

Reconciling Conflicting Accounts of Local Radiative Feedbacks in Climate Models

CHRISTOPHER HEDEMANN,^a THORSTEN MAURITSEN,^a JOHANN JUNGCLAUS,^a AND JOCHEM MAROTZKE^a

^a Max Planck Institute for Meteorology, Hamburg, Germany

(Manuscript received 12 July 2021, in final form 24 January 2022)

ABSTRACT: The literature offers conflicting findings about which regions contribute most to increases in the global radiative feedback after a forcing increase. This paper explains the disagreement by discriminating between two common definitions of the local feedback, which use either local temperature or global temperature as their basis. Although the two definitions of feedback have been previously compared in aquaplanet models with slab oceans, here the definitions are compared for the first time in an atmosphere–ocean general circulation model (MPI-ESM1.2) integrated over four doublings of atmospheric CO₂ concentrations. Large differences between the definitions can be seen in all feedbacks, but especially in the temperature and water vapor feedbacks. Differences of up to 10 W m⁻² K⁻¹ over the Southern Ocean can be explained by the pattern of surface warming, which weights the local feedbacks and reduces their contribution to the global mean. This finding is, however, dependent on the resolution of analysis, because the local-temperature definition is mathematically inconsistent across spatial scales. Furthermore, attempts to estimate the effect of “pattern weighting” by separating local feedbacks and warming patterns at the gridcell level fail, because the radiative change in key tropical regions is also determined by tropospheric stability via the global circulation. These findings indicate that studies of regional feedback change are more sensitive to methodological choices than previously thought, and that the tropics most likely dominate regional contributions to global radiative feedback change on decadal to centennial time scales.

SIGNIFICANCE STATEMENT: Radiative feedbacks are processes that either intensify or damp global surface warming. We compare two ways to define local radiative feedbacks in a climate model and find that the choice of definition drastically impacts the results. Differences in feedback between the definitions are up to 10 W m⁻² K⁻¹ over the Southern Ocean; by comparison, the estimate of the true global feedback is around -1 W m⁻² K⁻¹. Also, one of the definitions is mathematically inconsistent across different scales of spatial aggregation. Our findings matter because they help to reconcile disagreement in previous studies about which regions dominate global radiative feedback change in model simulations of global warming.

KEYWORDS: Feedback; Radiative fluxes; Radiative forcing; Climate models; General circulation models; Model comparison

1. Introduction

Future global surface warming is determined in part by Earth’s radiative feedbacks, which prescribe how much global surface warming must ensue to restore equilibrium after a radiative forcing is applied. The global radiative feedback parameter, which quantifies these feedbacks, is still uncertain (Sherwood et al. 2020), not least because this parameter is known to change with time and the climate state, and to differ between climate models. This paper investigates the definition

dependence of the feedback parameter locally, showing how different methodologies can lead to qualitatively different conclusions about local climate feedbacks.

Mounting evidence—from both climate models (Senior and Mitchell 2000; Gregory et al. 2004; Winton et al. 2010; Jonko et al. 2012; Li et al. 2013; Armour et al. 2013; Block and Mauritsen 2013; Geoffroy et al. 2013; Meraner et al. 2013; Rose et al. 2014; Andrews et al. 2015; Bloch-Johnson et al. 2015; Rugenstein et al. 2016; Armour 2017; Ceppi and Gregory 2017; Rugenstein et al. 2019, 2020; Dong et al. 2020) and observations or proxies (Huber and Caballero 2011; Hargreaves and Annan 2016; Royer 2016; Shaffer et al. 2016; Armour 2017; Dessler et al. 2018)—suggests that the global feedback parameter can change. Change in the parameter is thought to occur with time after a forcing is applied (e.g., Senior and Mitchell 2000; Winton et al. 2010; Armour et al. 2013; Gregory and Andrews 2016; Rugenstein et al. 2016; Haugstad et al. 2017; Paynter et al. 2018), and such time-dependent changes are sometimes referred to as “pattern effects” (Stevens et al. 2016) because of their connection to the changing

Mauritsen’s current affiliation: Stockholm University, Stockholm, Sweden.

Marotzke’s additional affiliation: Center for Earth System Research and Sustainability, Universität Hamburg, Hamburg, Germany.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Christopher Hedemann, christopher.hedemann@expertenrat-klima.de



This article is licensed under a Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

pattern of surface warming as the system equilibrates. The parameter is also thought to change with the equilibrium state of the climate, as represented by global mean surface temperature (e.g., Jonko et al. 2013; Caballero and Huber 2013; Meraner et al. 2013; Bloch-Johnson et al. 2015), or to change with both time and climate state (Rohrschneider et al. 2019).

Understanding the physical processes that drive feedback changes may help reduce uncertainty in future warming, which is why existing studies have attempted to locate the changes spatially. Some such studies have found that the change in the global feedback in the first 100 years or so after a forcing increase is driven primarily by the low to midlatitudes (Block and Mauritsen 2013; Andrews et al. 2015; Rugenstein et al. 2020; Dong et al. 2020). Particularly strong destabilizing increases in feedback are located over the eastern tropical Pacific Ocean (Andrews et al. 2015; Ceppi and Gregory 2017; Andrews and Webb 2018; Dong et al. 2020; Rugenstein et al. 2020). However, other studies have found instead that either the high-latitude regions drive the increase in global feedback (Armour et al. 2013), or specifically that the delayed warming over the Southern Ocean is the main driver (Senior and Mitchell 2000).

One of the many differences between these studies is the definition of the local feedback that they used. The studies that highlighted tropical regions used a definition of local feedbacks based on the global mean surface temperature. The two studies with alternate findings instead used a definition of feedback based on the local surface temperature (Armour et al. 2013) or the mean hemispheric temperature (Senior and Mitchell 2000).

What then are the consequences of using global versus local temperature in the definition of local feedbacks? Feldl and Roe (2013) compared these definitions in an aquaplanet setup at equilibrium, using an atmospheric model coupled to a slab mixed layer ocean component. No corresponding comparison has been undertaken for comprehensive coupled atmosphere–ocean general circulation models (AOGCMs) with a realistic continental setup, nor for shorter time scales such as in the decades to centuries after a forcing increase. However, realistic ocean heat uptake patterns and multidecadal time scales are arguably essential for understanding global feedback change (Armour et al. 2016; Rose and Rayborn 2016; Rugenstein et al. 2016). A range of other related questions posed in climate-modeling studies may be affected by the definition choice, including whether weighting of local feedbacks via the warming pattern drives changes in the global feedback (Armour et al. 2013; Ceppi and Gregory 2017; Colman and Hanson 2017), and which regions contribute most to intermodel differences in feedback (Crook et al. 2011; Zelinka and Hartmann 2012; Vial et al. 2013; Webb et al. 2013)?

In this study we compare for the first time the two feedback definitions in existing studies of AOGCMs (section 2) and complement this analysis by directly applying the definitions to output from a single AOGCM (section 4), showing how the choice of definition can lead to qualitatively different results in the calculation of feedbacks and feedback changes over time. We perform the analysis over four doublings of CO₂, to test the effect of forcing and signal-to-noise ratios on

our conclusions. We then discuss two potential scale-related problems that can arise when using the local-temperature definition of feedback. First, the local-temperature definition is mathematically inconsistent across spatial scales (section 5). Second, calculating the local-temperature feedbacks using gridcell data can lead to a statistically insignificant relationship between local surface warming and the change in top-of-atmosphere radiation budget, making the practice of calculating local-temperature feedbacks at these spatial scales problematic (section 6 and section 7).

2. Feedback definitions in existing studies

The global feedback parameter Λ can be estimated by linear regression of global change in top-of-atmosphere radiation R against global surface temperature change T after an abrupt forcing increase F is applied (Gregory et al. 2004):

$$R = F + \Lambda T, \quad (1)$$

which implies that $\Lambda = dR/dT$.

Alternatively, Eq. (1) can be rearranged to express the feedback as function of R normalized with respect to T by simple division (Murphy 1995). The choice of the regression method over the division method does not qualitatively affect our results, as discussed in section 4.

Although the regression can be applied to all available data points, it is also common practice to perform linear regression for different time periods, in order to investigate the change in feedbacks on different time scales (e.g., Gregory et al. 2004; Andrews et al. 2015; Ceppi and Gregory 2017; Rugenstein et al. 2020). In Fig. 1, these linear regressions are shown for the decadal, centennial, and millennial time periods in the MPI-ESM1.2 simulations. There are clear changes in slope between time periods, indicating that Λ is not constant but state dependent or time dependent, or both.

When moving from the global to the local scale, we can decompose the global feedback parameter into the spatial average of local contributions to the global feedback parameter, denoted for region i as λ_i : $\Lambda = \bar{\lambda}_i = \overline{dR_i}/dT$. Therefore, for a single region, the local feedback parameter can be defined with respect to the global surface temperature change:

$$\lambda_i = \frac{dR_i}{dT}. \quad (2)$$

Yet some studies have preferred to define this value with respect to the local surface temperature change:

$$\lambda_i^L = \frac{dR_i}{dT_i}, \quad (3)$$

where the superscript L denotes the local-temperature definition of the local feedback. Note that whereas λ_i can be spatially averaged to find the global feedback parameter Λ , this is not necessarily true for λ_i^L .

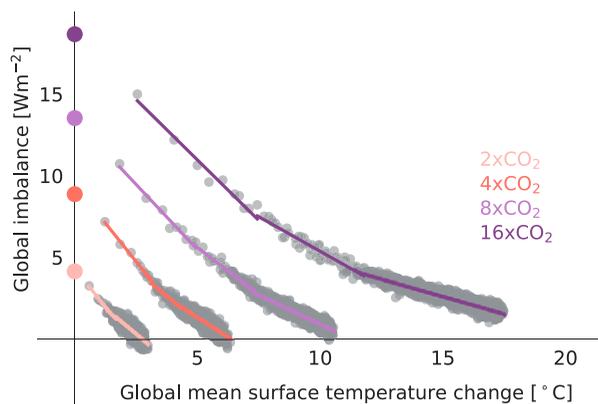


FIG. 1. Top-of-atmosphere radiation imbalance plotted against global mean surface temperature change, for abrupt forcing simulations of one or multiple CO_2 concentration doublings, integrated out to 1000 years with MPI-ESM1.2. Gray dots represent annual means of coupled model output, and colored lines represent least squares regression over the annual means for three periods: years 1–10, 11–100, and 101–1000. Large colored dots on the vertical axis represent effective radiative forcing estimated from fixed-SST experiments.

