

## Changes in Local and Global Climate Feedbacks in the Absence of Interactive Clouds: Southern Ocean–Climate Interactions in Two Intermediate-Complexity Models

PATRIK L. PFISTER<sup>a,b</sup> AND THOMAS F. STOCKER<sup>a,b</sup>

<sup>a</sup> *Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland*

<sup>b</sup> *Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland*

(Manuscript received 18 February 2020, in final form 6 October 2020)

**ABSTRACT:** The global-mean climate feedback quantifies how much the climate system will warm in response to a forcing such as increased CO<sub>2</sub> concentration. Under a constant forcing, this feedback becomes less negative (increasing) over time in comprehensive climate models, which has been attributed to increases in cloud and lapse-rate feedbacks. However, out of eight Earth system models of intermediate complexity (EMICs) not featuring interactive clouds, two also simulate such a feedback increase: Bern3D-LPX and LOVECLIM. Using these two models, we investigate the causes of the global-mean feedback increase in the absence of cloud feedbacks. In both models, the increase is predominantly driven by processes in the Southern Ocean region. In LOVECLIM, the global-mean increase is mainly due to a local longwave feedback increase in that region, which can be attributed to lapse-rate changes. It is enhanced by the slow atmospheric warming above the Southern Ocean, which is delayed due to regional ocean heat uptake. In Bern3D-LPX, this delayed regional warming is the main driver of the global-mean feedback increase. It acts on a near-constant local feedback pattern mainly determined by the sea ice–albedo feedback. The global-mean feedback increase is limited by the availability of sea ice: faster Southern Ocean sea ice melting due to either stronger forcing or higher equilibrium climate sensitivity (ECS) reduces the increase of the global mean feedback in Bern3D-LPX. In the highest-ECS simulation with 4 × CO<sub>2</sub> forcing, the feedback even becomes more negative (decreasing) over time. This reduced ice–albedo feedback due to sea ice depletion is a plausible mechanism for a decreasing feedback also in high-forcing simulations of other models.

**KEYWORDS:** Sea ice; Atmosphere–ocean interaction; Energy transport; Climate prediction; Climate models; Coupled models

### 1. Introduction

Energy balance models (EBMs) are invaluable diagnostic tools for assessing the behavior of more comprehensive climate models and their differences (e.g., Gregory et al. 2004; Winton et al. 2010; Geoffroy et al. 2013; Ceppi and Gregory 2019). The simplest EBM describes Earth's global mean energy budget (Gregory et al. 2004):

$$-\lambda \Delta T(t) = R_{\text{eff}} - N(t). \quad (1)$$

An effective radiative forcing  $R_{\text{eff}}$ , caused for example by an increased CO<sub>2</sub> concentration, causes a warming  $\Delta T(t)$  that is reduced and delayed by the ocean heat uptake (OHU)  $N(t)$ . There are physical and biogeochemical processes that modify an existing temperature perturbation both transiently and at equilibrium. These processes are referred to as climate feedbacks (IPCC 2013). Although these may be nonlinear with temperature, their overall effect is approximated by the linear feedback parameter  $\lambda < 0$  in the global EBM described by Eq. (1).

The global-mean behavior of many comprehensive climate models [general circulation models (GCMs)] can only be

reasonably emulated by Eq. (1) if a time dependence of the feedback is assumed [ $\lambda = \lambda(t)$ ]. Over time  $\lambda(t)$  becomes less negative in most GCMs (Geoffroy et al. 2013; Andrews et al. 2015; Gregory et al. 2015; Rugenstein et al. 2016a; Yoshimori et al. 2016; Rose and Rayborn 2016; Proistosescu and Huybers 2017; Sherwood et al. 2020; Dong et al. 2020), indicating an increasing climate sensitivity ( $= -1/\lambda$ ). This feedback time dependence has been attributed mainly to contributions of cloud and lapse-rate feedback changes (Rose et al. 2014; Andrews et al. 2015; Rose and Rayborn 2016; Zhou et al. 2016; Ceppi and Gregory 2017; Zhou et al. 2017; Andrews and Webb 2018; Dong et al. 2019).

In this study, we investigate the causes of the time dependence in the global mean feedback in two Earth system models of intermediate complexity (EMICs). These models were selected based on our earlier multimodel energy balance analysis (Pfister and Stocker 2018). Out of the 14 EMICs considered in that study, 8 lack interactive clouds (Eby et al. 2013). In only two out of these eight models,  $\lambda$  becomes substantially less negative with time (Pfister and Stocker 2018; Fig. S2 therein): in the Bern3D-LPX model (Ritz et al. 2011) and the LOVECLIM model (Goosse et al. 2010). As these models do not feature cloud feedbacks, and only one of them has a lapse-rate feedback (LOVECLIM), the dominant processes causing their global mean feedback time

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

Corresponding author: Patrik L. Pfister, patrik.pfister@climate.unibe.ch



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

DOI: 10.1175/JCLI-D-20-0113.1

dependence must differ from those found in GCMs. It is the purpose of this study to shed light on such additional processes in the framework of these two EMICs.

