The relative importance of forced and unforced temperature patterns in driving the time variation of low-cloud feedback

Yuan-Jen Lin[1,2], Grégory V. Cesana[1,2], Cristian Proistosescu[3], Mark D. Zelinka[4], and Kyle C. Armour[5]

[1]Center for Climate Systems Research, Columbia University
[2]NASA Goddard Institute for Space Studies
[3]Department of Atmospheric Sciences and Department of Geology, University of Illinois at Urbana-Champaign, Urbana-Champaign
[4]Lawrence Livermore National Laboratory
[5]School of Oceanography and Department of Atmospheric Sciences, University of Washington

January 18, 2024

The relative importance of forced and unforced temperature patterns in driving the time variation of low-cloud feedback

Yuan-Jen Lin[a,b] , Grégory V. Cesana[a,b] , Cristian Proistosescu[c] , Mark D. Zelinka[d] , and Kyle C. Armour[e]

[a] *Center for Climate Systems Research, Columbia University, New York, NY, USA*

[b] *NASA Goddard Institute for Space Studies, New York, NY, USA*

[c] *Department of Atmospheric Sciences and Department of Geology, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA*

[d] *Lawrence Livermore National Laboratory, Livermore, CA, USA*

[e] *School of Oceanography and Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA*

Corresponding author: Yuan-Jen Lin, yl5278@columbia.edu

ABSTRACT: Atmospheric models forced with observed sea-surface temperatures (SSTs) suggest more-stabilizing cloud feedback in recent decades, partly due to the surface cooling trend in the eastern Pacific (EP) and the warming trend in the western Pacific (WP). Here we show model evidence that the low-cloud feedback has contributions from both forced and unforced feedback components, and that its time variation arises in large part through changes in the relative importance of the two over time, rather than through variations in forced feedbacks as is often assumed. Initial-condition large ensembles (LEs) suggest that the SST patterns are dominated by unforced variations for 30-year windows ending prior to the 1980s. In general, unforced SSTs are representative of an ENSO-like pattern, which corresponds to weak low-level stability in the tropics and less-stabilizing low-cloud feedback. Since the 1980s, the forced signals have become stronger, outweighing the unforced signals for the 30-year windows ending after the 2010s. Forced SSTs are characterized by relatively uniform warming with an enhancement in the WP, corresponding to more-stabilizing low-cloud feedback in most cases. The time-evolving SST pattern due to this increasing importance of forced signals is the dominant contributor to the recent stabilizing shift of low-cloud feedback in the LEs. Observed SST patterns also suggest a reduction in the relative role of unforced ENSO-like variability since the 1980s. However, the observed SST patterns show strong WP warming and EP cooling trend, which actuates a shift in low-cloud feedback toward more-stabilizing values with a trend that lies outside the model ensembles.

1. Introduction

Projections of future warming in response to forcing depend on the magnitude of radiative feedbacks and, in particular, on how clouds will respond to changing climate conditions (Bony and Dufresne 2005; Sherwood et al. 2014; IPCC 2023). Previous research has shown that radiative feedbacks have considerable temporal variations (Andrews et al. 2015; Zhou et al. 2016; Andrews et al. 2018; Dong et al. 2020; Gregory et al. 2020; Rugenstein et al. 2020; Andrews et al. 2022), which adds to the uncertainty of climate prediction (Frey et al. 2017; Sherwood et al. 2020; Gjermundsen et al. 2021; Watanabe et al. 2021; IPCC 2023).

Radiative feedbacks vary over time in both the historical period (since around 1850) and future warming simulations. In most fully-coupled atmosphere-ocean general circulation models (AOGCMs) where the atmospheric CO_2 concentration is abruptly quadrupled and kept constant for the rest of the simulation, the net radiative feedback becomes less stabilizing over time (higher effective climate sensitivity) (Geoffroy et al. 2013; Andrews et al. 2015; Ceppi and Gregory 2017; Dong et al. 2020; Rugenstein et al. 2020). In the historical period, the feedback shows strong variability on decadal timescales. Most AOGCM historical simulations suggest a shift toward less-stabilizing net radiative feedback (higher effective climate sensitivity) over the past few decades (Gregory et al. 2020; Dong et al. 2021; Salvi et al. 2023). However, atmospheric general circulation models (AGCM) with prescribed observational SST and sea ice instead indicate a more-stabilizing net radiative feedback (lower effective climate sensitivity) during the same time period (Zhou et al. 2016; Gregory and Andrews 2016; Andrews et al. 2018, 2022). The time evolution of net radiative feedback has been interpreted through changes in SST patterns, also referred to as the pattern effect (Stevens et al. 2016; Zhou et al. 2017; Dong et al. 2019). The divergent trends of net radiative feedback between the abovementioned AOGCM and AGCM simulations in recent decades can be explained by discrepancies between the modeled and observed SST patterns (Dong et al. 2021).

The potential for radiative feedbacks to vary over time as the SST pattern evolves has largely been interpreted in terms of a forced climate response. For instance, as is seen most clearly under an abrupt CO_2 doubling or quadrupling, SST patterns and thus radiative feedbacks vary as the ocean adjusts on a range of timescales (Held et al. 2010; Winton et al. 2010; Armour et al. 2013; Geoffroy et al. 2013; Rose et al. 2014; Rose and Rayborn 2016; Rugenstein et al. 2016; Lin et al. 2019, 2021; Eiselt and Graversen 2023). Moreover, non-CO_2 forcing agents, such as anthropogenic aerosols or volcanic eruptions, can produce time-varying SST patterns and radiative feedbacks that are distinct from those from CO_2 forcing (Shindell 2014; Gregory et al. 2016; Marvel et al. 2016; Gregory et al. 2020; Günther et al. 2022; Salvi et al. 2023; Zhou et al. 2023). Another branch of literature has also shown that internal variability can influence radiative feedbacks through its influence on evolving SST patterns (Huber et al. 2014; Dessler et al. 2018; Gregory et al. 2020) and that, in general, the spatial patterns and magnitudes of radiative feedbacks under different modes of internal variability ("unforced feedbacks") are distinct from those induced by radiative forcing ("forced feedbacks") (Donohoe et al. 2014; Proistosescu et al. 2018; Wills et al. 2021; Uribe et al. 2022).

Here we investigate another contribution to the time variation of radiative feedbacks. In light of the fact that forced and unforced feedbacks have different magnitudes, it is possible that a portion of net radiative feedback time evolution may stem from a changing relative importance of internal variability and forced response – rather than through variations in the magnitude of forced feedbacks alone (as is often assumed). For instance, early in the historical record when radiative forcing is small, we might expect the net radiative feedback to largely reflect feedbacks associated with internal variability. However, later in the historical record and in the future when radiative forcing is strong, we might expect the net radiative feedback to largely reflect feedbacks induced by the forcing. A key question is, how important is such a shift in the relative importance of internal variability and forced response to the overall time variation of radiative feedbacks?

To answer the question, we begin by laying out a statistical framework to showcase how forced and unforced variations combine to yield a net global radiative feedback (Section 2). Results on the relative importance of the forced and unforced signals from three initial-condition large ensembles are then shown in Section 2. In Section 3, we focus on the time evolution of low-cloud feedback and decompose the feedback change into components related to changes in forced response, changes in unforced variability, and changes in their relative importance. Section 4 highlights the role of the SST pattern effect in connecting the time-evolving SST pattern and low-cloud feedback. A coherent analysis of the CMIP6 models and the observation data is shown in Section 5. In Section 6, we summarize our findings and discuss the further implications of our research.

2. Relative importance of forced and unforced responses

a. Initial-condition large ensembles

To isolate the forced responses from the unforced internal variability, we used single-model, initial-condition large ensembles, including Community Earth System Model Version 2 large ensemble (CESM2-LE; Rodgers et al. (2021)), Max Planck Institute for Meteorology Grand Ensemble (MPI-GE; Maher et al. (2019)), and simulations from the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS-LE; Kelley et al. (2020); Bauer et al. (2020); Miller et al. (2021)). The initial-condition large ensembles aim to create a large number of simulations with identical forcing and slightly different atmospheric and/or oceanic initial conditions. By taking the ensemble mean, the relative contribution of internal variability is expected to weaken to $1/\sqrt{N}$, where N is the size of the ensemble members (Gregory et al. 2020). Here CESM2-LE ($N = 100$), MPI-GE ($N = 100$), and GISS-LE ($N = 48$) all have relatively large N. Thus, any target field X in large ensemble simulations can be decomposed into two parts: (1) the ensemble-mean values of X (denoted as $\langle X \rangle$), which approximates the forced responses (X_f); and (2) the anomalies relative to the ensemble mean of X (denoted as X^*), which approximates unforced variability (X_u).

$$X = \langle X \rangle + X^*, \tag{2.1}$$

$$X_f \approx \langle X \rangle, \tag{2.2}$$

$$X_u \approx X^*. \tag{2.3}$$

b. Radiative feedback estimation

The evaluation of net radiative feedback often starts with the global-mean energy balance equation $N = F + R \approx F + \lambda T$ (Gregory et al. 2004), where N is net downward radiation at the top-of-atmosphere (TOA), F is effective radiative forcing, and R represents radiative responses (positive downwards). R is often approximated as λT, where T indicates the global-mean surface temperature responses that act to dampen or amplify R through stabilizing or destabilizing feedback processes, denoted as λ. Here the net radiative feedback (λ) is negative for a stable climate, thus a more negative (more-stabilizing) λ implies a less-sensitive climate.

In non-equilibrium climate states, such as for historical warming when the climate is still adjusting to forcing, λ is often quantified using the difference between two states, denoted as subscript $_1$ and $_2$ (Gregory et al. 2002; Dessler et al. 2018).

$$\lambda = \frac{(R_2 - R_1)}{(T_2 - T_1)} = \frac{(N_2 - N_1) - (F_2 - F_1)}{(T_2 - T_1)}. \tag{2.4}$$

At the same time, λ can be written in a differential form, where the derivatives can be estimated through linear regressions (Gregory et al. 2004; Rugenstein and Armour 2021).

$$\lambda = \frac{dR}{dT} = \frac{d(N - F)}{dT}. \tag{2.5}$$

c. Forced and unforced contributions to OLS regressions: Theory

The ordinary least squares (OLS) regression is widely used when examining the regression form of radiative feedback (e.g., Sherwood et al. (2020)). In a similar regression form, the SST pattern is usually calculated by regressing the map of regional SSTs against the global-mean values of surface temperature (e.g., Andrews et al. (2015)). The use of OLS regression relies on the assumption that the independent variable (i.e., the x-variable) is uncorrelated with the error term in the regression model, so the error term only considers unpredictable random error (i.e., the noise of the dependent variable). Thus, OLS regression estimates of radiative feedbacks may be biased when forced responses and unforced variability are tangled in both the independent variable (e.g., global-mean surface temperature) and the independent variable (e.g., radiation). In the following text, we will quantify the relative contribution of forced and unforced signals to OLS regressions. We will also show how the two components jointly drive the time variation of SST patterns and radiative feedbacks.

Take the regression of a given field X against global-mean surface temperature (T_g) in a given historical time period, for example. Both the independent variable (T_g) and the dependent variable (X) consist of two parts that evolve with time: a forced response to net radiative forcing (from greenhouse gases, aerosols, land use, etc.) and an unforced response related to internal variability.

We can express the two components as follows:

$$T_g = T_{g,f} + T_{g,u}, \tag{2.6}$$

$$X = X_f + X_u, \tag{2.7}$$

where the subscripts $_f$ and $_u$ indicate the forced and unforced response, respectively. By substituting the full response with the forced and unforced components, the regression-based estimate of $\frac{dX}{dT_g}$ can be written as:

$$\begin{aligned}\frac{dX}{dT_g} &= \frac{cov(X,T_g)}{var(T_g)} = \frac{cov(X_f + X_u, T_{g,f} + T_{g,u})}{var(T_{g,f} + T_{g,u})} \\ &= \frac{cov(X_f,T_{g,f}) + cov(X_f,T_{g,u}) + cov(X_u,T_{g,f}) + cov(X_u,T_{g,u})}{var(T_{g,f}) + 2cov(T_{g,f},T_{g,u}) + var(T_{g,u})},\end{aligned} \tag{2.8}$$

where $cov(x,y)$ is the covariance between the variable x and y, and $var(x)$ is the variance of x. Both are estimated within a given time period (e.g., a 30-year window). Since there is a general difference between the time evolution of forced and unforced response, namely, the former is more linear with time due to the accumulating radiative forcing, whereas the latter consists of internal variations across different timescales (mainly interannual to decadal oscillations for a 30-year window), we assume that the covariance between the forced and unforced response is small. Following this assumption, which we will test later, Equation 2.8 can then be expressed as:

$$\frac{dX}{dT_g} = \frac{cov(X_f,T_{g,f}) + cov(X_u,T_{g,u})}{var(T_{g,f}) + var(T_{g,u})} + \sigma, \tag{2.9}$$

where the residual (σ) accounts for the combined effect from the three covariances between the forced and unforced responses, including $cov(T_{g,f},T_{g,u})$, $cov(X_f,T_{g,u})$, and $cov(X_u,T_{g,f})$. By re-arranging Equation 2.9, the regression estimate can be decomposed into forced and unforced regressions as follows:

$$\begin{aligned}\frac{dX}{dT_g} &= \frac{cov(X_f,T_{g,f})}{var(T_{g,f})}\frac{var(T_{g,f})}{var(T_{g,f}) + var(T_{g,u})} + \frac{cov(X_u,T_{g,u})}{var(T_{g,u})}\frac{var(T_{g,u})}{var(T_{g,f}) + var(T_{g,u})} + \sigma \\ &= \frac{dX_f}{dT_{g,f}}\frac{var(T_{g,f})}{var(T_{g,f}) + var(T_{g,u})} + \frac{dX_u}{dT_{g,u}}\frac{var(T_{g,u})}{var(T_{g,f}) + var(T_{g,u})} + \sigma.\end{aligned} \tag{2.10}$$

Equation 2.10 suggests that the overall regression estimate is a linear combination of the forced and unforced regressions, with a specific weighting applied to each term. For forced regression ($\frac{dX_f}{dT_{g,f}}$), it is multiplied by the ratio of forced T_g variance ($var(T_{g,f})$) to the sum of forced T_g variance and unforced T_g variance ($var(T_{g,f})+var(T_{g,u})$). Similarly, unforced regression ($\frac{dX_u}{dT_{g,u}}$) is multiplied by the ratio of unforced T_g variance ($var(T_{g,u})$) to the sum of forced T_g variance and unforced T_g variance ($var(T_{g,f})+var(T_{g,u})$). We can simplify the equation further by writing it as:

$$\frac{dX}{dT_g} = \frac{dX_f}{dT_{g,f}}r + \frac{dX_u}{dT_{g,u}}(1-r)+\sigma, \tag{2.11}$$

$$r = \frac{var(T_{g,f})}{var(T_{g,f})+var(T_{g,u})}. \tag{2.12}$$

The ratio r and $(1-r)$ indicate, respectively, the relative importance of forced and unforced temperature variance during the interval over which the regression has been performed. When r is small, the influence of forced regression on overall regression is weak, and the regression is largely determined by the unforced variability of X and T_g. Vice versa for large r. If X is taken to be the net TOA radiation (R), the net radiative feedback can be written as a weighted sum of the feedback in response to forced variations and the feedback in response to unforced variations ($\lambda = \lambda_f r + \lambda_u(1-r)+\sigma$). Similarly, if X represents regional warming, then $\frac{dX}{dT_g}$ becomes the net warming pattern over the time interval, and it will like-wise be a weighted sum of forced and unforced components.

