Perturbing Parameters to Understand Cloud Contributions to Climate Change

Margaret L. Duffy,[a] Brian Medeiros,[a] Andrew Gettelman,[a,b] and Trude Eidhammer[a]

[a] *NSF National Center for Atmospheric Research, Boulder, Colorado*
[b] *Pacific Northwest National Laboratory, Richland, Washington*

ABSTRACT: The sensitivity of cloud feedbacks to atmospheric model parameters is evaluated using a CAM6 perturbed parameter ensemble (PPE). The CAM6 PPE perturbs 45 parameters across 262 simulations, 206 of which are used here. The spread in the total cloud feedback and its six components across the CAM6 PPE are comparable to the spread across the CMIP6 and AMIP ensembles, indicating that parametric uncertainty mirrors structural uncertainty. However, the high-cloud altitude feedback is generally larger in the CAM6 PPE than WCRP assessment, CMIP6, and AMIP values. We evaluate the influence of each of the 45 parameters on the total cloud feedback and each of the six cloud feedback components. We also explore whether the CAM6 PPE can be used to constrain the total cloud feedback, with inconclusive results. Further, we find that despite the large parametric sensitivity of cloud feedbacks in CAM6, a substantial increase in cloud feedbacks from CAM5 to CAM6 is not a result of changes in parameter values. Notably, the CAM6 PPE is run with a more recent version of CAM6 (CAM6.3) than was used for AMIP (CAM6.0) and has a smaller total cloud feedback ($0.56 \ \mathrm{W \ m^{-2} \ K^{-1}}$) as compared to CAM6.0 ($0.81 \ \mathrm{W \ m^{-2} \ K^{-1}}$) owing primarily to reductions in low clouds over the tropics and midlatitudes. The work highlights the large sensitivity of cloud feedbacks to both parameter values and structural details in CAM6.

KEYWORDS: Feedback; Climate sensitivity; Cloud parameterizations

1. Introduction

Constraining equilibrium climate sensitivity (ECS) is a crucial step to quantifying warming from anthropogenic emissions. However, doing so has proven difficult, with little change in the estimated ECS range from 1979 to the present (National Research Council 1979; Sherwood et al. 2020; Forster et al. 2021). Global climate models (GCMs) are useful for understanding the climate's response to forcing. However, the spread in ECS values across models remains large and has, in fact, increased between CMIP5 and CMIP6. From CMIP5 to CMIP6, the upper ECS estimate increased from 4.7 to 5.6 K and the multimodel mean ECS increased from 3.3 to 3.9 K. Further, 10 CMIP6 models have an ECS greater than that of the largest CMIP5 model (4.7 K) (Zelinka et al. 2020).

ECS depends on both the radiative forcing and the radiative feedback. The change in spread in ECS from CMIP5 to CMIP6 is the result of different combinations of radiative forcing and feedback (Zelinka et al. 2020). The radiative feedback includes contributions from both clear-sky and cloud feedbacks. The spread across models and observations is dominated by the spread in cloud feedbacks (Sherwood et al.

2020; Zelinka et al. 2020). Therefore, constraining cloud feedbacks is key to constraining ECS.

The relationship between ECS, forcing, and feedback can be conceptualized using the top-of-the-atmosphere (TOA) energy budget, which is given by

$$\Delta N = \Delta F + \lambda \Delta T, \tag{1}$$

where ΔN is the net top-of-the-atmosphere radiative flux anomaly (positive down), ΔF is the radiative forcing, λ is the total radiative feedback parameter, and ΔT is the global-mean near-surface air temperature response. We define ECS by assuming equilibrium ($\Delta N = 0$) and a radiative forcing which corresponds to a doubling of CO_2, given by

$$\mathrm{ECS} \equiv \Delta T_{2\mathrm{xCO_2}} = -\frac{\Delta F_{2\mathrm{xCO_2}}}{\lambda}. \tag{2}$$

From Eq. (2), it is clear that there is a close, inverse relationship between ECS and radiative feedbacks. In practice, calculating ECS requires running a $2\mathrm{xCO_2}$ or $4\mathrm{xCO_2}$ GCM simulation to equilibrium. However, other definitions of climate sensitivity also exist, such as effective climate sensitivity and transient climate sensitivity, and tend to be correlated with ECS across GCMs (Gregory et al. 2004; Rugenstein and Armour 2021). The radiative feedback λ can be further decomposed into cloud radiative feedbacks and the noncloud radiative feedbacks, given by

$$\lambda = \lambda_{\mathrm{cloud}} + \lambda_{\mathrm{noncloud}}. \tag{3}$$

The noncloud feedbacks include the Planck, water vapor, lapse rate, and surface albedo feedbacks. We focus on cloud

---

feedbacks because their uncertainty is larger than noncloud feedbacks (Sherwood et al. 2020).

The total cloud feedback is composed of various cloud changes with warming. Here, we separate cloud feedbacks into both their shortwave (SW) and longwave (LW) contributions and into six cloud feedback components as in Sherwood et al. (2020) [also called the World Climate Research Programme (WCRP) assessment]: high-cloud altitude, tropical marine low cloud, tropical anvil cloud area, land cloud amount, midlatitude marine low-cloud amount, and high-latitude low-cloud optical depth. These six cloud feedback components are the result of known processes, and Sherwood et al. (2020) assess their values and spreads using multiple lines of evidence. The high-cloud altitude feedback is a robustly positive feedback across lines of evidence (Sherwood et al. 2020) and across GCMs (Zelinka et al. 2022). The high-cloud altitude feedback results from cloud-top heights rising nearly isothermally which means that LW emission from cloud tops changes less than the increased surface emission, resulting in greater LW flux divergence and enhanced warming (i.e., a positive feedback) (Hartmann and Larson 2002; Zelinka and Hartmann 2010). The tropical marine low-cloud feedback is estimated to be a positive SW feedback resulting from reduced low-cloud amounts in the tropics with warming (Bony and Dufresne 2005; Klein et al. 2017). The tropical anvil cloud area feedback comes from a reduction in the area of tropical anvil clouds with warming. This reduced area has compensating LW and SW feedbacks, but observational evidence suggests the LW effect is greater than the SW effect for a net negative feedback (Hartmann et al. 2001; Williams and Pierrehumbert 2017). The land cloud feedback is expected to produce a positive feedback of smaller magnitude than any of the three feedback contributions already described. The land cloud feedback is the result of reduced cloudiness due to reduced relative humidity over land (Bretherton et al. 2014). Low-cloud reductions dominate the radiative effect of the land cloud feedback, resulting in a positive SW feedback (Kamae et al. 2016). Midlatitude low-cloud feedbacks are expected to provide a positive SW feedback due to reductions in low clouds (Kay et al. 2014; Zhai et al. 2015; McCoy et al. 2017). The high-latitude low-cloud optical depth feedback results from competing influences on high-latitude low-cloud optical depths with warming and therefore has an estimated feedback of 0 W m^{-2} K^{-1} but with substantial spread (Tan et al. 2016; Ceppi et al. 2016). These six components are separate from one another and sum to the total feedback if the unassessed component is small, which it is in most CMIP5 and CMIP6 models (Zelinka et al. 2022).

Focusing on GCMs, the spread in total cloud feedbacks has increased from CMIP5 to CMIP6, with an increase in spread in the midlatitude low-cloud amount (and scattering) feedback from CMIP5 to CMIP6 (Zelinka et al. 2020). The CMIP5 and CMIP6 multimodel mean cloud feedbacks (total cloud feedback and its six components) all fall within the WCRP assessed ranges, but there is substantial intermodel spread across the total cloud feedback and most of the six cloud feedback components. Further, the tropical marine low-cloud feedback tends to be smaller (a smaller positive feedback), and the tropical marine anvil cloud area feedback tends to be larger (a smaller negative feedback) than the WCRP assessed range. CMIP5 and CMIP6

models whose assessed feedbacks compared most favorably to the WCRP assessed values had total cloud feedback estimates of 0.4–0.6 W m^{-2} K^{-1} and ECS estimates of 3–4 K (Zelinka et al. 2022).

Despite the tremendous impact of clouds on Earth's climate, their subgrid size requires cloud processes to be parameterized in GCMs. Here, we evaluate the sensitivity of cloud feedbacks to atmospheric model parameter values in the Community Atmosphere Model version 6, CAM6 (Danabasoglu et al. 2020) using a perturbed parameter ensemble (PPE). An advantage of using a PPE is that we are able to use parameter sensitivity to identify processes setting the spread in cloud feedbacks. We use the CAM6 PPE and CMIP6 models to address the following questions:

1) How does the spread in cloud feedbacks across the CAM6 PPE compare to the spread in cloud feedbacks across CMIP6 models?
2) What is the sensitivity of cloud feedbacks to parameter values in CAM6?
3) Can the total cloud feedback be constrained using the CAM6 PPE?

Additionally, there have been substantial fluctuations in the cloud feedback and corresponding ECS values across generations of the Community Atmosphere Model (CAM). The cloud feedback and ECS values increased from CAM5 to CAM6 (Gettelman et al. 2019). However, we find here that the CAM6 PPE default simulation (CAM6.3) has lower cloud feedback and ECS values than previous CAM6 simulations (CAM6.0). Therefore, we also investigate the following questions:

4) Are changes in parameter values (tuning) responsible for the increase in cloud feedback from CAM5 to CAM6?
5) Why does the CAM6 PPE default simulation have a lower cloud feedback than previous CAM6 configuration?

The CMIP, AMIP, and PPE data are described in section 2. A description of the cloud radiative feedback calculation is also provided in section 2. Questions 1–5 are addressed sequentially in sections 3–7. For interested readers, we discuss the relationship between the CAM6 PPE used here and a version that is tuned based on paleoclimate evidence in section 8. We discuss and conclude in section 9.