Using a local rather than a global diagnostic EBM (section 4), two kinds of processes could explain the global mean feedback strengthening: 1) changing warming patterns acting on constant local feedbacks (Armour et al. 2013) and 2) changing local feedbacks (e.g., Rose et al. 2014; Rose and Rayborn 2016; Rugenstein et al. 2016a; Ceppi and Gregory 2017). Both processes can operate also in the absence of interactive clouds, and it is one of our aims to investigate their relative importance in the two analyzed EMICs.

Both warming and feedback patterns interact with the pattern of OHU. This interaction can be quantified by introducing an additional EBM parameter, the OHU efficacy (Winton et al. 2010). Alternatively to a time-dependent feedback, a time-dependent OHU efficacy can also explain the time dependence of the global mean climate sensitivity. Although the two perspectives were shown to be equivalent (Haugstad et al. 2017), we will also analyze the OHU efficacy of our two EMICs and its contributions from changing feedback patterns and warming patterns following Rose et al. (2014).

This paper is structured as follows. Section 2 describes the two models, performed simulations, and analysis. Our results are presented in sections 3–7. We show the global mean feedback time dependence in these EMICs (section 3) and investigate its causes in a regional EBM framework (section 4), also testing the assumption of constant local feedbacks (Armour et al. 2013) (section 5). We discuss the OHU efficacy perspective (section 6) and finally investigate the influence of higher forcings and climate sensitivities on the feedback time dependence (section 7). We conclude with section 8.

## 2. Models and methods

### a. Model descriptions

We analyze two EMICs, the Bern3D-LPX model (Ritz et al. 2011; Roth et al. 2014) and the LOVECLIM model version 1.2 (Goosse et al. 2010). Both models feature a thermodynamic sea ice component and a dynamic land vegetation model. The atmospheric and oceanic components are more comprehensive in LOVECLIM than in Bern3D-LPX. Both ocean components are three-dimensional dynamical models. The Bern3D-LPX includes a frictional geostrophic ocean model, LOVECLIM has a primitive equation, free surface ocean model.

The main difference between the two models is their atmospheric component. LOVECLIM features a three-level quasigeostrophic dynamical atmosphere, while Bern3D-LPX has a one-layer energy and moisture balance model (EMBM) with prescribed seasonal wind fields.

A common feature of both atmospheric components is their lack of interactive clouds. Cloud cover is prescribed seasonally in both models, and layer-wise in LOVECLIM. Therefore, these two models do not simulate a cloud feedback, which is their main difference from GCMs with regard to feedback patterns.

The only spatially dependent feedbacks explicitly parameterized in Bern3D's EMBM are the Planck, water vapor, and albedo feedbacks. The top-of-atmosphere (TOA) longwave

radiation is parameterized as an empirical polynomial function of both local temperature and local relative humidity (Ritz et al. 2011), which determines the spatial pattern of the combined Planck and water vapor feedback. The sum of the feedbacks that are not explicitly resolved within the model (e.g., lapse-rate and cloud feedbacks) is parameterized globally as  $\lambda_{\text{EMBM}}\overline{\Delta T}$ , where  $\overline{\Delta T}$  is the global mean warming and  $\lambda_{\text{EMBM}} = -0.7 \text{ W m}^{-2} \text{ K}^{-1}$  is tuned to achieve an equilibrium climate sensitivity (ECS) of 3°C in the standard version of the model, which is near the current best estimate based on multiple lines of observational evidence (Sherwood et al. 2020).

LOVECLIM does not feature such a global feedback parameterization, but explicitly simulates the Planck, water vapor, and albedo feedbacks as well. In addition, it explicitly simulates the lapse-rate feedback due to its multiple atmospheric layers. LOVECLIM's ECS is estimated at 2.1°C (section 3).

Local feedbacks can change with time based on two general mechanisms. First, there may be local nonlinearities in the feedbacks, for example a higher-order temperature dependence of the Planck and water vapor feedbacks, or local albedo feedback changes due to vegetation shifts or sea ice loss. Such nonlinearities may be simulated by both Bern3D-LPX and LOVECLIM. Second, there may be nonlocal contributions due to local atmospheric changes, such as lapse-rate changes due to atmospheric heat transport in the absence of surface warming (Po-Chedley et al. 2018; Dong et al. 2019). Such effects can only be simulated in models with a dynamical atmosphere, like LOVECLIM.

In Bern3D-LPX, horizontal sensible heat transport is parameterized as purely diffusive, that is, proportional to the atmospheric temperature gradient  $\nabla T$  with a latitudinally varying heat diffusivity. Simulated horizontal temperature gradients are smaller than observed gradients (Ritz et al. 2011). Latent heat transport is parameterized as a combination of diffusion (based on a separate moisture diffusivity) and advection (based on a fixed wind climatology). In LOVECLIM, both sensible and latent heat fluxes are computed from estimates of temperature, humidity, surface characteristics, and surface winds using standard bulk formulas (Goosse et al. 2010).