In summary, here we demonstrate how the forced and unforced signals jointly affect the strength of the OLS regression, which is widely used to calculate radiative feedback ($X = R$) and quantify SST patterns ($X = SST$). For each OLS regression, changes in either forced or unforced regression alter the strength of the overall regression. Even when both components are constant over time, changes in their relative importance (quantified as r and $(1-r)$) could lead to time variation in the overall regression.

d. Forced and unforced contributions to OLS regressions: Model results

Section 2c provides the theory of how the forced and unforced T_g variance determines their relative importance in OLS regressions. Figure 1 shows the model results that echo the theory. First, the forced and unforced T_g in the three LEs are shown (Figs. 1a-c). Note that to be more consistent with observational SSTs (Section 5), here we define T_g as the area-weighted average of near-global (60°S-60°N) surface temperature over the ocean. This T_g definition is different from the commonly used global-mean surface temperature due to the exclusion of land and polar regions, however, we highlight that the time evolution of the two is highly consistent and the different T_g definition does not change the conclusions. Before 1980, all three LEs suggest a relatively mild 60°S-60°N warming due to the forcing (less than 0.4 K increase in $T_{g,f}$ since 1850; approximately less than 0.1 K for 30-year intervals). For comparison, the range of unforced T_g variations over 30-year intervals is around 0.4-0.8 K. Since the 1980s, the forced component has strengthened, and so has the forced warming rate.

Figures 1d-f show the total, forced, and unforced T_g variance calculated in sliding 30-year windows, where the x-axis indicates the end year of each window. For 30-year windows ending before 1980 (i.e., before the first green line), the forced T_g variance is generally weaker than the unforced T_g variance. However, since the 1980s, the forced T_g variance has strengthened, while the unforced T_g variance remains of similar magnitudes. This different time evolution between forced and unforced T_g variance implies that forced responses have weighed more in OLS regressions after the 1980s.

Indeed, the ratio r (defined in Eq. 2.12), which quantifies the relative importance of forced signals, remains small for the 30-year windows ending before 1980 (Figs. 1g-i). The averaged r before 1980 (i.e., the average over multiple windows) is 0.18 ± 0.02 in CESM2-LE, 0.28 ± 0.03 in MPI-GE, and 0.15 ± 0.02 in GISS-LE, where the ensemble mean and 1 standard deviation (STD) across ensembles are shown. After 1980, r increases rapidly, in parallel with the rapid increase in GHG emissions. The forced and unforced T_g variances are comparable between the end year of the 1990s and the end year of the early 2000s ($r \sim 0.5$). As $var(T_{g,f})$ continues to strengthen, $var(T_{g,f})$ generally outweighs $var(T_{g,u})$ in the late 2000s ($r > 0.5$) and has become more and more dominant since then. In GISS-LE, the overtake of forced signals in the 2000s is less obvious than in the other two large ensembles. Take the end year of 2010 (1981-2010 window) for example, $r = 0.72 \pm 0.06$ in CESM2-LE, $r = 0.72 \pm 0.07$ in MPI-GE, and $r = 0.58 \pm 0.05$ in GISS-LE. Despite the weaker r in 1981-2010 window in GISS-LE, all three models show a pronounced increase in r between the 1951-1980 window and the 1981-2010 window (shown as the two vertical lines in

Fig. 1; see the numbers in Table 1), suggesting an increasing dominance of forced signals in SST patterns and radiative feedbacks over the past few decades.

While r is generally weak before the end year of 1980, we note that there are two local r maxima in all three models, which can be linked to major volcanic eruptions (Gregory et al. 2016). The first local maximum spans between the end year 1890 to the end year 1910 (i.e., the sliding 30-year windows from 1861-1890 to 1881-1910), arising from the volcanic eruption of Krakatau in 1883 and the Santa Maria eruption in 1902. The second local maximum exists around the end year of 1930 (1901-1930 window), which includes the influences from both the 1902 Santa Maria eruption and the 1912 Novarupta/Katmai eruption. Furthermore, there is a local minimum around the end year of 1960, consistent with the decrease in major volcanic eruptions between 1920-1960.

	CESM2-LE	MPI-GE	GISS-LE
$r\|_{1951-1980}$ (1)	0.24 ± 0.07	0.38 ± 0.08	0.25 ± 0.04
$r\|_{1981-2010}$ (1)	0.72 ± 0.06	0.72 ± 0.07	0.58 ± 0.05
C_{SST} ave. $(\frac{W}{m^2K})$	0.57 ± 0.04	0.37 ± 0.04	0.20 ± 0.04
$C_{SST,f}$ ave. $(\frac{W}{m^2K})$	0.49	0.50	0.44
$C_{SST,u}$ ave. $(\frac{W}{m^2K})$	0.59 ± 0.04	0.31 ± 0.05	0.14 ± 0.04
C_{EIS} ave. $(\frac{W}{m^2K})$	-0.05 ± 0.10	-0.57 ± 0.10	-0.63 ± 0.08
$C_{EIS,f}$ ave. $(\frac{W}{m^2K})$	-0.62	-0.86	-0.68
$C_{EIS,u}$ ave. $(\frac{W}{m^2K})$	0.14 ± 0.10	-0.39 ± 0.11	-0.62 ± 0.08
$IPWP\|_{1951-1980}$ (1)	0.44 ± 0.11	0.82 ± 0.07	0.78 ± 0.05
$IPWP_f\|_{1951-1980}$ (1)	1.09	1.04	0.97
$IPWP_u\|_{1951-1980}$ (1)	0.30 ± 0.10	0.70 ± 0.07	0.73 ± 0.04
$IPWP\|_{1981-2010}$ (1)	0.77 ± 0.10	0.96 ± 0.07	0.86 ± 0.05
$IPWP_f\|_{1981-2010}$ (1)	1.03	1.09	1.03
$IPWP_u\|_{1981-2010}$ (1)	0.32 ± 0.10	0.74 ± 0.08	0.71 ± 0.05

TABLE 1. Indices used to explain the time variation of low-cloud feedback (C) in the three initial-condition large ensemble simulations. r is the ratio of the forced T_g variance (Eq. 2.12). C_{SST} and C_{EIS} indicate the low-cloud feedback due to changes in SST patterns and EIS patterns, and the subscripts $_f$ and $_u$ denote the forced and unforced components (Eqs. 3.1-3.4). The "average" in rows 3-8 indicates the average of multiple 30-year sliding windows before the end year of 1980. Also, the Indo-Pacific Warm Pool warming ratio, noted as $IPWP$, is calculated as the regional average of $\frac{dSST}{dT_g}$ in the western Pacific convective regions (30°S-30°N, 50°E-160°W) over the tropical average of $\frac{dSST}{dT_g}$ (30°S-30°N). The pipe symbol (|) in this table is followed by the 30-year window that is used to calculate the targeted field. Note that the forced components are calculated based on the ensemble-mean fields, thus no spread across ensemble members is shown. For the total and unforced components, the ensemble-mean values and 1 STD across ensembles are shown.

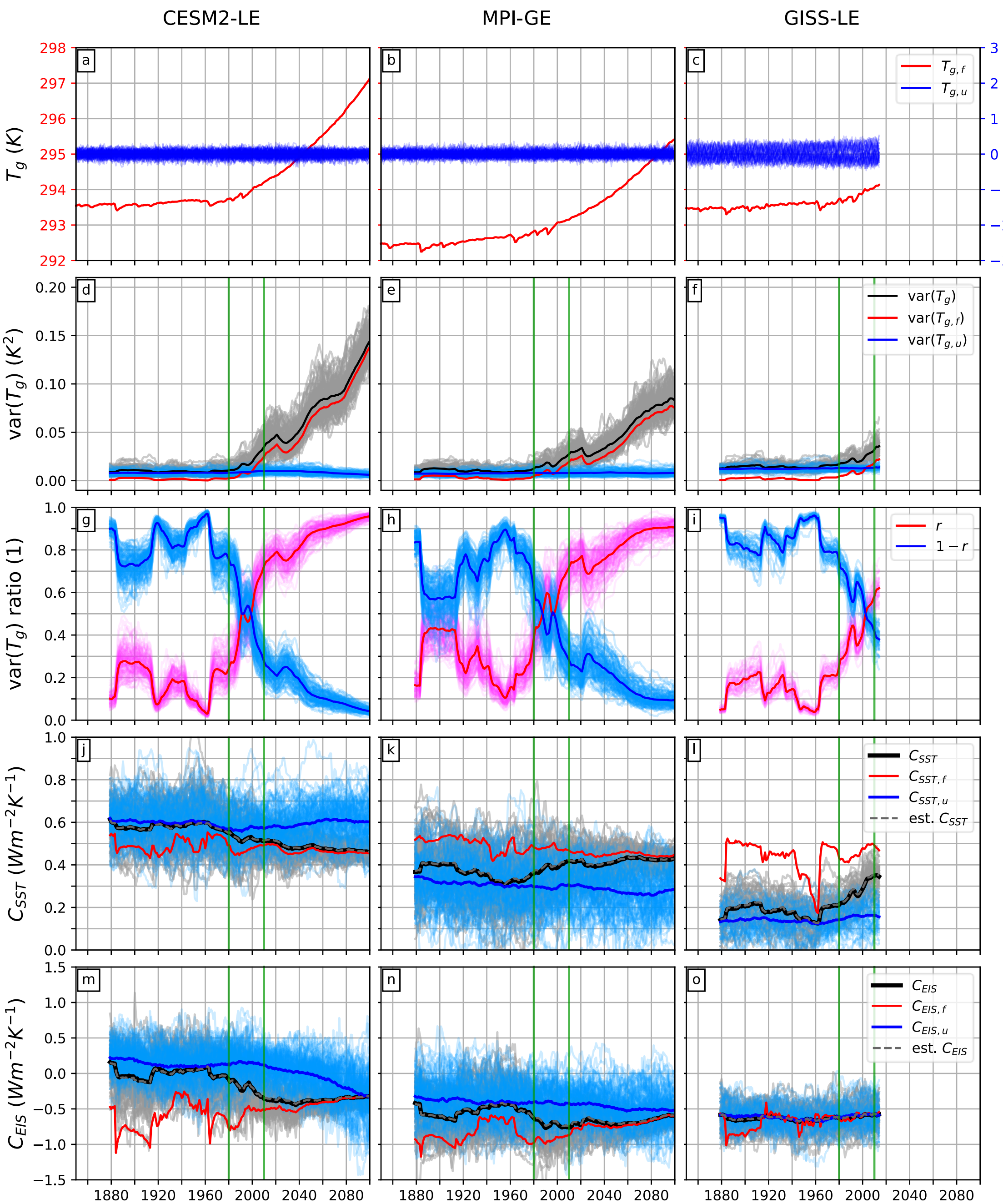

CESM2-LE
MPI-GE
GISS-LE
T_g (K)
$T_{g,f}$
$T_{g,u}$
var(T_g) (K^2)
var(T_g)
var($T_{g,f}$)
var($T_{g,u}$)
var(T_g) ratio (1)
r
$1-r$
C_{SST} $(Wm^{-2}K^{-1})$
C_{SST}
$C_{SST,f}$
$C_{SST,u}$
est. C_{SST}
C_{EIS} $(Wm^{-2}K^{-1})$
C_{EIS}
$C_{EIS,f}$
$C_{EIS,u}$
est. C_{EIS}

FIG. 1. Time evolution of forced (red) and unforced (blue) T_g in (a) CESM2-LE (b) MPI-GE, and (c) GISS-LE. Here T_g is the area-weighted average of surface temperature within 60°S-60°N over the ocean. (d-f) Time evolution of the variance of T_g (black), $T_{g,f}$ (red), and $T_{g,u}$ (blue) in the three LEs. The variation is calculated on a sliding 30-year window and the x-axis denotes the end year for each window. (g-i) The ratio of the $T_{g,f}$ variance (red; r defined in Eq. 2.12) and the ratio of the $T_{g,u}$ variance (blue; $1-r$) in the three LEs. (j-l) Time evolution of low-cloud feedback due to changes in full SST pattern (C_{SST}), forced SST pattern ($C_{SST,f}$; red), and unforced SST pattern ($C_{SST,u}$; blue) in the three LEs. Est. C_{SST} (dashed gray) is calculated as the ensemble mean of $rC_{SST,f}+(1-r)C_{SST,u}$, meaning that the difference between black and dashed gray line is the residual term for ensemble mean. (m-o) Similar to (j-l), but for changes in full EIS pattern (C_{EIS}; black), forced EIS pattern ($C_{EIS,f}$; red), and unforced EIS pattern ($C_{EIS,u}$; blue). Note that the range of y-axis in C_{EIS} is three times larger than in C_{SST}. For each figure, dark-colored lines indicate the ensemble mean and light-colored lines show each ensemble member. The end year of 1980 and 2010 is marked in (d)-(o), which respectively shows the results from 1951-1980 and 1981-2010 window that we focus on.