Although the global-temperature definition λ_i is more common, the local-temperature definition λ_i^L has been frequently used in the literature (e.g., Crook et al. 2011; Kay et al. 2012; Armour et al. 2013; Rose et al. 2014; Roe et al. 2015; Brown et al. 2016; Feldl and Bordoni 2016; Feldl et al. 2017; Frey et al. 2017; Colman and Hanson 2017; Bonan et al. 2018; Po-Chedley et al. 2018). Indeed, Feldl and Roe (2013) recommended λ_i^L as a more natural tool for explaining the spatial structure of feedback change, and for decomposing global feedback change into contributions from the warming pattern and local feedbacks. Colman and Hanson (2017) have also suggested that using λ_i^L would advance understanding of feedbacks derived from internal variability. More recently, however, Bloch-Johnson et al. (2020) suggested that using λ_i^L may overemphasize positive local cloud feedbacks derived from internal variability. What is the impact, if any, of using λ_i^L over λ_i in AOGCMs?

a. The local-temperature definition

The studies that have used the local-temperature definition λ_i^L have tended to find large feedback changes or intermodel variations in the Southern Ocean region or the high latitudes (see Table 1). The landmark study by Senior and Mitchell (2000) did not calculate λ_i^L directly, but instead used hemispheric warming to normalize cloud changes over the Southern Hemisphere. The authors found that these cloud changes over the Southern Hemisphere drove the change in the global feedback parameter. Armour et al. (2013) concluded that the high latitudes drive increases in the global feedback. Their results also show that, after a doubling of atmospheric CO_2 concentrations, the zonal-average feedbacks increase most strongly toward equilibrium in the subantarctic region (their Fig. 4c). The Southern Ocean is sometimes also found to have initially strong stabilizing

feedbacks in λ_i^L (Feldl and Bordoni 2016; Bonan et al. 2018), which is of interest because these feedbacks may weaken considerably as the Southern Ocean warms, leading to destabilizing increases in the global feedback parameter, as we show in section 4.

The Southern Ocean region is also thought to drive intermodel differences in global water vapor and lapse-rate feedbacks (Po-Chedley et al. 2018). And although Feldl and Bordoni (2016) did not explicitly examine intermodel differences, their Fig. 1 shows some of the largest differences between models over the Southern Ocean, especially in water vapor and lapse-rate feedbacks, which supports the results of Po-Chedley et al. (2018). Figure 1 in Bonan et al. (2018) also shows large intermodel variability in CMIP5 over the Southern Ocean. Furthermore, Frey et al. (2017) highlighted a sensitivity to cloud phase parameterization over the Southern Ocean. All of these studies used the λ_i^L to define local feedbacks.

However, some studies that used λ_i^L did not find the Southern Ocean to be so important for intermodel differences in feedback or warming. Crook et al. (2011) and Roe et al. (2015) used λ_i^L and found nevertheless that the tropics dominate intermodel differences in the global feedback parameter (Crook et al. 2011) and temperature response (Roe et al. 2015), respectively. But since these studies employed slab-ocean models, their setups would not capture the initially strong ocean heat uptake over the Southern Ocean region found in coupled models and observations (Armour et al. 2016).

Therefore, in existing studies that used a combination of the local-temperature definition λ_i^L and an AOGCM, the high latitudes or specifically the Southern Ocean appear to exhibit both large changes in local feedbacks, and large differences in local feedbacks between models.

b. The global-temperature definition

Studies that instead used the global-temperature definition λ_i have not emphasized the Southern Ocean or high-latitude feedbacks (see Table 2). Block and Mauritsen (2013) estimated the majority of feedback changes to occur in a wide band including the tropics and midlatitudes, but explicitly mentioned that the Southern Ocean does not contribute to the increase in feedback after a forcing increase. Andrews et al. (2015) and Rugenstein et al. (2020) partitioned the tropics differently (30°S – 30°N and 22°S – 22°N , respectively), but both arrived at around 60% contribution to global feedback change from the tropics on the centennial time scale after a forcing increase. Regardless of how the tropics are partitioned, studies using the global-temperature definition have repeatedly found that the strongest increases in feedback on decadal to centennial time scales are located over the tropical eastern Pacific (Andrews et al. 2015; Andrews and Webb 2018; Dong et al. 2020; Rugenstein et al. 2020). Indeed, Dong et al. (2020) showed explicitly that the tropical eastern Pacific is the area of strongest increase in feedback in both CMIP5 and CMIP6 models (see also Andrews et al. 2015; Ceppi and Gregory 2017; Zhou et al. 2017; Andrews and Webb 2018; Dong et al. 2019).

TABLE 1. Selection of studies that use the local temperature T_i to normalize or define local feedbacks λ_i^L . Boldface type indicates regions of significance, and italic type indicates relevant findings. Note that the relevant results or figures are not necessarily the main findings of the paper but rather are those relevant to this study. Model ensembles are given in parentheses. Label “atmos” signifies an atmosphere-only GCM usually forced by SSTs. Label SO signifies atmospheric models with a “slab” mixed layer ocean component. MEBM signifies the moist energy balance model.

Study	Spatial averaging	Forcing	Time scale (yr)	Model	Relevant results or figures
Senior and Mitchell (2000)	Hemisphere	$2 \times \text{CO}_2$	800	AOGCM	Southern Ocean changes in cloud, normalized by T_i , are responsible for <i>destabilizing increase</i> in global feedback
Crook et al. (2011)	Zonal; regional	$2 \times \text{CO}_2$	Equilibrium	SO	Tropical cloud feedbacks and water vapor and lapse-rate feedbacks dominate <i>model differences</i> in CMIP3 slab-ocean models and equilibrium warming
Armour et al. (2013)	Grid cell	$2 \times \text{CO}_2$	Multiple	AOGCM; SO	On the decadal to centennial time scale, temperature feedbacks in high latitudes contribute most to <i>destabilizing increase</i> in global feedback as these areas warm; change in <i>pattern weighting</i> drives global feedback change
Feldl and Roe (2013)	Zonal	$2 \times \text{CO}_2$	30	SO	The local definition increases the relative importance of feedbacks in the tropics relative to the high latitudes; local definition is more natural for understanding physical local processes and allows decomposition of spatial warming pattern
Roe et al. (2015)	Zonal	$2 \times \text{CO}_2$	Equilibrium	SO; MEBM	Tropical feedbacks are the <i>largest source of uncertainty</i> in global temperature response
Feldl and Bordoni (2016)	Zonal	$4 \times \text{CO}_2$	150	AOGCM (CMIP5); atmos	Their Fig. 1 shows large <i>intermodel variability</i> and the largest <i>stabilizing feedbacks</i> over the Southern Ocean in CMIP5
Ceppi and Gregory (2017)	Grid cell	$4 \times \text{CO}_2$	20; 150	AOGCM (CMIP5)	Mostly use λ_i^L ; use λ_i^L solely to estimate influence of <i>pattern weighting</i> and conclude that it has a negligible role in determining global feedback change in CMIP5 ensemble mean, with the exception of some models
Colman and Hanson (2017)	Zonal	Multiple	Multiple	AOGCM (CMIP5)	Use λ_i^L to estimate influence of <i>pattern weighting</i> and conclude that it may play an important role in determining global feedback change
Bonan et al. (2018)	Zonal	$4 \times \text{CO}_2$	100	AOGCM (CMIP5); MEBM	Their Fig. 1c suggests that the <i>largest stabilizing feedbacks</i> and high intermodel variability are found over the Southern Ocean in CMIP5
Po-Chedley et al. (2018)	Zonal	$4 \times \text{CO}_2$	120–140	AOGCM (CMIP5)	Particularly in the Southern Hemisphere , local feedbacks dominate the intermodel spread in global lapse-rate and water vapor feedbacks

The tropics are also thought to explain intermodel differences in CMIP3, CMIP5, and CMIP6 models when using λ_i (Zelinka and Hartmann 2012; Vial et al. 2013; Webb et al. 2013; Dong et al. 2020), even though the southern extratropics

do gain in importance in CMIP6 models relative to CMIP5 models (Dong et al. 2020; Zelinka et al. 2020).

Therefore, in studies that used the global-temperature definition λ_i , the tropics appear to exhibit both large changes in

TABLE 2. As in Table 1, but for the selection of studies that use the global temperature T to normalize or define local feedbacks λ_i .

Study	Forcing	Time scale (yr)	Model	Relevant results or figures
Zelinka and Hartmann (2012)	$\sim 3 \times \text{CO}_2$	Transient	AOGCM (CMIP3)	Pacific equatorial feedbacks are <i>net destabilizing</i> ; the largest source of <i>model differences</i> is shortwave cloud feedback over the tropics
Webb et al. (2013)	$2 \times \text{CO}_2$	100	SO	Tropical marine cloud feedbacks dominate the <i>intermodel differences</i> in global feedback
Block and Mauritsen (2013)	$4 \times \text{CO}_2$	20; 130	AOGCM	90% of <i>feedback change</i> driven by 50°S–60°N ; The Southern Ocean is stabilizing and so not responsible for <i>global feedback change</i>
Vial et al. (2013)	$4 \times \text{CO}_2$	130	AOGCM (CMIP5); atmos	The tropics dominate <i>intermodel differences</i> in global feedback parameter
Andrews et al. (2015)	$4 \times \text{CO}_2$	20; 150	AOGCM (CMIP5); atmos	The tropics (30°N–30°S) contribute 60% of total <i>increase in the global feedback</i> , particularly over the eastern Pacific ; primarily driven by clouds
Ceppi and Gregory (2017)	$4 \times \text{CO}_2$	20; 150	AOGCM (CMIP5)	Changes in tropospheric stability, particularly over the tropical east Pacific , dominate (but do not fully explain) <i>changes in global feedback</i>
Dong et al. (2019)	Idealized	—	Atmos; AOGCM	Warming pattern in the western tropical Pacific determines <i>change in global feedback</i> for historical and modeled warming
Dong et al. (2020)	$4 \times \text{CO}_2$	20; 150	AOGCM (CMIP5; CMIP6); atmos	<i>Feedback change</i> dominated by eastern tropical Pacific in both CMIP5 and CMIP6; <i>Intermodel differences</i> determined by the (for CMIP5) western tropical Pacific and (for CMIP6) southern extratropics and eastern tropical Pacific
Rugenstein et al. (2020)	Multiple	20, 150, and 1000	AOGCM (LongRunMIP)	58% of <i>feedback change</i> is driven by the tropics , defined as 22°N–22°S , between years 1–20 and 20–150, particularly over the eastern Pacific ; 23% <i>destabilizing increase</i> from 22°S to 60°S ; for <i>feedback change</i> between years 20–150 and 150–1000, these two latitudes are 47% and 42% respectively; <i>stabilizing decrease in feedback</i> in both time periods from southern high latitudes > 66°S

local feedbacks, and large differences in local feedbacks between models.

