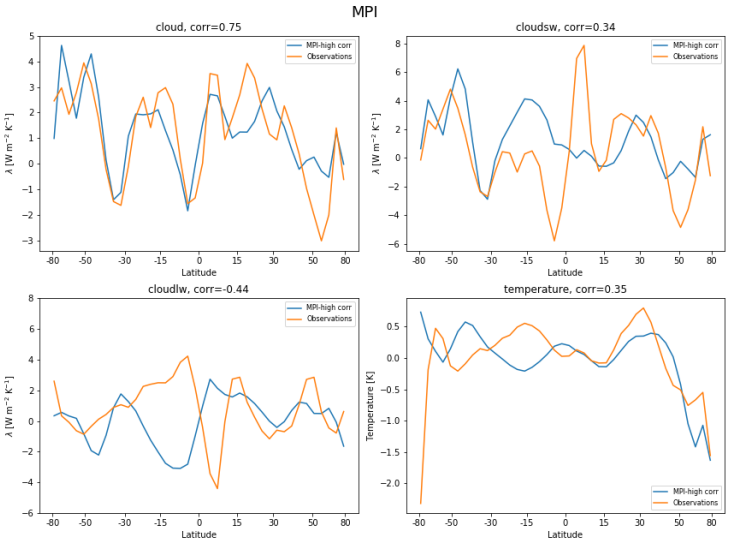

Figure 7: Zonal average of the differences between total, shortwave, and longwave cloud feedbacks and local vs global average temperature for the 10 combinations of high – low MPI ensemble members with the highest correlation between MPI zonal average total cloud feedback and observed zonal average total cloud feedback. The correlation in the title of each plot represents the correlation between the given value in MPI data and observed data.

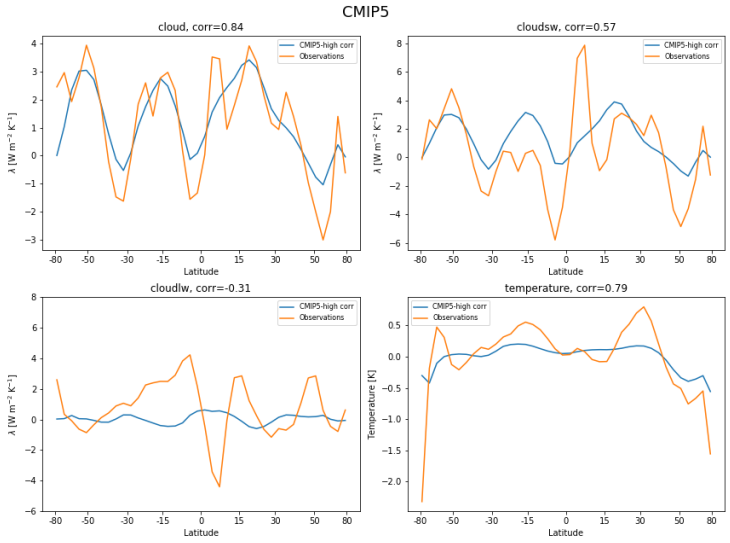

Figure 8: Zonal average of the differences between total, shortwave, and longwave cloud feedbacks and local vs global average temperature for the 10 combinations of high – low 10-year segments in each model with the highest correlation between CMIP5 piControl zonal average total cloud feedback and observed zonal average total cloud feedback, averaged across models. The correlation in the title of each plot represents the correlation between the given value in CMIP5 data and observed data.