2. Data description

a. PPE

The CAM6 PPE is available for use (Eidhammer et al. 2022), is described in detail in T. Eidhammer et al. (2023, unpublished manuscript), and is briefly summarized here. The CAM6 PPE consists of 262 CAM6 simulations.[1] The simulations are atmosphere-only simulations run with CAM6.3 with imposed climatological SST and sea ice cover (using the F2000climo component set of the model). The prescribed SST and sea ice fields have an annual cycle but no interannual

---

[1] In total, 263 simulations were run, but 1 was discarded due to numerical instability.
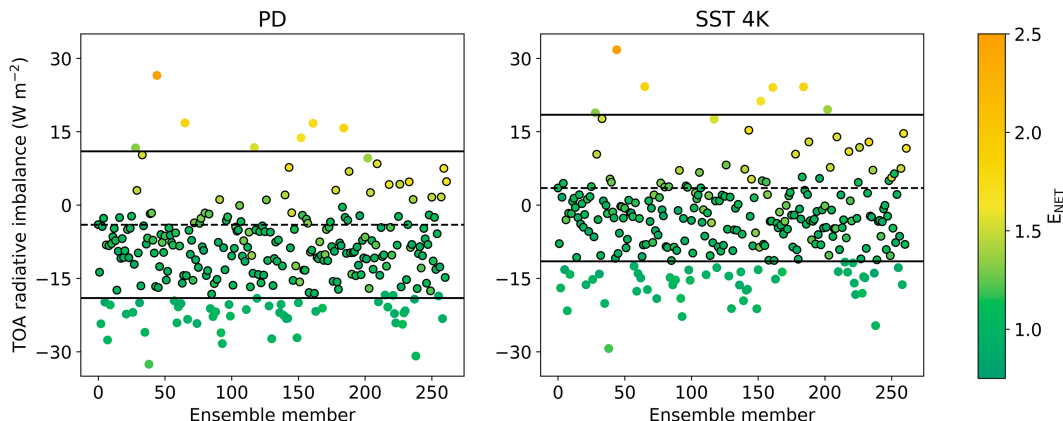
FIG. 1. Top-of-atmosphere radiative imbalance in each of the 262 PPE simulations for both the (left) PD and (right) SST4K experiments. The dashed horizontal line goes through the default value, and the black lines are 15 W m^{-2} away from the default. Dot colors correspond to the E_{NET} value in the PD simulation. The simulations in the subset used in the forthcoming analysis are encircled in black.

variability. One simulation is considered a "default" simulation because it uses the parameter values typical of CAM6. The remaining 261 simulations differ in 45 atmospheric parameter values across parameterizations for cloud and aerosol effects. The parameter ranges are chosen by expert judgment to span physical process limits. The chosen parameter ranges are intended to cover the plausible range of realistic values based on expert assessment and thus span the range of parametric uncertainty. Parameter values are selected using Latin hypercube sampling, meaning that the parameters are varied simultaneously, but randomly and independently from one another (McKay et al. 2000). Latin hypercube sampling ensures that for each parameter, the range of possible values is evenly sampled. Here, we show results from variations in 43 parameters.[2] For each of the 262 PPE simulations, we use two experiments: a control or present-day simulation, denoted "PD," and a warm simulation with a uniform 4-K SST warming, denoted "SST4K." All PPE simulations are run with the CFMIP observation simulator package (COSP) turned on, which produces joint histograms of cloud-top pressure and cloud optical depth in each simulation (Bodas-Salcedo et al. 2011). All PPE simulations are run for 3 years each. Monthly mean model output is used throughout.

The spread in feedbacks across the CAM6 PPE is sensitive to the range of each parameter value chosen in designing the CAM6 PPE. Although the parameter values are chosen by expert judgment to within physical process limits, combinations of different parameter values need not yield realistic climates (Stainforth et al. 2005). To address this, we take a subset of the 206 PPE simulations based on TOA, a radiative balance constraint. None of the CAM6 PPE simulations have been tuned based on TOA radiative imbalance. The default PD simulation has a TOA radiative imbalance of −4.0 W m^{-2}. The default SST4K simulation has a TOA radiative imbalance of 3.5 W m^{-2}. Here and throughout

the paper, we analyze only the 206 simulations whose TOA radiative imbalance falls within 15 W m^{-2} of the default simulation for both the PD and SST4K experiments. That is, we only keep simulations whose TOA imbalance is between −19 and 11 W m^{-2} in the PD experiment *and* whose TOA imbalance is between −11.5 and 18.5 W m^{-2} in the SST4K experiment. The TOA radiative imbalance for each simulation and experiment is shown in Fig. 1. The 15 W m^{-2} threshold is chosen based on the data in Fig. 1 somewhat subjectively, with the intention of eliminating data in the tails of the TOA imbalance distribution. We also evaluate our subset using the mean-state cloud error for each simulation's PD experiment. To do so, we use an aggregated scalar error metric of the impact of cloud errors on TOA radiation in the base state, E_{NET} (Klein et al. 2013), with details as in Zelinka et al. (2022). Reassuringly, we find that our subset eliminates the simulations with largest mean-state cloud errors (Fig. 1).

b. Feedback calculations

In the CAM6 PPE, the total radiative feedback is calculated as the ratio of net downward TOA flux ΔN to surface warming ΔT. We calculate cloud radiative feedbacks using the cloud radiative kernel technique with output from the PD and SST4K simulations. In general, cloud radiative kernels calculate the sensitivity of TOA radiation to perturbations in cloud fraction. Here, we calculate cloud radiative feedbacks using joint-histogram kernels from Zelinka et al. (2012). The kernels of Zelinka et al. (2012) are constructed from joint histograms of cloud-top pressure and optical depth to separate cloud-radiative effects into contributions by different cloud types.

The kernels calculate SW and LW cloud feedbacks separately for various cloud-top pressure and cloud optical depth values. This allows for decomposition of the cloud feedback into SW and LW contributions and into the six cloud feedback components of Sherwood et al. (2020), which are calculated in the CAM6 PPE using the methods of Zelinka et al. (2022). Obscuration effects, the hiding and revealing low

---

[2] Not 45 because two pairs of parameter values are perturbed simultaneously: CLUBB C6rt/C6thl and CLUBB C6rtb/C6thlb.
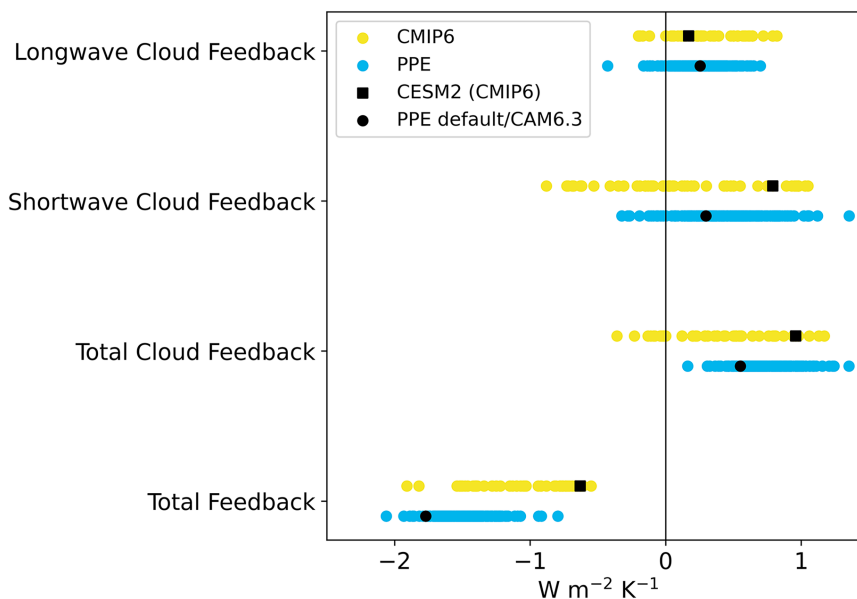
FIG. 2. Spread in the global-mean radiative feedback across CMIP6 models (yellow) and across PPE (blue). The total radiative feedbacks, the cloud-radiative feedbacks, SW cloud-radiative feedbacks, and LW cloud-radiative feedbacks are shown from bottom to top. The black square over the AMIP models denotes the CESM2 values, and the black dot over the CAM6 PPE indicates the CAM6 PPE default simulation value.

clouds by changes in high clouds, are accounted for. Ascent and descent regions are defined by the sign of vertical velocity at 500 hPa at that location in each simulation. The cutoff between high and low clouds is 680 hPa. The high-cloud altitude and land cloud amount feedbacks are calculated globally. The tropical marine low-cloud and tropical anvil cloud area feedbacks are calculated between 30°S and 30°N. The midlatitude low-cloud amount feedback is calculated between 30° and 60°N and between 30° and 60°S. The high-latitude low-cloud optical depth feedback is calculated between 40° and 70°N and between 40° and 70°S. Amount, altitude, and optical depth feedbacks linearly sum to the total feedback, so the cloud feedback components are separate from one another and sum to the total cloud feedback when the unassessed component is included.

c. CMIP6, AMIP, and assessed values

We compare total radiative feedbacks, total cloud feedbacks, and SW and LW cloud feedbacks in the CAM6 PPE to those in CMIP6 models. Feedbacks from coupled experiments are calculated using the CMIP6 preindustrial control ("piControl") and abrupt CO_2 quadrupling ("abrupt-4xCO2") experiments; we refer to these as CMIP6. Feedbacks from atmosphere-only experiments are calculated using "amip" (CMIP DECK) and "amip-p4K" (CFMIP, uniform 4-K warming) experiments for model years 1983–2008, referred to as AMIP.[3] Values for the

feedbacks are provided by Mark Zelinka's "cmip56_forcing_feedback_ecs" (https://github.com/mzelinka/cmip56_forcing_feedback_ecs) and "assessed_cloud_fbks" (https://github.com/mzelinka/assessed-cloud-fbks) GitHub repositories, as described by Zelinka et al. (2020, 2022). The corresponding calculations are repeated for the CAM6 PPE following the methods of Zelinka et al. (2020, 2022). CESM2 (CMIP) is not included in the six cloud feedback components of Zelinka et al. (2022) because one of the needed variables (clisccp) is not available for the piControl simulation.