### b. Experimental design and analysis

Our analysis encompasses three simulations for each model: Two perturbed simulations where the CO<sub>2</sub> concentration is abruptly doubled ( $2 \times \text{CO}_2$ ) or quadrupled ( $4 \times \text{CO}_2$ ) from the preindustrial concentration, and one unperturbed control simulation. Each simulation was run for 5000 years in Bern3D-LPX, and for 1000 years in LOVECLIM. For Bern3D-LPX, all simulations were carried out in three different model versions with the following ECS values: 3.0°C (standard version), 2.0°C (low-ECS version), and 6.0°C (high-ECS version). The model was tuned to these ECS values by varying the  $\lambda_{\text{EMBM}}$  parameter (see above).

Both models were analyzed analogously apart from three necessary differences. First, the calculation of anomalies for each quantity of interest between the abrupt simulations and the control simulation is done differently for the two models. We describe this by example of the temperature anomaly field (i.e., the warming pattern due to CO<sub>2</sub> forcing). For Bern3D-LPX, the temperature field of the 5000-yr control simulation is subtracted time step by time step from the temperature field of

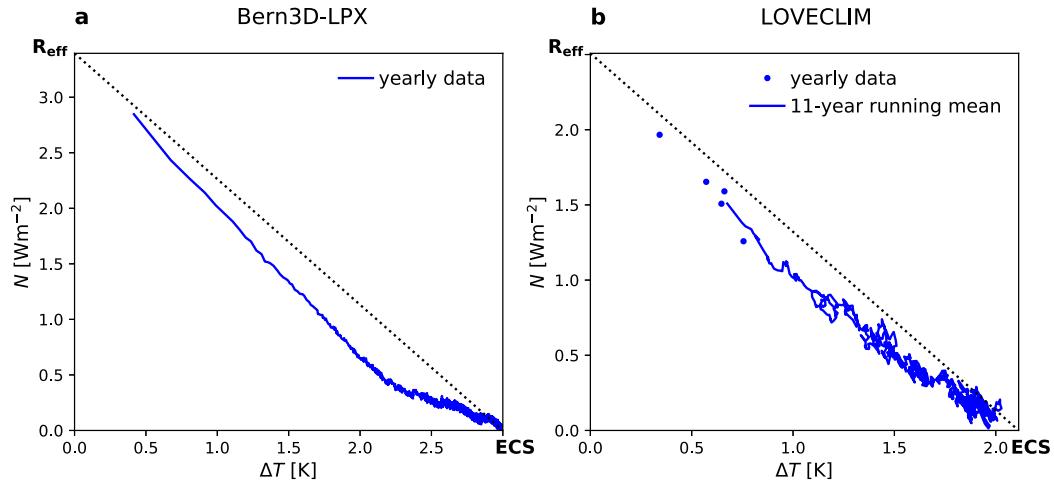


FIG. 1. (a),(b) Global mean heat uptake  $N$  plotted against global mean warming  $\Delta T$  in the  $2 \times \text{CO}_2$  simulations of both models. The time-varying slope of the model output corresponds to the feedback parameter  $\lambda$ . The quantities ECS and  $R_{\text{eff}}$  are explained in the text. Note that the continuous lines correspond to yearly data in (a), but to an 11-yr running mean in (b) where the yearly data of the first 5 years are shown by dots. This visualization is consistent with Fig. 9 and was chosen to minimize clutter in that figure. The dotted line indicates the hypothetical time-independent feedback.

the perturbed simulation. This is done to remove the artificially imposed variability that is equally present in all simulations, induced by the land vegetation model's 31-yr repeating climate pattern scaling (Stocker et al. 2013). For LOVECLIM, the mean temperature field of the 1000-yr control run is subtracted from the time steps of the perturbed simulation.

Second, as the LOVECLIM simulates a substantial interannual variability due to its dynamical atmosphere, a running mean was computed for all analyzed quantities and anomalies, to obtain the imprint of the idealized warming trend rather than short-term variations. A 31-yr running mean window was chosen for all visualizations except for Figs. 1 and 9, where an 11-yr window is used for better visibility of early changes. For Bern3D-LPX, which has little interannual variability, annual averages with no multiyear running mean are shown.

Finally, due to the different simulation lengths, equilibrium values are calculated differently in the two models. For Bern3D-LPX, the last simulation year (year 5000) is used to read out all equilibrium values. This is the true equilibrium, as equilibration takes roughly 4000 years (Fig. 7). For LOVECLIM, the last available 31-yr running mean (centered around year 985; i.e., the average over years 970–1000) is used to read out equilibrium patterns. As the model is not fully equilibrated after the simulation duration, the global mean warming has to be extrapolated to estimate ECS (section 3).