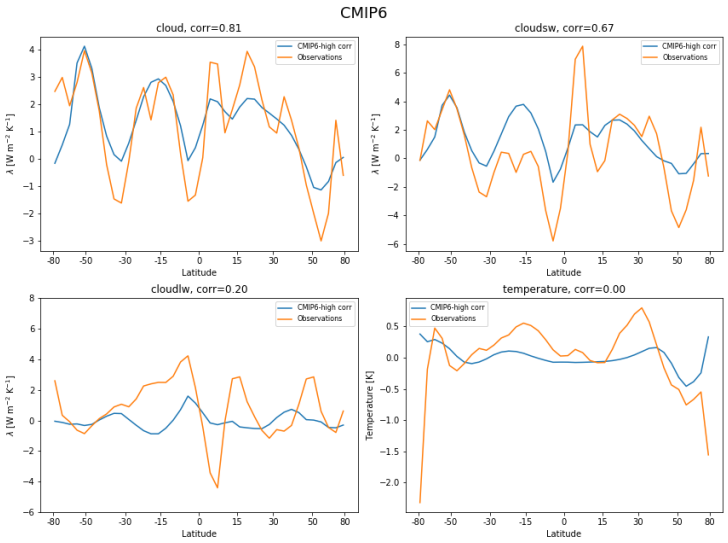

Figure 9: Zonal average of the differences between total, shortwave, and longwave cloud feedbacks and local vs global average temperature for the 10 combinations of high – low 10-year segments in each model with the highest correlation between CMIP6 piControl zonal average total cloud feedback and observed zonal average total cloud feedback, averaged across models. The correlation in the title of each plot represents the correlation between the given value in CMIP6 data and observed data.

Note that while there is good agreement for the total cloud feedback, the shortwave and longwave $\Delta\lambda_{cloud}$ show less correspondence between this combination of MPI ensembles and observations. Net cloud feedback is a better indicator of how low cloud changes (changes in high clouds tend to have LW and SW feedbacks that cancel, e.g., Zelinka et al., 2012), so this tells us that the model ensemble can produce similar patterns of low-cloud feedback variability, while doing a poorer job in high-cloud feedback variability.

Figures 8 and 9 show the same plots as Figure 9 for CMIP5 and CMIP6, respectively. For these two datasets, the 10 combinations of high – low 10-year segments with the highest correlation to observed $\Delta\lambda_{cloud}$ from each model were averaged, then each dataset was averaged across all models. Much like that of MPI, the zonal average $\Delta\lambda_{cloud}$ for both CMIP5 and CMIP6 closely follows the observed zonal average $\Delta\lambda_{cloud}$, with correlation coefficients of 0.84 and 0.81, respectively. While the correlation in the shortwave is higher in CMIP5 and CMIP6 than it is in

MPI, the longwave continues to show low correlation and the total $\Delta\lambda_{cloud}$ remains the most strongly correlated of all values.

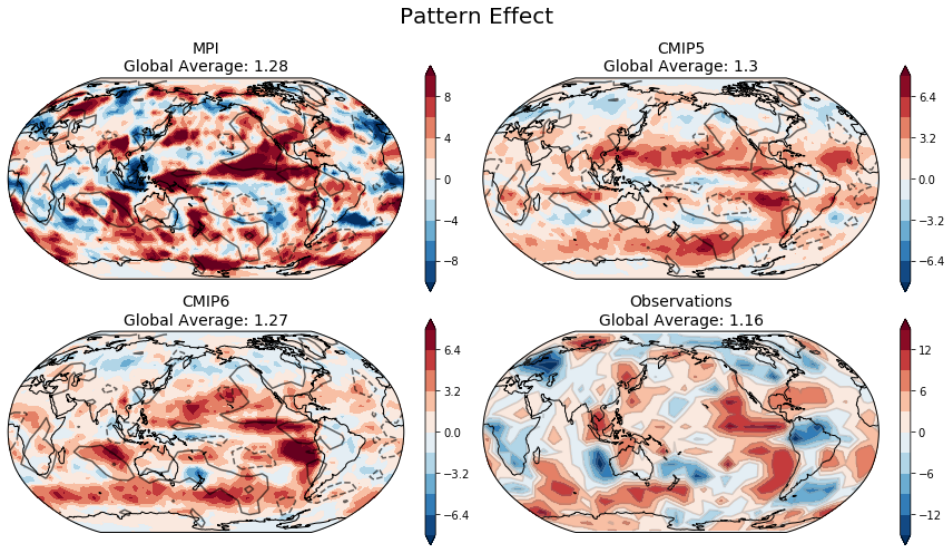

Figure 10: Difference between total cloud feedbacks for the 10 combinations of high – low segments within MPI, CMIP5 piControl, and CMIP6 piControl with the highest correlation between predicted zonal average total cloud feedback and observed zonal average total cloud feedback, averaged across models for CMIP5 piControl and CMIP6 piControl.

