Potential Problems Measuring Climate Sensitivity from the Historical Record

ANDREW E. DESSLER

Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

ABSTRACT

This study investigates potential biases between equilibrium climate sensitivity inferred from warming over the historical period (ECS_{hist}) and the climate system's true ECS (ECS_{true}). This paper focuses on two factors that could contribute to differences between these quantities. First is the impact of internal variability over the historical period: our historical climate record is just one of an infinity of possible trajectories, and these different trajectories can generate ECS_{hist} values 0.3 K below to 0.5 K above (5%–95% confidence interval) the average ECS_{hist}. Because this spread is due to unforced variability, I refer to this as the unforced pattern effect. This unforced pattern effect in the model analyzed here is traced to unforced variability in loss of sea ice, which affects the albedo feedback, and to unforced variability in warming of the troposphere, which affects the shortwave cloud feedback. There is also a forced pattern effect that causes ECS_{hist} to depart from ECS_{true} due to differences between today's transient pattern of warming and the pattern of warming at $2\times CO_2$ equilibrium. Changes in the pattern of warming lead to a strengthening low-cloud feedback as equilibrium is approached in regions where surface warming is delayed: the Southern Ocean, eastern Pacific, and North Atlantic near Greenland. This forced pattern effect causes ECS_{hist} to be on average 0.2 K lower than ECS_{true} ($\sim 8\%$). The net effect of these two pattern effects together can produce an estimate of ECS_{hist} as much as 0.5 K below ECS_{true}.

1. Introduction

Equilibrium climate sensitivity (ECS; i.e., the equilibrium warming in response to a doubling of CO_2) is one of the quantities that controls how much future warming we will experience in response to greenhouse gas emissions from anthropogenic activities. As such, it is frequently viewed as one of the most important numbers in climate science and much effort has been expended over decades attempting to constrain its value.

ECS can be calculated from observations or models as

$$ECS = -\frac{F_{2\times CO2}}{\lambda}, \qquad (1)$$

where $F_{2\times CO2}$ is the radiative forcing from doubled CO_2 and λ represents the top-of-atmosphere (TOA) flux change per degree of surface temperature change:

$$\lambda = \frac{\Delta R - \Delta F}{\Delta T_S}, \qquad (2)$$

where T_S is the global average surface temperature, R is the TOA flux, and F is the radiative forcing.

Some of the most influential estimates of ECS come from the observed warming during the historical period, between the mid-nineteenth century and today (referred to as ECS_{hist}). To estimate λ over this period (referred to as λ_{hist}), Δ in Eq. (2) represents the change between the mid-nineteenth century and the early twenty-first century. ECS_{hist} is then calculated using Eq. (1) and λ_{hist}.

There have been many estimates of ECS_{hist} from observations [summarized in Forster (2016); see also Knutti et al. (2017)]. These tend to be lower than ECS estimated from other sources and they anchor the lower end of the IPCC's canonical ECS range of 1.5–4.5 K. Recently, it has been argued that ECS_{hist} may not provide a good estimate of our climate system's true ECS (hereafter ECS_{true}). This is based on demonstrations in models that ΔR depends not just on how much warming occurs, but also on how that warming is distributed across the globe (Armour et al. 2013; Andrews et al. 2015; Zhou et al. 2016, 2017). In other words, two climate states with the same ΔT_s, but distributed differently, can have different values of ΔR, leading

Corresponding author: Andrew Dessler, adessler@tamu.edu

to different estimates of ECS_{hist} (Olson et al. 2013; Huber and Knutti 2014). Following standard practice, I will refer to this as the ''pattern effect'' (Stevens et al. 2016).

The pattern effect causes ECS_{hist} to depart from ECS_{true} if the aspects of the warming we experienced over the historical period differ from aspects of warming at $2 \times CO_2$ equilibrium. I intentionally leave vague what is meant by ''aspects'' as these will be investigated in detail in the paper. Basically, though, there are two different reasons why warming over the historical period may be different from the long-term warming. The first reason is that the historical observational record is just one member of an infinity of possible climate trajectories that Earth could have experienced over the last 150 years. Dessler et al. (2018, hereafter D18) used an ensemble of climate model runs to show that different trajectories could yield widely varying estimates of ECS_{hist}. These differences in ECS_{hist} are due to internal variability, so I will refer to this variability in ECS_{hist} as the ''unforced pattern effect.''

There is also a ''forced pattern effect.'' This is primarily related to the fact that the transient warming pattern over the twentieth century is expected to be different from the equilibrium pattern of warming; this will tend to make ECS_{true} larger than ECS_{hist} (Andrews et al. 2015; Armour 2017; Proistosescu and Huybers 2017; Ceppi and Gregory 2017). Previous analyses (e.g., Marvel et al. 2018; Andrews et al. 2018) have evaluated the combined forced and unforced pattern effects. In this paper, I analyze a large model ensemble to separately evaluate their magnitudes.

2. Model ensemble

We analyze output from various runs of the fully coupled Max Planck Institute Earth System Model version 1.1 (MPI-ESM1.1), collectively referred to as the Grand Ensemble. The MPI-ESM1.1 is a fully coupled climate model from the Max Planck Institute for Meteorology and consists of the ECHAM6.3 atmosphere and land model coupled to the MPI-OM ocean model. The Grand Ensemble is described in detail in Maher et al. (2019).

The MPI-ESM1.1 has a transient climate response of 1.78 K (Adams and Dessler 2019) and an effective climate sensitivity (calculated from a regression of the first 150 years of an abrupt $4 \times CO_2$ run) of 2.72 K. These values are near the middle of the CMIP5 ensemble range. We will analyze a large number of runs from this ensemble:

- A 100-member ensemble of runs with historical forcing (hereafter, the ''historical ensemble''). Each of the 100 members simulates the years 1850–2005 and uses identical historical natural and anthropogenic forcing. The ensemble members differ only in their initial conditions—each starts from a different state sampled from the preindustrial control simulation. This ensemble was used by D18 to characterize the impact of internal variability on ECS_{hist} and by Adams and Dessler (2019) to investigate the impact of internal variability on transient climate response. The ensemble produces a good simulation of the historical record, as seen in Fig. 2 of Maher et al. (2019).
- A 68-member ensemble of runs with CO_2 increasing at 1% per year (hereafter, the ''1% ensemble''). Each of the members is 150 years long and uses identical forcing. Like the historical ensemble, this ensemble's members differ only in their initial conditions.
- An abrupt $4 \times CO_2$ run. In this run, CO_2 is abruptly quadrupled from preindustrial values and then run for 2614 years. At that point, the model is nearly in equilibrium.
- A preindustrial control run. In this run, atmospheric conditions held at preindustrial values for 2000 years. This is the run from which all other runs branch from.

Effective radiative forcing used in this paper is calculated from fixed SST runs of the model. Historical effective radiative forcing (1850–2005) is $2.2 \, W \, m^{-2}$ (D18), while $2 \times CO_2$ and $4 \times CO_2$ forcing are 3.7 and $7.8 \, W \, m^{-2}$, respectively (Adams and Dessler 2019). In all calculations, surface temperature refers to 2-m air temperature.

3. Feedbacks in the historical ensemble

D18 showed that λ_{hist} from the historical ensemble ranged from -1.63 to $-1.17 \, W \, m^{-2} \, K^{-1}$ (5%–95%)—this spread is what I have designated as the unforced pattern effect. To gain physical insight into this, I decompose λ_{hist} into constituent feedbacks using the approach and radiative kernels of Soden et al. (2008), but using the feedback decomposition of Held and Shell (2012), in which the Planck and lapse-rate feedbacks assume constant relative humidity (RH). I will refer to these as the ''conventional'' feedbacks. For consistency with D18, I calculate those feedbacks by differencing relevant fields between the first and last decade of the runs. Picking different periods does not change the conclusions of this section.