### 3. Time-dependent global feedback

In this section, we illustrate the diagnosis of parameters for the global mean EBM given by Eq. (1), and show that the global mean feedback becomes less negative over time in the two analyzed EMICs. We only analyze the  $2 \times \text{CO}_2$  simulations in this section and the subsequent sections, and compare them to the  $4 \times \text{CO}_2$  simulations in section 7.

Figure 1 shows the global mean heat uptake  $N$  as a function of the global mean warming  $\Delta T$  in both models examined here. This relation yields two important diagnostic quantities. First,  $R_{\text{eff}}$  is the global mean heat imbalance at the onset of warming, that is, after fast radiative adjustments that do not affect  $\Delta T$  (e.g., Gregory et al. 2004). Following Andrews et al. (2015),  $R_{\text{eff}}$  is diagnosed for both models as the  $N$  intercept (i.e.,  $\Delta T = 0$ ) of the linear regression of simulation years 1–20. Second, the ECS is the equilibrium warming caused by a doubling of the  $\text{CO}_2$  concentration (e.g., Sherwood et al. 2020).

For Bern3D-LPX, ECS =  $3.0^\circ\text{C}$  is simply obtained as  $\Delta T$  at the end of the 5000-yr simulation. For LOVECLIM, an equilibrium is not fully reached at the end of the 1000-yr simulation, as indicated by a slight remaining heat imbalance ( $N > 0$ ). Therefore, ECS is diagnosed as the  $\Delta T$  intercept (i.e.,  $N = 0$ ) of the linear regression of simulation years 300–1000. This period is chosen because the regression slope is approximately constant after year 300. The resulting ECS =  $2.1^\circ\text{C}$  is only  $0.1^\circ\text{C}$  higher than the 1000-yr warming, because the remaining heat imbalance is small.

It is evident from Fig. 1 that the global mean feedback is not constant in either model [ $\lambda = \lambda(t)$ ]; otherwise, the model output should lie on the dotted line connecting  $R_{\text{eff}}$  and ECS [Eq. (1)]. In both models,  $\lambda$  becomes less negative over time, which means that climate sensitivity increases with time, in agreement with GCMs (see the introduction). In sections 4 and 5, we explore the characteristics and causes of this time dependence in these two models.

### 4. Time-dependent warming patterns and regional feedbacks

#### a. Local energy balance framework

The physical causes of the global mean time dependence can be investigated by generalizing Eq. (1) to a localized EBM

(Crook et al. 2011; Armour et al. 2013; Feldl and Roe 2013a,b; Rose et al. 2014; Rose and Rayborn 2016):

$$-\lambda(\mathbf{r}, t) \Delta T(\mathbf{r}, t) = R_{\text{eff}} - \Delta N_{\text{TOA}}(\mathbf{r}, t). \quad (2)$$

Here  $\mathbf{r}$  denotes the location on the sphere (longitude and latitude);  $\Delta T(\mathbf{r}, t)$  is the local warming at time  $t$  with reference to the preindustrial state;  $\Delta N_{\text{TOA}}(\mathbf{r}, t)$  is the local anomaly in the TOA energy imbalance (i.e., the difference from the preindustrial TOA energy imbalance pattern). Note that this EBM is equivalent to the one described by Rose et al. (2014), considering that  $\Delta N_{\text{TOA}}(\mathbf{r}, t) = \Delta N_{\text{OHU}}(\mathbf{r}, t) + \Delta(\nabla \cdot \mathbf{F}_{\text{atm}})(\mathbf{r}, t)$ , where  $\Delta N_{\text{OHU}}(\mathbf{r}, t)$  is the bottom-of-atmosphere heat uptake anomaly (dominated by OHU) and  $\Delta(\nabla \cdot \mathbf{F}_{\text{atm}})(\mathbf{r}, t)$  is the atmospheric heat flux divergence anomaly.

In the simulations analyzed here,  $R_{\text{eff}}$  is constant over time. In Bern3D-LPX, it is also globally uniform, as a global-mean forcing following Myhre et al. (1998) is prescribed. In LOVECLIM, the forcing is based on a Green's function (Goosse et al. 2010) and is spatially varying, that is,  $R_{\text{eff}}(\mathbf{r})$ . Unfortunately, LOVECLIM's  $\text{CO}_2$  forcing pattern has not been published, and a fixed-SST simulation to obtain it was not available to us. Therefore, we diagnose LOVECLIM's feedback assuming a globally uniform  $R_{\text{eff}}$ . To assess the influence a spatially varying feedback would have on the global feedback, we use a zonal pattern from a GCM, as described in section 4c.