3. Model and methods

To test both feedback definitions in a single AOGCM, four model runs using the Max Planck Institute Earth System Model version 1.2 in LR configuration (MPI-ESM1.2; Mauritsen et al. 2019) are integrated out to 1000 years. The LR configuration has approximately 200-km grid spacing with 47 vertical levels in the atmosphere component and approximately 150-km grid spacing with 40 vertical levels in the ocean component. Each run is started from a preindustrial control state, and atmospheric CO_2 concentrations are abruptly increased to either $2\times$, $4\times$, $8\times$, or $16\times$ the preindustrial concentration of 284.7 ppm.

The effective radiative forcing, used only for display purposes in Fig. 1, is determined from the top-of-atmosphere radiation imbalance in four experiments with ECHAM6.3, the atmospheric component of MPI-ESM1.2, in which the sea

surface temperature is held fixed but the CO_2 concentrations are increased to match each of the coupled runs (Hansen et al. 2005; Myhre et al. 2013). The small amount of land warming in these runs is corrected for in the forcing estimate, as suggested in Hansen et al. (2005); for this adjustment, the global feedback is deduced from linear regression between global top-of-atmosphere radiative imbalance and global surface temperature change in the coupled $4\times\text{CO}_2$ simulation.

The quantities required for calculating both feedback definitions are dR_i/dT , dR_i/dT_i , and dT_i/dT , where R_i is the local change in radiative imbalance at the top of the atmosphere, T_i is the change in local surface temperature, and T is the change in globally averaged surface temperature. The baseline for calculating the change in each variable is calculated from a time-mean of a 1500-yr control run. Then, a least squares linear regression is used to estimate the values dR_i/dT , dR_i/dT_i , and dT_i/dT from the regression slopes for three time periods: 1–10, 11–100, and 101–1000 years after the forcing increase.

For calculating discrete differences in these variables [only for Eq. (12), below], the difference in regression slope

between two periods is calculated. For example, to calculate $(dR_i/dT_i)\Delta(dT_i/dT)$ for Eq. (12), we first calculate the step change in the warming pattern by taking the difference in slopes of the $T_i - T$ regressions for the two periods,

$$\begin{aligned} \Delta \frac{dT_i}{dT}(t_1, t_2) &\approx \Delta \frac{dT_i}{dT}(t_1 + \delta t) \\ &= \frac{dT_i}{dT}(t_2) - \frac{dT_i}{dT}(t_1), \end{aligned} \quad (4)$$

which creates a central finite difference around the time $t + \delta t$, where $\delta t = 1/2(t_2 - t_1)$.

For the corresponding value of dR_i/dT_i also centered on $t + \delta t$, we simply average the slope coefficients of the $R - T_i$ regression:

$$\frac{dR_i}{dT_i}(t_1 + \delta t) \approx \frac{1}{2} \left[\frac{dR_i}{dT_i}(t_1) + \frac{dR_i}{dT_i}(t_2) \right]. \quad (5)$$

An additional experiment with identical forcing to the $4 \times \text{CO}_2$ was integrated out to 300 years with partial-radiative perturbation (PRP) diagnostics switched on (Wetherald and Manabe 1988; Colman and McAvaney 1997; Meraner et al. 2013). The PRP diagnostics enabled the radiative contributions of individual feedback types to be separated into temperature (including lapse-rate, Planck, and stratospheric-temperature feedbacks), water vapor, cloud, and albedo feedbacks. Usually, PRP calculations are performed between two years at near equilibrium, but here perturbations are calculated on 300-yr runs, thereby permitting the application of regression techniques. To this end, instantaneous snapshots of model variables were read out every 10 h so as to sample the full diurnal cycle every 5 days. These snapshots were then referenced to a preindustrial control run, which was integrated over 300 years, also with the diagnostics switched on. This method provides more accurate estimates of feedbacks than the commonly used radiative kernel technique, which has inaccuracies associated with the need to linearize otherwise state-dependent kernels (Block and Mauritsen 2013). This is particularly important for runs with strong forcing.

The error of the PRP method can be estimated by summing up all the radiative contributions, include those from atmospheric CO_2 , and comparing this with the actual change in the TOA imbalance. The error in the PRP diagnostic reaches a maximum at the end of the $16 \times \text{CO}_2$ integration of 0.025 W m^{-2} in longwave radiation and -0.008 W m^{-2} in shortwave radiation. This represents 0.2% and 0.1% of the total effective radiative forcing, respectively. In contrast, similar estimates for the kernel method suggest around 10% error (Jonko et al. 2012).

4. Feedback definitions in MPI-ESM1.2

As we show above, the choice of local-feedback definition appears to influence the magnitude of regional feedbacks and feedback variations in the existing literature. However, these studies encompass a wide range of models and setups, so there are almost certainly confounding effects. Furthermore,

there is the potential confounding effect of using either the regression (Gregory et al. 2004) or division method (Murphy 1995) to calculate feedbacks. The impact of definition choice can be put to the test without these confounding factors by directly comparing the two definitions in simulations with a single model, MPI-ESM1.2. The global values for radiative imbalance and surface warming from the simulations are shown in Fig. 1. We begin by examining zonally averaged feedbacks, which are widely used in the literature that examines λ_i^L (see Table 1).

The feedback definition clearly influences the estimated feedback change over the Southern Ocean (Figs. 2 and 3). In the local-temperature definition λ_i^L , the zonally averaged feedbacks around 60°S increase in the century after the abrupt forcing is applied (Figs. 2a and 3a). The global-temperature definition λ_i in this region instead decreases. For feedback changes around the equator, the choice of feedback definition has the opposite effect. In the case of $2 \times \text{CO}_2$ and $4 \times \text{CO}_2$, the feedback increases over the tropics are greater in λ_i than in λ_i^L (Fig. 2b). As a result, the increases in λ_i^L over the Southern Ocean rival the magnitude of changes in the tropics (Fig. 2a). The interpretations of model behavior are therefore not only quantitatively but also qualitatively different for the two feedbacks: in λ_i^L the Southern Ocean and the tropics have comparable feedback increases, whereas in λ_i the stabilizing changes in Southern Ocean feedback actually counteract the destabilizing changes in the tropics.

The difference between the two definitions becomes as large as $10 \text{ W m}^{-2} \text{ K}^{-1}$ over the Southern Ocean for the $4 \times \text{CO}_2$ simulation (Fig. 2b). Most of this difference is concentrated in the initial time period, where Southern Ocean warming is delayed and λ_i^L is considerably more negative than λ_i (Fig. 3b). This initial negative deviation appears to be caused in part by an initial warming rate over the Southern Ocean well below the global warming rate (Fig. 2c), which increases almost uniformly across all forcing strengths from the decadal to centennial period (Fig. 2d). Normalizing radiative changes at the top of atmosphere over a small warming increase would exaggerate local feedbacks, especially when those radiative changes are decoupled from the surface warming (see sections 5–7 for a more detailed investigation of this effect). Exaggerated feedbacks can also partially be seen over the equator for the first two doublings, where the warming rate also increases over time (cf. Figs. 2b,d). The initial negative deviation in λ_i^L may explain why studies using λ_i^L find particularly strong stabilizing feedbacks over the Southern Ocean (Feld and Bordoni 2016; Bonan et al. 2018).

Even at the hemispheric resolution (as applied in Senior and Mitchell 2000), choosing one definition of the local feedback over the other leads to contradictory interpretations of where the global feedback parameter is increasing most. In Fig. 3a, λ_i increases more in the Northern Hemisphere than the Southern Hemisphere. For λ_i^L , however, the feedback change in the Southern Hemisphere is greater (with the exception of the $2 \times \text{CO}_2$ scenario).

The physical origins of the differences in feedback definitions are found particularly in the temperature feedback (lapse-rate, Planck, and stratospheric-temperature feedbacks)

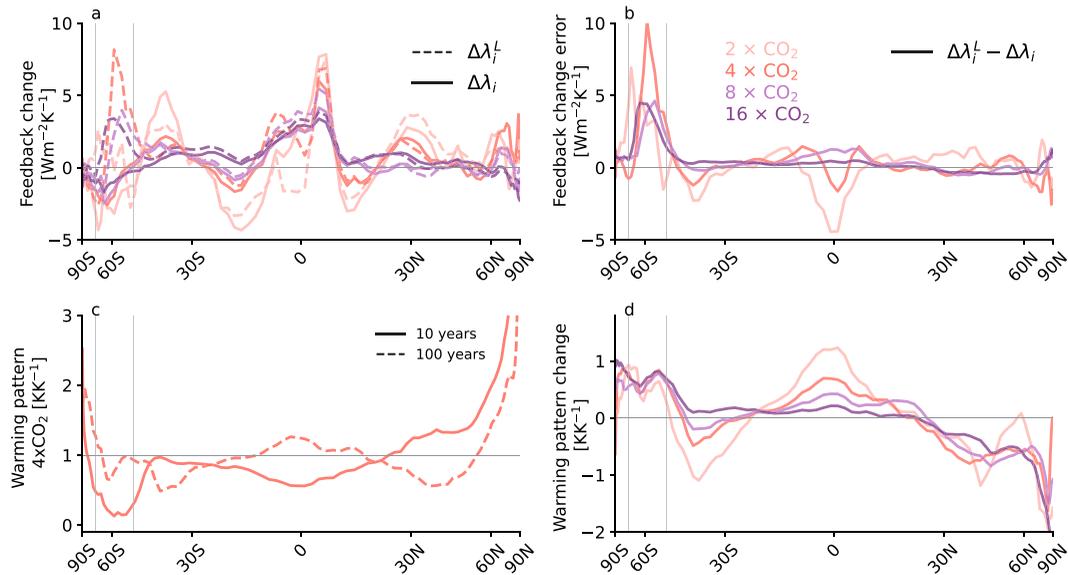


FIG. 2. Local feedbacks and warming patterns over four CO_2 doublings. Feedbacks are calculated using linear regression of local top-of-atmosphere radiation against either the local temperature λ_i^L or the global temperature λ_i over four CO_2 doublings. Shown are the (a) change between years 1–10 and years 11–100 for zonally averaged feedbacks, (b) difference between the local-temperature (λ_i^L) and global-temperature (λ_i) definitions for the feedback change shown in (a), (c) warming pattern dT_i/dT shown for the years 1–10 and years 11–100 in the $4 \times \text{CO}_2$ simulation, and (d) warming pattern change between years 1–10 and years 11–100 for all forcing strengths. Gray vertical lines show the Southern Ocean region, defined as $50\text{--}70^\circ\text{S}$.

and the water vapor feedback (see Fig. 4). The sign and magnitude of these feedbacks changes over both the tropics and the Southern Ocean region in the $4 \times \text{CO}_2$ simulation, depending on the feedback definition. The definition choice also influences the cloud feedback over the high latitudes, but the difference over the tropics is minimal. We note that changing from a regression approach (Gregory et al. 2004) to normalization by division (Murphy 1995) reduces the tropical feedback differences, but over the Southern Ocean it exacerbates the difference in the water vapor feedback and leaves the differences in temperature and cloud feedbacks relatively unchanged (see Fig. 5).