One disadvantage of the conventional approach is that the sum of the feedbacks may not equal λ_{hist}, leaving a residual that may be comparable in magnitude

to the pattern effect I am trying to diagnose. To address this, I also calculate feedbacks a second way, based on decomposing R into clear-sky and cloud radiative forcing (CRF) components:

$$R = R_{\text{clear sky LW}} + R_{\text{clear sky SW}} + R_{\text{CRF LW}} + R_{\text{CRF SW}}, \quad (3)$$

where LW and SW refer to longwave and shortwave fluxes, clear-sky fluxes refer to what the fluxes would be in the absence of clouds (leaving everything else the same), and CRF is the all-sky flux minus the clear-sky flux. The change in these fluxes (with the corresponding forcing subtracted off) divided by ΔT_S yields the individual feedbacks. For example, the clear-sky longwave feedback is

$$\lambda_{\text{clear sky LW}} = \frac{\Delta R_{\text{clear sky LW}} - \Delta F_{\text{clear sky LW}}}{\Delta T_S}. \quad (4)$$

The terms $\lambda_{\text{clear sky SW}}$, $\lambda_{\text{CRF LW}}$, and $\lambda_{\text{CRF SW}}$ are all calculated analogously. By construction, the sum of these feedbacks must equal λ_{hist}. I will refer to these as the CRF feedbacks.

Figure 1 shows a comparison between the ensemble-average feedbacks in the two breakdowns. I have grouped similar feedbacks together: $\lambda_{\text{Planck}} + \lambda_{\text{lapse rate}} + \lambda_{\Delta \text{RH}}$ with $\lambda_{\text{clear sky LW}}$, λ_{albedo} with $\lambda_{\text{clear sky SW}}$, $\lambda_{\text{LW cloud}}$ with $\lambda_{\text{CRF LW}}$, and $\lambda_{\text{SW cloud}}$ with $\lambda_{\text{CRF SW}}$. The feedback pairs do not agree exactly because of differences in the underlying physical processes. For example, $\lambda_{\text{clear sky SW}}$ disagrees with λ_{albedo} because $\lambda_{\text{clear sky SW}}$ contains a small fraction of the water vapor feedback caused by SW absorption by water vapor. Differences also arise from cloud-masking effects that mix the cloud and noncloud feedbacks in the CRF breakdown (Soden et al. 2004). A final difference arises because the conventional feedbacks do necessarily sum to λ_{hist}, but leave a small positive residual (λ_{hist} minus the sum of the feedbacks). The ensemble-average residual is $+0.40 \, \text{W m}^{-2} \, \text{K}^{-1}$, with 90% of the residuals falling between 0.31 and $0.50 \, \text{W m}^{-2} \, \text{K}^{-1}$. Overall, though, the two feedback breakdowns give a similar picture of the breakdown of λ_{hist}. Comparisons to the CMIP5 ensemble average also show reasonable agreement (Fig. 1).

Figure 2 shows the latitude distribution of the average and standard deviation of the ensemble. In agreement with observations, the historical ensemble simulates the largest warming in the Northern Hemisphere (NH), although it overestimates warming there [Fig. 2a; see also Fig. 2 of Adams and Dessler (2019)] and simulates least warming in the Southern Hemisphere (SH).
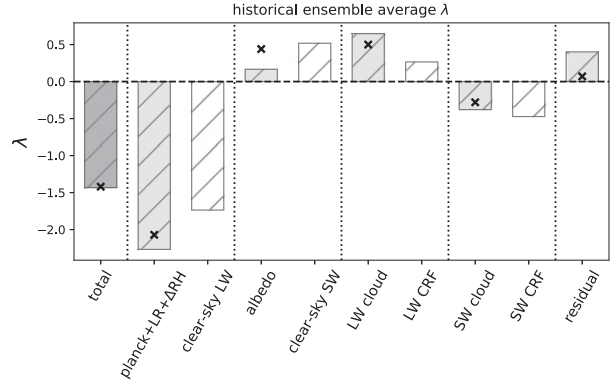

FIG. 1. Comparison of ensemble-average λ $(\text{W m}^{-2} \, \text{K}^{-1})$ from the historical ensemble using two different feedback breakdowns. Gray bars represent the conventional feedbacks, while white bars show the CRF feedbacks; similar feedbacks are grouped together. λ_{hist} is the total feedback, estimated using Eq. (2). Residual is the difference between λ_{hist} and the sum of the feedbacks from the conventional breakdown. By construction, the sum of the CRF feedbacks sum to λ_{hist}. All feedbacks are calculated by differencing the first decade from the last. The bold crosses are the average of the CMIP5 ensemble, taken from Table S1 of Ceppi and Gregory (2017).

The term $\lambda_{\text{clear sky LW}}$ (Fig. 2c) is basically a mirror image of the surface warming pattern, showing that regions with more warming radiate more energy back to space. The term $\lambda_{\text{clear sky SW}}$ is primarily driven by loss of sea ice, so it maximizes in the polar regions (Fig. 2d). Note that $\lambda_{\text{cloud LW}}$ is larger than $\lambda_{\text{CRF LW}}$ at almost all latitudes (Fig. 2e) due to cloud masking effects (Soden et al. 2004) and $\lambda_{\text{cloud SW}}$ and $\lambda_{\text{CRF SW}}$ are similar except in the Arctic (Fig. 2f), where cloud masking effects are also important.

4. Quantifying the unforced pattern effect

D18 calculated λ_{hist} in the historical ensemble and found that the 5%–95% spread in λ_{hist} is -1.63 to -1.17. Given that λ_{hist} is equal to the sum of individual feedbacks, variability in λ_{hist} must be driven by variability in the underlying feedbacks. Figure 3 summarizes this by plotting the average feedbacks in the 10 ensemble members with the highest ECS_{hist} minus the average in the 10 lowest ECS_{hist} ensemble members. This shows that 55% of the unforced pattern effect is due to differences in the $\lambda_{\text{clear sky SW}}$, with most of the remainder, 39%, due to differences in $\lambda_{\text{CRF SW}}$. Differences in the LW feedbacks contribute ~5%.