3. Spread in cloud feedbacks

We begin by comparing the distribution of total cloud radiative feedbacks between the subset of 206 CAM6 PPE simulations and the CMIP6 models. The spread across the CMIP6 models represents structural uncertainty; each model solves different equations with different numerical methods. On the other hand, the spread across the CAM6 PPE represents parametric uncertainty; the only difference between members is the parameter values. The total feedback, along with the total cloud feedback and its SW and LW components, is shown in Fig. 2. The CAM6 PPE default simulation has a smaller (more negative) total feedback and a smaller (less positive) total cloud feedback than the CESM2-coupled experiment (Fig. 2; compare black square and black circle). This is an important and unexpected finding which we discuss further in section 7. We also find that the spread across total cloud feedbacks in the CAM6 PPE and CMIP6 models are comparable to one another. There is no a priori expectation for the CAM6 PPE spread to be similar to the CMIP6 spread. The

---

[3] We use 1983–2008 values instead of the full 1979–2014 because that is what is used in Zelinka et al. (2022), which was done for comparison with observations.
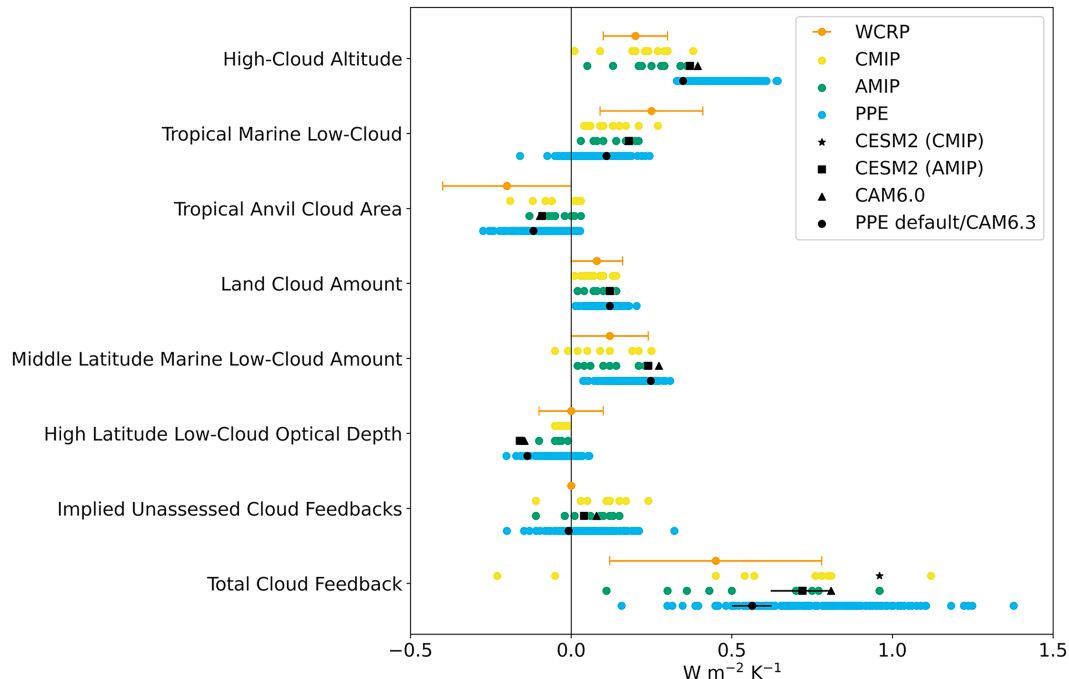
FIG. 3. Comparison of total cloud feedbacks, unassessed cloud feedbacks, and cloud feedback components in WCRP assessment (orange), CMIP6 models (yellow), AMIP models (green), and the CAM6 PPE (blue). The range of the WCRP-assessed feedbacks represents the 1σ spread. The black star in the CMIP total cloud feedback row denotes the CESM2 (CMIP) value. The black squares in the AMIP rows denote the CESM2 (AMIP) values, and the bar going through it in the total cloud feedback row covers the interannual spread. The black triangle in the AMIP total feedback row denotes the simulation of CAM6.0 forced with F2000climo SST (see text for details). The black dots in the CAM6 PPE rows denote the CAM6 PPE default simulation values, which use CAM6.3 and F2000climo, and the black bar going through it in the total cloud feedback row covers the interannual spread in a 12-yr version of the CAM6 PPE default simulation.

two ensembles are entirely different with different numbers of members and different sources of spread (i.e., systematic parameter changes in the CAM6 PPE and opportunistic and structural differences across CMIP6).

To compare the processes setting the comparable spreads in cloud feedbacks across the CAM6 PPE simulations and CMIP6 models, we further decompose cloud feedbacks into 1) their shortwave (SW) and longwave (LW) contributions and 2) into the six cloud feedback components of Sherwood et al. (2020). Other decompositions are possible; we use the SW and LW decomposition because of its simplicity and the six cloud feedback components of Sherwood et al. (2020) for comparison with the recent work (e.g., Zelinka et al. 2022). The SW and LW contributions to the cloud feedback are shown in Fig. 2 and compare favorably to one another. A bar graph with the mean and standard deviation of these data is shown in Fig. S1 in the online supplemental material.

The total cloud and the six components of Sherwood et al. (2020) are shown in Fig. 3 for the CMIP6 models, AMIP models, and CAM6 PPE. The 1σ range of each component in the WCRP assessment by Sherwood et al. (2020) is also included for comparison. A bar graph with the mean and standard deviation of these data is shown in Fig. S2. The AMIP models are more similar to the CAM6 PPE simulations than the

CMIP6 models because the AMIP and CAM6 PPE simulations are all atmosphere-only with imposed uniform 4-K warming. Therefore, we focus on the comparison between CAM6 PPE and AMIP simulations.

We begin by comparing the mean of the cloud feedback components between the CAM6 PPE and the AMIP models. The mean total cloud feedback is larger in the CAM6 PPE than in the AMIP models. This larger total cloud feedback is the consequence of much larger high-cloud altitude feedback, which is partially compensated by a smaller mean tropical marine low-cloud feedback. The other four feedbacks have mean values which fall within the assessed 1σ range of Sherwood et al. (2020). Additionally, the mean unassessed cloud feedback component is close to zero in the PPE, but this component is the least understood, and it is not known whether or not it should be zero. The AMIP models have very similar mean cloud feedbacks to the CMIP6 models.

We next compare the spread in the cloud feedbacks across the CAM6 PPE and AMIP models. The range (largest minus smallest value) in the total cloud feedback is comparable between AMIP (0.85 W m^{-2} K^{-1}) and the CAM6 PPE (1.22 W m^{-2} K^{-1}). This is reflected in each of the six cloud feedback components, which have similar ranges across the two ensembles. The AMIP models have very similar spreads in cloud feedbacks to the

CMIP6 models. However, the CMIP6 range for the high-latitude low-cloud optical depth feedback is much smaller than the CAM6PPE range and also much smaller than the assessed 1σ range of Sherwood et al. (2020).

Overall, the comparable spreads in the total cloud feedback between the CAM6 PPE, CMIP6, and AMIP models is the result of comparable spreads in SW and LW cloud feedbacks and of comparable spreads in the six cloud feedback components. Further, no individual component of the cloud feedback is controlling the spread in any of the three ensembles, and these different feedback components contribute about the same to the total cloud feedback spread in the three ensembles. Again, we emphasize that there is no a priori reason to expect a single model's estimated parametric uncertainty to reproduce the spread in cloud feedbacks in CMIP6 or AMIP.

Why might the parametric and structural uncertainties be comparable? Why does it matter? That the spreads in cloud feedbacks due to parametric and structural uncertainties are similar suggests that the processes controlling cloud feedbacks can be influenced by either structural or parametric differences. For example, if some cloud property is contributing to a feedback and that property varies with structural and parametric differences which are similarly uncertain, then the portion of the cloud feedbacks set by that property would have similar spreads across CMIP6 and the CAM6 PPE. The CAM6 PPE is advantageous because the only difference between ensemble members is parameter values and each parameter controls a process in the model. Therefore, in section 4, we evaluate the influence of various parameters on feedbacks.

Notably, the total cloud feedback is smaller in the CAM6PPE default simulation than in the CESM2 (AMIP) simulation, which is smaller than the CESM2 (CMIP) simulation. We investigate these differences in section 7.

4. Parameters

a. Which parameters set the spread in cloud feedbacks?

An advantage of using a perturbed parameter ensemble is the ability to identify the influence of various parameters on the cloud feedbacks. Parameters correspond to physical processes, so understanding which parameters control the spread in a given feedback offers some insights into the processes controlling that feedback in CAM6. We use regression models to quantify the influence of each parameter on the spread in the total cloud feedback, each of its six components, and the unassessed cloud feedback. Scatterplots of each parameter against each of the cloud feedbacks studied here indicate that it is reasonable to assume linearity. The linear regression model is given by

$$\hat{\lambda}_j = a_{0,j} + \sum_{i=0}^{45} a_{i,j}\tilde{p}_i, \tag{4}$$

where $\hat{\lambda}_j$ is the estimated cloud feedback component, $a_{0,j}$ is the intercept for a cloud feedback component j, $a_{i,j}$ are the regression coefficients for each parameter i and each cloud

feedback component j, and $\tilde{p}_i$ are the parameter values that are scaled to fall between 0 and 1 according to

$$\tilde{p}_i = \frac{p_i - p_{i,\min}}{p_{i,\max} - p_{i,\min}}, \tag{5}$$

where p_i are the unscaled parameter values, $p_{i,\min}$ is the minimum value in p_i, and $p_{i,\max}$ is the maximum value in p_i. In practice, we rescale the parameters using Python's sklean.preprocessing.MinMaxScaler function. To avoid overfitting due to the large number of parameters (43) relative to feedbacks (206), we use a least absolute shrinkage and selection operator (LASSO) regression, which imposes an L1 penalty to yield a sparse model (i.e., a model in which some parameters are eliminated by setting their coefficient to 0). The degree of sparseness is set by a tuning parameter. Here, the tuning parameter is calculated using an automated, fivefold cross-validation function (sklearn.linear_model.LassoCV in Python with default function parameters). While it is common to use machine learning emulation where we use a LASSO regression (e.g., Dagon et al. 2020), we find using 20-member test datasets that a LASSO regression model has a comparable mean-squared error to the Gaussian process and random forest machine learning models for the cloud feedbacks studied here. See Watson-Parris et al. (2022) for more details about machine learning emulators.