3. Time-evolving low-cloud feedback

a. Forced and unforced contributions to low-cloud feedback

Previous research has suggested that the responses of marine low clouds are the primary source of inter-model spread in climate sensitivity estimates (e.g., Bony and Dufresne (2005)), and that the change in cloud radiative feedback is responsible for the time evolution of net radiative feedback (e.g. Zhou et al. (2016)). Therefore, here we focus on the time evolution of low-cloud radiative feedback by combining the changes in SST and estimated inversion strength (EIS; Wood and Bretherton (2006)) per unit warming with observation-based meteorological cloud radiative kernels (Scott et al. 2020; Myers et al. 2021), as illustrated below.

$$C_{SST} = \frac{\partial R}{\partial SST}\frac{dSST}{dT_g}, \tag{3.1}$$

$$C_{EIS} = \frac{\partial R}{\partial EIS}\frac{dEIS}{dT_g}. \tag{3.2}$$

C_{SST} and C_{EIS} indicate the low-cloud radiative feedback due to changes in SST and EIS pattern, respectively. $\frac{\partial R}{\partial SST}$ and $\frac{\partial R}{\partial EIS}$ are the meteorological cloud radiative kernels that quantify low-cloud radiative responses to local SST and EIS perturbations, developed by Scott et al. (2020). Note that the meteorological cloud radiative kernels are evaluated separately in four different observational datasets, and we adopt the average of four kernels due to their similar patterns and overall magnitudes (Scott et al. 2020; Myers et al. 2021). More importantly, since the meteorological cloud radiative kernels are time-invariant and model-independent, any time dependence of the low-cloud radiative feedback (C) analyzed here arises from the time evolution of SST or EIS patterns. The inter-model spread in C can also be fully attributed to the spread in SST or EIS patterns.

To evaluate the relative contributions from forced and unforced patterns of SST and EIS to the time-evolving C, we combine Equation 2.11 with Equations 3.1-3.2:

$$C_{SST} = \frac{\partial R}{\partial SST}[\frac{dSST_f}{dT_{g,f}}r + \frac{dSST_u}{dT_{g,u}}(1-r)+\sigma] = C_{SST,f}r + C_{SST,u}(1-r)+\epsilon, \tag{3.3}$$

$$C_{EIS} = \frac{\partial R}{\partial EIS}[\frac{dEIS_f}{dT_{g,f}}r + \frac{dEIS_u}{dT_{g,u}}(1-r)+\sigma] = C_{EIS,f}r + C_{EIS,u}(1-r)+\epsilon, \tag{3.4}$$

where $C_{SST,f}$ is the forced component and $C_{SST,u}$ is the unforced component of the low-cloud feedback that arises from the local impact of the SST pattern. Similarly, $C_{EIS,f}$ and $C_{EIS,u}$ indicate the forced and unforced component of the EIS-related low-cloud feedback, which can be mostly attributed to the remote impact of the SST patterns. ϵ represents the residual and is simply σ (Eq. 2.11) multiplied by time-invariant meteorological cloud radiative kernels.

The main advantage of our low-cloud feedback evaluation is to isolate the influences of SST and EIS patterns on low-cloud feedback from other factors, such as the inter-model spread of time-evolving radiative forcing (F; Pincus et al. (2016)) and the uncertainty of low-cloud radiative responses to SST and EIS perturbations. Moreover, using meteorological cloud radiative kernels, we have constrained observationally the dependence of low-cloud radiative effects on meteorology. However, the caveat is that the low-cloud feedback evaluated here could be different from the low-cloud feedback estimated exclusively in the models (i.e., allowing for model-specific coefficients).

b. Similarity and disparity among models

Following Equations 3.3-3.4, here we review the time variation of low-cloud feedback from each AOGCM large ensembles and explain the shift in low-cloud feedback over the past few decades. First, most of the ensemble members in CESM2-LE suggest a trend toward more negative C_{SST} and C_{EIS} between the end year of 1960 and the end year of 2010, with EIS having a stronger trend (black lines in Figs. 1j, m). Despite this negative trend in both C_{SST} and C_{EIS}, we can barely see the corresponding change in either forced or unforced components. Instead, we find that the negative trend of C_{SST} and C_{EIS} is driven by changes in the relative importance of forced and unforced components. Between the end year of 1960 and the end year of 2010, there is a transition from being dominated by unforced signals (small r) to being dominated by forced signals (large r; Fig. 1g). When unforced signals dominate, the overall C is largely determined by its unforced component, thus the two have similar magnitudes (closer blue and black lines when r is small). Similarly, when forced signals dominate, the overall C is largely determined by the forced component (closer red and black lines when r is large). For both C_{SST} and C_{EIS}, since the unforced component is generally more positive than the forced component, the decreasing importance of unforced feedback (i.e., the increasing importance of forced feedback) in recent decades gives rise to a more-negative (more-stabilizing) low-cloud feedback during this time.

Similar explanations can be applied to MPI-GE and GISS-LE. For example, if the unforced feedback component is more positive than the forced feedback component, such as C_{EIS} in CESM2-LE and MPI-GE and C_{SST} in CESM2-LE (Table 1), the increasing importance of forced signals implies a negative trend of the overall low-cloud feedback (Figs. 1j,m,n). If the forced feedback component is more positive than the unforced feedback component (e.g., C_{SST} in MPI-GE and GISS-LE; Table 1), the increasing importance of forced signals then implies a positive trend of the overall feedback (Figs. 1k,l). If, in the last case, the forced and unforced feedback have similar values (e.g., C_{EIS} in GISS-LE; Table 1), the overall feedback would barely change while r varies over time (Fig. 1o). For all three models, the ensemble-mean residual (ϵ) is negligible, shown as the difference between the gray dashed lines and the black lines in Figures. 1j-o. The result indicates small covariances of $cov(T_{g,f}, T_{g,u})$, $cov(X_f, T_{g,u})$, and $cov(X_u, T_{g,f})$ with X as SST or EIS, and that the assumption made in Section 2c is generally valid.

By comparing the forced and unforced feedback among the three large ensembles, we also find that the inter-model spread of the C_{SST} and C_{EIS} arises mostly from the unforced component instead of the forced component. For SST contribution, $C_{SST,f}$ is 0.49, 0.50, and 0.44 $\frac{W}{m^2K}$ in CESM2-LE, MPI-GE, and GISS-LE, respectively. However, $C_{SST,u}$ is 0.59 ± 0.04, 0.31 ± 0.05, and 0.14 ± 0.04 $\frac{W}{m^2K}$ at the same model order. As for the EIS contribution, the spread of unforced feedback is even larger to the extent that the sign is also uncertain. $C_{EIS,u}$ is positive (0.14 ± 0.10 $\frac{W}{m^2K}$) in CESM2-LE, while it is negative in MPI-GE and GISS-LE (-0.39 ± 0.11 and -0.62 ± 0.08 $\frac{W}{m^2K}$, respectively). At the same time, the forced components have the same negative sign and similar magnitudes ($C_{EIS,f} = -0.62$, -0.86, and -0.68 $\frac{W}{m^2K}$ in CESM2-LE, MPI-GE, and GISS-LE, respectively).

c. Attribution of time-evolving low-cloud feedback

As shown in Equations 3.3-3.4, the temporal evolution of low-cloud feedback (C_{SST} and C_{EIS}) can be driven by three possible components: (1) changes in the forced low-cloud feedback ($C_{SST,f}$ and $C_{EIS,f}$) (2) changes in the unforced low-cloud feedback ($C_{SST,u}$ and $C_{EIS,u}$), and (3) changes in the relative importance between the forced and unforced signals, expressed as the ratio r (Eq. 2.12). Since the derivation for C_{SST} and C_{EIS} is identical, we will drop the subscripts and write the general form for low-cloud feedback as:

$$C(t) = C_f(t)r(t) + C_u(t)[1 - r(t)] + \epsilon(t). \tag{3.5}$$

Here t indicates a given 30-year window used to calculate the feedback and the ratio. For the next 30-year window, we can write the same form with $t = t + 1$. The change in C between the two adjacent 30-year windows is then expressed as:

$$\delta C(t) = \delta C_f(t)\overline{r(t)} + \delta C_u(t)[1 - \overline{r(t)}] + \delta r(t)[\overline{C_f(t)} - \overline{C_u(t)}] + \delta\epsilon(t), \tag{3.6}$$

$$\delta X(t) = X(t+1) - X(t), \tag{3.7}$$

$$\overline{X(t)} = \frac{X(t+1) + X(t)}{2}. \tag{3.8}$$

X can be C, C_f, C_u, or r. To attribute the C difference between the two non-adjacent windows, for example, the 30-year windows 1951-1980 and 1981-2010, we can sum all the $\delta C(t)$ between the two:

$$\sum_{t=1951-1980}^{t=1980-2009} \delta C(t) = \sum_{t=1951-1980}^{t=1980-2009} \{\delta C_f(t)\overline{r(t)} + \delta C_u(t)[1 - \overline{r(t)}] + \delta r(t)[\overline{C_f(t)} - \overline{C_u(t)}] + \delta\epsilon(t)\}. \tag{3.9}$$

For clarity, we omit the time index and re-write Equation 3.9 into a more general form:

$$\Delta C = \sum \delta C_f \overline{r} + \sum \delta C_u (1 - \overline{r}) + \sum \delta r (\overline{C_f} - \overline{C_u}) + \Delta\epsilon. \tag{3.10}$$

Δ denotes the C difference given two windows and is simply the sum of all the differences from adjacent windows between the two. Using Equation 3.10, we attribute the change in low-cloud feedback between any two windows to the contribution of forced feedback changes (the first term on the RHS), followed by the contribution of unforced feedback changes and the contribution of the ratio changes (the second and the third term on the RHS). $\Delta\epsilon$ again indicates the residual, which is associated with the combined effects from the covariance between forced and unforced signals (see Section 2c for more details).

1) Historical period

Figure 2 shows the ΔC decomposition between the 30-year window 1951-1980 and 1981-2010. In both CESM2-LE and MPI-GE, ΔC_{EIS} is stronger than ΔC_{SST} and suggests a negative shift of low-cloud feedback during this period, consistent with Figure 1. The decomposition also reveals that the change in r is the primary reason for the negative shift of C_{EIS}, shown as strong negative values of Δ_3. In CESM2-LE, all ensemble members agree that the increasing importance of the forced signals (increasing r) leads to negative ΔC_{EIS} (Fig. 2d). More than 75% of the ensemble members in MPI-GE agree with the above result (Fig. 2e). Meanwhile, the change in the forced component gives rise to a small increase in low-cloud feedback (Δ_1). The influence of the unforced feedback change varies among ensembles and has no robust contribution to ΔC in recent decades (Δ_2). In GISS-LE, Δ_3 contributes to the weak positive shift of C_{SST} (Fig. 2c). As for ΔC_{EIS}, the strength of the forced and unforced components is similar, therefore the contribution of Δr is weak and insignificant (Fig. 2f).

By decomposing the low-cloud feedback change between 1951-1980 and 1981-2010, we summarize that the increasing importance of forced signals (increasing r) is the main cause for the shift in low-cloud feedback over the past few decades in CESM2-LE and MPI-GE.

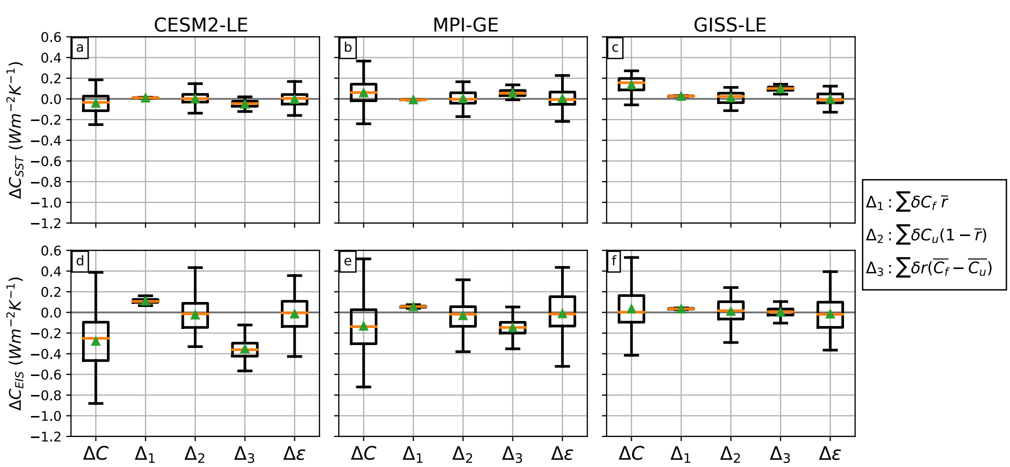


Fig. 2. Decomposition of ΔC_{SST} between the 30-year window 1951-1980 and 1981-2010 in (a) CESM2-LE (b) MPI-GE, and (c) GISS-LE. (d-f) Same as (a-c), but for ΔC_{EIS}. Δ_1, Δ_2, and Δ_3 denote the first, the second, and the third term RHS of Equation 3.10. In each box plot, the orange line indicates the median and the green triangle indicates the ensemble mean.