High ECS_{hist} ensemble members have a larger fraction of warming in the extratropics (and less in the tropics) than the low ECS_{hist} models (Fig. 4a). The difference in the LW clear-sky feedback (Fig. 4c),
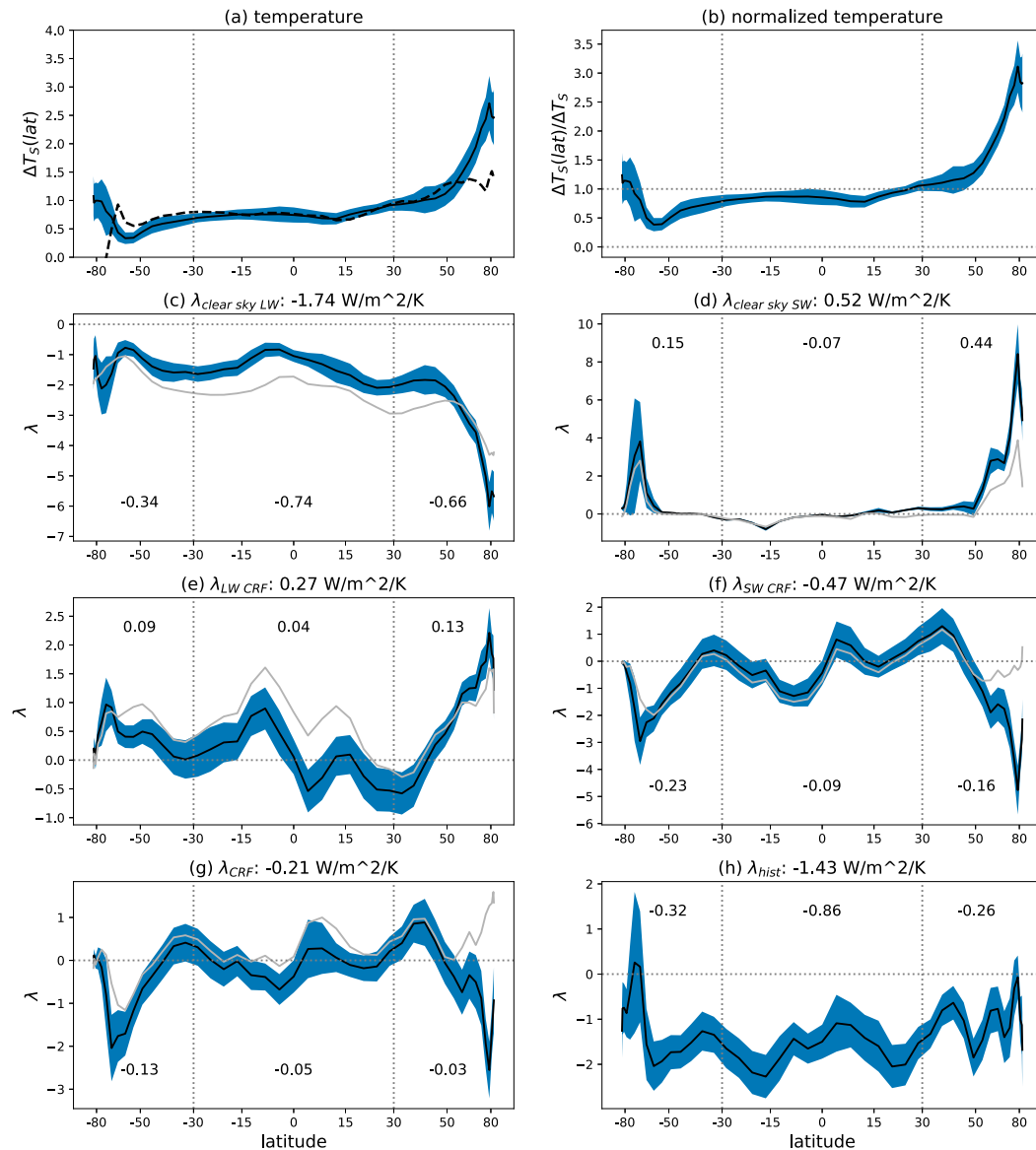
FIG. 2. (a) Zonal- and ensemble-average average ΔT_S (solid line) and the Berkeley Earth surface temperature observations (dashed line) (Rohde et al. 2013). The model has been masked to match Berkeley coverage. (b) Zonal- and ensemble-average normalized warming in the historical ensemble (ΔT_S at each latitude divided by the global average ΔT_S; the global average of this quantity is one by construction). (c)–(h) Zonal- and ensemble-average feedbacks in the historical ensemble. Black lines are from the CRF breakdown; gray lines are the corresponding feedback from conventional breakdown (where appropriate). All are plotted against area-weighted latitude. The three numbers on each panel show the contributions to the feedback of the latitude ranges 90°–30°S, 30°S–30°N, and 30°–90°N to the global average; the sum of these numbers is equal to the global average. In each panel, the shaded band shows the ±1σ range of the ensemble around the average.

$\Delta\lambda_{\text{clear sky LW}}$, basically mirrors the temperature difference, with positive values (meaning the feedback is less negative) in regions with lower warming fractions (the tropics) and negative values in regions with larger warming fractions (the extratropics). Integrating over latitude, these differences cancel and so the global-average $\Delta\lambda_{\text{clear sky LW}}$ is basically zero. This seems

likely to be generally true, so it might be expected that this feedback should generally contribute little to the pattern effect.

The term $\Delta\lambda_{\text{clear sky SW}}$ reflects changes in the surface albedo feedback and Fig. 3 shows that it is contributing the majority to the unforced pattern effect. Figure 4d shows that this is arising almost entirely from the
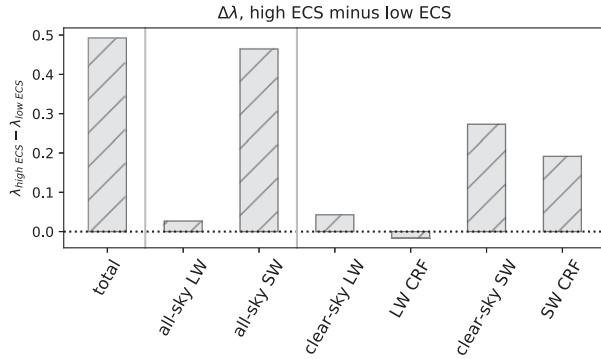
FIG. 3. Average feedbacks of the 10 historical ensemble members with the highest ECS$_{hist}$ (least negative λ_{hist}) minus the average of the 10 members with the lowest ECS$_{hist}$ (most negative λ_{hist}).

Antarctic region. Thus, while there is a strong ensemble-average $\lambda_{clear\ sky\ SW}$ in the Arctic (Fig. 2d), there is little variability within the ensemble in this feedback, so $\Delta\lambda_{clear\ sky\ SW}$ there is close to zero.

The all-sky LW feedback difference ($\Delta\lambda_{all\ sky\ LW} = \Delta\lambda_{LW\ CRF} + \Delta\lambda_{clear\ sky\ LW}$) in Fig. 4g shows that the near-zero difference in the LW feedback comes from cancellation between positive differences in the tropics and negative differences at high latitudes. The all-sky SW feedback difference ($\Delta\lambda_{all\ sky\ SW} = \Delta\lambda_{SW\ CRF} + \Delta\lambda_{clear\ sky\ SW}$) in Fig. 4h, which is responsible for most of the unforced pattern effect, reveals that on average 45% of the difference is coming from the Southern Hemisphere extratropics (5%–95% of individual ensemble members range have values of 36%–56%), 20% is coming from the tropics (14%–44%), and 35% is coming from the Northern Hemisphere extratropics (12%–39%).

5. Causes of the unforced pattern effect

a. Sea ice

The pattern of $\Delta\lambda_{clear\ sky\ SW}$, with the maximum located in the Antarctic region (Fig. 4d), strongly suggests that variability in sea ice loss among the ensembles is responsible for the spread in this feedback in the ensemble. Indeed, I find a strong correlation between the decrease of sea ice over the historical period in each ensemble member and the surface albedo feedback in that member (Figs. 5a,b). More quantitatively, variability in the loss of sea ice explains about half of the variance in λ_{hist} (Fig. 5c).

I find some connection between variability in the ocean circulation and variability in sea ice. In particular, changes in the Atlantic multidecadal oscillation (AMO) index and the South Atlantic multidecadal oscillation (SAMO) index correlate with λ_{albedo} from the same

hemisphere (Fig. 6) (plots using $\lambda_{clear\ sky\ SW}$ look nearly identical). The AMO and SAMO indices are defined as the average of detrended SST over the North Atlantic (0°–60°N, 0°–80°W) and South Atlantic (60°S–0°, 60°W–40°E), respectively (B. Yao et al. 2019, unpublished manuscript). There is little correlation between these indices and the albedo feedback in the other hemisphere ($|r| < 0.13$). I have looked at other indices [the Pacific decadal oscillation (PDO), South Pacific decadal oscillation (SPDO), interdecadal Pacific oscillation (IPO), tripole index (TPI), and Indian Ocean dipole (IOD)] and also find weak correlations between any of them and albedo feedback variability.