The results presented here are different from those of Feldl and Roe (2013), who found that the local-temperature definition reduced the importance of the high-latitude feedbacks relative to the tropical feedbacks in a slab-ocean, aquaplanet setup. Our results indicate that the opposite is true for a comprehensive AOGCM. The Southern Ocean high-latitude feedbacks gain in relative importance over low-latitude feedbacks when the local-temperature definition is used. The differences in our findings are likely due to the delayed warming over the Southern Ocean in AOGCMs (Armour et al. 2016; Rose and Rayborn 2016; Rugenstein et al. 2016), which is not present in a slab-ocean model (see section 6 for further details).

Last, we apply the two definitions to model output using data at the gridcell resolution (Fig. 6), to examine potential differences masked by zonal averaging. The definitions differ most noticeably over the tropics, not the Southern Ocean, especially in the first two doublings, $2 \times \text{CO}_2$ and $4 \times \text{CO}_2$. Increases in λ_i over the

tropical Pacific and decreases in λ_i over the “Maritime Continent” are absent from λ_i^L (Figs. 6a,d). These differences decrease for higher forcing strengths. Furthermore, although zonally averaged values indicate strong increases in λ_i^L over the Southern Ocean latitudes (Fig. 2a), at the gridcell resolution, there are merely scattered regional increases and decreases in λ_i^L (Fig. 6). Therefore, depending on whether we use data at the gridcell resolution or zonally averaged data to calculate λ_i^L , the relative importance of feedbacks over the Southern Ocean changes drastically, indicating an inconsistency across spatial scales.

5. Inconsistency across spatial scales

A direct comparison of λ_i^L calculated from gridcell data and zonally averaged data shows that the diagnosed feedbacks are considerably different (Fig. 7). The large difference between feedback definitions over the Southern Ocean in the zonally averaged data all but disappears for the gridcell data. Instead, $\Delta\lambda_i^L$ is greater than $\Delta\lambda_i$ over parts of the tropics and northern midlatitudes. For $\Delta\lambda_i$, the spatial resolution of the regressions has no effect on the outcome.

This resolution dependency in λ_i^L violates the naive assumption that the zonally averaged feedbacks are indeed equivalent to the zonal average of the gridcell feedbacks. The reason for this violation is that λ_i^L is not *associative*, that is, the order of the steps taken to calculate λ_i^L —spatial averaging and regression—matters for the outcome.

To demonstrate this effect, we first perform the local regressions before spatial averaging, which returns the following expression:

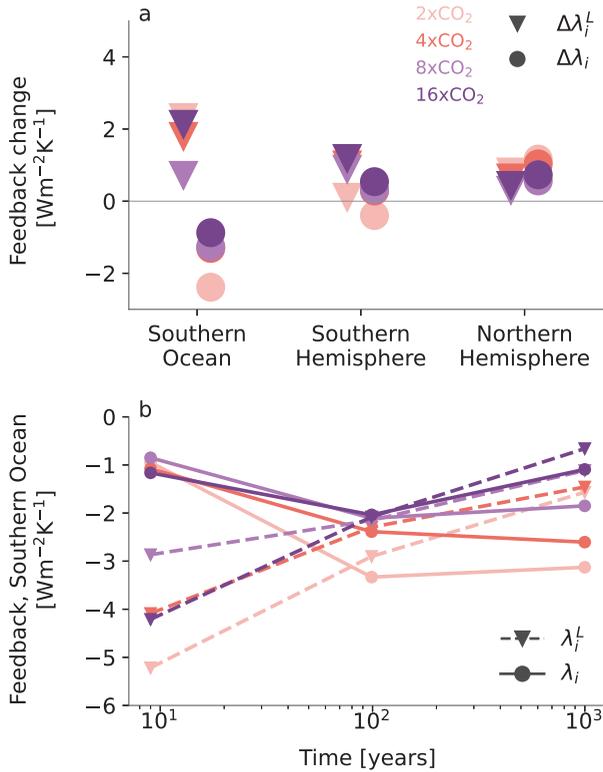


FIG. 3. Local feedbacks and feedback change over four CO₂ doublings for two definitions of the local feedback. Feedbacks are calculated using linear regression of local top-of-atmosphere radiation against either the local temperature λ_i^L or the global temperature λ_i over four CO₂ doublings, for the periods covered by years 1–10, 11–100, and 101–1000. Shown are (a) the change between years 1–10 and years 11–100, calculated over spatially aggregated data for the Southern Ocean region, and for the Southern and Northern Hemispheres, and (b) both feedback definitions over all three time periods for the Southern Ocean region.

$$\begin{aligned}\overline{\lambda_i^L} &= w_1 \lambda_1^L + w_2 \lambda_2^L + \dots \\ &= w_1 \frac{dR_1}{dT_1} + w_2 \frac{dR_2}{dT_2} + \dots,\end{aligned}\quad (6)$$

where w_i represents the spatial weighting in region i . If we perform the spatial averaging first, an alternative expression is returned:

$$\begin{aligned}\overline{\lambda_i^{L*}} &= \frac{d\overline{R_i}}{d\overline{T_i}} \\ &= \frac{d(w_1 R_1 + w_2 R_2 + \dots)}{d(w_1 T_1 + w_2 T_2 + \dots)}.\end{aligned}\quad (7)$$

Equations (6) and (7) are not equivalent. Consider a simple two-box example, where the global value for λ_i^L must be calculated from data at the hemispheric resolution (SH = Southern Hemisphere; NH = Northern Hemisphere). The hemispheres have equal area so that $w_{SH} = w_{NH} = w = 1/2$. Assume also that equal changes in radiation occur in both hemispheres,

$R_{SH} = R_{NH} = R$, and that the regression step can be replaced by a division (recall that R and T are defined as anomalies). This process yields

$$\begin{aligned}\overline{\lambda_i^L} &= w \lambda_{SH}^L + w \lambda_{NH}^L = w \frac{R}{T_{SH}} + w \frac{R}{T_{NH}} \\ &= R \left(\frac{1}{2T_{SH}} + \frac{1}{2T_{NH}} \right)\end{aligned}\quad (8)$$

for Eq. (6) and

$$\begin{aligned}\overline{\lambda_i^{L*}} &= \frac{\overline{R_i}}{\overline{T_i}} = \frac{wR + wR}{wT_{SH} + wT_{NH}} \\ &= R \left(\frac{2}{T_{SH} + T_{NH}} \right)\end{aligned}\quad (9)$$

for Eq. (7). The two expressions are not equivalent. Particularly large differences arise between Eqs. (8) and (9) when one of the warming rates T_i is close to zero, when they are markedly different in magnitude, or when they are of opposite sign. In coupled model experiments with abrupt forcing, the Southern Ocean and the tropical Pacific have warming rates initially well below unity and locally close to zero (Fig. 8c), which is why λ_i^L changes when calculated with zonal or gridcell data. If λ_i^L can change so dramatically dependent on the scale of analysis, the value of displaying and analyzing λ_i^L in studies of local feedback at all becomes questionable.

Only in the case in which $T_i = T$ for all regions do the above expressions become equivalent, which is the case for the global-temperature definition λ_i . The values for λ_i are therefore independent of resolution (Fig. 7).

6. The effect of the warming pattern

Bearing in mind the resolution dependence of the local-temperature definition, we now ask what causes the differences between the two feedback definitions. They can be related as follows:

$$\lambda_i = \frac{dR_i}{dT} = \frac{dR_i}{dT_i} \frac{dT_i}{dT} = \lambda_i^L \frac{dT_i}{dT}.\quad (10)$$

The global-temperature definition of local feedbacks can be written as the product of the local-temperature definition and the warming pattern. Any difference between λ_i^L and λ_i should therefore lie in the weighting of the local feedbacks via the warming pattern term – called *pattern weighting* here.

Pattern weighting has been considered by some as the primary cause of change in the global feedback (Winton et al. 2010; Armour et al. 2013), or at least thought to play an important role (Colman and Hanson 2017). Pattern weighting has since been disregarded as a minor effect in the CMIP5 ensemble mean (Ceppi and Gregory 2017), and has also been disregarded as the cause of intermodel differences in water vapor and lapse-rate feedbacks, which are argued to be driven by changes in λ_i^L instead (Po-Chedley et al. 2018). Therefore, although there seems to be agreement that the pattern of surface warming impacts the global feedback

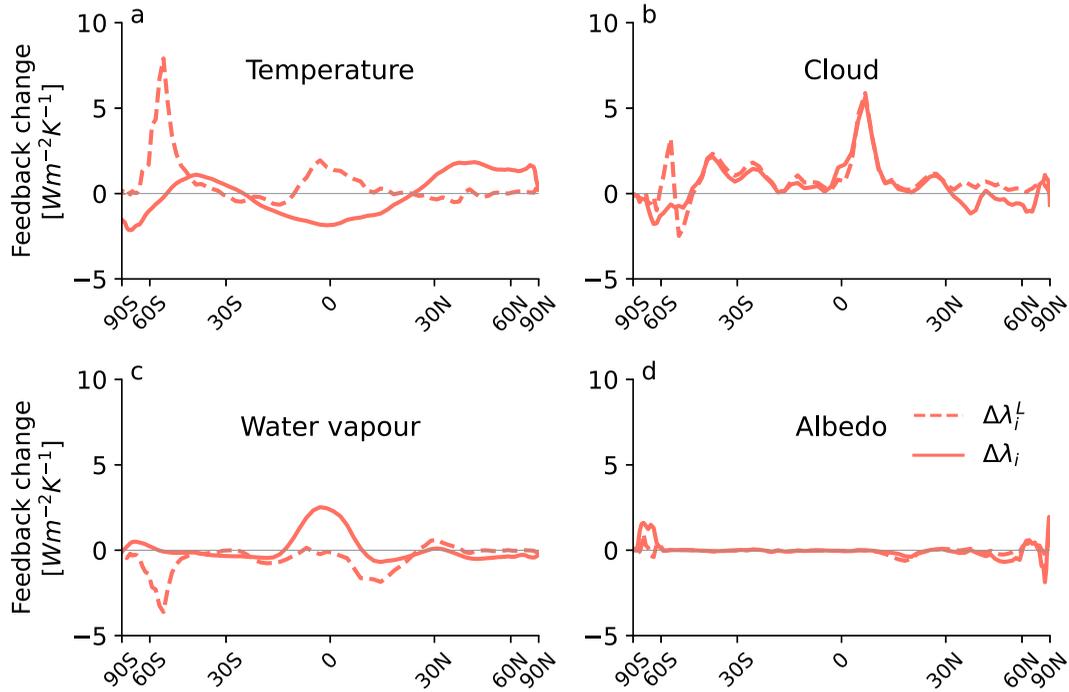


FIG. 4. Local feedback change between years 1–10 and years 11–100 for $4\times\text{CO}_2$, separated into contributions from (a) temperature, (b) clouds, (c) water vapor, and (d) albedo using the partial-radiative-perturbations method. Local feedback changes are calculated using a linear regression against either global temperature change (solid line; $\Delta\lambda_i$) or zonally averaged surface temperature change (dotted line; $\Delta\lambda_i^L$).

parameter (Winton et al. 2010; Armour et al. 2013; Andrews et al. 2015; Zhou et al. 2016; Ceppi and Gregory 2017; Haugstad et al. 2017; Andrews et al. 2018; Dong et al. 2019), there is little evidence to suggest that this impact consistently occurs via the changing pattern weighting of constant local feedbacks.