To analyze the feedback pattern at a given time  $t$ , we calculate  $\lambda(\mathbf{r}, t)$  using Eq. (2), given the global mean  $R_{\text{eff}}$  (section 2b) as well as  $\Delta T(\mathbf{r}, t)$  and  $\Delta N_{\text{TOA}}(\mathbf{r}, t)$  from the model output. The separate longwave and shortwave feedbacks,  $\lambda_{\text{LW}}$  and  $\lambda_{\text{SW}}$ , are calculated similarly.  $\Delta N_{\text{TOA}}$  is replaced by the contributions  $\Delta N_{\text{LW,TOA}}$  and  $\Delta N_{\text{SW,TOA}}$ , respectively. The longwave radiative forcing  $R_{\text{eff}}$  is accounted for only in the  $\lambda_{\text{LW}}$  calculation.

### b. Analysis and comparison of Bern3D-LPX and LOVECLIM

Figures 2 and 3 show the three patterns  $\lambda(\mathbf{r}, t)$ ,  $\Delta T(\mathbf{r}, t)$ , and  $\Delta N_{\text{TOA}}(\mathbf{r}, t)$  relevant for Eq. (2) for Bern3D-LPX and LOVECLIM, respectively. In addition, the longwave and shortwave feedback patterns  $\lambda_{\text{LW}}(\mathbf{r}, t)$  and  $\lambda_{\text{SW}}(\mathbf{r}, t)$  are also shown. Each pattern is given at the end of the simulation (equilibrium or near-equilibrium; first column) and at year 100 (transient; second column). Their difference is shown in the third column (equilibrium minus transient). For  $\Delta N_{\text{TOA}}(\mathbf{r}, t)$ , only the difference is shown in Figs. 2o and 3o, while the first two columns show the corresponding sea ice extent (Figs. 2m,n and 3m,n).

Figure 4 shows the zonal averages for the same quantities, but with an additional panel row for OHU. Four different time steps up to year 1000 are shown for both models, and for Bern3D-LPX also the equilibrium (year 5000) is shown. In the following, we describe effects seen in Figs. 2–4 simultaneously, referencing the panels that best visualize these effects.

Both models simulate a strong equilibrium warming over the Southern Ocean (Figs. 2b and 3b), which develops more slowly than in the rest of the world (Figs. 4a,b). This delay in Southern Ocean warming is due to the strong OHU in the Southern Ocean (Figs. 4k,l). It is in agreement with observations

(Armour et al. 2016) and GCM simulations (Winton et al. 2010; Armour et al. 2013; Andrews et al. 2015; Proistosescu and Huybers 2017; Rugenstein et al. 2020; Dong et al. 2020).

The changes in local feedbacks  $\lambda(\mathbf{r}, t)$  are also strongest over the Southern Ocean (Figs. 2f and 3f). It is evident that these changes are dominated by shortwave effects in Bern3D-LPX (Fig. 2l) and by longwave effects in LOVECLIM (Fig. 3i), as detailed below.

In Bern3D-LPX, the feedback pattern is mainly determined by the shortwave feedback pattern, which in turn reflects the regions of sea ice melting. In equilibrium, the sea ice cover in the Southern Ocean is almost entirely melted away (Fig. 2n). Therefore, the feedback is strongest near the Antarctic coast (Figs. 2e,k) where annual-mean preindustrial sea ice cover was largest. In contrast, 100 years after the abrupt  $\text{CO}_2$  forcing, sea ice melting has not yet progressed to the Ross and Weddell Seas, but is strongest near the border of the Southern Ocean sea ice extent (Fig. 2m). There, the feedback is much stronger than in equilibrium (Figs. 2d,j), because temperature continues to rise toward the equilibrium whereas sea ice is already virtually gone after 100 years.

The Bern3D-LPX longwave feedback pattern is much more uniform than the shortwave pattern (Figs. 2g,h) and is determined by the model's empirically based longwave parameterization that combines the Planck and water vapor feedbacks. Local changes in  $\lambda_{\text{LW}}$  amount to  $0.3 \text{ W m}^{-2} \text{ K}^{-1}$  or less, and mostly follow a zonal structure (Fig. 2i). In the lower latitudes, between roughly  $40^\circ\text{S}$  and  $60^\circ\text{N}$ , the longwave feedback becomes more negative toward equilibrium. This is what is expected from the Planck feedback: as Earth warms, more longwave radiation is lost to space. This is reflected in a negative  $\Delta N_{\text{TOA}}$  apart from the sea ice regions (Fig. 2o). As this increase in outgoing radiation is strongest in the high latitudes because the warming is most pronounced there, we also expect the Planck feedback to become more negative in this region. Oppositely, the total longwave feedback becomes less negative in the high latitudes (Fig. 2i), which must thus be due to the water vapor feedback. Relative humidity increases over the Arctic Ocean and (more pronouncedly) over Antarctica, while decreasing over midlatitude continents (not shown). These changes are in line with the  $\lambda_{\text{LW}}$  pattern changes, and apparently dominate the temperature-induced (Planck) feedback changes in the high latitudes.