The six cloud feedback components are separate from one another and approximately sum to the total cloud feedback. Therefore, the sum of the regression models given by Eq. (4) for each of the six cloud feedback components yields an estimate of the total cloud feedback, as follows:

$$\lambda_{\text{cloud}} \approx \sum_j a_{0,j} + \sum_{i=0}^{45} \left(\sum_j a_{i,j}\right)\tilde{p}_i, \tag{6}$$

where j are the cloud feedback components. Conveniently, the regression coefficients $a_{i,j}$ can be interpreted as the spread of the estimated cloud feedback across the CAM6 PPE due to variations in the corresponding parameter. This nice property exists because the parameter values $\tilde{p}_i$ are scaled to fall between 0 and 1. More specifically, a given parameter's influence on the cloud feedback component j is estimated by $a_{i,j}\tilde{p}_i$ for some parameters with index i. Since the smallest $\tilde{p}_i$ value is always 0 and the largest $\tilde{p}_i$ value is always 1, the difference in λ_j due to variations in parameter $\tilde{p}_i$ is $a_{i,j}$. For example, if a parameter has a regression coefficient of 0.1, all else equal, the estimated cloud feedback component will be 0.1 W m^{-2} K^{-1} greater when the parameter value is 1 than when it is 0. Therefore, a regression coefficient of 0.1 means that variations in that parameter are contributing a spread of 0.1 W m^{-2} K^{-1} to the estimated cloud feedback component. Further, the contribution of a given parameter's variation to the total cloud feedback is approximately equal to $\sum_j a_{i,j}\tilde{p}_i$. In short, larger regression coefficients indicate more contributions to the spread across the CAM6 PPE.

Table 1 shows the regression coefficients for each cloud feedback component for the three parameters most highly correlated with the total cloud feedback: CLUBB_C8, MG2_DCS,

TABLE 1. LASSO regression coefficient for each parameter, multiplied by 100 by visual clarity; Hi Alt = high-cloud altitude, Trop Lo = tropical marine low cloud, Trop Anvil = tropical anvil cloud area, Land = land cloud amount, Mid Lo = middle latitude marine low-cloud amount, and Hi Lo = high-latitude low-cloud optical depth.

Scheme	Parameter name	Description	Total	Sum assessed	Hi Alt	Trop Lo	Trop Anvil	Land	Mid Lo	Hi Lo	Unassessed
CLUBB	C8	Coef No. 1 in C8 skewness equation	14.04	9.71	−0.74	6.35	0	1.66	1.54	0.91	3.03
MG2/PUMAS	DCS	Autoconversion size threshold ice-snow	25.01	16.33	4.72	0.31	8.72	0.96	2.64	−1.02	8.12
ZM	capelmt	Triggering threshold for ZM convection	16.68	14.45	0.23	−0.74	7.82	−1.15	0.96	7.33	1.95

and ZM_capelmt. CLUBB_C8 is a skewness coefficient associated with the third moment of vertical velocity. Larger CLUBB_C8 values correspond to thicker, more reflective clouds. MG2_DCS is the ice–snow autoconversion size threshold; a larger value favors more ice crystals over snow in the atmosphere. ZM_capelmt is a CAPE threshold value for triggering deep convection. Therefore, these processes have the largest influence on the total cloud feedback in the CAM6 PPE. Table S1 shows the regression coefficients for each cloud feedback component for all 43 parameters. Comparing the regression coefficients allows us to see 1) how much spread in the total feedback is associated with each parameter and 2) how each of the six cloud feedback components contributes to the total spread. More specifically, according to this simple model, about 0.14 W m^{-2} K^{-1} of the spread in the total cloud feedback is attributed to variations in CLUBB_C8, and that about 0.06 W m^{-2} K^{-1} of that spread is from the tropical marine low-cloud feedback. About 0.25 W m^{-2} K^{-1} of the spread in the total cloud feedback is attributed to variations in MG2_DCS, and that about 0.09 W m^{-2} K^{-1} of that spread is from the tropical anvil cloud area feedback and about 0.05 W m^{-2} K^{-1} of that spread is from the high-cloud altitude feedback. Further, ZM_capelmt contributes about 0.17 W m^{-2} K^{-1} of the spread in the total cloud feedback, and about 0.08 W m^{-2} K^{-1} is from the tropical anvil cloud area feedback and about 0.07 W m^{-2} K^{-1} is from the high-latitude low-cloud optical depth feedback. The sum of the six cloud feedback component regression coefficients may differ from the total cloud feedback regression coefficient because of 1) unassessed cloud feedback and 2) statistical model errors. We also show the regression coefficients for the unassessed cloud feedback, and they partially explain the discrepancy for many of the parameters (Table 1 and Table S1).

The simple LASSO regression approach is very useful for estimating the influence of each parameter on cloud feedbacks but is subject to some limitations. The regression model does not capture nonlinear relationships between the parameter and feedback, nor the influence of interactions between parameters on the feedback. Nonetheless, given the comparable spreads in cloud feedbacks between CMIP models and the CAM6 PPE, the processes associated with these three influential parameters may be promising areas of future model development in order to constrain cloud feedbacks.

b. Influence of parameters on mean-state cloud errors and total cloud feedback

We plot each parameter as a function of the mean-state cloud error and feedbacks. We evaluate the total cloud feedback as compared with 1) mean-state cloud errors and 2) the WCRP-assessed estimate of the feedback components. We use E_{NET} to evaluate mean-state cloud errors and an aggregated difference metric, RMSD, to compare with the WCRP estimate. The E_{NET} metric is described in section 2a and in Klein et al. (2013) and Zelinka et al. (2022). Following Zelinka et al. (2022), the difference metric used to compare with WCRP estimates is the root-mean-square deviation (RMSD) of the estimated feedback as compared to the central WCRP-assessed value for each of the six assessed feedback components. The aggregated metric is simply the square root of the average of the six squared differences. Therefore, the difference metric is minimized when all six of the assessed cloud feedbacks are close to their assessed value. We plot the parameter value as a function of the mean-state cloud error E_{NET} and an aggregated distance metric for feedback RMSD for the three parameters most highly correlated with cloud feedback: MG2_DCS, ZM_capelmt, and CLUBB_C8 (Fig. 4) and for the two parameters most highly correlated with E_{NET}: MG2_accre_enhan_fact and ZM_cldfrac_dp2 (Fig. 5). These parameters were selected because their correlation coefficients are greater than 0.2. The parameter MG2_accre_enhan_fact is an accretion enhancing factor, with larger values corresponding to more accretion of cloud droplets onto rain. The ZM_cldfrac_dp2 is a parameter for deep convective cloud fraction as a function of convective mass flux, with larger values corresponding to a higher deep convective cloud fraction. The equivalent figures for the remaining parameters are shown in Figs. S3–S7. The parameters most correlated with the total cloud feedback are different from the parameters most correlated with E_{NET}, which is not surprising given the lack of correlation between E_{NET} and the total cloud feedback (section 5 and Fig. 6c).

5. CAM6 PPE constraint on the total cloud feedback

Given the large parametric sensitivity of cloud feedbacks in the CAM6 PPE, we ask whether the CAM6 PPE provides a constraint on the total cloud feedback. To that end, we further compare the CAM6 PPE to mean-state cloud errors using E_{NET} and to WCRP estimates of cloud feedbacks using
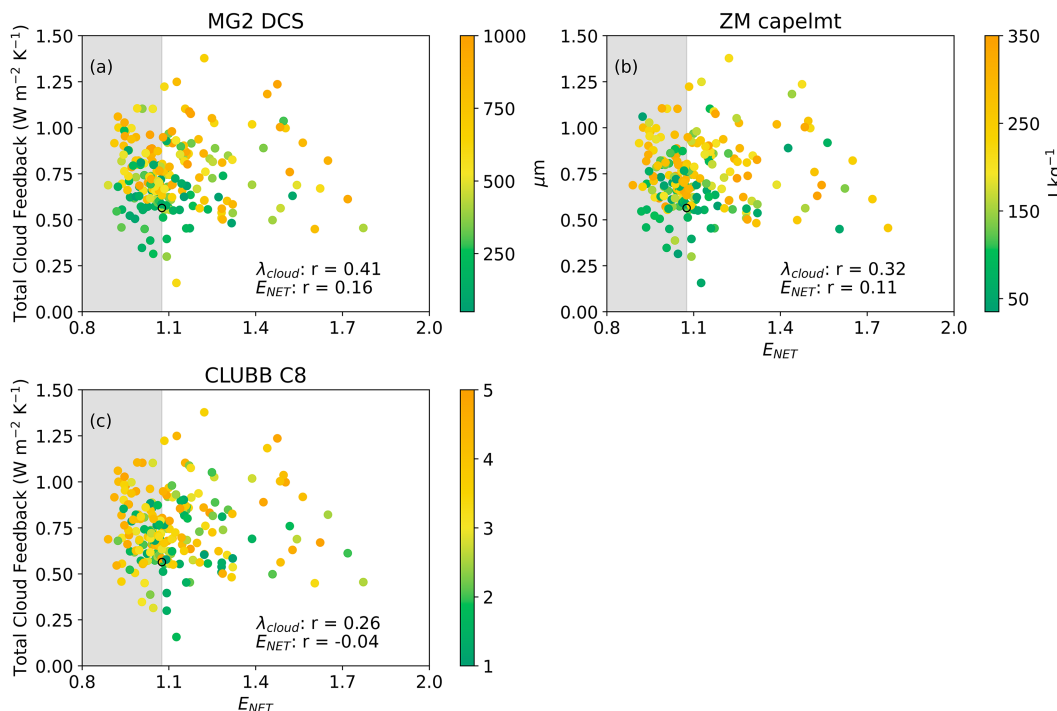
FIG. 4. Scatterplot of the total cloud feedback vs the cloud feedback E_{NET} in the CAM6 PPE with color of dots corresponding to the parameter value for three parameters. Each panel corresponds to a different parameter; the three most influential parameters with respect to the *cloud feedback* are represented. Gray shading covers the region where E_{NET} values are equal to or smaller than those of the default simulation. The default simulation is encircled in black.