Of particular note, the slopes of the fits in Fig. 6a and 6b are 0.07 and 0.03 W m^{-2} K^{-1}, respectively. This means that the albedo feedback response to a unit change in the SAMO is more than twice the response to the AMO. This again emphasizes the key role the Southern Hemisphere has in varying λ_{hist}. It is worth noting that CMIP5-era models do not always do a great job of simulating the details of Antarctic sea ice (Turner et al. 2013), so verifying this result with other approaches, preferably tied to observations, should be a priority.

b. Shortwave clouds

The latitudinal pattern of $\Delta\lambda_{cloud\ SW}$ (Fig. 4f) does not point to a clear physical mechanism. However, previous work (Zhou et al. 2016, 2017; Andrews and Webb 2018; Ceppi and Gregory 2017; Fueglistaler 2019) has pointed toward atmospheric stability as key for regulating the cloud feedback and D18 identified ΔT_{500}, 500-hPa tropical (30°N–30°S) temperature, as providing a fundamental control on planetary energy balance. I find that variability in $\Delta T_{500}/\Delta T_S$, warming of the tropical atmosphere per unit global-average surface warming, explains much of the variability in $\lambda_{cloud\ SW}$ (Fig. 7a). This leads to $\Delta T_{500}/\Delta T_S$ having a strong correlation with λ_{hist} (Fig. 7b).

The slope of the line in Fig. 7a indicates that greater warming of the troposphere makes the SW cloud feedback more negative. If the slope were due mainly to the low-cloud feedback, then I would expect to also see a similar or stronger correlation with the net cloud feedback ($\lambda_{cloud} = \lambda_{cloud\ SW} + \lambda_{cloud\ LW}$) because net cloud feedback is a better indicator of low cloud changes (changes in high clouds tend to have LW and SW feedbacks that cancel; e.g., Zelinka et al. 2012). However, correlations between λ_{cloud} (and also λ_{CRF}) versus $\Delta T_{500}/\Delta T_S$ yield low correlations ($r = -0.08$). This suggests that mid- and high-level clouds are also playing a role in the variation of the SW cloud feedback in Fig. 7a.

I have also correlated $\lambda_{cloud\ SW}$ with other indicators of stability, such as estimated inversion strength
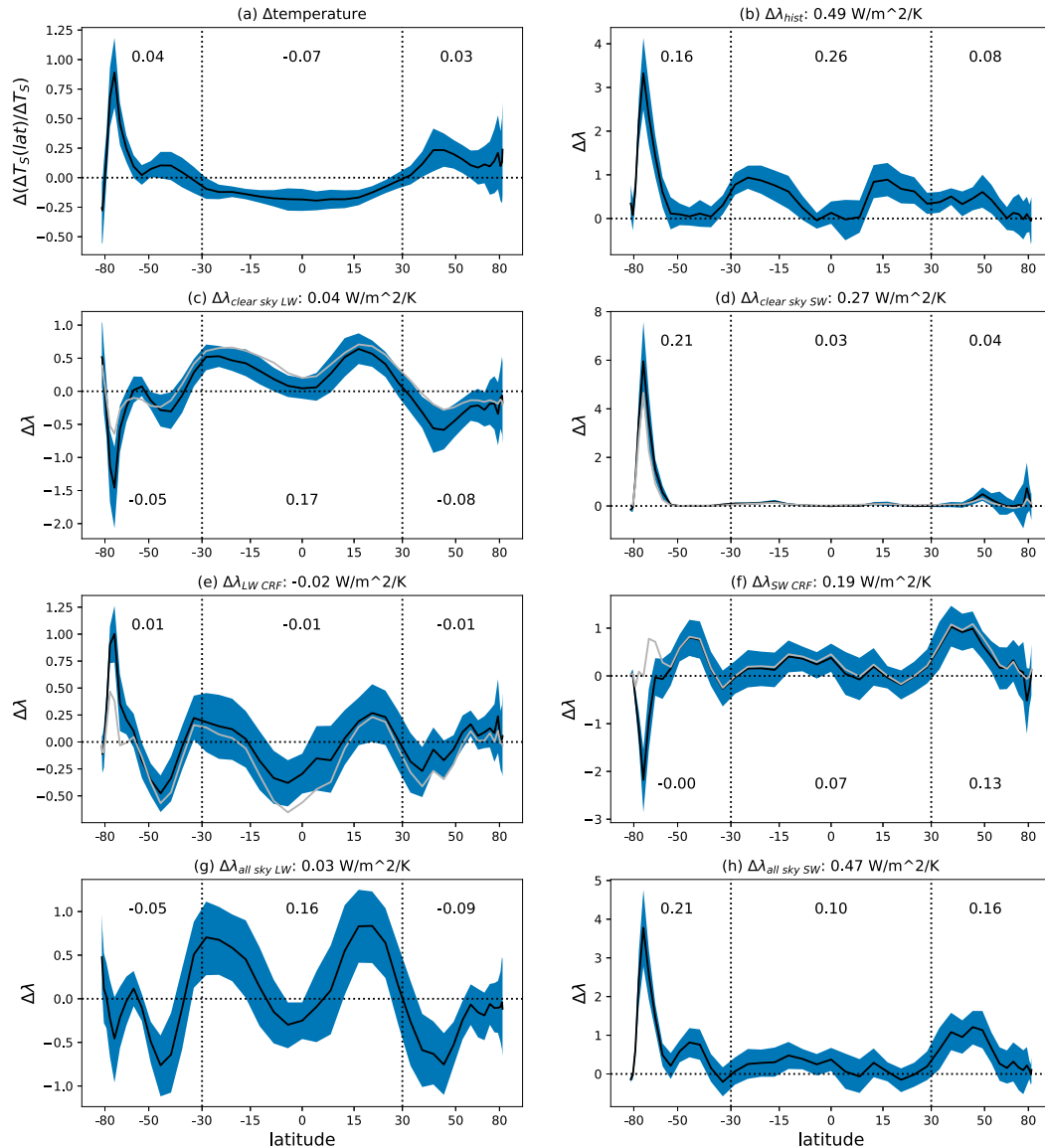
FIG. 4. As in Fig. 2, but for the change between the average of the 10 ensemble members with the highest ECS and the average of the 10 with the lowest ECS. The blue shaded region shows the 5%–95% uncertainty of the difference.

(Wood and Bretherton 2006) and find that ensemble members whose atmosphere becomes more stable also have a more negative $\lambda_{\mathrm{cloud\ SW}}$ (not shown). However, the correlation ($r = -0.54$) is not as good as with $\Delta T_{500}/\Delta T_S$. This is again consistent with the signal in Fig. 7a having a nontrivial contribution from mid- and high-level clouds. Investigating the altitude distribution of clouds driving the unforced pattern effect in models and observations should be a high priority for future work.