2) Future warming scenarios

In addition to the shift in low-cloud feedback (C) over the past few decades, we notice that C also evolves with time in future warming projections. In general, we find that C_{EIS} becomes more positive (less stabilizing) throughout the century in the SSP370 simulations of CEMS2-LE and the RCP8.5 simulations of MPI-GE, while the change in C_{SST} is relatively weak (Fig. 1). The result is consistent with previous studies suggesting a less-stabilizing net radiative feedback over time due to EIS changes in CO_2-increasing simulations (Rose and Rayborn 2016; Ceppi and Gregory 2017, 2019; Dong et al. 2020; Lin et al. 2021).

To quantify and attribute the change in C_{EIS}, we decompose ΔC_{EIS} between the current climate (i.e., 1981-2010) and the projected climate at the end of the century (i.e., 2071-2099; Fig. 3). More than 50% ensemble members in CESM2-LE and more than 75% ensemble members from MPI-GE show positive ΔC_{EIS} in response to future warming. More importantly, this long-term positive change in C_{EIS} arises mainly from changes in the forced component, shown as strong positive values of Δ_1 (Figs. 3c,d). Changes in either unforced component (Δ_2) or the relative importance between forced and unforced signals (Δ_3) instead lead to a more-negative C_{EIS} for most of the ensemble members.

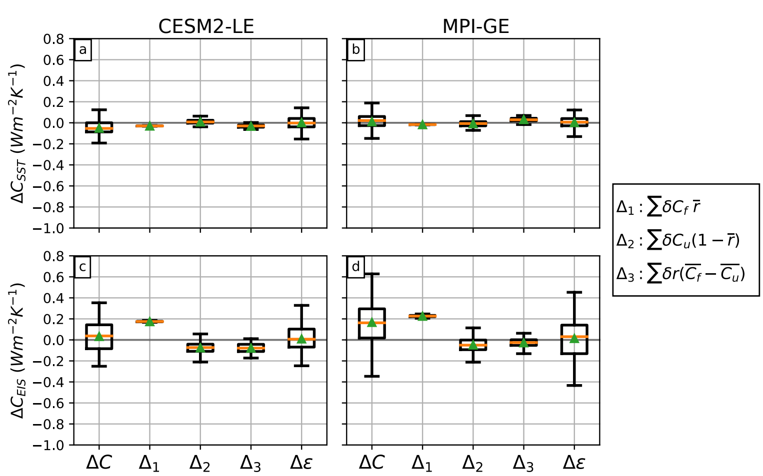


FIG. 3. Decomposition of ΔC_{SST} between the 30-year window 1981-2010 and 2070-2099 in (a) CESM2-LE and (b) MPI-GE. (c-d) Same as (a-b), but for ΔC_{EIS}. Δ_1, Δ_2, and Δ_3 denote the first, the second, and the third term RHS of Equation 3.10. In each box plot, the orange line indicates the median and the green triangle indicates the ensemble mean.

4. The role of SST pattern effect

a. Overview

In this research, the time variation of low-cloud feedback (C_{SST} and C_{EIS}) solely depends on changes in SST and EIS regressions, including their forced and unforced components. To build a physical understanding that connects the two, we compare the time-evolving SST and EIS regressions among the three large ensembles (Figs. 4-5) and highlight the role of SST pattern effect in setting the time-evolving C. The spatial patterns $\frac{dSST}{dT_g}$ and $\frac{dEIS}{dT_g}$ are obtained by regressing local SST and EIS onto global-mean temperature change and separating into forced and unforced components following Equation 2.11.

For all three large ensembles, the overall SST pattern ($\frac{dSST}{dT_g}$; first column of Fig. 4) is determined by both the forced ($\frac{dSST_f}{dT_{g,f}}$; second column) and the unforced component ($\frac{dSST_u}{dT_{g,u}}$; third column), depending on their relative importance indicated by r (Fig. 1; Table 1). The residual (σ; fourth column of Fig. 4) remains small throughout different time periods, again indicating that the assumption made in Section 2c is valid. The forced SST pattern characterizes more uniform warming per unit increase in $T_{g,f}$ while the unforced SST pattern is more heterogeneous per unit change in $T_{g,u}$. All three models commonly show this fundamental difference between the forced and unforced SST patterns from the preindustrial era to the end of the 21st century.

b. Historical period

Within the large ensembles, the unforced SST pattern is similar to the SST anomalies from the prevailing climate variability in interannual timescales - El Niño Southern Oscillation (ENSO), which we first illustrate for CESM2-LE. Per unit increase in T_g, there is an enhanced surface warming in the EP and surface cooling in the WP (Fig. 4, first row, third column). Due to the small r in the 30-year window of 1951-1980 ($r = 0.24 \pm 0.07$), the overall SST pattern ($\frac{dSST}{dT_g}$) also holds the ENSO-like SST features (Fig. 4, first row, first column). The surface cooling in the WP convective regions leads to an overall cooling in the free troposphere in the tropics, which destabilizes the low-level troposphere. The destabilization is particularly strong in the EP because of the substantial contrast between the free-tropospheric cooling and the surface enhanced warming (Fig. 5, first row, first and third column). Also, the low-level destabilization acts to decrease the marine stratocumulus cloud over the EP (Wood and Bretherton 2006), accounting for more-positive low-cloud feedback during this time.

In the 30-year window of 1981-2010, on the other hand, the overall SST pattern is largely affected by the forced signals ($r = 0.72 \pm 0.06$). Compared to the unforced SST pattern, the forced SST pattern is more spatially uniform, with slightly enhanced warming in the Northern Hemisphere (NH) and reduced warming in the Southern Hemisphere (SH). The surface warming in the WP convective regions is also stronger than that in the EP stratocumulus cloud regions (Fig. 4, second row, second column). As a result, the overall SST pattern in 1981-2010 is less ENSO-like compared to that in 1951-1980 (Fig. 4, second row, first column), corresponding to a more-stable low-level troposphere in the EP (Fig. 5, second row, first column). The 1981-2010 meteorology condition favors low-cloud formation and more-stabilizing low-cloud feedback in comparison to that in 1951-1980.

The above-mentioned mechanism applies to all three LEs, highlighting the distinct forced and unforced SST patterns that jointly shape the overall SST pattern with time-dependent weighting (r) for each component. We particularly focus on the warming contrast between the WP and EP, which explains the time variation of low-cloud feedback through modifying low-level stability (C_{EIS}), while C_{SST} that quantifies local SST impacts plays a minor role in adjusting the C variations in recent decades. To further quantify this radiatively essential SST pattern, we define the Indo-Pacific Warm Pool warming ratio ($IPWP$) as the regional average of $\frac{dSST}{dT_g}$ in the western Pacific convective regions (30°S-30°N, 50°E-160°W) over the tropical average of $\frac{dSST}{dT_g}$ (30°S-30°N), consistent with the quantification proposed in Dong et al. (2019) and Wills et al. (2022).

In CESM2-LE, the $IPWP$ index becomes larger from 0.44 ± 0.11 in 1951-1980 to $= 0.77 \pm 0.10$ in 1981-2010 (Table 1), associated with more-stabilizing low-cloud feedback. The increased $IPWP$ index can be explained by the increase in r, along with small unforced $IPWP$ indices and large forced $IPWP$ indices. In MPI-GE and GISS-LE, we also observe an increase in the $IPWP$ index but both with weaker magnitudes (Table 1). Given the similar time evolution of r among models and the large inter-model spread of unforced C_{EIS}, we propose that the inter-model spread of $IPWP$ time evolution arises mostly from the unforced $IPWP$ index.

Indeed, all three models produce similar forced $IPWP$ index, ranging from 0.97 to 1.09 in 1951-1980 and 1.03 to 1.09 in 1981-2010. The close-to-unity values in both time periods indicate that the WP warming is similar to the overall warming in the tropics (i.e., spatially uniform forced SST responses). However, the unforced $IPWP$ index varies widely among models. In 1951-1980, $IPWP_u$ is 0.30 ± 0.10 in CESM2-LE, associated with the ENSO-like unforced SST pattern in which surface cooling occurs in the WP convective regions (Fig. 4, first row, third column). Meanwhile, this WP cooling is much more limited and weaker in MPI-GE despite the model still projecting an ENSO-like unforced SST pattern (Fig. 4, fourth row, third column). In GISS-LE, there is barely any cooling in the WP convective regions. Surface warming is strong in both the tropical WP and EP regions (Fig. 4, seventh row, third column). This weak-to-no cooling in the WP region would correspond to a higher $IPWP$ index in both models (0.70 ± 0.07 in MPI-GE and 0.73 ± 0.04 in GISS-LE) in 1951-1980, limiting the $IPWP$ increase in 1981-2010 when forced signals has become more dominated. In summary, the large inter-model spread of unforced WP relative warming ($IPWP_u$) echoes the large spread of unforced $C_{EIS,u}$ (Fig. 1), which is responsible not only for the inter-model spread of C_{EIS}, but also the diverse time evolution of C_{EIS} among models.

c. Future warming scenarios

The relative importance of the forced response (r) outweighs the unforced variability for the 30-year window ending around the 2010s and has become increasingly dominant over time since then. At the end of the 21st century, r reaches 0.85-0.95, depending on the models and warming scenarios (Fig. 1). Comparing the SST patterns between the window 1981-2010 and 2071-2099, we find that the changes in the overall SST pattern arise mostly from the changes in the forced component, except that some of the tropical regions are still influenced by unforced variability in the earlier period. The result is expected since r has been large, indicating a weak influence from the unforced variability. The change in forced SST pattern features delayed warming in the southeastern Pacific and the Southern Ocean (Fig. 4), corresponding to a less-positive EIS in the two regions (Fig. 5) and a less-stabilizing C_{EIS}. The time-evolving SST and EIS patterns explain the forced contribution to the positive ΔC_{EIS} shown in Figure 3, highlighting the role of the pattern effect.

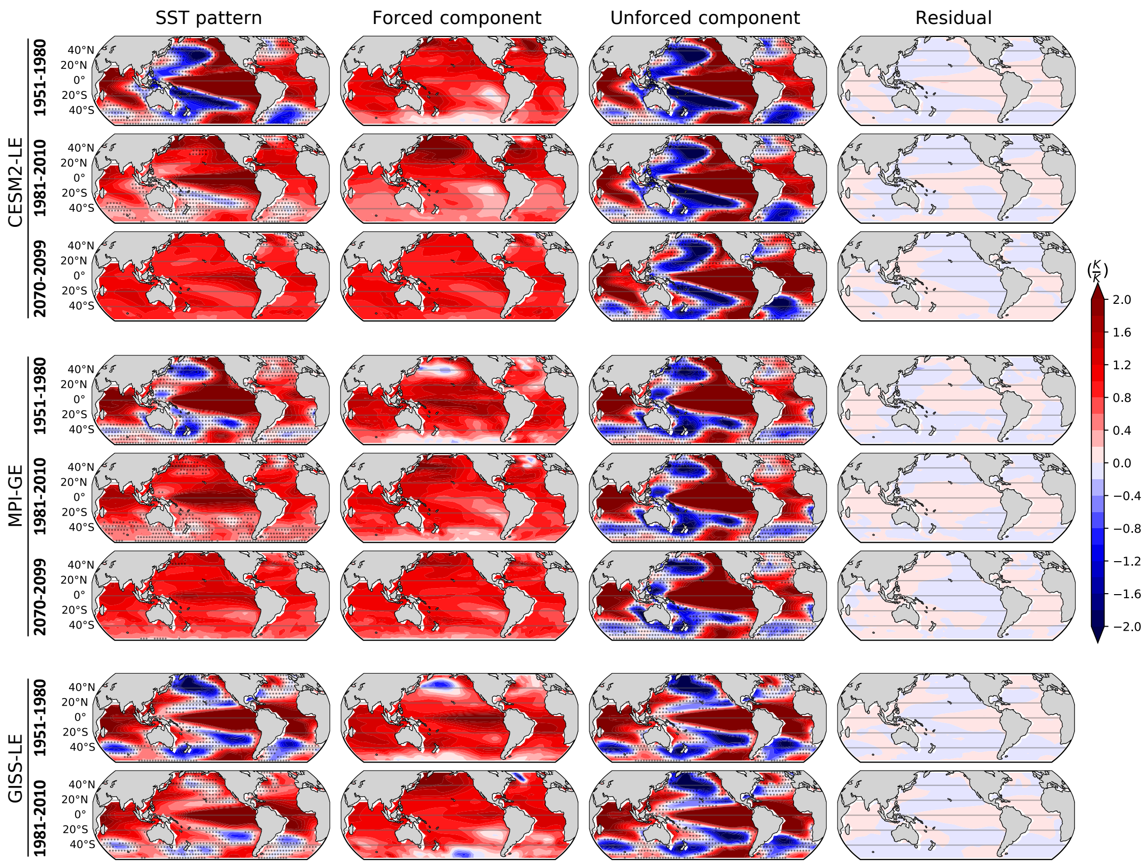


FIG. 4. (from left to right) Overall SST pattern ($\frac{dT}{dT_g}$), the forced component ($\frac{dT_f}{dT_{g,f}}$), the unforced component ($\frac{dT_u}{dT_{g,u}}$), and the residual term (σ) calculated from different 30-year windows in the three AOGCM large ensembles. For each panel, the 30-year window and the large ensembles used are labeled on the left. In the overall and unforced SST patterns, contours indicate the ensemble mean and stippling indicates the regions where the 1 STD calculated across ensembles is larger than the ensemble-mean values. For the forced SST pattern, no stippling is shown since it is calculated based on the ensemble-mean fields before applying the OLS regression.