RMSD. The six assessed feedbacks nearly sum to the total cloud feedback, so the total cloud feedback must approach the assessed value as the RMSD approaches 0 unless there is a large unassessed component of the feedback. Importantly, it should be kept in mind that the assessed cloud feedback values are only estimates; they do not represent the true cloud feedbacks. In fact, several of the assessed values have been updated by more recent studies (e.g., Wall et al. 2022; McKim et al. 2023). This difference metric is only as valuable as the

assessed value. Another limitation of this approach is that the difference metric only uses the mean of the assessed range and neglects the estimated spread associated with each of the six cloud feedback components. For these reasons, we put more emphasis on the constraint made using mean-state cloud errors than on the constraint made by comparing with WCRP values.

We plot the total cloud feedback as a function of RMSD in Fig. 6a and as a function of E_{NET} in Fig. 6b. For reference, the
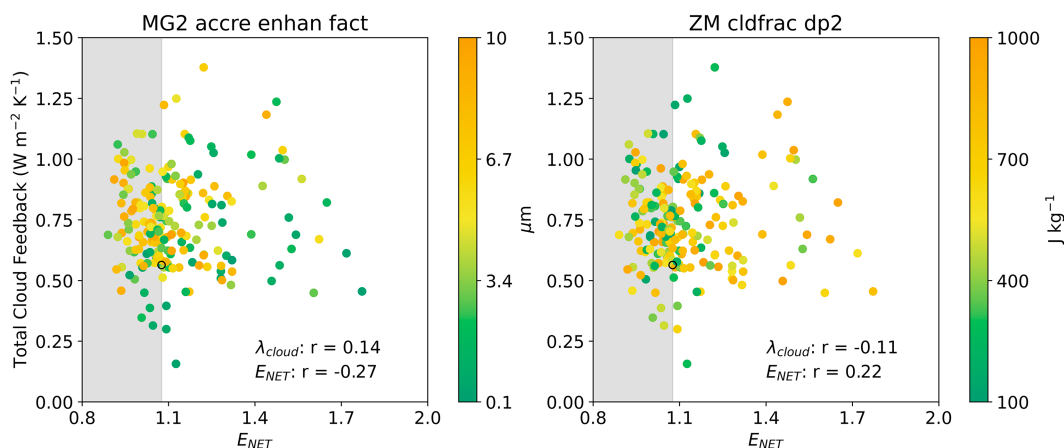

FIG. 5. As in Fig. 4, but the two most influential parameters with respect to E_{NET} are shown.
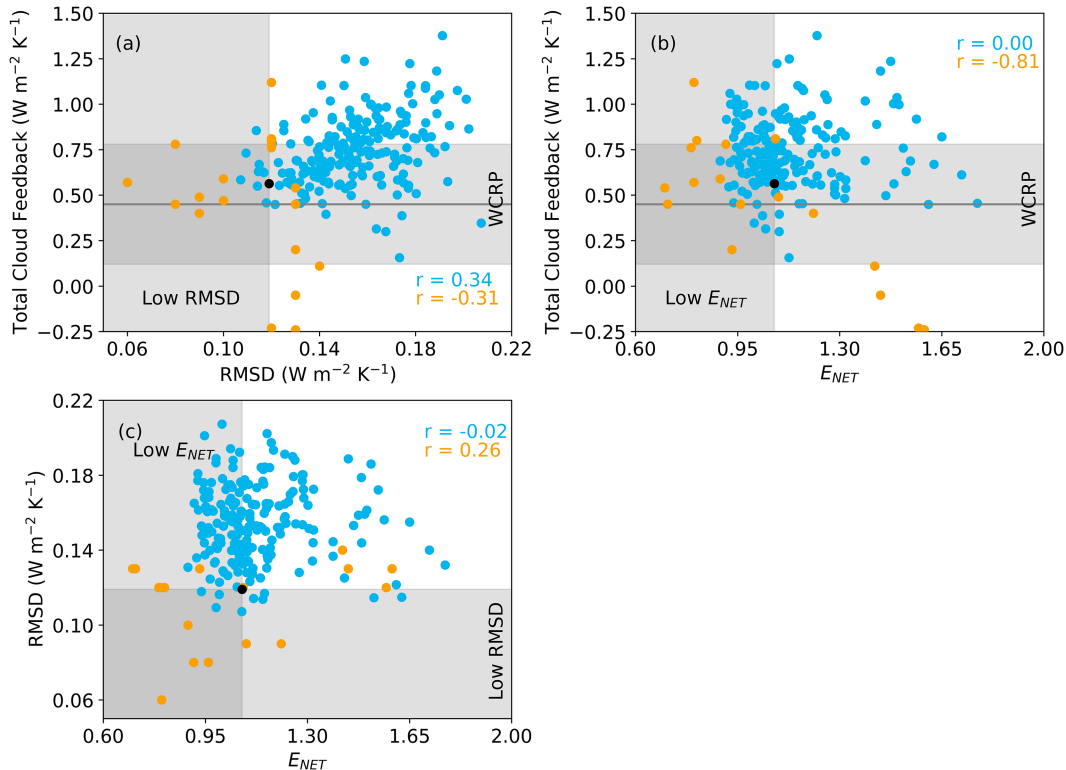
FIG. 6. Scatterplot of the total cloud feedback vs the cloud feedback (a) RMSD and (b) E_{NET} in the CAM6 PPE simulations (blue) and the CMIP5 and CMIP6 models (orange). The black dot denotes the default PPE simulation. In (a) and (b), the dark gray horizontal line is the WCRP-assessed cloud feedback value, and the horizontal gray shaded region covers the 1σ WCRP range. The vertical gray shaded region covers the region where RMSD and E_{NET} values are equal to or smaller than those of the default PPE simulation. (c) A scatterplot of RMSD vs E_{NET}. The horizontal gray shaded region covers RMSD values equal to or smaller than those of the CAM6 PPE default simulation, and the vertical gray shaded region covers E_{NET} values equal to or smaller than those of the CAM6 PPE default simulation.

CMIP5 and CMIP6 values are also included in the plot and are the same as those from Figs. 3 and 4 of Zelinka et al. (2022). Looking at RMSD in Fig. 6a, we find that the CAM6 PPE simulations exhibit some similar behavior to the CMIP models: simulations with a small RMSD tend to have more moderate cloud feedbacks than those with large RMSD. The large-RMSD models include both anomalously small and large cloud feedbacks. However, there are a number of differences in the behavior between the CAM6 PPE and the CMIP models of Zelinka et al. (2022). The CAM6 PPE members tend to have larger RMSD values than the CMIP models. The smallest RMSD value among the CAM6 PPE members is about 0.11 W m^{-2} K^{-1}, whereas there are seven CMIP models with smaller RMSD values. Further, the distribution of total cloud feedbacks is shifted toward larger values in the CAM6 PPE as compared to CMIP models. Looking at Fig. 3, both of these differences may be the result of comparatively large high-cloud altitude feedbacks in the CAM6 PPE.

Looking at Fig. 6b, the mean-state cloud errors tell a different story than that of the RMSD. In the CMIP models, the total cloud feedback is anticorrelated with E_{NET}, suggesting E_{NET} constrains cloud feedbacks to be large. On the other hand, in the CAM6 PPE, the total cloud feedback is uncorrelated with E_{NET}, and there is a large spread in cloud feedbacks among the small E_{NET} ensemble members, suggesting that mean-state cloud errors do not provide an effective constraint on the total cloud feedback in the CAM6 PPE. Figure 6c plots the relationship between mean-state cloud errors (E_{NET}) and WCRP difference metric (RMSD), and we find that, consistent with the CMIP models, there is no relationship between the two metrics in the CAM6 PPE.

We conclude that we are unable to robustly constrain the total cloud feedback using the CAM6 PPE with available information. We leave further understanding of the lack of relationship between E_{NET} and RMSD to future work. We also leave further understanding of the inconsistent relationship between E_{NET} and the cloud feedback across the CAM6 PPE and CMIP6 ensembles to future work.