Given the role played by ΔT_{500} in regulating $\lambda_{\mathrm{cloud\ SW}}$, a natural question is whether ENSO is playing a role. Figure 7c shows that ΔENSO, the change in ENSO3.4 index between the first and last decade, does indeed correlate to some extent with $\lambda_{\mathrm{cloud\ SW}}$. I also find that ΔAMO has about the same magnitude correlation (Fig. 7d). ΔPDO (not shown) also correlates with $\lambda_{\mathrm{cloud\ SW}}$, but ΔENSO and ΔPDO are strongly correlated ($r = 0.73$), so I do not consider them independent regressors.

c. Putting it all together

The magnitude of the unforced pattern effect is affected by the exact periods selected, as I discuss in some
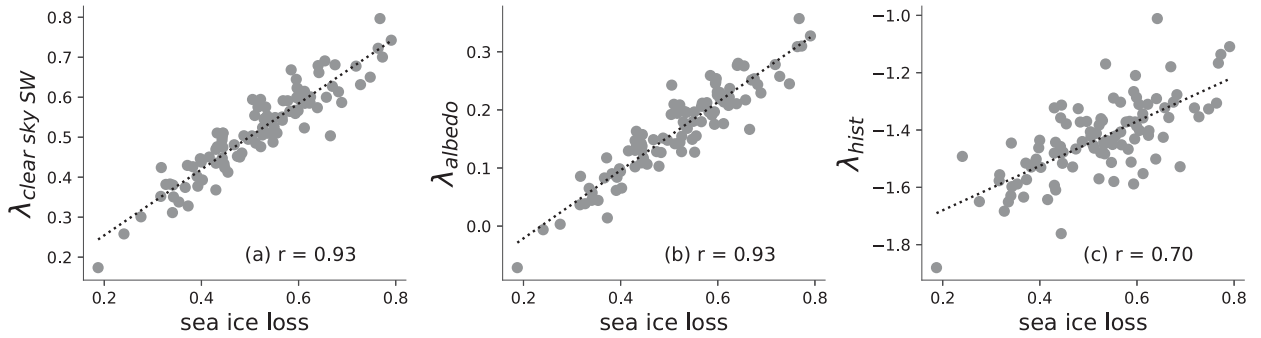
FIG. 5. Scatterplots of the loss of sea ice vs (a) $\lambda_{\text{clear sky SW}}$, (b) λ_{albedo}, and (c) λ_{hist}. Each dot represents the value from one member of the historical ensemble. The change in sea ice is the change in percent ice coverage (e.g., a decrease in sea ice coverage from 4% of the globe to 3.5% would yield a change of 0.5%). As with the feedbacks, the change in sea ice is the average difference between the first and last decade of the run. The dotted line is a linear least squares fit; r values of the fit are shown in each panel.

detail later in the paper. Most investigators use 1859–82 as a base period due to the lack of volcanic activity during those years. Using that base period, the 5%–95% spread in λ_{hist} in the ensemble is -1.46 to $-1.14\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$, with an ensemble average of $-1.33\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$ (changing the base and end periods does not change any of the previous results). This corresponds to an ensemble spread of ECS_{hist} of 2.53 to 3.24 K, with an ensemble average of 2.79 K. Thus, the unforced pattern effect can lead to ECS_{hist} having a bias of -0.26 to $+0.46\,\text{K}$ (-9% to $+16\%$) relative to the ensemble average ECS_{hist}.

I also show that unforced variability in two key parameters, sea ice loss and $\Delta T_{500}/\Delta T_S$, largely control the unforced pattern effect in this ensemble. These two parameters are correlated ($r = -0.56$) because sea ice loss exposes relatively warm ocean water, so members with more sea ice loss also have higher ΔT_S. This means that there is also a positive correlation between λ_{albedo}

and $\lambda_{\text{cloud SW}}$ ($r = 0.40$), so that variability in these two feedbacks work in the same direction to generate large variability in λ_{hist}.

A key caveat to our conclusions is the question of whether the model accurately simulates unforced variability. While we have not analyzed the fidelity of this particular model, previous work has pointed up some potential deficiencies in CMIP5-era models' simulation of unforced variability (e.g., Zhou et al. 2016; Hedemann et al. 2017; Kajtar et al. 2019). Determining whether other models—and nature—show this same unforced pattern effect should be a high priority.

6. The forced pattern effect

Magnitude

To determine the magnitude of the forced pattern effect, I begin by averaging λ_{hist} over all members of
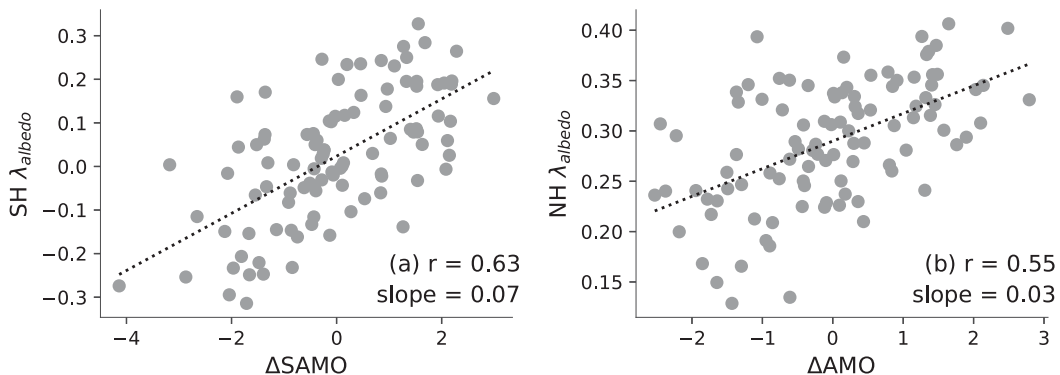

FIG. 6. (a) λ_{albedo} averaged over the Southern Hemisphere vs ΔSAMO; (b) λ_{albedo} averaged over the Northern Hemisphere vs ΔAMO. ΔSAMO and ΔAMO are the change in the respective index between the first and last decade of the historical period in each ensemble member. Each dot represents the value from one member of the historical ensemble and the dotted line is a linear least squares fit. The slope and correlation coefficient of the fit are displayed in each panel.
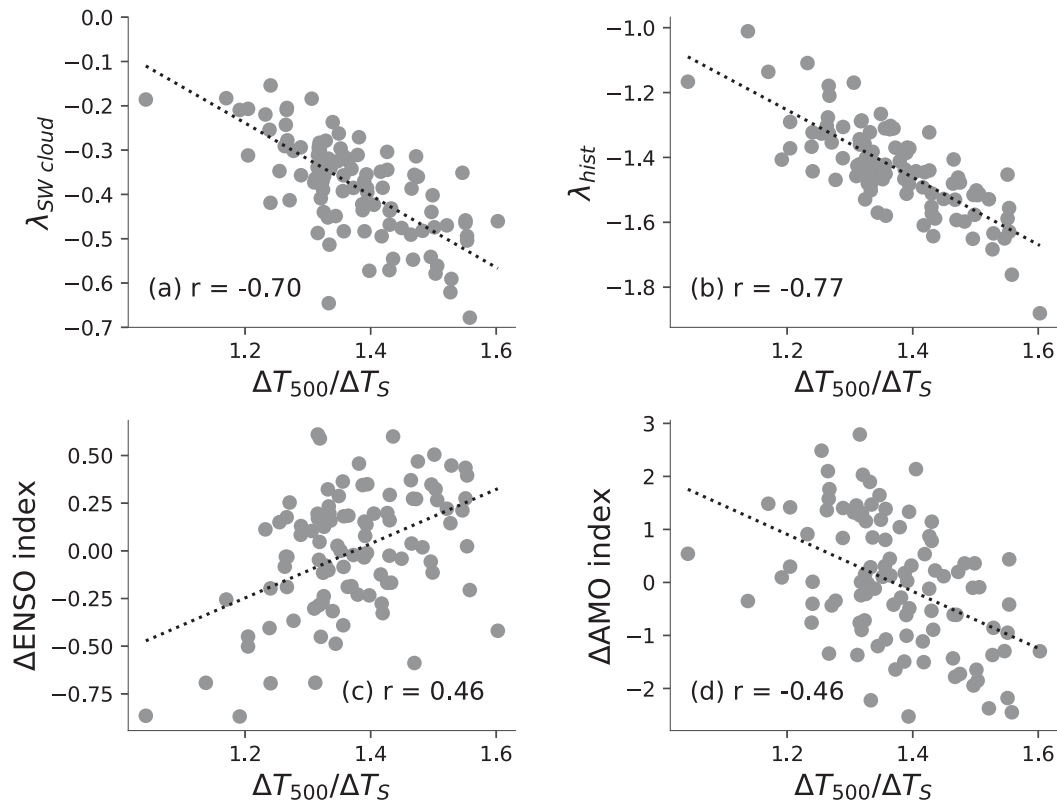
FIG. 7. (a) $\lambda_{\text{SW cloud}}$ vs $\Delta T_{500}/\Delta T_S$, where T_{500} is the 500-hPa tropical (30°N–30°S) temperature, (b) λ_{hist} vs $\Delta T_{500}/\Delta T_S$, (c) ΔENSO index vs $\Delta T_{500}/\Delta T_S$, and (d) ΔAMO index vs $\Delta T_{500}/\Delta T_S$. In all panels Δ represents the change between first and last decade of the runs. Each point in each panel is derived from one of the historical ensemble members and the dotted line is a linear least squares fit.