In LOVECLIM, the longwave feedback pattern is much less uniform (Figs. 3g,h) and defines the total feedback pattern in combination with local shortwave feedback maxima in sea ice regions, as well as some continental regions (Figs. 3j,k). The strongest shortwave feedback peaks suggest snow and ice loss in the Himalayas and vegetation losses or shifts in tropical Africa and Australia, but this is not investigated further here. Apart from local shortwave feedback changes in these regions and over Antarctic sea ice (Fig. 3l), the changing local feedback pattern is dominated by changes in the longwave feedback (Fig. 3i). Most notable is a very strongly negative feedback over the Southern Ocean (mainly in the Pacific sector), which gets less negative toward equilibrium (Figs. 3f,i and 4d,f). This is at least partially due to a lapse-rate effect, as investigated in section 4c. In a small region in the Pacific sector of the Southern

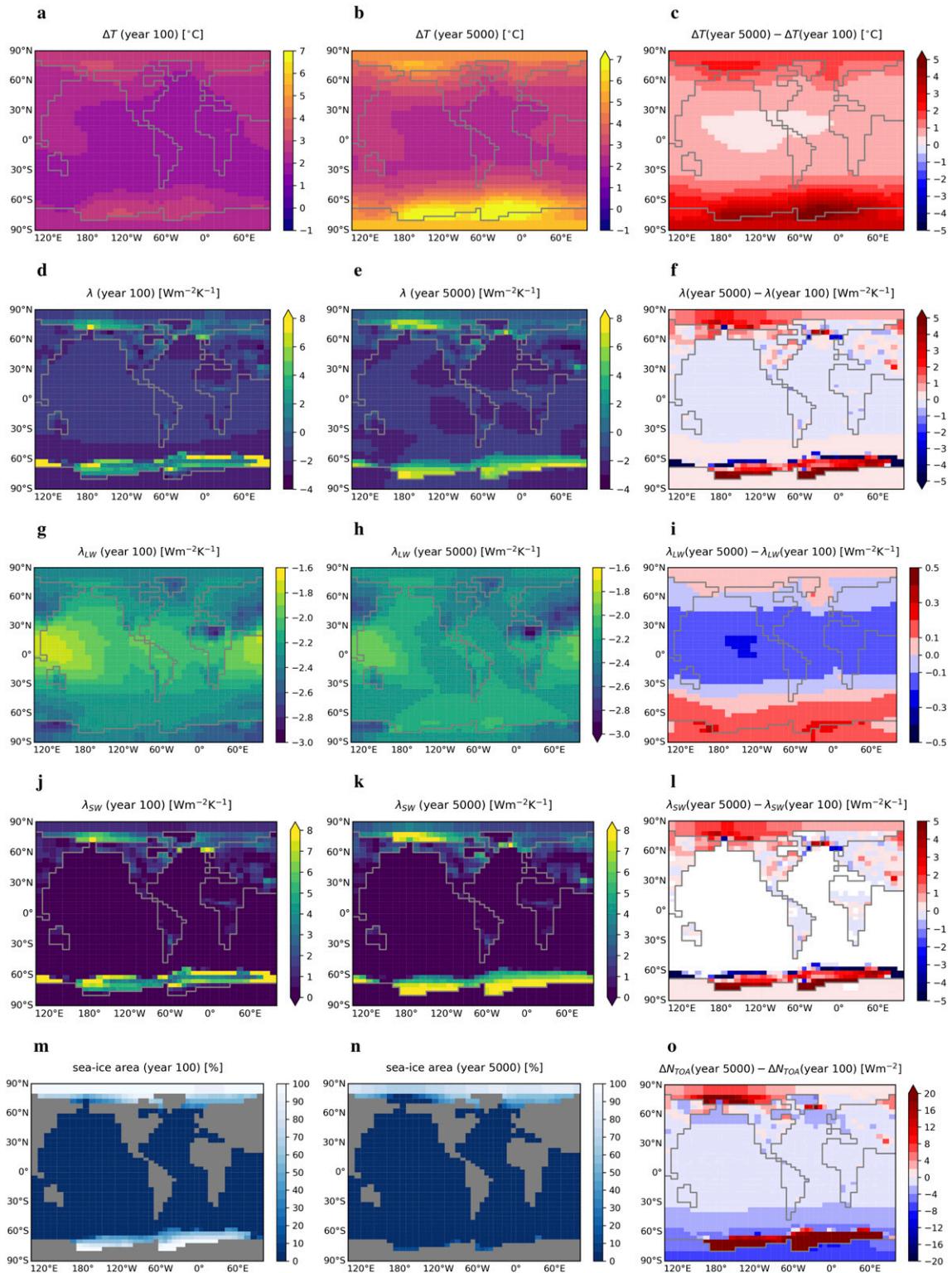


FIG. 2. (a)–(c) Local warming, (d)–(i) feedbacks (total, longwave, and shortwave), (m), (n) sea ice, and (o) TOA radiation change in the  $2 \times \text{CO}_2$  simulation of the Bern3D-LPX model. (left) Transient patterns (year 100), (center) equilibrium patterns (year 5000), and (right) their difference (equilibrium – transient). Note that the color range in (i) is much smaller than in (f) and (l). In white regions in (l), changes are smaller than  $0.01 \text{ W m}^{-2} \text{ K}^{-1}$ .