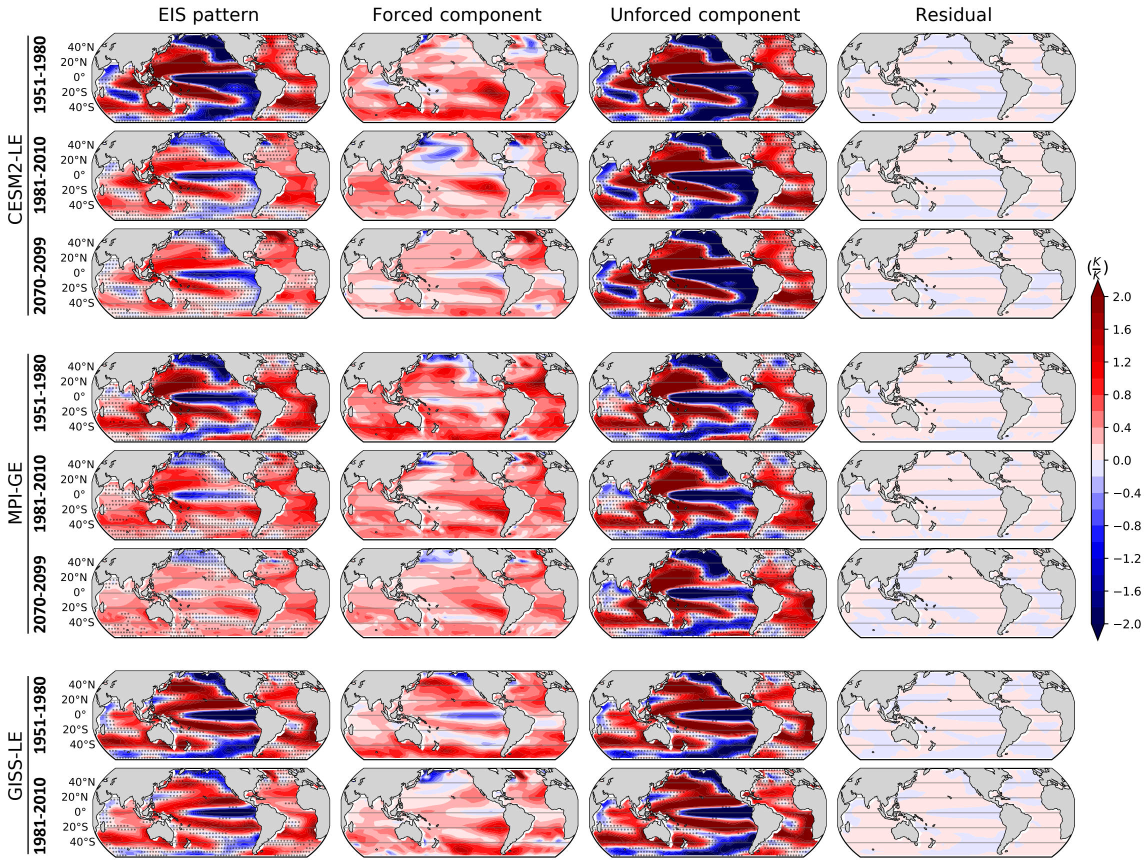


FIG. 5. Same as Figure 4, but for overall EIS pattern ($\frac{dEIS}{dT_g}$), the forced component ($\frac{dEIS_f}{dT_{g,f}}$), the unforced component ($\frac{dEIS_u}{dT_{g,u}}$), and the residual term (σ) calculated from each 30-year window in each AOGCM large ensembles.

5. Signal partition based on linear trends

a. Overview

In previous sections, we separate forced and unforced signals through initial-condition large ensembles. However, the real world has only one realization with mixed forced and unforced signals. Also, not all the models conduct initial-condition large ensembles. Therefore, we developed a method to approximately discriminate the contributions from forced and unforced signals that can be applied to both observations and models (when using only one ensemble member). In particular, we can write:

$$X = X_{tr} + X_{de}, \tag{5.1}$$

where X_{tr} indicates the linear-trend part of the responses calculated in a given 30-year window and X_{de} represents the linearly detrended part. The trend component captures relatively long-term responses and is more likely to be driven by accumulative forcings, such as greenhouse gas emissions. However, X_{tr} does not equal forced responses ($X_f \neq X_{tr}$). For example, linear trends hardly capture the responses to volcanic eruptions, which have short-term yet strong impacts on both temperature and radiation. Other natural or anthropogenic forcings might have nonlinear influences as well, which won't be fully represented by linear trends. At the same time, while X_{de} might be able to capture interannual-to-decadal unforced variations (e.g., ENSO), it does not capture multi-decadal variability that fluctuates on a timescale longer than 30 years since the data are detrended on sliding 30-year windows. In other words, multi-decadal variability will be included as part of the trend component, so X_{de} does not equal unforced responses as well ($X_u \neq X_{de}$). The separation of forced and unforced signals in the observation or in a single model realization has remained a difficult issue in multiple research fields. While trend and detrend components can not be directly approximated to forced and unforced components as in large ensembles, they might serve as *proxies* of forced and unforced components as they still show the basic differences between the two. Similar to the equation derived in Section 2c, the impacts of trend and detrend components on OLS regressions can be evaluated as follows.

$$\frac{dX}{dT_g} = \frac{dX_{tr}}{dT_{g,tr}} r_{tr} + \frac{dX_{de}}{dT_{g,de}} (1 - r_{tr}) + \sigma, \tag{5.2}$$

$$r_{tr} = \frac{var(T_{g,tr})}{var(T_{g,tr}) + var(T_{g,de})}. \tag{5.3}$$

The above equations are the same as Equations 2.11-2.12, with X_f replaced by X_{tr} and X_u replaced by X_{de}. The residual (σ) here accounts for the combined effect of $cov(T_{g,tr}, T_{g,de})$, $cov(X_{tr}, T_{g,de})$, and $cov(X_{de}, T_{g,tr})$. To distinguish r from the two different approaches, we express the relative importance of the trend component as r_{tr}, as opposed to r which indicates the relative importance of the forced component.

One way to validate the trend/detrend components as proxies of forced/unforced responses is to compare these two approaches of signal partition in the same large ensembles. Figure 6 shows C_{SST} and C_{EIS} in the three large ensembles with their trend/detrend components and the relative importance of each component. We find that the trend and detrend components can reproduce a number of key features of the forced and unforced components, respectively. For example, both r and r_{tr} have increased rapidly over the past few decades, and the two indices remain similar throughout the century (compare Figs. 6d-f with Figs. 1g-i). The consistency between the two approaches underscores the recent intensification of forced responses, which will dominate the sliding linear trends.

The major difference between r and r_{tr} occurs for the 30-year windows ending prior to the 1990s. r_{tr} is generally weaker than r. Here we attribute the difference between r and r_{tr} mainly to volcanic forcings based on the following reasons. First, the climate impacts of volcanic eruptions are strong yet relatively short-term. The temperature variation within a short time period is hardly captured by 30-year sliding linear trends, leading to a weaker r_{tr}. Also, the r_{tr} time variation does not include characteristics related to volcanic eruptions as in r (Section 2d), namely, r_{tr} does not strengthen with more-frequent eruptions and supposedly larger temperature variations, and does not weaken with less-frequent eruptions as well (e.g., 1940-1960). As a result, we explore accounting for volcanic eruptions in the trend and detrend components (method described in Text S1). We find that r_{tr} becomes larger and contains time-varying features related to eruptions, as in r. However, because our main focus is on recent decades when the increasing importance of forced responses is already well represented by the trend component without accounting for volcanic eruptions, the conclusion of this study is unchanged with and without accounting for volcanic eruptions. For simplicity, here we only show the results based on sliding linear trends (trend component) and the anomalies (detrend component) without accounting for volcanic forcings, and leave the results with volcanic forcings accounted in the Supplemental Material (Figs. S1-S4; Tables S2-S3).

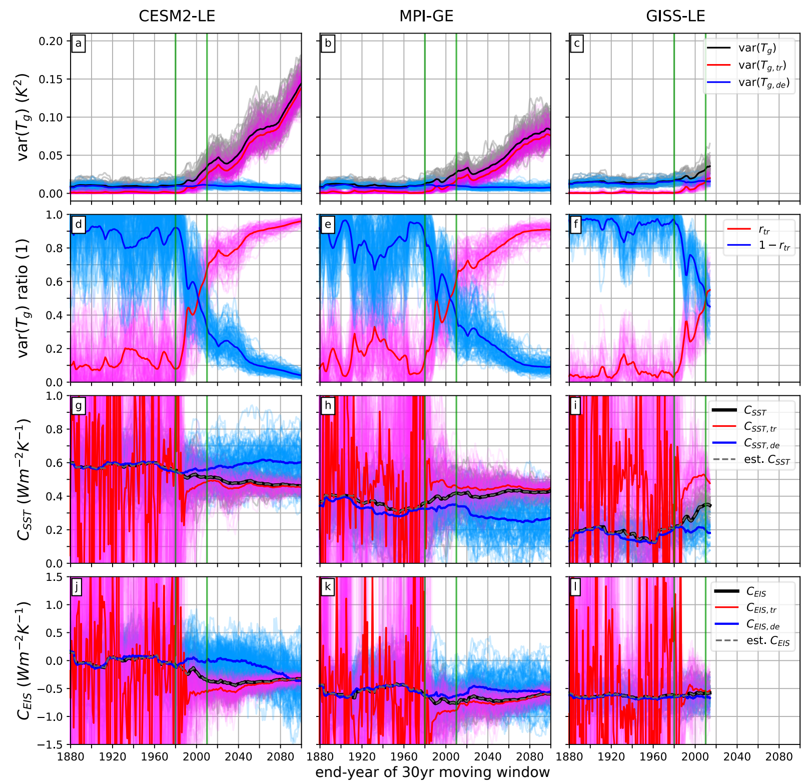


FIG. 6. Same as Figs. 1d-o, but for the overall (black), trend component (red), and detrend component (blue) of each field in three large ensemble simulations.

b. Relative importance of trend and detrend components

Having validated the application of trend and detrend components, we proceed to examine 23 atmosphere-ocean coupled models participating in Coupled Model Intercomparison Project phase 6 (CMIP6; Eyring et al. (2016)), along with three observational datasets. For each CMIP6 model, only one ensemble member is used for the historical and SSP370 simulations. In observations, we analyze time-varying SST patterns from version 1.1 of the Met Office Hadley Centre sea ice and sea surface temperature data set (HadISST1.1; Rayner et al. (2003)), NOAA Extended Reconstruction SSTs Version 5 (ERSSTv5; Huang et al. (2017)), and NOAA-CIRES-DOE Twentieth Century Reanalysis version 3 (20CRv3). Time-varying patterns of EIS can also be calculated from NOAA-CIRES-DOE 20CRv3.

The relative importance of the trend component (r_{tr}) remains small for the 30-year windows ending prior to the 1980s, and has increased rapidly over the past few decades in CMIP6 models and observations (Figs. 7c-d). The result is consistent with both r and r_{tr} in large ensembles, highlighting the agreement between models and observations on the robust strengthening of forced and linear responses over the past few decades. Still, we note that the r_{tr} values are slightly weaker than r in NOAA-CIRES-DOE and ERSSTv5, and are even weaker in the CMIP6 models and HadISST, partly due to the lack of volcanic eruptions in the trend component (compare Table 2 with Table 1).

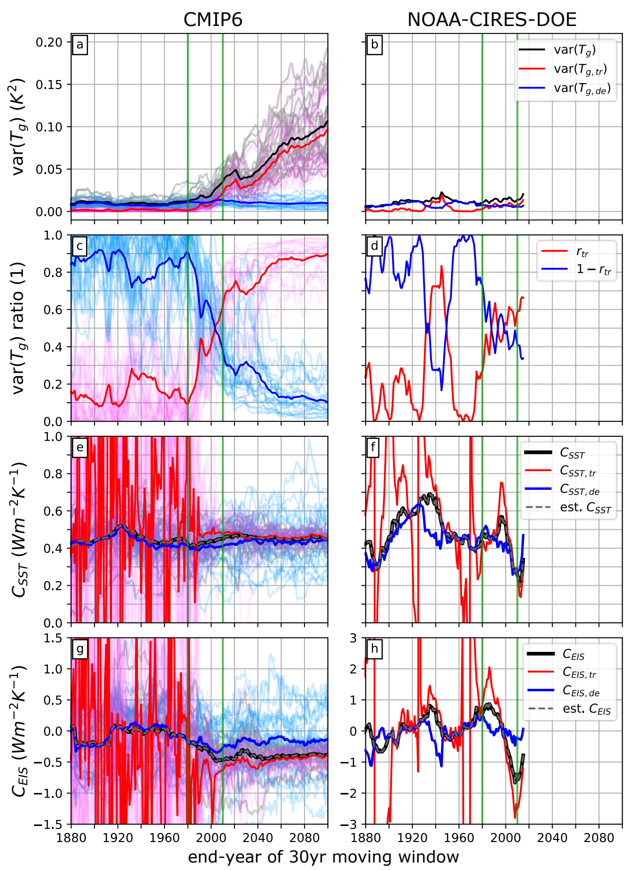


FIG. 7. Time variation of the variance of T_g (black), $T_{g,tr}$ (red), and $T_{g,de}$ (blue) in (a) CMIP6 models and (b) NOAA-CIRES-DOE 20CRv3. Variance is calculated on a sliding 30-year window. (c-d) Similar to (a-b), but for the ratio of the T_g variance (defined in Eq. 5.3). (e-f) Time variation of low-cloud feedback due to changes in the overall SST pattern (C_{SST}) along with its trend (red) and detrend (blue) components in (e) CMIP6 models and (f) NOAA-CIRES-DOE 20CRv3. Est. C_{SST} (dashed gray) is calculated as $r_{tr}C_{SST,tr} + (1-r_{tr})C_{SST,de}$, meaning that the difference between black and dashed gray line is the residual term. (g-h) Same as (e-f), but for the low-cloud feedback due to changes in EIS patterns (C_{EIS}). Dark-colored lines indicate the ensemble mean and light-colored lines show each ensemble member.