the historical ensemble, yielding $\lambda_{\text{hist-average}}$. The main challenge with this calculation is that $\lambda_{\text{hist-average}}$ is affected by the choice of averaging periods used in Eq. (2)—using different base and end periods can lead $\lambda_{\text{hist-average}}$ to vary by a factor of 2, from -0.8 to $-1.6 \, \text{W m}^{-2} \text{K}^{-1}$ (Fig. 8).

Previous investigators have attempted to get around this problem by picking periods unaffected by volcanic eruptions and I will follow that approach here. Picking volcanically unperturbed base periods (1860–69 or 1870–79) and end periods (1970–79 or 1996–2005) produce estimates of $\lambda_{\text{hist-average}}$ between -1.33 and $-1.27 \, \text{W m}^{-2} \text{K}^{-1}$, with an average of $-1.32 \, \text{W m}^{-2} \text{K}^{-1}$.

To verify this estimate, I have also analyzed a 68-member ensemble forced by CO_2 increasing at 1% per year (and no volcanoes). Figure 9 shows λ derived from the ensemble average of these 1% runs, hereafter $\lambda_{1\%\text{-average}}$. This is derived using Eq. (2) with ensemble average fluxes and temperatures; Δ is the difference between the average of the first 10 years of the run and the average of a sliding 10-yr window. While there is some decadal variability, it is clear that $\lambda_{1\%\text{-average}}$ has much less variation with time than

$\lambda_{\text{hist-average}}$, reflecting more uniform forcing, particularly the lack of volcanoes. Over the entire 150-yr run, the median value of $\lambda_{1\%\text{-average}}$ is $-1.33 \, \text{W m}^{-2} \text{K}^{-1}$ (5%–95% of the values range from -1.29 to $-1.41 \, \text{W m}^{-2} \text{K}^{-1}$), very close to $\lambda_{\text{hist-average}}$ for nonvolcanic periods.
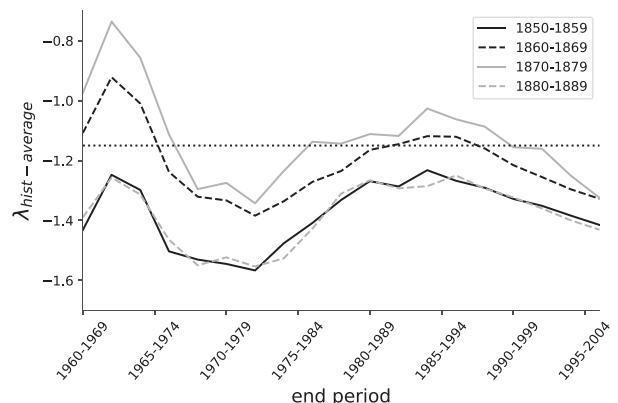

FIG. 8. Estimates of $\lambda_{\text{hist-average}}$ ($\text{W m}^{-2} \text{K}^{-1}$) with various base and end periods. Each line has a different base period, as indicated in the figure legend. The x axis indicates the end period. The horizontal dotted line is the value of $\lambda_{4 \times CO2}$.
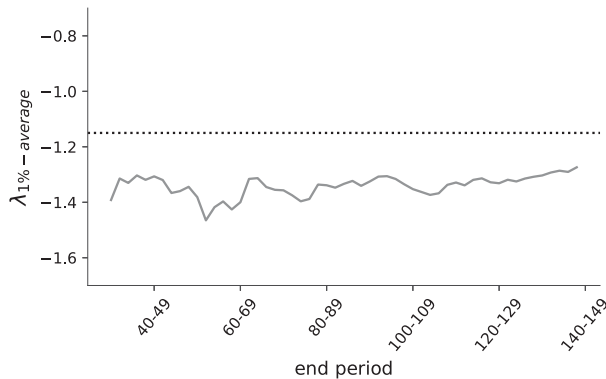
FIG. 9. Estimates of ensemble-average $\lambda_{1\%}$ ($\mathrm{W\,m^{-2}\,K^{-1}}$) as a function of end period, the value on the x axis. The y axis has the same range as Fig. 8 to make comparison easier. All values are calculated with a base period of the first 10 years of the ensemble. The horizontal dotted line is the value of $\lambda_{4\times CO2}$.

To estimate the forced pattern effect, I also need an estimate of λ from a more strongly forced equilibrium run, for which I use the 2614-yr abrupt $4\times CO_2$ run. I estimate $\lambda_{4\times CO2}$ using Eq. (2), with Δ representing the difference between the time average of the 2000-yr preindustrial control run and the time average of the last 500 years of the run, which covers a period nearly in equilibrium (the trend in T_S over this period is 0.02 K century^{-1}). This calculation yields a value of λ_{4xCO2} of $-1.15\,\mathrm{W\,m^{-2}\,K^{-1}}$.

Thus, $\lambda_{4\times CO2}$ is about 15% less negative than $\lambda_{\text{hist-average}}$. However, the forced pattern effect should be the difference between the historical ensemble average and a $2\times CO_2$ run, but the Grand Ensemble does not have appropriate $2\times CO_2$ runs. Mauritsen et al. (2019) analyzed both $2\times CO_2$ and $4\times CO_2$ runs of the MPI-ESM 1.2 model, a model closely related to the one used here. Using data from Fig. 12 and Table 5 of that paper, I estimate that $\lambda_{4\times CO2}$ is about 7% less negative than $\lambda_{2\times CO2}$. Previous work on this (Meraner et al. 2013; Mauritsen et al. 2019) suggests that increasing λ with warming is due to increasing strength of the water vapor feedback, related to an increase in height of the tropopause, and an increasingly positive cloud feedback.

I therefore conclude that ECS_{true} (ECS in the $2\times CO_2$ run) is about 8% larger, corresponding to 0.2 K, than ensemble-average ECS_{hist}. This estimate of the forced pattern effect is smaller than suggested by previous analyses (Armour 2017; Proistosescu and Huybers 2017) but close to values found by other analyses (Mauritsen and Pincus 2017; Lewis and Curry 2018).

Estimates of ECS_{true} require long forced runs to near-equilibrium conditions. Because such runs are relatively

rare, most previous estimates of the pattern effect have used effective climate sensitivity (where λ_{eff} is estimated from a Gregory regression of an abrupt $4\times CO_2$ run) instead of true climate sensitivity. For this model, λ_{eff} derived from regression of the first 150 years of the $4\times CO_2$ run is $-1.36\,\mathrm{W\,m^{-2}\,K^{-1}}$, meaning that effective climate sensitivity is more negative than $\lambda_{\text{hist-average}}$—meaning, in turn, that $ECS_{true} < ECS_{hist}$ and implying a *negative* forced pattern effect. However, the result is quite sensitive to the period selected for the regression. Regressing years 20–150 yields a λ_{eff} of -1.09, a forced pattern effect larger than found here. I expect this may be quite different for different models, so one should be cautious interpreting estimates of the pattern effect based on Gregory regressions over arbitrary periods.