The global average $\Delta\lambda_{cloud}$ for all three ensembles when only considering the high-correlation pairs of segments is very similar in magnitude to that of the observations. A global average $\Delta\lambda_{cloud}$ of 1.28 W/m^2/K is measured in MPI, while CMIP5 and CMIP6 report a global average $\Delta\lambda_{cloud}$ of 1.3 W/m^2/K and 1.27 W/m^2/K, respectively. Comparing these results to the observed value of 1.16 W/m^2/K shows that this method of filtering model data can simulate observations effectively.

The full spatial patterns of $\Delta\lambda_{cloud}$ for the combinations of high – low pairs with the highest zonal average correlation for MPI, CMIP5, CMIP6, and the observations are plotted in Figure 10. While differences still exist, this comparison is superior to the one in Figure 3, which shows the $\Delta\lambda_{cloud}$ from the average of the pairs of the appropriate 10 highest and lowest segments for each dataset. This is confirmed via the Taylor diagram in Figure 11, which shows that the

spatial patterns of $\Delta\lambda_{cloud}$ resulting from using the average of difference pairs highly correlated to the observations more accurately represent the spatial patterns of the observations than the spatial patterns resulting from using the average of all differences between pairs of highest and lowest segments.

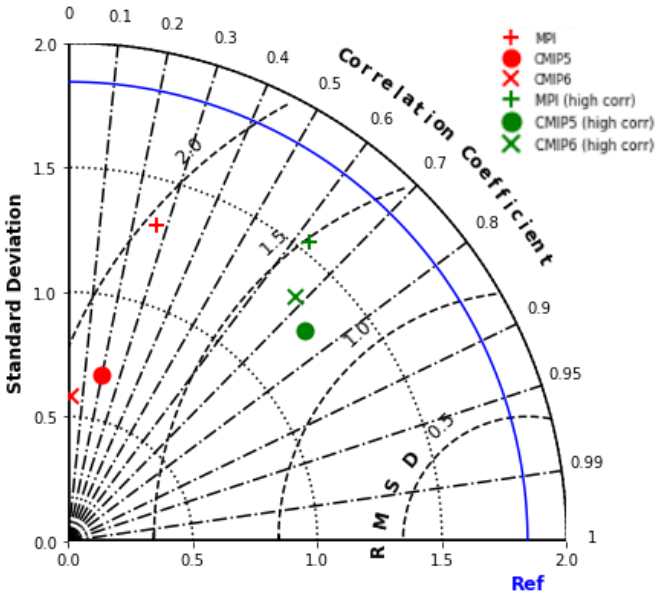

Figure 11: Taylor diagram showing the correlation of the spatial patterns of $\Delta\lambda_{cloud}$ when using all pairs vs pairs with high correlation for all datasets.

CONCLUSIONS

Overall, the methods used in this study produced spatial patterns of model $\Delta\lambda_{cloud}$ that resembled that of observations from 2000 – 2017 to varying degrees of success. The results tended to be more accurate when the methods involved a further refinement of the data, implying that model $\Delta\lambda_{cloud}$ over a long period of time is likely not the best predictor of future $\Delta\lambda_{cloud}$. However, filtering model output by taking the difference between segments with high and low λ_{cloud}, then further by correlation to observations, resulted in model outputs that were much more similar to observations. This means that model ensembles like MPI, CMIP5, and CMIP6 are capable of producing realistic predictions of the cloud feedback parameter under the right conditions.

In an age when climate change is one of the most pressing social issues, having a good grasp on the processes driving Earth's climate system is vital to understanding what the future might look like. While λ_{cloud} is just one component of the total λ, it has historically been the component associated with the most uncertainty. Any process that might reduce that uncertainty in λ_{cloud} would be a step towards reducing the overall uncertain of the total λ and therefore ECS. Once more accurate and reliable estimates of ECS can be made, we can make more authoritative estimates of the future of Earth's climate, an important component in the mitigation of climate change.

REFERENCES

Andrews, T., and Coauthors, 2018: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity. *Geophysical Research Letters*, **45,** 8490-8499.