	CMIP6	NOAA-CIRES-DOE	ERSSTv5	HadISST
$r_{tr}\|_{1951-1980}$ (1)	0.10 ± 0.12	0.27	0.22	0.03
$r_{tr}\|_{1981-2010}$ (1)	0.62 ± 0.15	0.59	0.61	0.56

TABLE 2. The relative importance of the trend component (r_{tr}; Eq. 5.3) in CMIP6 models and three different observational SST datasets. The pipe symbol (|) is followed by the 30-year window that is used to calculate r_{tr}. For CMIP6 models, the multimodel-mean values and 1 STD across models are shown.

c. Time-evolving low-cloud feedback in historical period

Most of the CMIP6 models and the observation suggest a more-stabilizing C_{EIS} in the recent 30-50 years while the change in C_{SST} stays weaker (Figs. 7e-h; Fig. 8), consistent with the results obtained from large ensemble simulations (CESM2-LE and MPI-GE, less consistent with GISS-LE). Regarding the cause of the C_{EIS} change between 1951-1980 and 1981-2010, most CMIP6 models agree that the increasing importance of trend component is the main contributor (strong negative Δ_3 in Fig. 8c), with the other two terms also slightly contributing to the negative shift of C_{EIS}. The result is aligned with the three large ensembles (Fig. 2). However, the NOAA-CIRES-DOE reanalysis (NOAA-CIRES-DOE 20CRv3) reveals that the change in trend component is the dominant contributor to the overall negative shift of C_{EIS} (strong negative Δ_1 in Fig. 8d), inconsistent with the ΔC_{EIS} decomposition from the CMIP6 models and the large ensembles.

Similar LEs-observations differences can be found in the corresponding SST patterns. In 1951-1980, all three observations analyzed here show an enhanced warming in the Southern Hemisphere (SH) and a reduced warming in the Northern Hemisphere (NH) for the trend component (Fig. 9, second column), while the large ensembles show a more uniform warming with slight enhancement in the NH extratropics and the WP region for the forced component (Fig. 4, second column). Since both r and r_{tr} are small during this time, the overall SST/EIS patterns are less affected by the forced or trend component. In 1981-2010 when r and r_{tr} are large, the difference between the models (forced component) and the observations (trend component) is essential in interpreting the models-observations differences. The major difference between the two is the large-scale, triangle-shaped Eastern Pacific cooling in the observation (trend component; Fig. 9) that is not reproduced in the forced responses of AOGCM's large ensemble simulations (Fig. 4). The ultimate reason for the contrasting SST patterns between the models and observations is not well understood at this point, but below we provide two plausible explanations.

First, the difference between models and observations might come from the fact that the trend/detrend components do not fully represent the forced/unforced responses. Despite the modification based on major eruptions that has been made, other issues such as nonlinear forced responses (included in detrend component) or multi-decadal natural variability (included in trend component) would also affect the partition between the trend and detrend components. Pacific Decadal Oscillation (PDO), for example, is expected to affect the global SST pattern. During the 1980s to the 2010s, the PDO index shows a negative trend, which is aligned with the large-scale Eastern Pacific cooling in the trend component of the three observations.

Second, it is possible that the models fail to project the correct forced responses as in the observation, leading to biased SST patterns and weak negative to near-zero trends of low-cloud feedback over the past few decades, consistent with previous research pointing out the systematic model biases on recent surface warming patterns (Dong et al. 2021; Wills et al. 2022).

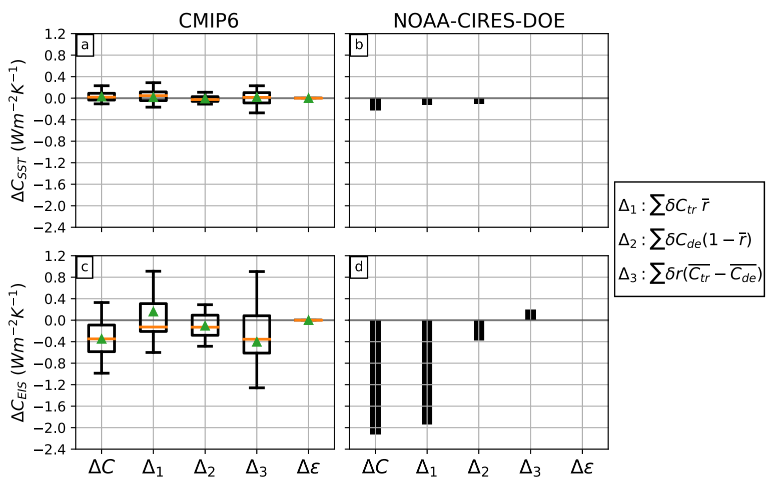


FIG. 8. Attribution of the change in C_{SST} between the 30-year window 1951-1980 and 1981-2010 in (a) CMIP6 models and (b) NOAA-CIRES-DOE 20CRv3. (c-d) Same as (a-c), but for the change in C_{EIS}. In CMIP6, the orange line indicates the median and the green triangle indicates the multimodel mean.

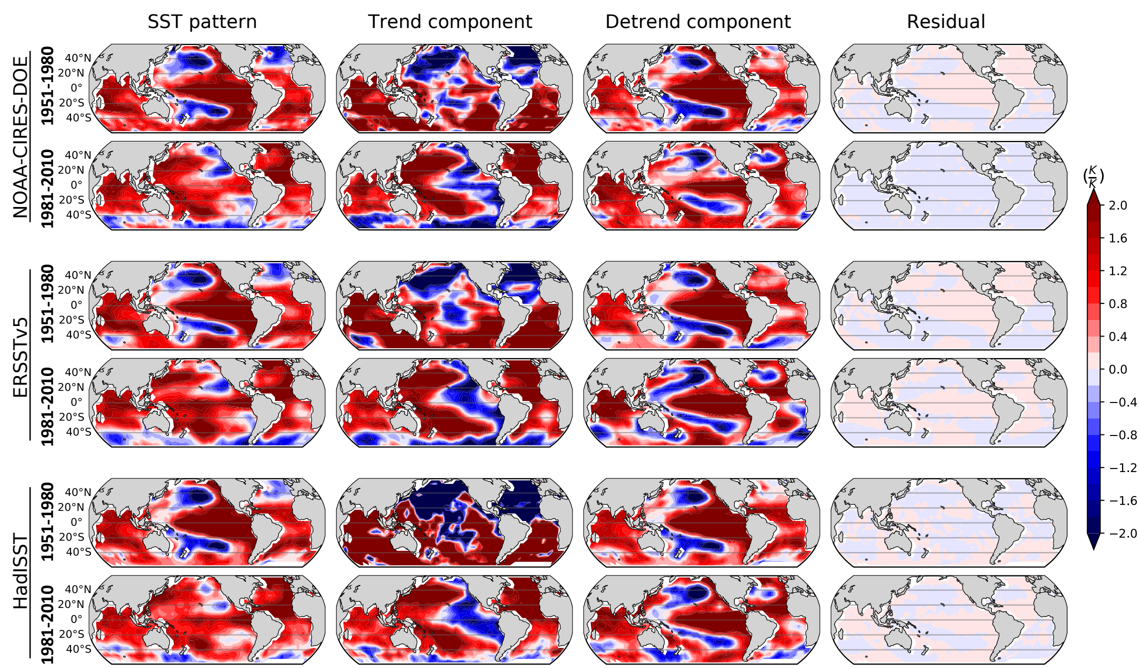


FIG. 9. (from left to right) Overall SST pattern ($\frac{dT}{dT_g}$), the trend component ($\frac{dT_{tr}}{dT_{g,tr}}$), the detrend component ($\frac{dT_{de}}{dT_{g,de}}$), and the residual term (σ) calculated from different 30-year windows in the three observations. For each panel, the 30-year window and the observation used are labeled on the left.

6. Summary and discussion

This research examines the role of the SST pattern effect in driving the time-varying low-cloud feedback (C), with a particular focus on the relative importance between forced responses and unforced variability. We provide evidence that the time variation of C estimated via OLS regressions can be attributed to three main contributors: changes in its forced component, changes in its unforced component, and changes in the relative importance between the forced and unforced components (see Sections 2c and 3c for more details).

Using initial-condition large ensembles, we find that the unforced signals outweigh the forced signals for 30-year windows ending prior to the 1980s (Figs. 1d-f), thus the overall SST and EIS patterns are strongly influenced by the unforced components, characterizing ENSO-like surface conditions (Figs. 4-5). For 30-year windows ending after the 1980s, the forced signals have strengthened, surpassing unforced signals around the 2010s (Figs. 1d-f), in parallel with the rapid increase of external forcings. Since the forced SST patterns are relatively uniform (the second column of Fig. 4), the overall SST patterns after the 1980s have become less heterogeneous (the first column of Fig. 4). The time-evolving SST pattern gives rise to changes in low-cloud feedback directly (C_{SST}) and through modifying low-level stability (C_{EIS}). C_{SST} and C_{EIS} are the low-cloud radiative feedbacks due to changes in SST and EIS, respectively. Most of the ensemble members in CESM2-LE and MPI-GE agree on the stabilizing shift of C_{EIS} in the recent 30-50 years, with a magnitude larger than the change in C_{SST}. More importantly, we find that the increasing importance of forced signals is the dominant contributor to the negative shift of C_{EIS} over the past few decades (Fig. 2).

These results highlight the crucial role of strengthening forced responses relative to unforced variations in modifying C, especially within recent decades when the overall radiative feedback shifts from being dominated by unforced signals to being dominated by forced signals. This shift can lead to large apparent time variations in feedbacks that are distinct from the type of pattern-effect mechanisms related to ocean heat uptake that are invoked to explain time-varying feedbacks in CO_2 doubling or quadrupling simulations. Rather, the time-evolving pattern described here arises from the fact that OLS estimates have a time-varying mix of forced and unforced SST patterns and feedbacks. Thus, a “pattern-effect” may arise even if both forced and unforced patterns are themselves time-invariant. This influence has not been clearly quantified or demonstrated to our knowledge. Therefore, we suggest the forced and unforced component of radiative feedbacks should be evaluated separately. When available, large ensembels should be used. When large ensembles are not available, computing the ratio of trends in radiation and temperature rather than regressing radiation against temperature can help filter out unforced high-frequency variability, as well as volcanic events.

To conduct a coherent analysis on multiple CMIP6 models and observations for which large ensembles are not available, we develop a method based on linear trends, aiming to approximately isolate forced responses from unforced variations (which roughly correspond to trend and detrend components). Consistent with the large ensembles, the detrend component dominates the trend component for 30-year windows ending prior to the 1980s. The trend component has strengthened since then, taking over the detrend component recently in most CMIP6 models and the observations (Figs. 7c-d). Despite the similar time evolution of the relative importance of forced/unforced and trend/detrend components, we find a much stronger stabilizing shift of C_{EIS} in the observation that lies outside the model ensembles over the past few decades (Fig. 8). The change in the trend component is the main factor causing a strong and negative shift of C_{EIS} in the observation, which is inconsistent with the large ensembles that highlight the increasing importance of the forced component in driving the negative change of C_{EIS} (compare Fig. 8 with Fig. 2). The discrepancy between the observation and large ensembles arises from the observed SST trends in the recent decades that are not included in the model ensembles (compare Fig. 9 with Fig. 4).

Acknowledgments. The work of YJL, GVC, CP, and KCA was supported by U.S. Department of Energy (DOE) Regional and Global Model Analysis program grant no. DE-SC0022110. The work of MDZ was supported by the U.S. Department of Energy (DOE) Regional and Global Model Analysis program area and was performed under the auspices of the U.S. DOE by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. KCA was supported by National Science Foundation (NSF) Award AGS-1752796, National Oceanic and Atmospheric Administration (NOAA) MAPP Program Award NA20OAR4310391, and a Calvin Professorship in Oceanography.

Data availability statement. The CESM2-LE dataset was made available by the CESM2 Large Ensemble Community Project and supercomputing resources from IBS Center for Climate Physics, which can be downloaded from `https://www.cesm.ucar.edu/community-projects/lens2/data-sets`. The MPI-GE data can be downloaded from `https://www.mpimet.mpg.de/en/grand-ensemble/`. The GISS-LE data and the CMIP6 AOGCM outputs can be downloaded from the ESGF portal (`https://esgf-node.llnl.gov/search/cmip6/`). HadISST1.1 data can be downloaded from `https://www.metoffice.gov.uk/hadobs/hadisst/`. ERSSTv5 data can be downloaded from the NOAA National Centers for Environmental Information website (`https://www1.ncdc.noaa.gov/pub/data/cmb/ersst/v5/netcdf/`). NOAA/CIRES/DOE 20th Century Reanalysis V3 can be downloaded from the NOAA ESRL Physical Sciences Division website (`http://www.esrl.noaa.gov/psd/data/`).

References

Andrews, T., 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. *Journal of Climate*, **28 (4)**, 1630–1648, https://doi.org/10.1175/Jcli-D-14-00545.1.

Andrews, T., and Coauthors, 2018: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity. *Geophysical Research Letters*, **45 (16)**, 8490–8499, https://doi.org/10.1029/2018gl078887.

Andrews, T., and Coauthors, 2022: On the Effect of Historical SST Patterns on Radiative Feedback. *Journal of Geophysical Research: Atmospheres*, **127 (18)**, https://doi.org/

10.1029/2022jd036675.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2013: Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, **26 (13)**, 4518–4534.