6. CAM5 to CAM6

Given the substantial influence of parameters on cloud feedbacks in CAM6, and the large increase in ECS from CAM5 to CAM6, we next evaluate whether changes in parameter values (parameter tuning) between CAM5 and CAM6 contribute to
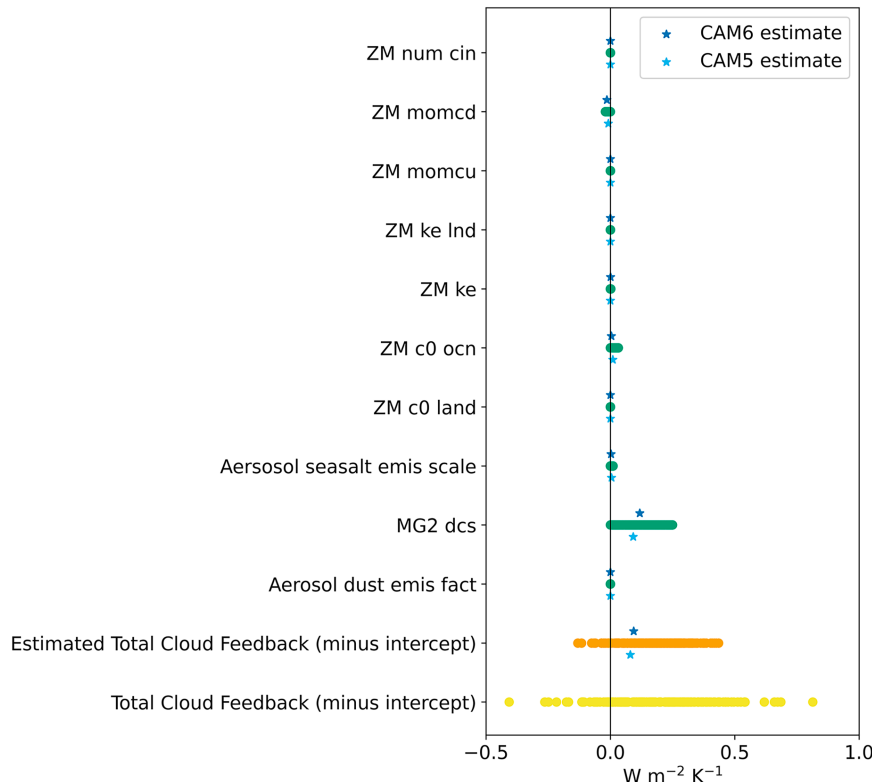
Fig. 7. Total cloud feedback in each PPE simulation (yellow), the LASSO regression estimate of the total cloud feedback (orange), and the contribution to that estimate from the 10 parameters that 1) appear in both CAM5 and CAM6 and 2) changed from CAM5 to CAM6 (green). Stars denote the estimated total cloud feedback and the contributions from each parameter using CAM5 (light blue, below) and CAM6 (dark blue, above) parameter values. The regression intercept has been subtracted from both the estimated and actual cloud feedback values.

the increase in the total cloud feedback and ECS (Gettelman et al. 2019).

To address this question, we develop a simple linear regression model of total cloud feedback as a function of each parameter value. We again use a LASSO regression to avoid overfitting. The regression is similar to those of Eqs. (4) and (6), except here the total cloud feedback is estimated directly instead of as the sum of the cloud feedback components. The simple regression model has a root-mean-square error (RMSE) of 0.13 W m^{-2} K^{-2}. Figure 7 compares the regression model estimate to the actual cloud feedback (minus the intercept). Looking at Fig. 7, it is clear that this simple approach struggles to capture the very large and very small total cloud feedback values. Still, the RMSE value and Fig. 7 provide enough confidence in our simple model estimate of the total feedback to proceed.

To use this simple model to evaluate the role of parameter values in changes in the total cloud feedback between CAM5 and CAM6, we create "CAM5" and "CAM6" parameter sets. The CAM6 parameter set is the default parameter set. On the other hand, the CAM5 parameter set uses the CAM5 values for the 10 parameters which 1) appear in both CAM5 and CAM6 and 2) change from CAM5 to CAM6 and default

parameter values for the remaining parameters. Each parameter set is scaled to fall between 0 and 1 according to Eq. (5). Figure 7 shows the contribution of each of these 10 parameters to the estimated total cloud feedback, which is the parameter value multiplied by its corresponding regression coefficient. Total cloud feedback values are then estimated by plugging the CAM5 and CAM6 parameter sets into the regression model. The CAM5 and CAM6 estimates and their contributions from each parameter are shown in Fig. 7.

Examining Fig. 7, we learn that there is a negligible change in the estimated total cloud feedback due to differences in parameters from CAM5 to CAM6: only a 0.01 W m^{-2} K^{-1} increase. For reference, Gettelman et al. (2019) found an increase of 0.27 W m^{-2} K^{-1} from CAM5 to CAM6. This analysis points to structural changes to the model as the cause of the increase in cloud feedbacks, not parameter tuning.

7. Differences in total cloud feedback across CESM2 experiments

From the "total cloud feedback" row of Fig. 3, it is clear that the CESM2 (CMIP) simulation has a larger total cloud feedback (0.96 W m^{-2} K^{-1}) than the CESM2 (AMIP)

simulation (0.72 W m^{-2} K^{-1}), which is larger still than that of the CAM6 PPE default simulation (0.56 W m^{-2} K^{-1}).[4] These discrepancies are quite large. To demonstrate that these discrepancies in cloud feedback are substantial, we solve for their corresponding ECS values using CESM2 (CMIP) values of ΔF and $\lambda_{noncloud}$. We find substantially different corresponding ECS values of 5.1 K in CESM2 (CMIP), 3.7 K in CESM2 (AMIP), and 3.1 K in CAM6 PPE default.

The difference in total cloud feedback between CESM2 (CMIP) and CESM2 (AMIP) may be attributable, not exhaustively, to 1) different abilities of these simulations to represent the "pattern" effect and/or 2) differences in high-latitude optical depth feedbacks. In GCMs, the pattern effect describes a time evolution of radiative feedbacks, including cloud radiative feedbacks, which is largely attributable to the evolution of SST patterns over time (Dong et al. 2019). The CESM2 (CMIP) cloud feedback is calculated using 150-yr piControl and abrupt-4xCO$_2$ simulations. In contrast, the CESM2 (AMIP) cloud feedback is calculated using 25-yr amip and amip_p4K (i.e., uniform warming) simulations. Therefore, the pattern effect influences the cloud feedback in the CESM2 (CMIP) simulation but is not represented by the CESM2 (AMIP) simulation because it has uniform 4-K warming. We note that the pattern effect is large in CESM2 as compared to other models and an observation-based estimate (Andrews et al. 2022). Another possible explanation for the larger cloud feedback in CESM2 (CMIP) than CESM2 (AMIP) is the evolution of a high-latitude optical depth feedback. This feedback is negative and goes to zero with warming and is the result of decreases in ice clouds and increases in liquid cloud over the Southern Ocean with warming. However, as the planet warms, there are fewer ice clouds and more liquid clouds, so this feedback approaches zero. This feedback increases from -1.25 W m^{-2} K^{-1} in the first 15 years to -0.02 W m^{-2} K^{-1} in the last 15 years of a 150-yr fully coupled abrupt-4xCO$_2$ simulation of CESM2 (Bjordal et al. 2020).[5] The value of this feedback in CESM2 (AMIP) is -0.97 W m^{-2} K^{-1}. We do not have the necessary output to calculate this feedback in CESM2 (CMIP), but we hypothesize that it is likely larger (less negative) in the 150-yr abrupt-4xCO$_2$ CESM2 (CMIP) simulation than in the CESM2 (AMIP) simulation. Therefore, this high-latitude optical depth feedback is another potential explanation for the discrepancy in cloud feedbacks between the CESM2 (AMIP) and CESM2 (CMIP).

On the other hand, the difference in total cloud feedback between the default CAM6 PPE simulation (0.56 W m^{-2} K^{-1}) and the CESM2 (AMIP) simulation (0.72 W m^{-2} K^{-1}) is surprising. Differences between the CESM2 (AMIP) and CAM6 PPE default simulations include the following: 1) the CAM6 PPE default simulation is only 3 years long, while the CESM2 (AMIP) simulation is 25 years long; 2) the CAM6 PPE default simulation is forced with different SSTs than the CESM2 (AMIP) simulation (CAM6 PPE SSTs do not have interannual variability, while CESM2 (AMIP) SSTs do); and 3) CESM2 (AMIP) uses CAM6.0, whereas the CAM6 PPE default simulations use CAM6.3.[6] To evaluate whether simulation length or interannual variability can explain the discrepancy in total cloud feedback, we calculate the interannual spread in total cloud feedback for each of the 25 years of the CESM2 (AMIP) simulation and for a 12-yr version of the CAM6 PPE default simulation (Fig. 3). We find that neither spread explains the difference in total cloud feedback.

To further evaluate which of these differences explains the discrepancy in total cloud feedback, we run a 3-yr simulation using CAM6.0 [the version of CAM used in CESM2 (AMIP)] forced with F2000climo SSTs (the same SSTs as CAM6 PPE). This simulation, which we refer to as CAM6.0, differs from the CAM6 PPE default (CAM6.3) simulation only in model generation. The total feedback in CAM6.0 (F2000climo) is 0.81 W m^{-2} K^{-1}, which is even larger than that in the CESM2 (AMIP) simulation (0.72 W m^{-1}), falling near the high end of interannual variability (cf. the black square and its whiskers with the black triangle in Fig. 3). Note that the difference between the CESM2 (AMIP) simulation and CAM6.0 (F2000climo) simulation is likely attributable to the difference in SST between the two. The AMIP SSTs show a strong pattern effect over the duration of the historical period which is not captured by F2000climo, which is cyclic (Andrews et al. 2022). Finally, by comparing the black circle with the black triangle in Fig. 3, it appears that modifications to CAM6 made between CAM6.0 and CAM6.3 have reduced the climate sensitivity. Parameter values did not change between these two experiments, so this difference is attributable entirely to structural modifications to CAM6.

Having identified that structural changes between model versions impact the total cloud feedback, we next evaluate the difference between the six cloud feedback components in CAM6.0 and CAM6.3 (Fig. 3). Figure 3 reveals that the difference in the cloud feedback is primarily due to differences in high-cloud altitude, tropical marine low-cloud, and unassessed cloud feedbacks. We further evaluate the difference between cloud feedbacks between CAM6.0 and CAM6.3 by plotting the high-, midlevel-, and low-cloud contributors to both the SW and LW cloud feedbacks (Fig. 8). From Fig. 8, it is clear that the decrease in cloud feedbacks is primarily due to the SW contribution, with small decreases due to the LW contribution. Further, the SW contribution is dominated by the low-cloud contribution in the tropics and midlatitudes. To the extent that the LW cloud feedback decreases is mostly due to changes in the high-cloud contribution in the tropics and subtropics. We leave further evaluation of the difference in the cloud feedback between CAM6.0 and CAM6.3 to future work.