7. Causes of the forced pattern effect

Figure 10 shows global-average feedbacks from the 1% ensemble and abrupt $4\times CO_2$ run, as well as the differences between them. The forced pattern effect shown here is almost entirely due to SW cloud feedbacks, also noted by Andrews et al. (2015).

Figure 11 shows the spatial pattern of the difference in the total cloud feedback. I plot the total cloud feedback (LW + SW) because that feedback correlates better with low-cloud changes. And I plot the cloud feedback rather than the CRF feedback because the CRF feedback has large values in the polar regions associated with cloud masking of changes in surface albedo rather than changes in clouds.

Maxima in $\Delta\lambda_{\text{cloud}}$ occur in regions where warming is delayed: the Southern Ocean, eastern Pacific, and Atlantic Ocean south of Greenland. As the surface in these regions eventually warms in the future, the stability of the atmosphere decreases, leading to a reduction in low clouds, thereby increasing the magnitude of the cloud feedback as the climate warms (Senior and Mitchell 2000; Ceppi and Gregory 2017; Andrews et al. 2018).

8. Conclusions

In this paper, I have addressed the question: Is ECS estimated from historical observations (ECS_{hist}) a good estimate of the true ECS (ECS_{true}) of our climate system? I have investigated two reasons why the answer may be "no." First, the historical observational record is just one member of an infinity of theoretical climate trajectories for the Earth since preindustrial. Different climate trajectories over this period yield estimates of ECS_{hist} that can differ from the ensemble
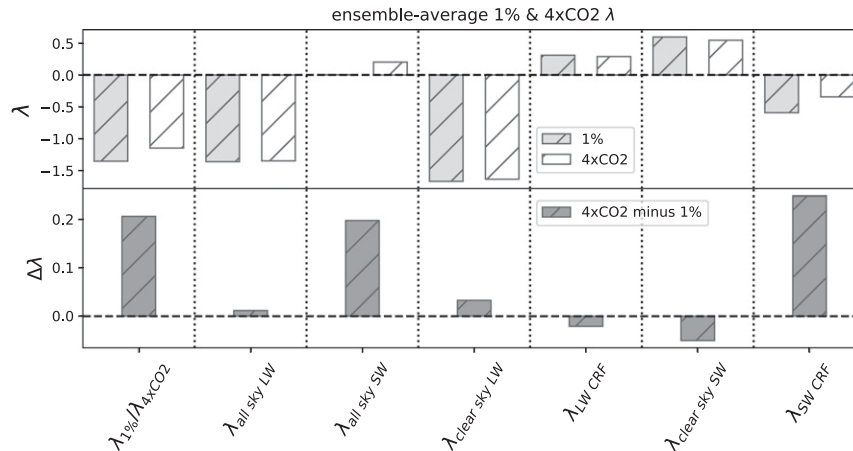
FIG. 10. (top) λ (W m^{-2} K^{-1}) averaged over the 1% ensemble and from the abrupt $4\times CO_2$ run. (bottom) The difference ($4\times CO_2$ minus 1% ensemble average, in W m^{-2} K^{-1}) between these two model experiments. The 1%-ensemble feedbacks are calculated by differencing the average of the first 10 years and the average of years 30–110 of the ensemble average fluxes and temperatures.

average ECS_{hist} by -0.26 to $+0.46$ K, corresponding to -9% to $+16\%$ (5%–95% confidence interval).

This unforced pattern effect arises mainly from two sources: unforced variability in loss of sea ice, particularly in the Antarctic, which leads to variability in the surface albedo feedback, and unforced variability in tropical tropospheric warming, which leads to variability in the cloud feedbacks. Variability in these two parameters correlates with well-known climate indices (e.g., ENSO, AMO), suggesting that the unforced pattern effect is controlled by known modes of internal variability. This may give us a way to evaluate where the observed historical record lies within the envelope of all possible records. Doing so should obviously be a high priority for the community.

The second reason that ECS_{hist} may not be a good estimate of ECS_{true} is that the average transient warming pattern over the twentieth century is expected to be different from the equilibrium pattern of warming for doubled CO_2, and this can also lead to differences between ECS_{hist} and ECS_{true}. Because this effect is related to forced warming, I refer to this as the forced pattern effect.

I estimate that ECS_{true} is 8% larger than ensemble-average ECS_{hist}, corresponding to a bias of 0.2–0.3 K. This forced pattern effect is mainly due to a less negative low-cloud feedback at equilibrium arising in oceanic regions of where warming is delayed, namely the Southern Ocean, east Pacific, and the Atlantic south of Greenland (Senior and Mitchell 2000; Andrews et al. 2015; Ceppi and Gregory 2017). Attempting to estimate the magnitude of the forced pattern effect with respect to effective climate sensitivity (estimated from regressions of the first 150 years or so of a $4\times CO_2$ run)

can yield widely varying results and those estimates should be treated with caution.

There are similarities between the forced and unforced pattern effect. Both operate primarily in the SW, and in both cases $\Delta\lambda_{cloud\ SW} \approx 0.2$ W m^{-2} K^{-1} (Figs. 3 and 10). The main difference is that the unforced pattern effect has an additional contribution from $\Delta\lambda_{albedo}$, which does not play a significant role in the forced pattern effect. Another difference is that we can have high confidence in the direction of the forced pattern effect: ensemble-average ECS_{hist} should be lower than ECS_{true}. For the unforced pattern effect, the sign is less clear, although previous analyses (Gregory and Andrews 2016; Marvel et al. 2018; Andrews et al. 2018) have argued that the observed climate trajectory of
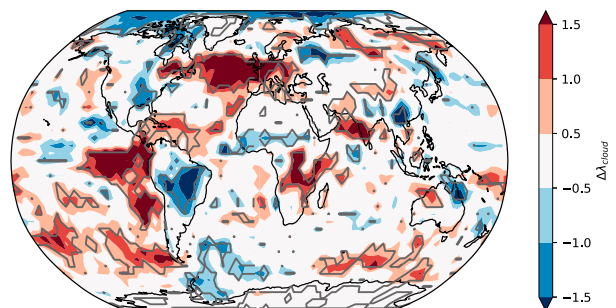

FIG. 11. Cloud feedback in the $4\times CO_2$ run minus the ensemble average cloud feedback in the 1% ensemble. The 1% ensemble feedbacks are calculated by differencing the average of the first 10 years and the average of years 30–110 of the ensemble average fluxes and temperatures. The gray lines show where the ensemble-average difference has the same sign as difference between the $4\times CO_2$ feedback and the feedback in 90% of the individual ensemble members.

the twentieth century produces a ECS$_{hist}$ that is lower than ensemble-average ECS$_{hist}$. If so then the forced and unforced pattern effects would add, leading to an ECS$_{hist}$ from observations that could be 0.5 K or 17% less than ECS$_{true}$.

Any bias from this calculation will add to other biases in the calculation. For example, incomplete and changing spatial coverage of the surface temperature record, as well as the fact that historical surface temperature measurements are blends of SST over ocean and air temperature over land and ice (Cowtan and Way 2014; Cowtan et al. 2015) also bias ECS$_{hist}$ low. These temperature biases can by themselves lead to a 20%–30% low bias for ECS$_{hist}$ (Richardson et al. 2016; Adams and Dessler 2019). Combining this with a bias from the pattern effect could lead to very large low biases in ECS$_{hist}$. Considering all of the potential biases between ECS$_{hist}$ and ECS$_{true}$, it seems premature to argue that climate models are overestimating ECS$_{true}$ based on comparisons to observational estimates of ECS$_{hist}$.