Bauer, S. E., and Coauthors, 2020: Historical (1850–2014) Aerosol Evolution and Role on Climate Forcing Using the GISS ModelE2.1 Contribution to CMIP6. *Journal of Advances in Modeling Earth Systems*, **12 (8)**, https://doi.org/10.1029/2019ms001978.

Bony, S., and J. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, **32 (20)**.

Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proceedings of the National Academy of Sciences*, https://doi.org/10.1073/pnas.1714308114.

Ceppi, P., and J. M. Gregory, 2019: A refined model for the Earth's global energy balance. *Climate Dynamics*, https://doi.org/10.1007/s00382-019-04825-x.

Dessler, A. E., T. Mauritsen, and B. Stevens, 2018: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmospheric Chemistry and Physics*, **18 (7)**, 5147–5155, https://doi.org/10.5194/acp-18-5147-2018.

Dong, Y., K. C. Armour, M. D. Zelinka, C. Proistosescu, D. S. Battisti, C. Zhou, and T. Andrews, 2020: Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models. *Journal of Climate*, **33 (18)**, 7755–7775, https://doi.org/10.1175/jcli-d-19-1011.1.

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. *Journal of Climate*, **32 (17)**, 5471–5491.

Dong, Y., and Coauthors, 2021: Biased Estimates of Equilibrium Climate Sensitivity and Transient Climate Response Derived From Historical CMIP6 Simulations. *Geophysical Research Letters*, **48 (24)**, https://doi.org/10.1029/2021gl095778.

Donohoe, A., K. C. Armour, A. G. Pendergrass, and D. S. Battisti, 2014: Shortwave and longwave radiative contributions to global warming under increasing CO 2. *Proceedings of the National Academy of Sciences*, **111 (47)**, 16 700–16 705, https://doi.org/10.1073/pnas.1412190111.

Eiselt, K.-U., and R. G. Graversen, 2023: On the Control of North-Hemispheric Feedbacks by AMOC, Evidence from CMIP and Slab-Ocean Modeling. *Journal of Climate*, 1–43.

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

Frey, W. R., E. A. Maroon, A. G. Pendergrass, and J. E. Kay, 2017: Do Southern Ocean Cloud Feedbacks Matter for 21st Century Warming? *Geophysical Research Letters*, **44 (24)**, https://doi.org/10.1002/2017gl076339.

Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. Olivié, and S. Tytéca, 2013: Transient climate response in a two-layer energy-balance model. Part II: Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs. *Journal of Climate*, **26 (6)**, 1859–1876.

Gjermundsen, A., A. Nummelin, D. Olivié, M. Bentsen, O. Seland, and M. Schulz, 2021: Shutdown of Southern Ocean convection controls long-term greenhouse gas-induced warming. *Nature Geoscience*, **14 (10)**, 724–731, https://doi.org/10.1038/s41561-021-00825-x.

Gregory, J., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31 (3)**.

Gregory, J. M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, **43 (8)**, 3911–3920, https://doi.org/10.1002/2016gl068406.

Gregory, J. M., T. Andrews, P. Ceppi, T. Mauritsen, and M. Webb, 2020: How accurately can the climate sensitivity to $$\hbox {CO} _ {2} $$ be estimated from historical climate change? *Climate Dynamics*, **54 (1-2)**, 129–157.

Gregory, J. M., T. Andrews, P. Good, T. Mauritsen, and P. M. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dynamics*, **47 (12)**, 3979–3991, https://doi.org/10.1007/s00382-016-3055-1.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *Journal of Climate*, **15 (22)**, 3117–3121, https://doi.org/10.1175/1520-0442(2002)015⟨3117:aobeot⟩2.0.co;2.

Günther, M., H. Schmidt, C. Timmreck, and M. Toohey, 2022: Climate feedback to stratospheric aerosol forcing: the key role of the pattern effect. *Journal of Climate*, **35 (24)**, 7903–7917.

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *Journal of Climate*, **23 (9)**, 2418–2427.

Huang, B., and Coauthors, 2017: NOAA extended reconstructed sea surface temperature (ERSST), version 5. *NOAA National Centers for Environmental Information*, **30 (8179-8205)**, 25.

Huber, M., U. Beyerle, and R. Knutti, 2014: Estimating climate sensitivity and future temperature in the presence of natural climate variability. *Geophysical Research Letters*, **41 (6)**, 2086–2092, https://doi.org/10.1002/2013gl058532.

IPCC, 2023: The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity. *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 923–1054, https://doi.org/10.1017/9781009157896.009, URL https://www.cambridge.org/core/books/climate-change-2021-the-physical-science-basis/earths-energy-budget-climate-feedbacks-and-climate-sensitivity/AE57C97E588FF3060C7C7E47DD4F3C6E.

Kelley, M., and Coauthors, 2020: GISS-E2.1: Configurations and Climatology. *Journal of Advances in Modeling Earth Systems*, **12 (8)**, https://doi.org/10.1029/2019ms002025.

Lin, Y., Y. Hwang, P. Ceppi, and J. M. Gregory, 2019: Uncertainty in the evolution of climate feedback traced to the strength of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*, **46 (21)**, 12 331–12 339, https://doi.org/10.1029/2019GL083084.

Lin, Y., Y. Hwang, J. Lu, F. Liu, and B. E. J. Rose, 2021: The Dominant Contribution of Southern Ocean Heat Uptake to Time-Evolving Radiative Feedback in CESM. *Geophysical Research Letters*, **48 (9)**, https://doi.org/10.1029/2021GL093302.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability. *Journal of Advances in Modeling Earth Systems*, **11 (7)**, 2050–2069, https://doi.org/10.1029/2019ms001639.

Marvel, K., G. A. Schmidt, R. L. Miller, and L. S. Nazarenko, 2016: Implications for climate sensitivity from the response to individual forcings. *Nature Climate Change*, **6 (4)**, 386–389.

Miller, R. L., and Coauthors, 2021: CMIP6 Historical Simulations (1850–2014) With GISS-E2.1. *Journal of Advances in Modeling Earth Systems*, **13 (1)**, https://doi.org/10.1029/2019ms002034.

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. *Nature Climate Change*, **11 (6)**, 501–507, https://doi.org/10.1038/s41558-021-01039-0.

Pincus, R., P. M. Forster, and B. Stevens, 2016: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6. *Geosci. Model Dev.*, **9 (9)**, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016.

Proistosescu, C., A. Donohoe, K. C. Armour, G. H. Roe, M. F. Stuecker, and C. M. Bitz, 2018: Radiative Feedbacks From Stochastic Variability in Surface Temperature and Radiative Imbalance. *Geophysical Research Letters*, **45 (10)**, 5082–5094, https://doi.org/10.1029/2018gl077678.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, **108 (D14)**, https://doi.org/10.1029/2002jd002670.

Rodgers, K. B., and Coauthors, 2021: Ubiquity of human-induced changes in climate variability. *Earth System Dynamics*, **12 (4)**, 1393–1411, https://doi.org/10.5194/esd-12-1393-2021.

Rose, B. E. J., K. C. Armour, D. S. Battisti, N. Feldl, and D. D. B. Koll, 2014: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophysical Research Letters*, **41 (3)**, 1071–1078, https://doi.org/10.1002/2013gl058955.

Rose, B. E. J., and L. Rayborn, 2016: The Effects of Ocean Heat Uptake on Transient Climate Sensitivity. *Current Climate Change Reports*, **2 (4)**, 190–201, https://doi.org/10.1007/s40641-016-0048-4.

Rugenstein, M., and Coauthors, 2020: Equilibrium Climate Sensitivity Estimated by Equilibrating Climate Models. *Geophysical Research Letters*, **47 (4)**, https://doi.org/10.1029/2019gl083898.

Rugenstein, M. A. A., and K. C. Armour, 2021: Three Flavors of Radiative Feedbacks and Their Implications for Estimating Equilibrium Climate Sensitivity. *Geophysical Research Letters*, **48 (15)**, https://doi.org/10.1029/2021gl092983.

Rugenstein, M. A. A., K. Caldeira, and R. Knutti, 2016: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophysical Research Letters*, **43 (18)**, 9877–9885, https://doi.org/10.1002/2016gl070907.

Salvi, P., J. M. Gregory, and P. Ceppi, 2023: Time-evolving radiative feedbacks in the historical period. *Journal of Geophysical Research: Atmospheres*, https://doi.org/10.1029/2023jd038984.

Scott, R. C., T. A. Myers, J. R. Norris, M. D. Zelinka, S. A. Klein, M. Sun, and D. R. Doelling, 2020: Observed Sensitivity of Low-Cloud Radiative Effects to Meteorological Perturbations over the Global Oceans. *Journal of Climate*, **33 (18)**, 7717–7734, https://doi.org/10.1175/jcli-d-19-1028.1.

Sherwood, S. C., S. Bony, and J.-L. Dufresne, 2014: Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, **505 (7481)**, 37–42.

Sherwood, S. C., and Coauthors, 2020: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. *Reviews of Geophysics*, **58 (4)**, https://doi.org/10.1029/2019rg000678.

Shindell, D. T., 2014: Inhomogeneous forcing and transient climate sensitivity. *Nature Climate Change*, **4 (4)**, 274–277, https://doi.org/10.1038/nclimate2136.

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 (11)**, 512–522, https://doi.org/10.1002/2016ef000376.

Uribe, A., F. A. Bender, and T. Mauritsen, 2022: Observed and CMIP6 Modeled Internal Variability Feedbacks and Their Relation to Forced Climate Feedbacks. *Geophysical Research Letters*, **49 (24)**, https://doi.org/10.1029/2022gl100075.

Watanabe, M., J.-L. Dufresne, Y. Kosaka, T. Mauritsen, and H. Tatebe, 2021: Enhanced warming constrained by past trends in equatorial Pacific sea surface temperature gradient. *Nature Climate Change*, **11 (1)**, 33–37, https://doi.org/10.1038/s41558-020-00933-3.

Wills, R. C. J., K. C. Armour, D. S. Battisti, C. Proistosescu, and L. A. Parsons, 2021: Slow Modes of Global Temperature Variability and Their Impact on Climate Sensitivity Estimates. *Journal of Climate*, **34 (21)**, 8717–8738, https://doi.org/10.1175/jcli-d-20-1013.1.

Wills, R. C. J., Y. Dong, C. Proistosecu, K. C. Armour, and D. S. Battisti, 2022: Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change. *Geophysical Research Letters*, **49 (17)**, https://doi.org/10.1029/2022gl100011.

Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, **23 (9)**, 2333–2344, https://doi.org/10.1175/2009jcli3139.1.

Wood, R., and C. S. Bretherton, 2006: On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of climate*, **19 (24)**, 6425–6432.

Zhou, C., M. Wang, M. D. Zelinka, Y. Liu, Y. Dong, and K. C. Armour, 2023: Explaining Forcing Efficacy With Pattern Effect and State Dependence. *Geophysical Research Letters*, **50 (3)**, https://doi.org/10.1029/2022gl101700.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, **9 (12)**, 871.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2017: Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *Journal of Advances in Modeling Earth Systems*, **9 (5)**, 2174–2189.

Supplemental Material of

"The relative importance of forced and unforced temperature patterns in driving the time variation of low-cloud feedback"

Yuan-Jen Lin[a,b] , Grégory V. Cesana[a,b] , Cristian Proistosescu[c] , Mark D. Zelinka[d] , and Kyle C. Armour[e]

[a] *Center for Climate Systems Research, Columbia University, New York, NY, USA*

[b] *NASA Goddard Institute for Space Studies, New York, NY, USA*

[c] *Department of Atmospheric Sciences and Department of Geology, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA*

[d] *Lawrence Livermore National Laboratory, Livermore, CA, USA*

[e] *School of Oceanography and Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA*

Corresponding author: Yuan-Jen Lin, yl5278@columbia.edu

S1. Modification based on volcanic eruptions

Large volcanic eruptions can influence global climate and can be considered as external forcings. However, since their impacts are strong but relatively short-term (compared with other forcing agents), they are barely captured by 30-year linear trends. Here, by estimating the impacts of major volcanic eruptions and redistributing them into the trend components, we are able to improve consistency between r and r_{tr}.

The procedures to modify the trend/detrend components are described as follows. First, we identify six major volcanic eruptions and assume their impacts on global temperature will last for around five years (*Impact Period* in Table S1). For the 30-year windows that include these impact periods, we compute the linear-trend part of the target field X without these impact periods ($X_{tr,noVE}(t)$), thus those linear components outside the impact periods are less affected by major volcanic eruptions.

$$\text{Modified } X_{tr}(t) = \begin{cases} X_{tr,noVE}(t) & \text{if } t \notin \text{impact period,} \\ X_{tr,noVE}(t) + \kappa[X(t) - X_{tr,noVE}(t)] & \text{if } t \in \text{impact period.} \end{cases} \quad \text{(S1)}$$

Second, within the impact periods, we calculate the differences between full responses and linear responses of X (i.e., $X(t) - X_{tr,noVE}(t)$) and suggest that the differences, to a varying extent, are related to major volcanic eruptions and thus should be included in our modified trend components to better represent the forced responses. Third, we multiplied these nonlinear anomalies with an eruption-dependent, time-invariant κ that indicates the relative importance of volcanic eruptions among all other nonlinear factors, such as internal variability, responses to aerosol forcing, etc. Here we design the values of κ based on the strength of ENSO as it is the dominant interannual climate variability and has global climate impacts. κ is weaker when the major volcanic eruption coincides with strong ENSO events. Fourth and the last step, we compute modified $X_{de}(t)$ by subtracting modified $X_{tr}(t)$ from $X(t)$. The modified $X_{tr}(t)$ and $X_{de}(t)$ can then be used to calculate modified r_{tr}, SST/EIS patterns, and low-cloud feedback in sliding windows as usual.