---

[4] The cloud feedback in amip-future4K is 0.50 W m^{-2} K^{-1}, lower than any of the three simulations we focus on in this paper. This may be the result of very different SST patterns, especially in the tropical Pacific.

[5] Bjordal et al. (2020) calculate the high-latitude low-cloud optical depth feedback slightly differently from Sherwood et al. (2020) and elsewhere in this paper. Further, the details of the CESM2 simulations of Bjordal et al. (2020) are slightly different from those of CESM2 (CMIP).

---

[6] More specifically, CESM2 (AMIP) uses cam6-0-026 and the CAM6 PPE is run with cam6-3-026.
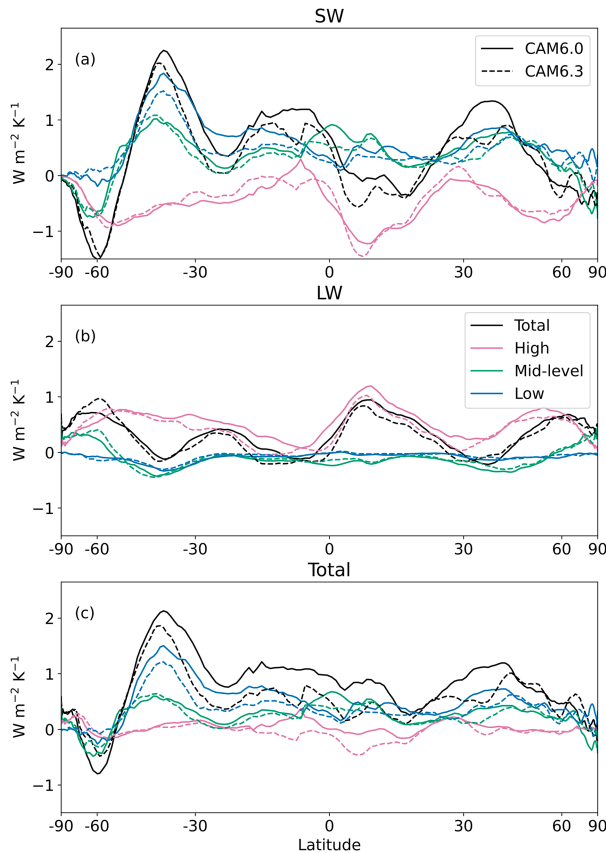
FIG. 8. Zonal-mean cloud feedbacks in CAM6.0 (solid) and in CAM6.3 (dashed). (c) The total cloud feedback is decomposed into (a) SW and (b) LW contributions. The total cloud feedback (black) is also decomposed into high- (pink), midlevel- (green), and low-cloud (blue) contributions. The x axis is scaled as the sine of latitude.

8. Comparison with PaleoCalibr

To further understand our results, we compare the CAM6 PPE used here to another PPE which used the Paleoclimate-calibrated CAM6 (CAM6-PaleoCalibr), which has a smaller ECS than the standard version of CAM6 (Zhu et al. 2022). CAM6-PaleoCalibr removed an inappropriate limiter on the cloud ice number that is present in other versions of CAM6 and uses a shorter time step in the MG2 microphysics scheme. CAM6-PaleoCalibr also uses a coarser resolution of $1.9° \times 2.5°$ than the CAM6 PPE simulations, which have a resolution of $0.9° \times 1.25°$. As a result of differences in horizontal resolution, a few of the default parameters are tuned differently, including a smaller MG2_DCS value, a smaller dust emission factor, a smaller CLUBB_gamma, and a larger seasalt emission scaling in the coarser version. In addition, the PaleoCalibr PPE uses year 1850 levels of atmospheric constituents and control climate SSTs, whereas the CAM6 PPE uses year 2000 levels. Finally, the PaleoCalibr PPE uses a larger range of CLUBB_C8 and CLUBB_C11b to explore their influence on the state dependence of the cloud feedback. The PaleoCalibr default

simulation has a cloud feedback of 0.36 W m^{-2} K^{-2}, which is smaller than that of the CAM6 PPE used above (0.56 W m^{-2} K^{-2}). Figure S8 compares the assessed cloud feedbacks in the PaleoCalibr PPE to those in Fig. 3. Although the PaleoCalibr PPE default simulation has a smaller total cloud feedback than the CAM6 PPE default simulation, the relatively large high-cloud altitude feedback persists in the PaleoCalibr PPE, indicating that the changes made to reduce the total cloud feedback in the PaleoCalibr version do not address this discrepancy.

9. Discussion and conclusions

The CAM6 PPE provides a tool for evaluating the sensitivity of cloud feedbacks to uncertain model parameters. Surprisingly, the CAM6 PPE nearly captures the large spread across CMIP6 models. We did not expect this a priori, although a similar spread was found in the Met Office's atmospheric GCM (Rostron et al. 2020). Further, the decomposition of these cloud feedbacks into 1) their SW and LW components and 2) the cloud feedback components of Sherwood et al. (2020) shows similar spreads across the two ensembles. Taking the spread across CMIP6 models to represent structural uncertainty and the spread across the CAM6 PPE to represent parametric uncertainty, these results indicate that the spread in cloud feedbacks due to parametric uncertainty is similar to that due to structural uncertainty, though this does not imply that changes in parameters explain differences across models. An advantage of using a PPE is that each parameter corresponds to a physical process, and the spread in cloud feedbacks due to each parameter and corresponding process can be estimated. The parameters with the greatest influence on cloud feedbacks across the CAM6 PPE are CLUBB_C8, MG2_DCS, and ZM_capelmt. These parameters control aspects of cloud thickness, cloud ice, and deep convection, respectively.

We further use the CAM6 PPE to identify the parameters with the most influence on total cloud feedback and mean-state cloud errors. We find that the most influential parameters (by correlation) with mean-state cloud errors are MG2 accretion enhancement factor (MG2_accre_enhan_factor) and ZM deep convective cloud fraction (ZM_cldfrac_dp2). We find that the most influential parameters (by correlation) with total cloud feedback are MG2 ice–snow autoconversion size threshold (MG2_DCS), ZM triggering threshold for convection (ZM_capelmt), and a CLUBB coefficient in the C8 skewness equation (CLUBB_C8) and point out the values which are associated with the smallest cloud feedback errors. Therefore, an improved representation of the processes corresponding to these parameters in models is a promising avenue for constraining simulated cloud feedbacks in CAM6 and perhaps other models too. However, we intentionally stop short of recommending these values for model *tuning*. First, climate model simulations of warmer climates should not be tuned to a desired "answer" since there are no observations with which to tune. Additionally, as mentioned above, the assessment of Sherwood et al. (2020), while the most comprehensive to date, does not represent the truth. Therefore, climate models should not be developed or tuned to match these values. Further, the model may be subject to equifinality; that is,

different combinations of parameters may yield the same cloud feedback error. Moreover, cloud feedbacks are not the only simulated feature of the model of interest but are the only subject of this analysis. Finally, it is clear from this analysis that ECS and cloud feedbacks are very sensitive to both parameter values and structural differences in the model, so would not provide a strong constraint for model tuning.

The range of a high-cloud altitude feedback in the CAM6 PPE sits at larger values than that of CMIP6 models, AMIP models, and the WCRP assessment. More specifically, the CAM6 PPE simulations tend to have larger high-cloud altitude feedbacks than the assessed 1σ range and than most CMIP6 and AMIP models, and CESM2 has the largest high-cloud altitude feedback for any CMIP6 model. In contrast, Sherwood et al. (2020) consider the high-cloud altitude feedback to be among the more certain of the six cloud feedback components. Their assessed value reflects the mean and standard deviation across GCMs, with support from observations, cloud-resolving models, and large eddy simulations. In the CAM6 PPE, the parameters most correlated with the high-cloud altitude feedback are the convective parcel temperature perturbation parameter in the convection scheme (ZM_tiedke_add) and a scaling for subgrid velocity of ice activation in the microphysics scheme (microp_aero_wsubi_scale). Further investigation of the relatively large high-cloud altitude feedback in CESM2 should be a priority for future investigations.

We use the CAM6 PPE to attempt to constrain the total cloud feedback by comparing 1) a measure of the cloud feedback discrepancy (as compared with Sherwood et al. 2020) and 2) a measure of mean-state cloud errors in each CAM6 PPE member to the total cloud feedback. We find inconsistent results between the two metrics. Comparing with CMIP models, we find that WCRP-assessed values provide a similar constraint in the CAM6 PPE as in the CMIP models (moderate cloud feedback). However, we find that despite the strong anticorrelation between mean-state cloud errors and total cloud feedback in CMIP models, mean-state cloud errors and total cloud feedback are uncorrelated in the CAM6 PPE, which is broadly consistent with the results of Tsushima et al. (2020), who used a PPE created with the Met Office's atmospheric GCM. The mean-state cloud errors and the difference from WCRP-assessed values are uncorrelated with one another in both the CMIP ensemble and CAM6 PPE. While the comparison with WCRP-assessed values provides similar constraints in both the CMIP ensemble and CAM6 PPE, this analysis is subject to several important limitations, both here and in Zelinka et al. (2022). Sherwood et al. (2020) estimate the uncertainty of each cloud feedback component, and this estimated uncertainty is not accounted for in the error metric used here. Further, the assessment of Sherwood et al. (2020) has limitations, and more recent work has updated the assessed values (e.g., Myers et al. 2021). Finally, since the six cloud feedback components roughly sum to the total, the total cloud feedback estimate must approach that of Sherwood et al. (2020) as RMSD approaches zero, provided that the unassessed feedbacks are small. Therefore, we find our analysis which uses the CAM6 PPE to constrain total cloud feedback inconclusive.