This analysis highlights potential areas of future research. First, this analysis uses a single model ensemble. Do other model ensembles confirm these results? Second, how can we estimate the impact of the unforced pattern effect on the actual historical climate record? This will require a better understanding of how internal variability impacts the important climate parameters (in our analysis, sea ice and tropospheric temperature) as well as an estimate of how these have changed over the historical period.

In the short term, these uncertainties in estimates of ECS$_{hist}$ may be difficult to reduce or eliminate. In that case, better estimates of ECS may rely on developing alternative methods of estimating ECS that are less impacted by these pattern effects, such as using short-term internal variability over the last few decades to estimate ECS$_{true}$ (e.g., Dessler and Forster 2018).

REFERENCES

Adams, B. K., and A. E. Dessler, 2019: Estimating transient climate response in a large-ensemble global climate model simulation. *Geophys. Res. Lett.*, **46**, 311–317, https://doi.org/10.1029/2018GL080714.

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, https://doi.org/10.1175/JCLI-D-17-0087.1.

——, J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, **28**, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1.

——, and Coauthors, 2018: Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophys. Res. Lett.*, **45**, 8490–8499, https://doi.org/10.1029/2018GL078887.

Armour, K. C., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nat. Climate Change*, **7**, 331–335, https://doi.org/10.1038/nclimate3278.

——, C. M. Bitz, and G. H. Roe, 2013: Time-varying climate sensitivity from regional feedbacks. *J. Climate*, **26**, 4518–4534, https://doi.org/10.1175/JCLI-D-12-00544.1.

Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proc. Natl. Acad. Sci.*, **114**, 13 126–13 131, https://doi.org/10.1073/pnas.1714308114.

Cowtan, K., and R. G. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quart. J. Roy. Meteor. Soc.*, **140**, 1935–1944, https://doi.org/10.1002/qj.2297.

——, and Coauthors, 2015: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophys. Res. Lett.*, **42**, 6526–6534, https://doi.org/10.1002/2015GL064888.

Dessler, A. E., and P. M. Forster, 2018: An estimate of equilibrium climate sensitivity from interannual variability. *J. Geophys. Res. Atmos.*, **123**, 8634–8645, https://doi.org/10.1029/2018JD028481.

——, 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, https://doi.org/10.5194/acp-18-5147-2018.

Forster, P. M., 2016: Inference of climate sensitivity from analysis of Earth's energy budget. *Annu. Rev. Earth Planet. Sci.*, **44**, 85–106, https://doi.org/10.1146/annurev-earth-060614-105156.

Fueglistaler, S., 2019: Observational evidence for two modes of coupling between sea surface temperatures, tropospheric temperature profile, and shortwave cloud radiative effect in the tropics. *Geophys. Res. Lett.*, **46**, 9890–9898, https://doi.org/10.1029/2019GL083990.

Gregory, J. M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophys. Res. Lett.*, **43**, 3911–3920, https://doi.org/10.1002/2016GL068406.

Hedemann, C., T. Mauritsen, J. Jungclaus, and J. Marotzke, 2017: The subtle origins of surface-warming hiatuses. *Nat. Climate Change*, **7**, 336–339, https://doi.org/10.1038/nclimate3274.

Held, I., and K. M. Shell, 2012: Using relative humidity as a state variable in climate feedback analysis. *J. Climate*, **25**, 2578–2582, https://doi.org/10.1175/JCLI-D-11-00721.1.

Huber, M., and R. Knutti, 2014: Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nat. Geosci.*, **7**, 651–656, https://doi.org/10.1038/ngeo2228.

Kajtar, J. B., M. Collins, L. M. Frankcombe, M. H. England, T. J. Osborn, and M. Juniper, 2019: Global mean surface temperature response to large-scale patterns of variability in observations and CMIP5. *Geophys. Res. Lett.*, **46**, 2232–2241, https://doi.org/10.1029/2018GL081462.

Knutti, R., M. A. A. Rugenstein, and G. C. Hegerl, 2017: Beyond equilibrium climate sensitivity. *Nat. Geosci.*, **10**, 727–736, https://doi.org/10.1038/ngeo3017.

Lewis, N., and J. Curry, 2018: The impact of recent forcing and ocean heat uptake data on estimates of climate sensitivity. *J. Climate*, **31**, 6051–6071, https://doi.org/10.1175/JCLI-D-17-0667.1.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *J. Adv. Model. Earth Syst.*, **11**, 2050–2069, https://doi.org/10.1029/2019MS001639.

Marvel, K., R. Pincus, G. A. Schmidt, and R. L. Miller, 2018: Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophys. Res. Lett.*, **45**, 1595–1601, https://doi.org/10.1002/2017GL076468.

Mauritsen, T., and R. Pincus, 2017: Committed warming inferred from observations. *Nat. Climate Change*, **7**, 652–655, https://doi.org/10.1038/nclimate3357.

——, and Coauthors, 2019: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *J. Adv. Model. Earth Syst.*, **11**, 998–1038, https://doi.org/10.1029/2018MS001400.

Meraner, K., T. Mauritsen, and A. Voigt, 2013: Robust increase in equilibrium climate sensitivity under global warming. *Geophys. Res. Lett.*, **40**, 5944–5948, https://doi.org/10.1002/2013GL058118.

Olson, R., R. Sriver, W. Chang, M. Haran, N. M. Urban, and K. Keller, 2013: What is the effect of unresolved internal climate variability on climate sensitivity estimates? *J. Geophys. Res. Atmos.*, **118**, 4348–4358, https://doi.org/10.1002/jgrd.50390.

Proistosescu, C., and P. J. Huybers, 2017: Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Sci. Adv.*, **3**, e1602821, https://doi.org/10.1126/sciadv.1602821.

Richardson, M., K. Cowtan, E. Hawkins, and M. B. Stolpe, 2016: Reconciled climate response estimates from climate models and the energy budget of Earth. *Nat. Climate Change*, **6**, 931–935, https://doi.org/10.1038/nclimate3066.

Rohde, R., and Coauthors, 2013: A new estimate of the average Earth surface land temperature spanning 1753 to 2011. *Geoinf. Geostat. Overview*, **1** (1), https://doi.org/10.4172/2327-4581.1000101.

Senior, C. A., and J. F. B. Mitchell, 2000: The time-dependence of climate sensitivity. *Geophys. Res. Lett.*, **27**, 2685–2688, https://doi.org/10.1029/2000GL011373.

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, https://doi.org/10.1175/1520-0442(2004)017<3661:OTUOCF>2.0.CO;2.

——, 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, https://doi.org/10.1175/2007JCLI2110.1.

Stevens, B., S. C. Sherwood, S. Bony, and M. J. Webb, 2016: Prospects for narrowing bounds on Earth's equilibrium climate sensitivity. *Earth's Future*, **4**, 512–522, https://doi.org/10.1002/2016EF000376.

Turner, J., T. J. Bracegirdle, T. Phillips, G. J. Marshall, and J. S. Hosking, 2013: An initial assessment of Antarctic sea ice extent in the CMIP5 models. *J. Climate*, **26**, 1473–1484, https://doi.org/10.1175/JCLI-D-12-00068.1.

Wood, R., and C. S. Bretherton, 2006: On the relationship between stratiform low cloud cover and lower-tropospheric stability. *J. Climate*, **19**, 6425–6432, https://doi.org/10.1175/JCLI3988.1.

Zelinka, M. D., 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.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nat. Geosci.*, **9**, 871–874, https://doi.org/10.1038/ngeo2828.

——, ——, and ——, 2017: Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *J. Adv. Model. Earth Syst.*, **9**, 2174–2189, https://doi.org/10.1002/2017MS001096.