The importance of the warming pattern in our simulations can be seen in Fig. 8, which shows λ_i , λ_i^L , and dT_i/dT for all three time periods for $4\times\text{CO}_2$. There is a close alignment between the areas where λ_i^L and λ_i differ and where the local warming is considerably slower than the global average. For example, during the first decade, when the tropical Pacific experiences almost no warming, there are strongly negative values in λ_i that are missing in λ_i^L for this region. Likewise for the centennial period, discrepancies between the two definitions are strongest over the western Pacific and Maritime Continent, where local warming rates are well below the global average.

To quantify the pattern-weighting effect in MPI-ESM1.2, we partition the effect of changes in the warming pattern and changes in λ_i^L as follows. The derivative of Eq. (10) with respect to T yields the following expression:

$$\frac{d\lambda_i}{dT} = \frac{dT_i}{dT} \frac{d\lambda_i^L}{dT} + \lambda_i^L \frac{d^2T_i}{dT^2}. \quad (11)$$

Simple step changes in λ_i are analogously expressed by the discrete difference equivalent of Eq. (11):

$$\frac{\Delta\lambda_i}{\Delta T} = \frac{1}{\Delta T} \left(\frac{dT_i}{dT} \Delta\lambda_i^L + \lambda_i^L \Delta \frac{dT_i}{dT} \right) \quad \text{and}$$

$$\Delta\lambda_i = \underbrace{\frac{dT_i}{dT} \Delta\lambda_i^L}_{\text{FC}} + \underbrace{\lambda_i^L \Delta \frac{dT_i}{dT}}_{\text{PC}}. \quad (12)$$

The change in the total local feedback $\Delta\lambda_i$ can be divided into two parts: The feedback change part (FC) describes the local-temperature definition of feedback change $\Delta\lambda_i^L$, weighted by the average warming pattern. The pattern change part (PC) describes the changes in the warming pattern weighted by constant λ_i^L .

We should be able to reconstruct $\Delta\lambda_i$ from $\Delta\lambda_i^L$ by consecutively including the effect of pattern weighting in both parts FC and PC. Any remaining difference between the reconstruction and the true value of $\Delta\lambda_i$ can be viewed as an error caused by the process of decomposition. Figure 9a shows the magnitude of the two parts for each latitude in $4\times\text{CO}_2$: first, including the impact of constant pattern weighting to produce a weighted feedback change (FC), and second, the impact of the *change* in pattern weighting (PC). The greatest impact of these effects is over the Southern Ocean. There, FC reduces the error somewhat, but PC is required to reduce the error to almost zero.

Pattern weighting is important over the Southern Ocean because of the slow initial warming there. The boundary layer is initially convectively decoupled from the free

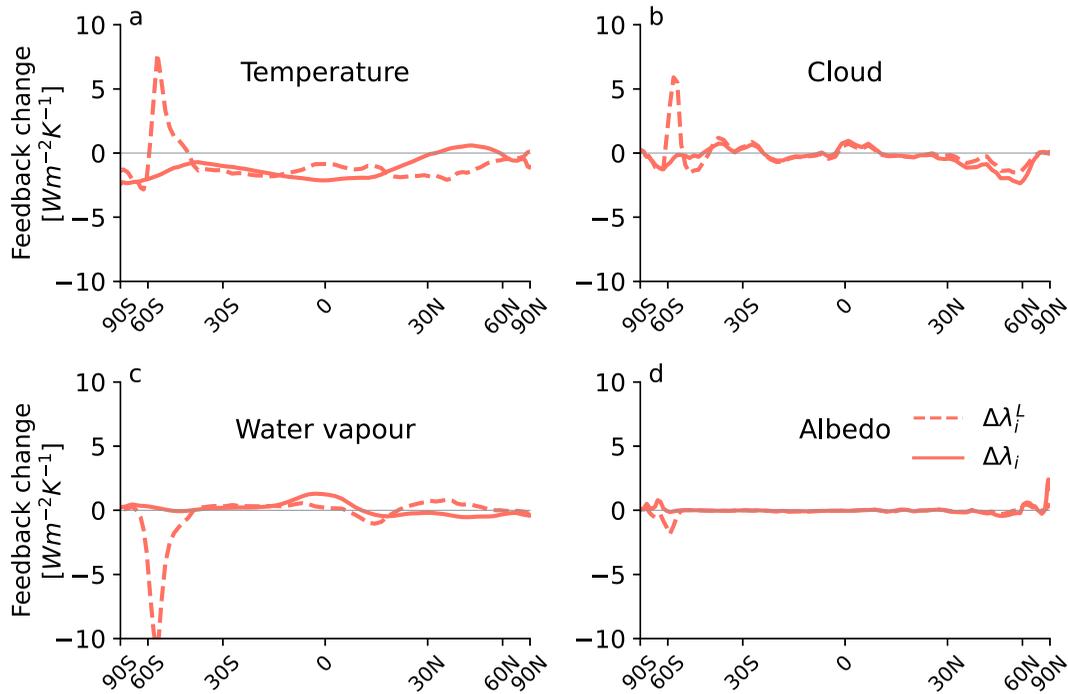


FIG. 5. As in Fig. 4, but local feedback changes are calculated by dividing top-of-atmosphere radiative changes by either global temperature change (solid line; $\Delta\lambda_i$) or zonally averaged surface temperature change (dotted line; $\Delta\lambda_i^z$).

troposphere, which allows increases in outgoing radiation to occur despite the slow surface warming (for a conceptual model, see Bloch-Johnson et al. 2020), creating strongly negative λ_i^z (cf. Feldl and Bordoni 2016 and Bonan et al. 2018). When the local warming accelerates, the radiative changes are normalized by a larger warming trend, resulting in the impression that the local feedback increases.

This effect could also exaggerate intermodel differences in λ_i^z , depending on the warming rate over the Southern Ocean. Indeed, the intermodel differences in warming rate are largest in the extratropics, where initial warming rates can differ between less than 0 and greater than 2 K K^{-1} among CMIP5 models (Andrews et al. 2015).

Pattern weighting is not only important regionally but also for the change in global feedback parameter, as shown in Fig. 9b. Using zonally averaged data, the global spatial average of $\Delta\lambda_i^z$ is larger than the global feedback parameter for all forcing strengths—in the case of $8 \times \text{CO}_2$ by as much as a factor of 2. In all forcing strengths except $2 \times \text{CO}_2$, including pattern weighting to produce FC reduces the error by around half, while adding PC further reduces the remaining error, leaving only a small residual difference. Therefore, the change in weighting via the warming pattern is crucial for feedback change both regionally and globally, when seen from the zonal perspective. These findings may place MPI-ESM1.2 in a minority of models, as Ceppi and Gregory (2017) have suggested, which show a substantial pattern-weighting effect.

However, it is also possible that previous attempts to decompose pattern weighting and local feedback are highly sensitive to methodological choices. Notice that for the

gridcell resolution in Fig. 9b, the sum of parts FC and PC does not accurately reconstruct the global feedback parameter. For $2 \times \text{CO}_2$, not even zonally averaged data can reconstruct the global feedback. Except for the extreme case of $16 \times \text{CO}_2$, the error in the reconstruction for grid cell data (and for zonally averaged data in $2 \times \text{CO}_2$) is on the same order of magnitude as the global feedback parameter itself. The magnitude of the error would prohibit drawing any conclusions from a decomposition into warming pattern and local feedback at the gridcell resolution, which is the resolution that Ceppi and Gregory (2017) used to test the pattern weighting hypothesis. In the next section, we explore why decomposition into warming pattern and λ_i^z fails on the gridcell resolution, and how such a decomposition might be improved.

7. Surface warming alone cannot predict local radiation budget

To evaluate the effectiveness of the decomposition into warming pattern and λ_i^z on the gridcell resolution, we attempt to reconstruct the feedback from the individual components. Focusing on the feedback itself rather than feedback change simplifies the analysis at different time scales.

The poor decomposition is evident in Fig. 8, in which the regressions used to calculate each parameter are subject to a two-tailed t test with a 2.5%–97.5% confidence interval that the regression slope is not zero. Trends shaded with stippling fail this test and are not significantly different from zero. For the decadal time period, almost all gridcell values of λ_i and λ_i^z fail the test. For these areas, there is not sufficient correlation