Andrews, T., and M. J. Webb, 2018: The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming. *J Climate*, **31,** 641-654.

Dessler, A. E., 2013: Observations of Climate Feedbacks over 2000–10 and Comparisons to Climate Models. *J Climate*, **26,** 333-342.

Dessler, A. E., 2020: Potential Problems Measuring Climate Sensitivity from the Historical Record. *J Climate*, **33,** 2237-2248.

Dessler, A. E., and P. M. Forster, 2018: An Estimate of Equilibrium Climate Sensitivity From Interannual Variability. *Journal of Geophysical Research: Atmospheres*, **123,** 8634-8645.

Dessler, A. E., T. Mauritsen, and B. Stevens, 2018: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmos. Chem. Phys.*, **18,** 5147-5155.

Dufresne, J.-L., and S. Bony, 2008: An Assessment of the Primary Sources of Spread of Global Warming Estimates from Coupled Atmosphere–Ocean Models. *J Climate*, **21,** 5135-5144.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9,** 1937-1958.

Forster, P. M., 2016: Inference of Climate Sensitivity from Analysis of Earth's Energy Budget. *Annual Review of Earth and Planetary Sciences*, **44,** 85-106.

Gates, W. L., 1976: The Numerical Simulation of Ice-Age Climate with a Global General Circulation Model. *Journal of the Atmospheric Sciences*, **33,** 1844-1873.

Gregory, J. M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, **43,** 3911-3920.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *J Climate*, **15,** 3117-3121.

Hersbach, H., and Coauthors, 2020: The ERA5 Global Reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **n/a**.

Huang, Y., Y. Xia, and X. Tan, 2017: On the pattern of CO2 radiative forcing and poleward energy transport. *Journal of Geophysical Research: Atmospheres*, **122,** 10,578-510,593.

IPCC, 2014: *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kato, S., and Coauthors, 2018: Surface Irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Data Product. *J Climate*, **31,** 4501-4527.

Kiehl, J. T., 1994: On the Observed Near Cancellation between Longwave and Shortwave Cloud Forcing in Tropical Regions. *J Climate*, **7,** 559-565.

Lewis, N., and J. A. Curry, 2015: The implications for climate sensitivity of AR5 forcing and heat uptake estimates. *Climate Dynamics*, **45,** 1009-1023.

Loeb, N. G., and Coauthors, 2018: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *J Climate*, **31,** 895-918.

Loeb, N. G., and Coauthors, 2020: New Generation of Climate Models Track Recent Unprecedented Changes in Earth's Radiation Budget Observed by CERES. *Geophysical Research Letters*, **47,** e2019GL086705.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability. *Journal of Advances in Modeling Earth Systems*, **11,** 2050-2069.

Marvel, K., R. Pincus, G. A. Schmidt, and R. L. Miller, 2018: Internal Variability and Disequilibrium Confound Estimates of Climate Sensitivity From Observations. *Geophysical Research Letters*, **45,** 1595-1601.

Otto, A., and Coauthors, 2013: Energy budget constraints on climate response. *Nature Geoscience*, **6,** 415-416.

Shell, K. M., J. T. Kiehl, and C. A. Shields, 2008: Using the Radiative Kernel Technique to Calculate Climate Feedbacks in NCAR's Community Atmospheric Model. *J Climate*, **21,** 2269-2282.

Soden, B. J., A. J. Broccoli, and R. S. Hemler, 2004: On the Use of Cloud Forcing to Estimate Cloud Feedback. *J Climate*, **17,** 3661-3665.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying Climate Feedbacks Using Radiative Kernels. *J Climate*, **21,** 3504-3520.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93,** 485-498.

Zelinka, M. D., S. A. Klein, and D. L. Hartmann, 2012: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part II: Attribution to Changes in Cloud Amount, Altitude, and Optical Depth. *J Climate,* **25,** 3736-3754.

Zhou, C., M. D. Zelinka, A. E. Dessler, and P. Yang, 2013: An Analysis of the Short-Term Cloud Feedback Using MODIS Data. *J Climate,* **26,** 4803-4815.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience,* **9,** 871-874.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2017: Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *Journal of Advances in Modeling Earth Systems,* **9,** 2174-2189.