Table S2 shows the skill improvement of the modified trend/detrend components as proxies of forced/unforced components. For example, in CESM2-LE, modified $r_{tr} = 0.25 \pm 0.14$ in 1951-1980 ($r = 0.24 \pm 0.07$) and $r_{tr} = 0.67 \pm 0.11$ in 1981-2010 ($r = 0.72 \pm 0.06$). The improvement of r_{tr} also helps to reduce the number of extreme feedback values (compare Fig. 6 with Fig. S1), thus the trend/detrend components of low-cloud feedback become more comparable to the forced/unforced components.

Despite the improvement of using trend/detrend components as proxies of forced/unforced components, the outcome remains consistent whether major volcanic eruptions are considered or not. To maintain simplicity, we present the results excluding major volcanic eruptions in the main paper and reserve the corresponding discussion here.

Major Eruptions	Impact Period	κ	Strong/super ENSO
Krakatau	1883-1887	0.7	NA
Santa Maria	1902-1906	0.5	1902-03 (MEI> 2)
Novarupta/Katmai	1912-1916	0.5	1916-17 (MEI< −2)
Agung	1963-1967	0.7	NA
El Chichón	1982-1986	0.3	1982-83 (MEI> 3)
Pinatubo	1991-1995	0.5	1991-92 (MEI> 2)

TABLE S1. Six historical major volcanic eruptions considered in the modification of trend/detrend components. Impact period is approximated to five years (starting at the eruption year). For each impact period, the relative importance of the major volcanic eruption to other nonlinear responses (e.g., natural variability) is parameterized to κ. Here κ is set based on the strength of ENSO. We use Multivariate ENSO Index Version 2 (MEI.v2) and define *strong* ENSO as MEI > 2 (< −2) and *super* ENSO as MEI > 3 (< −3). When there is no strong ENSO (−2 <MEI< 2), κ is set to 0.7. When there is one or more strong (super) ENSO, κ is set to 0.5 (0.3).

	CESM2-LE	MPI-GE	GISS-LE
$r_{tr}\|_{1951-1980}$ (1)	0.08 ± 0.11	0.14 ± 0.12	0.03 ± 0.04
$r_{tr}\|_{1981-2010}$ (1)	0.67 ± 0.11	0.61 ± 0.13	0.49 ± 0.13
mod $r_{tr}\|_{1951-1980}$ (1)	0.25 ± 0.14	0.32 ± 0.13	0.20 ± 0.08
mod $r_{tr}\|_{1981-2010}$ (1)	0.73 ± 0.09	0.67 ± 0.13	0.54 ± 0.14

TABLE S2. The relative importance of the trend component before modification (r_{tr}) and after modification (mod r_{tr}). The modification is made by accounting for the influences of the historical major volcanic eruptions. Both are calculated in the windows 1951-1980 and 1981-2010 in the three large ensembles, with the ensemble-mean value and 1 standard deviation across ensembles are shown.

	CMIP6	NOAA-CIRES-DOE	ERSSTv5	HadISST
mod $r_{tr}\|_{1951-1980}$ (1)	0.28 ± 0.13	0.34	0.29	0.10
mod $r_{tr}\|_{1981-2010}$ (1)	0.70 ± 0.15	0.66	0.60	0.51

TABLE S3. The relative importance of the modified trend component (mod r_{tr}) in CMIP6 models and three different observational SST datasets. The pipe symbol (|) is followed by the 30-year window that is used to calculate modified r_{tr}. For CMIP6 models, the multimodel-mean value and 1 standard deviation across models are shown.

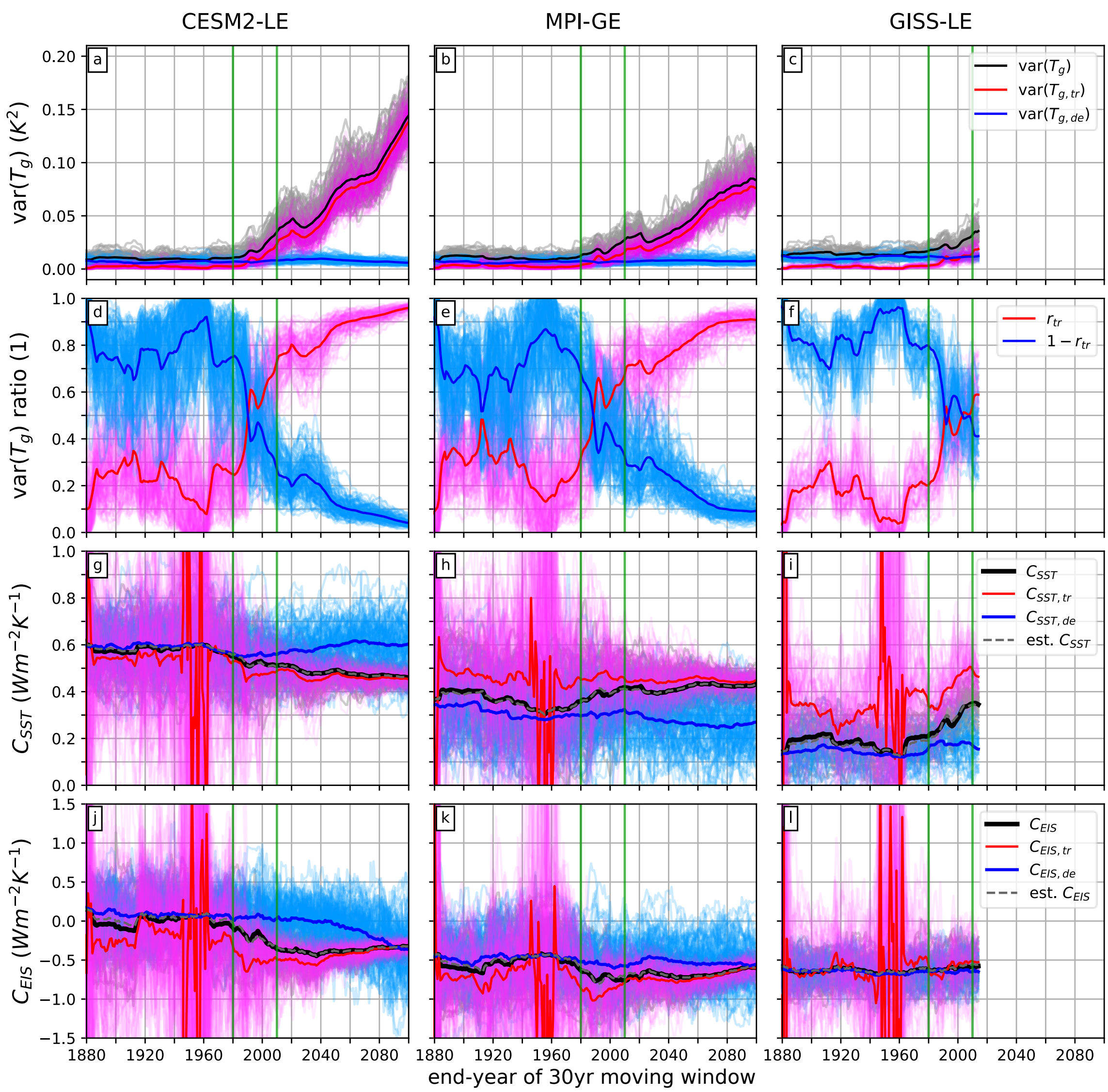
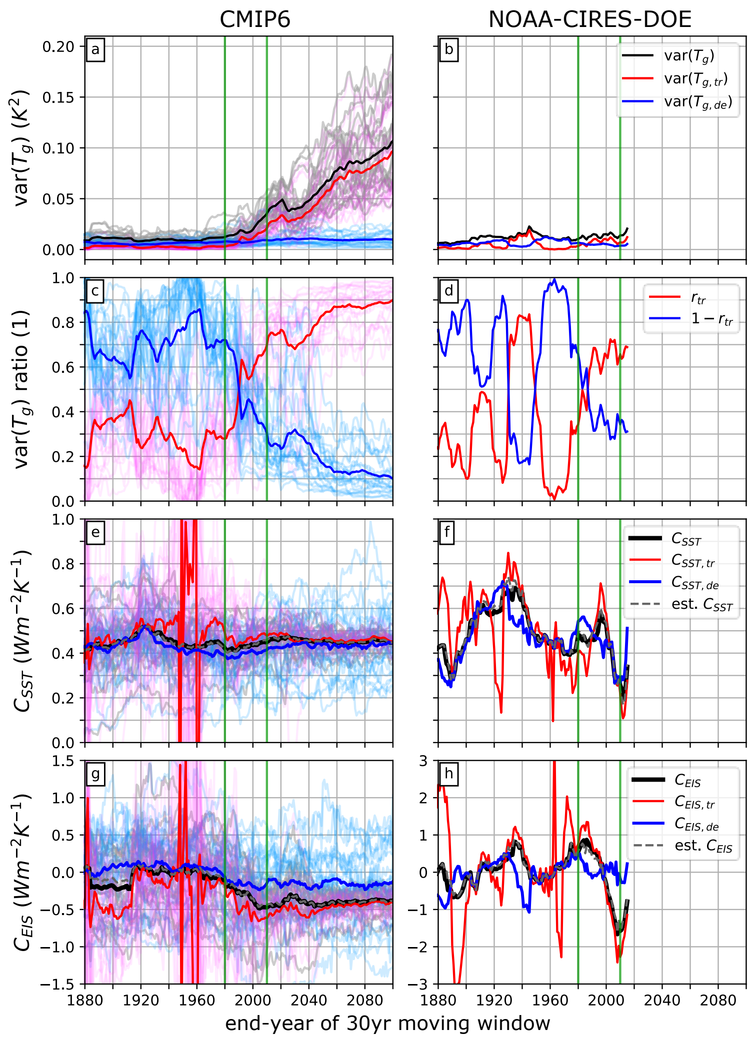


FIG. S1. Same as Figure 6, but with modification made by accounting for the influences of the historical major volcanic eruptions.



63 FIG. S2. Same as Figure 7, but with modification made by accounting for the influences of the historical major
64 volcanic eruptions.

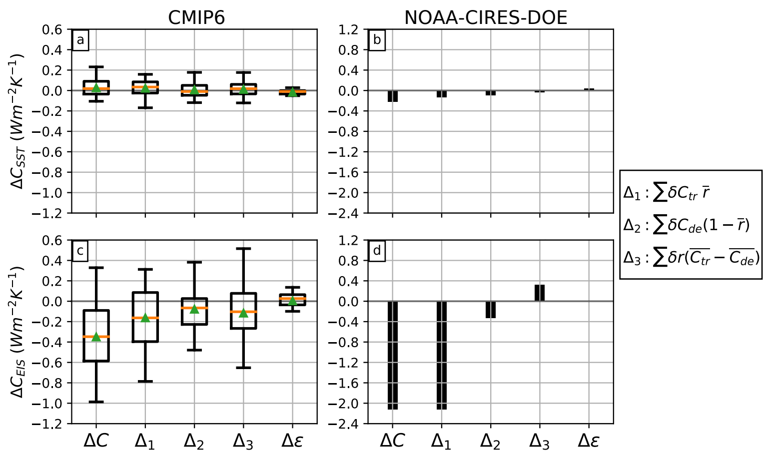


FIG. S3. Same as Figure 8, but with modification made by accounting for the influences of the historical major volcanic eruptions.

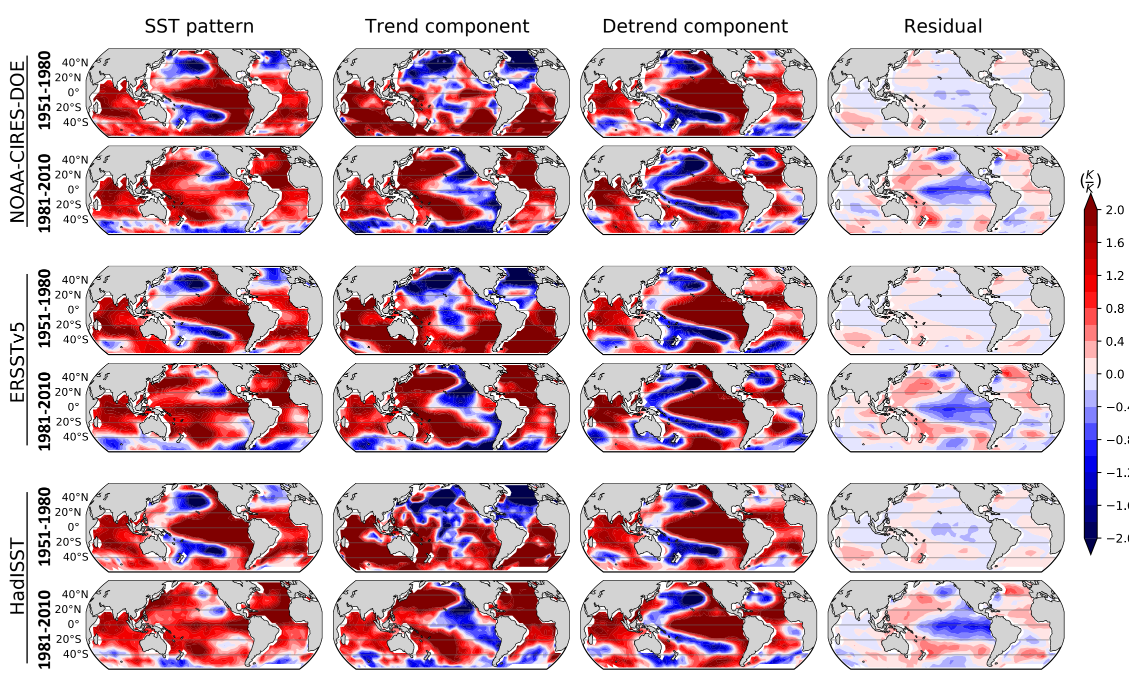


FIG. S4. Same as Figure 9, but with modification made by accounting for the influences of the historical major volcanic eruptions.