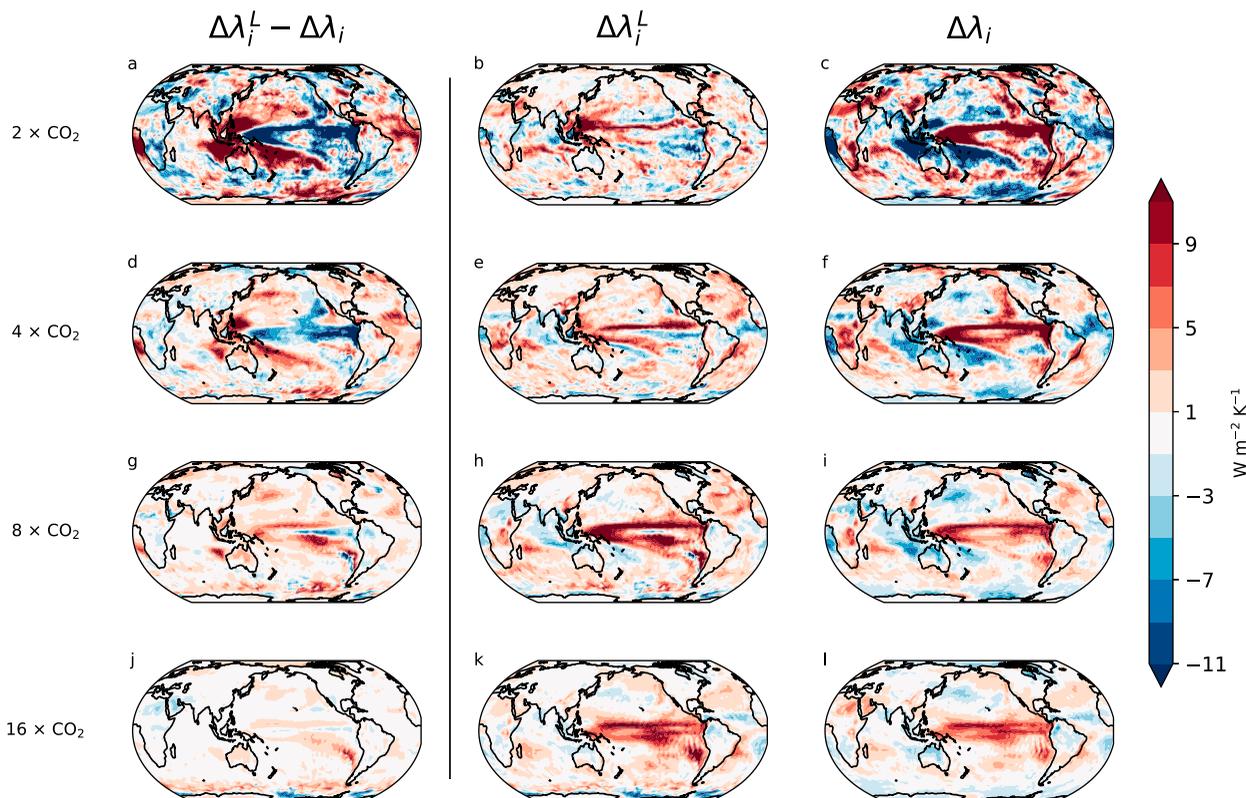


FIG. 6. Local feedback change between years 1–10 and years 11–100 over four CO_2 doublings, for two definitions of the local feedback: (a),(d),(g),(j) difference between the two definitions; (b),(e),(h),(k) local feedback changes calculated using linear regression against local temperature $\Delta\lambda_i^L$; and (c),(f),(i),(l) local feedback changes calculated using linear regression against global temperature $\Delta\lambda_i$.

between top-of-atmosphere radiation change and either local or global surface warming to justify a regression and therefore a feedback parameter. On the centennial time period, regions of strong negative feedback over the Maritime Continent and Atlantic pass the significance test in λ_i , but to a lesser extent

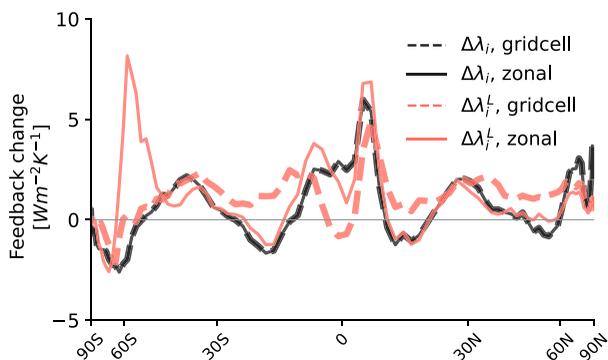


FIG. 7. Feedback change between years 1–10 and years 11–100 for the $4\times\text{CO}_2$ simulation, according to the local-temperature definition ($\Delta\lambda_i^L$; orange) and global-temperature definition ($\Delta\lambda_i$; black). Solid lines represent regressions performed over zonally averaged data; dotted lines represent regressions over gridcell data and then zonally averaged.

in λ_i^L . The southern midlatitudes fail the test in both feedback definitions.

Just as the decomposition of feedbacks on the gridcell resolution is poor, so is the reconstruction. The product of dR_i/dT_i and dT_i/dT should return λ_i , but it does not do so for any of the time periods (Fig. 10). In the first decade, strong negative feedbacks over the tropics are missing from the reconstruction (cf. Figs. 10a and 10b). Negative feedbacks are also underestimated in the Southern Ocean regions and the Atlantic. For the centennial time period, negative feedbacks are underestimated over the Maritime Continent, the Atlantic, the midlatitudes, and the Southern Ocean, which are all regions where the local warming is delayed by ocean heat uptake (Figs. 8c,f). For the millennial time period, large areas of the feedback in the mid- and high latitudes are not accurately reconstructed. These results indicate that surface warming is not sufficient to explain the local top-of-atmosphere radiation change, especially in the decadal time period, but also for large parts of the globe in the centennial and millennial time periods.

The decomposition and reconstruction fails locally because the surface temperature cannot account for top-of-atmosphere radiation changes. In some regions, upper-tropospheric warming—convectively decoupled from the surface warming and driven instead by the global circulation—impacts the radiation

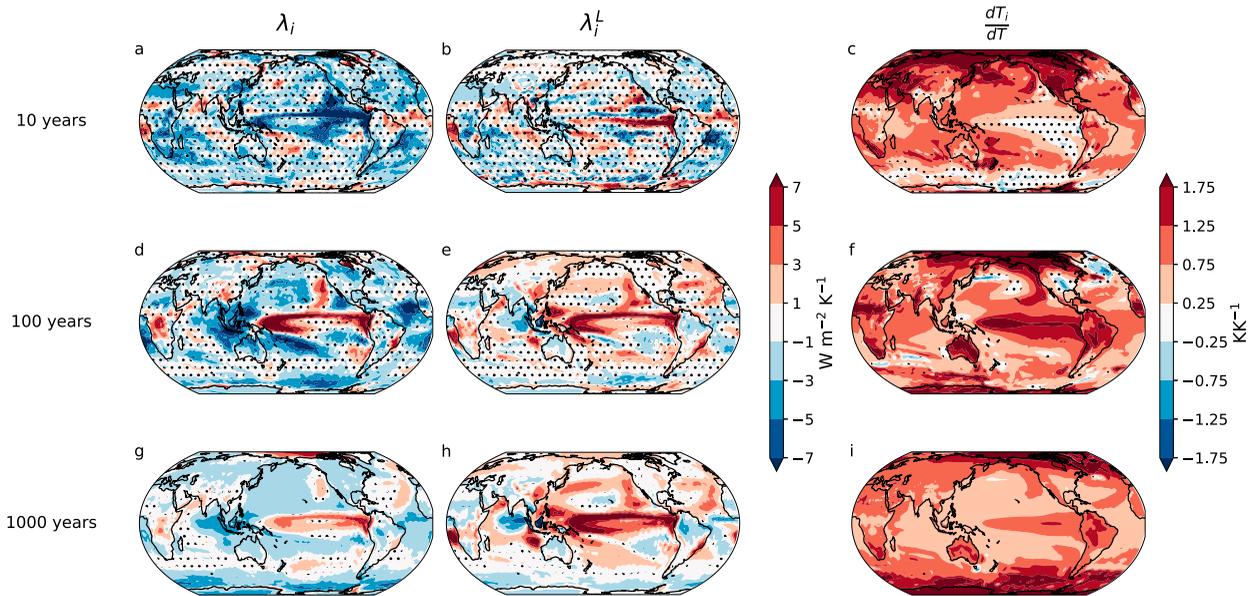


FIG. 8. Local feedbacks according to both definitions and the warming pattern in the $4 \times \text{CO}_2$ simulation over all three time periods (1–10, 11–100, and 101–1000 years). Stippling represents failure of a two-sided t test with 95% confidence that the regression slope is not equal to zero.

budget more than surface warming. In this case, the decomposition into surface warming pattern and λ_i^L would fail because there is no causal connection or correlation between the surface warming and the top-of-atmosphere radiation changes.

Alternative explanatory variables for radiation change have included ocean heat uptake (Winton et al. 2010; Rose et al. 2014; Rugenstein et al. 2016), nonlocal effects (Zhou et al. 2017; Dong et al. 2019, 2020; Bloch-Johnson et al. 2020), midtropospheric temperature (Dessler et al. 2018), and tropospheric stability (Zhou et al. 2016; Ceppi and Gregory 2017; Andrews et al. 2018; Ceppi and Gregory 2019). Since ocean heat uptake arguably impacts feedback change via the surface warming pattern (Haugstad et al. 2017), including ocean heat uptake would not improve the decomposition. Nonlocal effects are proposed to act principally via changes in tropospheric warming in ascent regions (Zhou et al. 2017; Dong et al. 2019; Bloch-Johnson et al. 2020). Tropospheric stability is also an important controlling factor for low cloud feedbacks (Klein et al. 2017; Myers et al. 2021). Therefore, by including tropospheric stability, we might effectively capture both the nonlocal framework and midtropospheric temperature with a single local variable. Ceppi and Gregory (2019) improved estimates of the global energy balance [Eq. (1)] by including a measure for tropospheric stability S :

$$R = F + \Delta T + \sigma S, \quad (13)$$

where σ is a feedback parameter relating the large-scale tropospheric stability to changes in outgoing radiation at the top of atmosphere. If S is important for the global change, it might also contribute to the local model accordingly:

$$R_i = F_i + \lambda_i^L T_i + \sigma_i S_i. \quad (14)$$

Derivation with respect to global surface warming T yields

$$\begin{aligned} \lambda_i &= \frac{dR_i}{dT} = \frac{dR_i}{dT_i} \frac{dT_i}{dT} + \frac{dR_i}{dS_i} \frac{dS_i}{dT} \\ &= \lambda_i^L \frac{dT_i}{dT} + \sigma_i \frac{dS_i}{dT}. \end{aligned} \quad (15)$$

Ceppi and Gregory (2019) measured tropospheric stability using the large-scale estimated inversion strength (Wood and Bretherton 2006), a quantity that is spatially averaged over ocean regions equatorward of 50° and is based on the moist adiabatic lapse rate and the lower-troposphere stability—defined in turn as the difference in potential temperature between 700 hPa and the surface (Klein and Hartmann 1993). For our local analysis we use the lower-troposphere stability as a measure of tropospheric stability, calculated for any region i .

The result of including this additional term can be seen in Fig. 10, where the third column shows the decomposition according to the pattern of S_i . In the decadal period, S_i adds many of the feedbacks that were missing due to a decomposition using surface warming alone (cf. Figs. 10a and 10d).