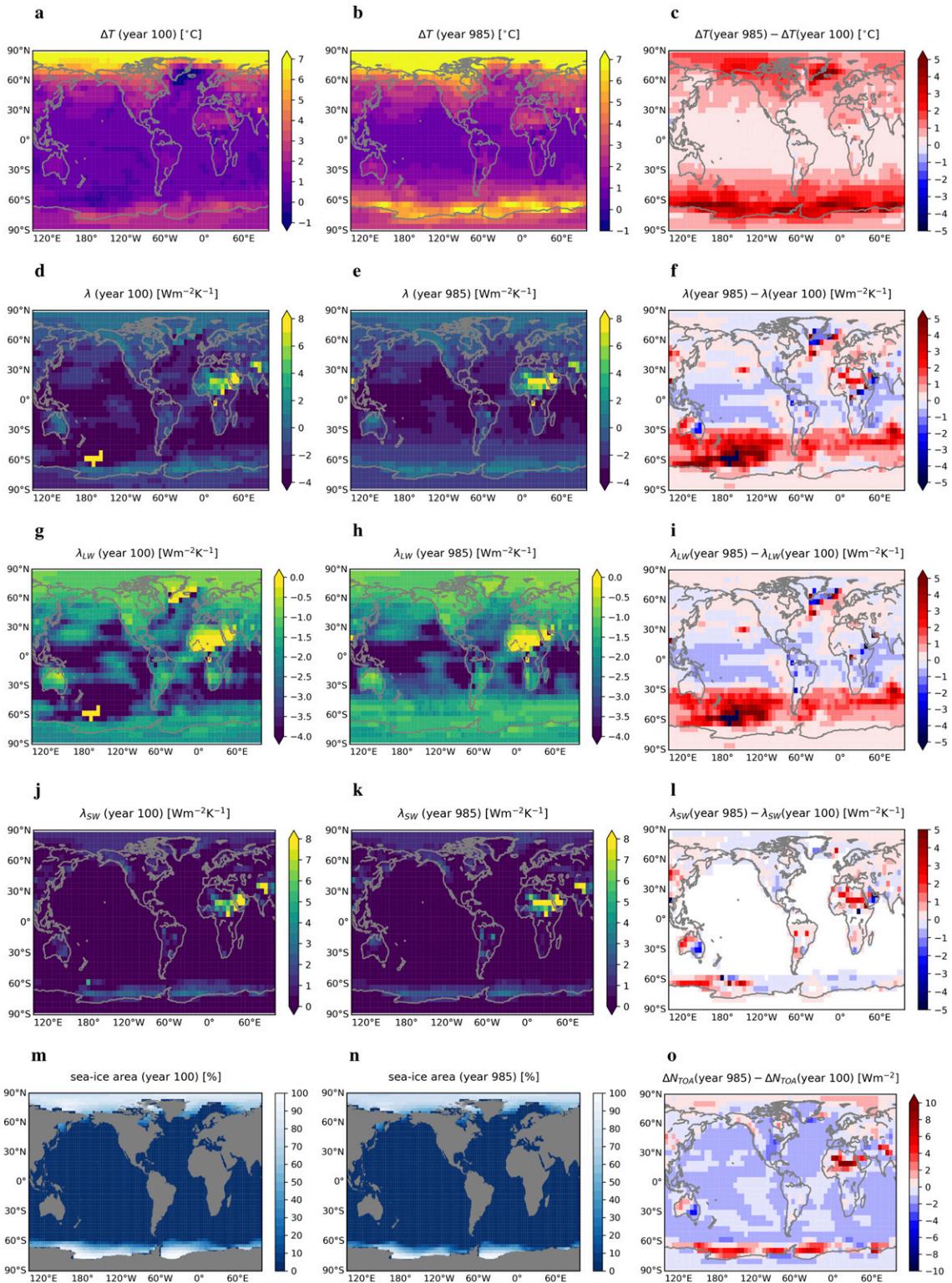


FIG. 3. As Fig. 2, but for the LOVECLIM model. Here, the center column is the last available 31-yr running mean (years 970–1000), which is used as the forced equilibrium reference for LOVECLIM (see text). Note that the color ranges in (g), (h), (i), and (o) differ from the corresponding Bern3D-LPX panels.

Ocean, the feedback becomes strongly positive instead (yellow in Fig. 3d), corresponding to a slight local cooling (see also section 4c). However, this is unimportant for the zonal or regional mean feedback (Figs. 4d,f).