Given the substantial influence of parameters on cloud feedbacks in the CAM6 PPE, we ask whether parameter tuning is responsible for the increase in the cloud feedback and thus ECS from CAM5 to CAM6. Using a simple regression model, we estimate that changes in parameters between CAM5 and CAM6 have little influence on the total cloud feedback and therefore are not responsible for the increase in the cloud feedback between generations. This indicates that structural changes to the model between CAM5 and CAM6 are responsible for the increase in total cloud feedback and ECS.

We evaluate the smaller cloud feedback in the CAM6 PPE default simulation as compared to CESM2 (AMIP). We find that changes to the model between CAM6.0 [used for CESM2 (AMIP)] and CAM6.3 (used for the CAM6 PPE) explain the reduced cloud feedback. It is unclear whether the decrease in the cloud feedback between CAM6.0 and CAM6.3 foreshadows smaller cloud feedbacks in future versions of CESM because ultimately the model's cloud feedback will depend heavily on the role of coupling and further changes during model development.

Overall, our analysis highlights the large sensitivity of ECS estimates to both changes in parameters and structural changes in CAM6, which may help explain why it has been so difficult to constrain.

REFERENCES

Andrews, T., and Coauthors, 2022: On the effect of historical SST patterns on radiative feedback. *J. Geophys. Res. Atmos.*, **127**, e2022JD036675, https://doi.org/10.1029/2022JD036675.

Bjordal, J., T. Storelvmo, K. Alterskjær, and T. Carlsen, 2020: Equilibrium climate sensitivity above 5°C plausible due to state-dependent cloud feedback. *Nat. Geosci.*, **13**, 718–721, https://doi.org/10.1038/s41561-020-00649-1.

Bodas-Salcedo, A., and Coauthors, 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, **92**, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1.

Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**, L20806, https://doi.org/10.1029/2005GL023851.

Bretherton, C. S., P. N. Blossey, and C. Stan, 2014: Cloud feedbacks on greenhouse warming in the superparameterized climate model SP-CCSM4. *J. Adv. Model. Earth Syst.*, **6**, 1185–1204, https://doi.org/10.1002/2014MS000355.

Ceppi, P., D. T. McCoy, and D. L. Hartmann, 2016: Observational evidence for a negative shortwave cloud feedback in middle to high latitudes. *Geophys. Res. Lett.*, **43**, 1331–1339, https://doi.org/10.1002/2015GL067499.

Dagon, K., B. M. Sanderson, R. A. Fisher, and D. M. Lawrence, 2020: A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5. *Adv. Stat. Climatol. Meteor. Oceanogr.*, **6**, 223–244, https://doi.org/10.5194/ascmo-6-223-2020.

Danabasoglu, G., and Coauthors, 2020: The Community Earth System Model version 2 (CESM2). *J. Adv. Model. Earth Syst.*, **12**, e2019MS001916, https://doi.org/10.1029/2019MS001916.

Dong, Y., C. Proistosescu, K. C. Armour, and D. S. Battisti, 2019: Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's function approach: The preeminence of the western Pacific. *J. Climate*, **32**, 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1.

Eidhammer, T., A. Gettelman, and K. Thayer-Calder, 2022: CESM2.2-CAM6 perturbed parameter ensemble (PPE). UCAR/NCAR, accessed December 2022, https://doi.org/10.26024/BZNE-YF09.

Forster, P., and Coauthors, 2021: The Earth's energy budget, climate feedbacks, and climate sensitivity. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 923–1054.

Gettelman, A., and Coauthors, 2019: High climate sensitivity in the Community Earth System Model version 2 (CESM2). *Geophys. Res. Lett.*, **46**, 8329–8337, https://doi.org/10.1029/2019GL083978.

Gregory, J. M. and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, **31**, L03205, https://doi.org/10.1029/2003GL018747.

Hartmann, D. L., and K. Larson, 2002: An important constraint on tropical cloud-climate feedback. *Geophys. Res. Lett.*, **29**, 1951, https://doi.org/10.1029/2002GL015835.

——, L. A. Moy, and Q. Fu, 2001: Tropical convection and the energy balance at the top of the atmosphere. *J. Climate*, **14**, 4495–4511, https://doi.org/10.1175/1520-0442(2001)014<4495:TCATEB>2.0.CO;2.

Kamae, Y., T. Ogura, M. Watanabe, S.-P. Xie, and H. Ueda, 2016: Robust cloud feedback over tropical land in a warming climate. *J. Geophys. Res. Atmos.*, **121**, 2593–2609, https://doi.org/10.1002/2015JD024525.

Kay, J. E., B. Medeiros, Y.-T. Hwang, A. Gettelman, J. Perket, and M. G. Flanner, 2014: Processes controlling Southern Ocean shortwave climate feedbacks in CESM. *Geophys. Res. Lett.*, **41**, 616–622, https://doi.org/10.1002/2013GL058315.

Klein, S. A., Y. Zhang, M. D. Zelinka, R. Pincus, J. Boyle, and P. J. Gleckler, 2013: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator. *J. Geophys. Res. Atmos.*, **118**, 1329–1342, https://doi.org/10.1002/jgrd.50141.

——, A. Hall, J. R. Norris, and R. Pincus, 2017: Low-cloud feedbacks from cloud-controlling factors: A review. *Surv. Geophys.*, **38**, 1307–1329, https://doi.org/10.1007/s10712-017-9433-3.

McCoy, D. T., R. Eastman, D. L. Hartmann, and R. Wood, 2017: The change in low cloud cover in a warmed climate inferred from AIRS, MODIS, and ERA-Interim. *J. Climate*, **30**, 3609–3620, https://doi.org/10.1175/JCLI-D-15-0734.1.

McKay, M. D., R. J. Beckman, and W. J. Conover, 2000: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **42**, 55–61, https://doi.org/10.1080/00401706.2000.10485979.

McKim, B., S. Bony, and J.-L. Dufresne, 2023: Physical and observational constraints on the anvil cloud area feedback. *ESS Open Archive*, preprint, https://doi.org/10.22541/au.167769953.39966398/v2.

Myers, T. A., R. C. Scott, M. D. Zelinka, S. A. Klein, J. R. Norris, and P. M. Caldwell, 2021: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity. *Nat. Climate Change*, **11**, 501–507, https://doi.org/10.1038/s41558-021-01039-0.

National Research Council, 1979: *Carbon Dioxide and Climate: A Scientific Assessment*. National Academies Press, 34 pp., https://doi.org/10.17226/12181.

Rostron, J. W., D. M. H. Sexton, C. F. McSweeney, K. Yamazaki, T. Andrews, K. Furtado, M. A. Ringer, and Y. Tsushima, 2020: The impact of performance filtering on climate feedbacks in a perturbed parameter ensemble. *Climate Dyn.*, **55**, 521–551, https://doi.org/10.1007/s00382-020-05281-8.

Rugenstein, M. A. A., and K. C. Armour, 2021: Three flavors of radiative feedbacks and their implications for estimating equilibrium climate sensitivity. *Geophys. Res. Lett.*, **48**, e2021GL092983, https://doi.org/10.1029/2021GL092983.

Sherwood, S. C., and Coauthors, 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.*, **58**, e2019RG000678, https://doi.org/10.1029/2019RG000678.

Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406, https://doi.org/10.1038/nature03301.

Tan, I., T. Storelvmo, and M. D. Zelinka, 2016: Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, **352**, 224–227, https://doi.org/10.1126/science.aad5300.

Tsushima, Y., M. A. Ringer, G. M. Martin, J. W. Rostron, and D. M. H. Sexton, 2020: Investigating physical constraints on climate feedbacks using a perturbed parameter ensemble. *Climate Dyn.*, **55**, 1159–1185, https://doi.org/10.1007/s00382-020-05318-y.

Wall, C. J., T. Storelvmo, J. R. Norris, and I. Tan, 2022: Observational constraints on Southern Ocean cloud-phase feedback. *J. Climate*, **35**, 5087–5102, https://doi.org/10.1175/JCLI-D-21-0812.1.

Watson-Parris, D., and Coauthors, 2022: ClimateBench v1.0: A benchmark for data-driven climate projections. *J. Adv. Model. Earth Syst.*, **14**, e2021MS002954, https://doi.org/10.1029/2021MS002954.

Williams, I. N., and R. T. Pierrehumbert, 2017: Observational evidence against strongly stabilizing tropical cloud feedbacks. *Geophys. Res. Lett.*, **44**, 1503–1510, https://doi.org/10.1002/2016GL072202.

Zelinka, M. D., and D. L. Hartmann, 2010: Why is longwave cloud feedback positive? *J. Geophys. Res.*, **115**, D16117, https://doi.org/10.1029/2010JD013817.

——, S. A. Klein, and D. L. Hartmann, 2012: Computing and partitioning cloud feedbacks using cloud property histograms. Part I: Cloud radiative kernels. *J. Climate*, **25**, 3715–3735, https://doi.org/10.1175/JCLI-D-11-00248.1.

——, T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor, 2020: Causes of higher climate sensitivity in CMIP6 models. *Geophys. Res. Lett.*, **47**, e2019GL085782, https://doi.org/10.1029/2019GL085782.

——, S. A. Klein, Y. Qin, and T. A. Myers, 2022: Evaluating climate models' cloud feedbacks against expert judgment. *J. Geophys. Res. Atmos.*, **127**, e2021JD035198, https://doi.org/10.1029/2021JD035198.

Zhai, C., J. H. Jiang, and H. Su, 2015: Long-term cloud change imprinted in seasonal cloud variation: More evidence of high climate sensitivity. *Geophys. Res. Lett.*, **42**, 8729–8737, https://doi.org/10.1002/2015GL065911.

Zhu, J., and Coauthors, 2022: LGM paleoclimate constraints inform cloud parameterizations and equilibrium climate sensitivity in CESM2. *J. Adv. Model. Earth Syst.*, **14**, e2021MS002776, https://doi.org/10.1029/2021MS002776.