What is occurring when S_i succeeds but T_i fails to reproduce λ_i ? In these regions, radiative changes at the top of atmosphere are caused by warming in the tropospheric column independently of the surface. Convective decoupling of troposphere and surface warming can lead to a very weak relationship between outgoing radiation changes and surface warming, so that regression over this relationship produces values for λ_i^L that are indistinguishable from zero. The reconstruction $(dR_i/dS_i)(dS_i/dT)$ succeeds in these areas by virtue

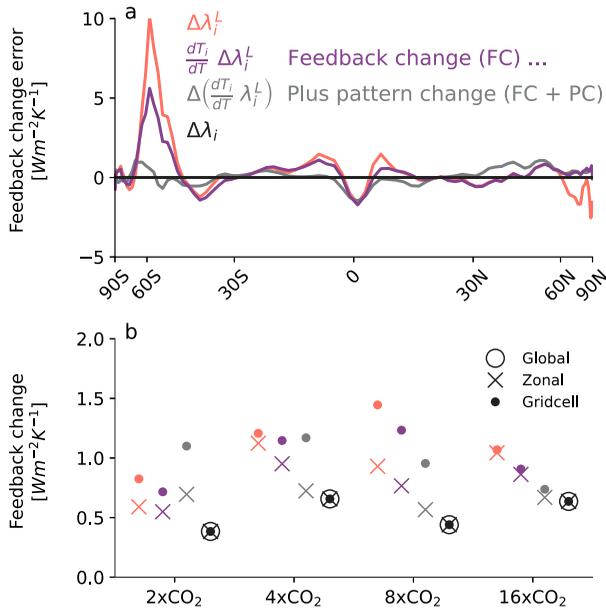


FIG. 9. Feedback change between time periods 1–10 and 11–100, according to the local-temperature definition ($\Delta\lambda_i^L$; orange) and global-temperature definition ($\Delta\lambda_i$; black), with a decomposition of individual components as derived in Eq. (12). Including the weighting of $\Delta\lambda_i^L$ by the warming pattern yields the feedback change component (FC; purple). Additionally including the pattern change component (FC + PC; gray) should yield the full global-temperature definition $\Delta\lambda_i$. The results are shown (a) expressed as difference from $\Delta\lambda_i$ for the $4\times\text{CO}_2$ scenario in zonal resolution and for (b) each of the four components calculated at the zonal and gridcell resolutions and then spatially averaged. The global feedback parameter in (b) (black open circles) is a regression of globally averaged R over T .

of the fact that S_i has a causal relationship to the radiative changes.

For the centennial time period, the S_i component continues to contribute important negative feedbacks over the tropics

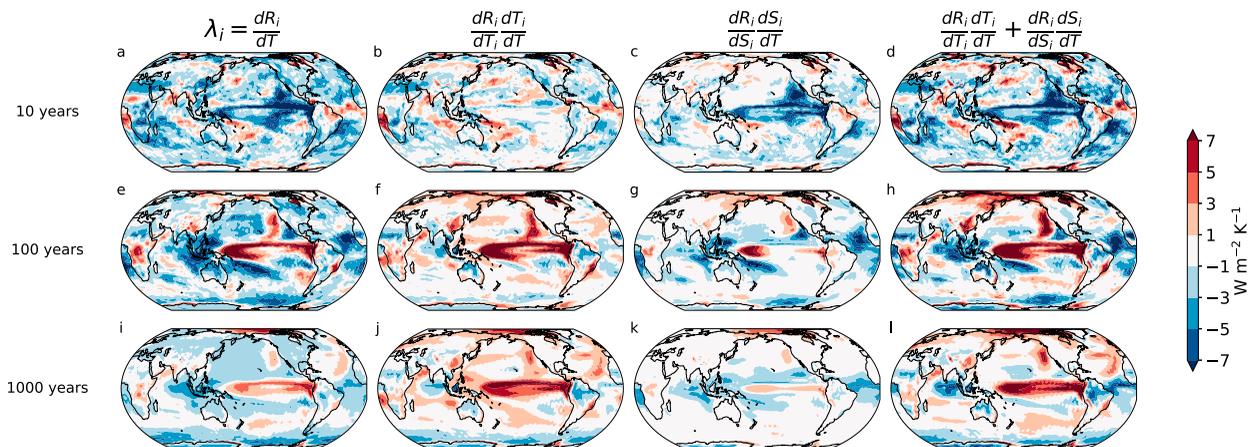


FIG. 10. Local feedback and its decomposition in the $4\times\text{CO}_2$ simulation over all three time periods (1–10, 11–100, 101–1000 years): (a),(e),(i) the global-temperature definition of local feedback λ_i , followed by reconstructions of λ_i based on a decomposition (b),(f),(j) with respect to local temperature (T_i) and (c),(g),(k) with respect to local lower-troposphere stability (S_i). Also shown is (d),(h),(l) the sum of both reconstructions.

and Maritime Continent for the centennial time period (cf. Figs. 10e and 10h). However, for the millennial period, the new reconstruction (Fig. 10l) performs poorly relative to the true local feedback (Fig. 10i).

Therefore, whereas Ceppi and Gregory (2017) find that global tropospheric stability can explain patterns in feedback change, we find that both the local surface warming and local tropospheric stability changes are required to adequately explain local feedbacks. For the millennial time period, however, the decomposition into local warming and tropospheric stability breaks down.

This calls into question previous attempts to decompose local feedbacks into λ_i^L and the warming pattern at the level of the grid cell, and therefore makes assessing pattern weighting at this resolution problematic or at least sensitive to methodological choices. Considering that the success of the decomposition is sensitive to scale, sensitive to forcing, and presumably sensitive to the regression or division methods (cf. Figs. 4 and 5), it is perhaps unsurprising that some studies have found support for pattern weighting (Winton et al. 2010; Armour et al. 2013) while others have found more support for alternative hypotheses (Ceppi and Gregory 2017; Po-Chedley et al. 2018). Knowing the true effect of pattern weighting would require a more comprehensive and systematic approach that ensures that the decomposition successfully reconstructs the global feedback parameter.

8. Conclusions

There are two ways to define the local radiative feedback in the existing literature, either with global surface temperature or local surface temperature. To date, the effect of choosing one definition over the other has not been analyzed in AOGCMs. We find that the definition choice can influence the sign and magnitude of feedbacks in different regions, leading to qualitatively different conclusions about regional feedback strength and change.

There are several aspects of our results that lead to this conclusion. First, studies that use the local-temperature definition tend to find large feedbacks or feedback variations in the high latitudes or over the Southern Ocean. Studies that use the global-temperature definition instead tend to find large feedbacks and feedback variations in the tropics. This does not mean that all of these studies are seeking to understand changes in the global feedback—their research questions and their findings vary—but it does indicate a broad correspondence between methodological choice and the interpretation of local feedbacks and feedback changes. Second, when we analyze the results of a single AOGCM over four CO₂ doublings, we find that the local-temperature definition exaggerates Southern Ocean feedbacks, misconstruing their importance for the global feedback. We show that this bias could also lead to exaggeration of intermodel differences in the Southern Ocean region, since warming rates differ most between models in the extratropics as compared with other regions. Third, we show analytically that the global-temperature definition of local feedback change is more complete than the local-temperature definition, because it directly accounts for the weighting of feedbacks by the warming pattern. Differences between both feedback definitions can be mostly accounted for by incorporating the change in the warming pattern and the resulting change in weighting of local feedbacks, at least for zonally averaged data and forcing strengths above $2 \times \text{CO}_2$.

We additionally discuss for the first time the sensitivity of local feedbacks to the resolution of analysis and explain the reasons for this sensitivity. First, the local-temperature definition is not associative, so that zonal averages of feedbacks calculated from gridcell properties are mathematically different from feedbacks calculated from zonally averaged properties. The global-temperature definition of local feedbacks is, however, equivalent across spatial scales and can be linearly integrated to yield the true global feedback. Second, at the gridcell resolution, the local surface warming fails to explain radiation changes over regions in which tropospheric warming or vertical stability drives the outgoing radiation balance. Therefore, assessments of local-temperature feedbacks and pattern weighting at the gridcell level (such as in Ceppi and Gregory 2017) may fail to reconstruct the global feedback parameter. We show that decompositions into feedback and warming pattern can be more successful with zonal data for forcing strengths larger than $2 \times \text{CO}_2$. For data at gridcell resolution, we show that incorporating tropospheric stability changes into the conceptual model can help reconstruct the top-of-atmosphere radiative changes for decadal and centennial time scales.

In summary, our findings show that the local-temperature definition of feedbacks should be interpreted with caution in AOGCMs, and that the global-temperature definition of local feedbacks is more robust. We conclude that the strong feedback effects in high latitudes and over the Southern Ocean in existing literature are an artifact resulting from this methodological choice. When we discount studies that use the local-temperature definition of feedback to locate regional drivers of the global feedback, the literature speaks overwhelmingly for the tropics as the region that contributes most to the increase in the global

feedback parameter in the century after an abrupt forcing increase.

Acknowledgments. The authors thank Maria Rugenstein and three anonymous reviewers for their critical and insightful reviews of the paper and Bjorn Stevens for helpful discussions about feedbacks and state dependence. This work was supported by the Max-Planck-Gesellschaft (MPG). Authors Hedemann and Marotzke were partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2037 ‘CLICCS - Climate, Climatic Change, and Society’ – Project Number: 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg. Author Mauritsen acknowledges funding from the European Research Council (ERC) (Grant Agreement 770765) and the European Union's Horizon 2020 research and innovation program (Grant Agreement 820829). Computational resources were made available by the Deutsches Klimarechenzentrum (DKRZ).

Data availability statement. The MPI-ESM1.2 model version (release 1.2.01p5 ‘CMIP6p5’) used to generate simulations in this study is available at <https://code.mpimet.mpg.de/versions/477> after first registering at <https://mpimet.mpg.de/en/science/modeling-with-icon/code-availability/mpi-esm-users-forum>. Model simulations were run at DKRZ on the mistral computer. Computer code used in the postprocessing of raw model output as well as zonal and global means of data used in this study has been deposited with the Max Planck Society (<http://hdl.handle.net/21.11116/0000-0009-D2D1-D>).

REFERENCES

- 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>.
- , J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.*, **9**, 549–554, <https://doi.org/10.1038/ngeo2731>.
- Bloch-Johnson, J., R. T. Pierrehumbert, and D. S. Abbot, 2015: Feedback temperature dependence determines the risk of

- high warming. *Geophys. Res. Lett.*, **42**, 4973–4980, <https://doi.org/10.1002/2015GL064240>.
- , M. Rugenstein, and D. S. Abbot, 2020: Spatial radiative feedbacks from internal variability using multiple regression. *J. Climate*, **33**, 4121–4140, <https://doi.org/10.1175/JCLI-D-19-0396.1>.
- Block, K., and T. Mauritsen, 2013: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO₂. *J. Adv. Model. Earth Syst.*, **5**, 676–691, <https://doi.org/10.1002/jame.20041>.
- Bonan, D. B., K. C. Armour, G. H. Roe, N. Siler, and N. Feldl, 2018: Sources of uncertainty in the meridional pattern of climate change. *Geophys. Res. Lett.*, **45**, 9131–9140, <https://doi.org/10.1029/2018GL079429>.
- Brown, P. T., W. Li, J. H. Jiang, and H. Su, 2016: Unforced surface air temperature variability and its contrasting relationship with the anomalous TOA energy flux at local and global spatial scales. *J. Climate*, **29**, 925–940, <https://doi.org/10.1175/JCLI-D-15-0384.1>.
- Caballero, R., and M. Huber, 2013: State-dependent climate sensitivity in past warm climates and its implications for future climate projections. *Proc. Natl. Acad. Sci. USA*, **110**, 14162–14167, <https://doi.org/10.1073/pnas.1303365110>.
- Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proc. Natl. Acad. Sci. USA*, **114**, 13126–13131, <https://doi.org/10.1073/pnas.1714308114>.
- , and —, 2019: A refined model for the Earth's global energy balance. *Climate Dyn.*, **53**, 4781–4797, <https://doi.org/10.1007/s00382-019-04825-x>.
- Colman, R., and B. J. McAvaney, 1997: A study of general circulation model climate feedbacks determined from perturbed sea surface temperature experiments. *J. Geophys. Res.*, **102**, 19383–19402, <https://doi.org/10.1029/97JD00206>.
- , and L. Hanson, 2017: On the relative strength of radiative feedbacks under climate variability and change. *Climate Dyn.*, **49**, 2115–2129, <https://doi.org/10.1007/s00382-016-3441-8>.
- Crook, J. A., P. M. Forster, and N. Stuber, 2011: Spatial patterns of modeled climate feedback and contributions to temperature response and polar amplification. *J. Climate*, **24**, 3575–3592, <https://doi.org/10.1175/2011JCLI3863.1>.
- Dessler, A. E., T. Mauritsen, and B. Stevens, 2018: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmos. Chem. Phys.*, **18**, 5147–5155, <https://doi.org/10.5194/acp-18-5147-2018>.
- Dong, Y., C. Proistosescu, K. C. Armour, and D. S. Battisti, 2019: Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's function approach: The preeminence of the western Pacific. *J. Climate*, **32**, 5471–5491, <https://doi.org/10.1175/JCLI-D-18-0843.1>.
- , 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. *J. Climate*, **33**, 7755–7775, <https://doi.org/10.1175/JCLI-D-19-1011.1>.
- Feldl, N., and G. H. Roe, 2013: Four perspectives on climate feedbacks. *Geophys. Res. Lett.*, **40**, 4007–4011, <https://doi.org/10.1002/grl.50711>.
- , and S. Bordoni, 2016: Characterizing the Hadley circulation response through regional climate feedbacks. *J. Climate*, **29**, 613–622, <https://doi.org/10.1175/JCLI-D-15-0424.1>.
- , —, and T. M. Merlis, 2017: Coupled high-latitude climate feedbacks and their impact on atmospheric heat transport. *J. Climate*, **30**, 189–201, <https://doi.org/10.1175/JCLI-D-16-0324.1>.
- Frey, W. R., E. A. Maroon, A. G. Pendergrass, and J. E. Kay, 2017: Do Southern Ocean cloud feedbacks matter for 21st century warming? *Geophys. Res. Lett.*, **44**, 12447–12456, <https://doi.org/10.1002/2017GL076339>.
- Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. J. L. Olivie, and S. Tytca, 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. *J. Climate*, **26**, 1859–1876, <https://doi.org/10.1175/JCLI-D-12-00196.1>.
- 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>.
- , and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, **31**, L03205, <https://doi.org/10.1029/2003GL018747>.
- Hansen, J., and Coauthors, 2005: Efficacy of climate forcings. *J. Geophys. Res.*, **110**, D18104, <https://doi.org/10.1029/2005JD005776>.
- Hargreaves, J. C., and J. D. Annan, 2016: Could the Pliocene constrain the equilibrium climate sensitivity? *Climate Past*, **12**, 1591–1599, <https://doi.org/10.5194/cp-12-1591-2016>.
- Haugstad, A. D., K. C. Armour, D. S. Battisti, and B. E. Rose, 2017: Relative roles of surface temperature and climate forcing patterns in the inconstancy of radiative feedbacks. *Geophys. Res. Lett.*, **44**, 7455–7463, <https://doi.org/10.1002/2017GL074372>.
- Huber, M., and R. Caballero, 2011: The early Eocene equable climate problem revisited. *Climate Past*, **7**, 603–633, <https://doi.org/10.5194/cp-7-603-2011>.
- Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu, 2012: Climate feedbacks in CCSM3 under changing CO₂ forcing. Part I: Adapting the linear radiative kernel technique to feedback calculations for a broad range of forcings. *J. Climate*, **25**, 5260–5272, <https://doi.org/10.1175/JCLI-D-11-00524.1>.
- , —, —, and —, 2013: Climate feedbacks in CCSM3 under changing CO₂ forcing. Part II: Variation of climate feedbacks and sensitivity with forcing. *J. Climate*, **26**, 2784–2795, <https://doi.org/10.1175/JCLI-D-12-00479.1>.
- Kay, J. E., M. M. Holland, C. M. Bitz, E. Blanchard-Wrigglesworth, A. Gettelman, A. Conley, and D. Bailey, 2012: The influence of local feedbacks and northward heat transport on the equilibrium Arctic climate response to increased greenhouse gas forcing. *J. Climate*, **25**, 5433–5450, <https://doi.org/10.1175/JCLI-D-11-00622.1>.
- Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. Climate*, **6**, 1587–1606, [https://doi.org/10.1175/1520-0442\(1993\)006<1587:TSCOLS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2).
- , A. Hall, J. R. Norris, and R. Pincus, 2017: Low-cloud feedbacks from cloud-controlling factors: A review. *Surv. Geophys.*, **38**, 1307–1329, <https://doi.org/10.1007/s10712-017-9433-3>.
- Li, C., J.-S. Storch, and J. Marotzke, 2013: Deep-ocean heat uptake and equilibrium climate response. *Climate Dyn.*, **40**, 1071–1086, <https://doi.org/10.1007/s00382-012-1350-z>.
- Mauritsen, T., and Coauthors, 2019: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂. *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>.
- Murphy, J. M., 1995: Transient response of the Hadley Centre coupled ocean-atmosphere model to increasing carbon-dioxide. Part III: Analysis of global-mean response using simple-models. *J. Climate*, **8**, 496–514, [https://doi.org/10.1175/1520-0442\(1995\)008<0496:TROTHC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<0496:TROTHC>2.0.CO;2).
- Myers, T. A., R. C. Scott, M. D. Zelinka, S. A. Klein, J. R. Norris, and P. M. Caldwell, 2021: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity. *Nat. Climate Change*, **11**, 501–507, <https://doi.org/10.1038/s41558-021-01039-0>.
- Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis*, T. Stocker et al., Eds., Cambridge University Press, 659–740, <https://doi.org/10.1017/CBO9781107415324.018>.
- Paynter, D., T. L. Frölicher, L. W. Horowitz, and L. G. Silvers, 2018: Equilibrium climate sensitivity obtained from multimillennial runs of two GFDL climate models. *J. Geophys. Res. Atmos.*, **123**, 1921–1941, <https://doi.org/10.1002/2017JD027885>.
- Po-Chedley, S., K. C. Armour, C. M. Bitz, M. D. Zelinka, B. D. Santer, and Q. Fu, 2018: Sources of intermodel spread in the lapse rate and water vapor feedbacks. *J. Climate*, **31**, 3187–3206, <https://doi.org/10.1175/JCLI-D-17-0674.1>.
- Roe, G. H., N. Feldl, K. C. Armour, Y.-T. T. Hwang, and D. M. W. Frierson, 2015: The remote impacts of climate feedbacks on regional climate predictability. *Nat. Geosci.*, **8**, 135–139, <https://doi.org/10.1038/ngeo2346>.
- Rohrschneider, T., B. Stevens, and T. Mauritsen, 2019: On simple representations of the climate response to external radiative forcing. *Climate Dyn.*, **53**, 3131–3145, <https://doi.org/10.1007/s00382-019-04686-4>.
- Rose, B. E. J., and L. Rayborn, 2016: The effects of ocean heat uptake on transient climate sensitivity. *Curr. Climate Change Rep.*, **2**, 190–201, <https://doi.org/10.1007/s40641-016-0048-4>.
- , 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. *Geophys. Res. Lett.*, **41**, 1071–1078, <https://doi.org/10.1002/2013GL058955>.
- Royer, D. L., 2016: Climate sensitivity in the geologic past. *Annu. Rev. Earth Planet. Sci.*, **44**, 277–293, <https://doi.org/10.1146/annurev-earth-100815-024150>.
- Rugenstein, M. A. A., K. Caldeira, and R. Knutti, 2016: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophys. Res. Lett.*, **43**, 9877–9885, <https://doi.org/10.1002/2016GL070907>.
- Rugenstein, M., and Coauthors, 2019: LongRunMIP: Motivation and design for a large collection of millennial-length AOGCM simulations. *Bull. Amer. Meteor. Soc.*, **100**, 2551–2570, <https://doi.org/10.1175/BAMS-D-19-0068.1>.
- , and Coauthors, 2020: Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophys. Res. Lett.*, **47**, e2019GL083898, <https://doi.org/10.1029/2019GL083898>.
- 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>.
- Shaffer, G., M. Huber, R. Rondanelli, and J. O. Pepke Pedersen, 2016: Deep time evidence for climate sensitivity increase with warming. *Geophys. Res. Lett.*, **43**, 6538–6545, <https://doi.org/10.1002/2016GL069243>.
- Sherwood, S. C., and Coauthors, 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.*, **58**, e2019RG000 678, <https://doi.org/10.1029/2019RG000678>.
- 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>.
- Vial, J., J.-L. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dyn.*, **41**, 3339–3362, <https://doi.org/10.1007/s00382-013-1725-9>.
- Webb, M. J., F. H. Lambert, and J. M. Gregory, 2013: Origins of differences in climate sensitivity, forcing and feedback in climate models. *Climate Dyn.*, **40**, 677–707, <https://doi.org/10.1007/s00382-012-1336-x>.
- Wetherald, R. T., and S. Manabe, 1988: Cloud feedback processes in a general circulation model. *J. Atmos. Sci.*, **45**, 1397–1416, [https://doi.org/10.1175/1520-0469\(1988\)045<1397:CFPIAG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1397:CFPIAG>2.0.CO;2).
- Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate*, **23**, 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. *J. Climate*, **19**, 6425–6432, <https://doi.org/10.1175/JCLI3988.1>.
- Zelinka, M. D., and D. L. Hartmann, 2012: Climate feedbacks and their implications for poleward energy flux changes in a warming climate. *J. Climate*, **25**, 608–624, <https://doi.org/10.1175/JCLI-D-11-00096.1>.
- , T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor, 2020: Causes of higher climate sensitivity in CMIP6 models. *Geophys. Res. Lett.*, **47**, e2019GL085782, <https://doi.org/10.1029/2019GL085782>.
- 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>.