In both models, the initial atmospheric warming has a minimum at roughly 40°–60°S (Figs. 4a,b), which has been simulated in many models of different model generations (Kushner et al. 2001; Andrews et al. 2015). This minimum is stronger in LOVECLIM, due in part to its negative local feedback (section 4c) and in part to its stronger local OHU (Figs. 4k,l), which is enhanced by the dynamical atmosphere. As consistently simulated in models with dynamical atmospheres (Fyfe and Saenko 2006; Russell et al. 2006; Sen Gupta et al. 2009), the wind stress over the Southern Ocean increases due to the warming (by about 13% averaged over 40°–80°S), enhancing the Ekman circulation. This delays the increase of the sea surface temperature while the atmosphere above is warming faster, leading to strong local OHU (e.g., Armour et al. 2016). In turn, this strong OHU delays the local atmospheric warming  $\Delta T(\mathbf{r}, t)$ . On another note, the temperature gradients in Bern3D-LPX are generally smaller, suggesting that the diffusive heat transport parameterization smooths out temperature gradients faster than LOVECLIM's dynamical atmosphere.

One of the striking differences between the two models' warming patterns is the polar amplification in the Northern Hemisphere (Figs. 2a, 3a, and 4a,b). This amplification is stronger in LOVECLIM than in any other EMIC from the EMIC-AR5 intercomparison, while the amplification in Bern3D-LPX is among the lower half of those EMICs (Eby et al. 2013, their Fig. 4). Most of the northern high-latitude warming of LOVECLIM occurs already during the first 20–150 simulation years (Fig. 4b). It may thus contribute to the initial time dependence in the global mean feedback of LOVECLIM (Fig. 7a, roughly up to year 150). But the long-term feedback time dependence up to year 1000, which is the main focus of this paper, is mainly explained by the aforementioned slower warming in the southern high latitudes. Therefore, we do not further investigate the northern polar amplification here.

Sea ice melting is stronger in the Bern3D-LPX than in LOVECLIM, although the surface warming is weaker in the northern high latitudes and comparable in the southern high latitudes. This is because the high-latitude temperatures of the preindustrial control simulation are much lower for LOVECLIM than for Bern3D-LPX, and therefore not as close to the sea ice melting point. Compared to the high-latitude temperatures from present-day observations, the Bern3D-LPX and LOVECLIM control simulations have a low and high bias, respectively. Averaged over the latitudes  $\sim 60^\circ$ – $80^\circ$ S, where sea ice is located, the bias amounts to  $+2.7^\circ\text{C}$  for Bern3D-LPX and  $-3.3^\circ\text{C}$  for LOVECLIM. This comparison is not entirely accurate, as the control simulations correspond to preindustrial conditions, while we averaged years 1958–2001 of the ERA-40 reanalysis data (Uppala et al. 2005) for the observational reference. The positive (negative) bias of Bern3D-LPX/LOVECLIM would therefore be somewhat larger (smaller), but this correction may be minor because the Southern Ocean has warmed less than the global mean over the observational period (Armour et al. 2016).

For a more direct comparison with LOVECLIM, we have also analyzed the warming and radiative patterns in Bern3D-

LPX in year 1000 (gray lines in Fig. 4) instead of year 5000 (black lines). The patterns are qualitatively similar; the statements made above would therefore also hold if year 1000 were used as equilibrium reference, as in LOVECLIM. Nevertheless, we note some differences: About 20% of the preindustrial Southern Ocean sea ice cover remains until year 1000 (not shown), which vanishes completely until year 5000. Consequently, the Southern Ocean warming is further amplified compared to the rest of the world (Fig. 4a), and the albedo-induced peak in  $\Delta N_{\text{TOA}}(\mathbf{r}, t)$  and  $\lambda(\mathbf{r}, t)$  extends to the highest latitudes bordering on Antarctica (Figs. 4g,i). Furthermore, the enhanced OHU over the Southern Ocean diminishes toward equilibrium. For LOVECLIM, it is expected that some more sea ice would melt until equilibrium, but a complete meltdown is highly unlikely as the extrapolated additional global warming from year 1000 to equilibrium is only  $0.1^\circ\text{C}$ .

### c. Local longwave feedback changes in LOVECLIM

In this subsection, we take a closer look at the aforementioned longwave feedback time dependence in LOVECLIM (Fig. 3i). We investigate two questions: First, is the atmospheric temperature evolution compatible with an increasing lapse-rate feedback, which could account for the longwave feedback becoming less negative? Note that the global-mean lapse-rate feedback can be positive or negative; we refer to an “increasing” lapse-rate feedback when the lapse-rate feedback changes contribute to a less-negative total longwave feedback. Second, how is the longwave feedback pattern affected by our assumption of a uniform forcing  $R_{\text{eff}}$ ?

Greenhouse gas forcings do not only increase surface temperatures globally, but also alter the vertical temperature gradients in the atmosphere. The change in vertical temperature with height is referred to as the lapse rate. As the lapse rate changes, the longwave radiation to space is different than what we would expect if looking only at surface temperatures. For example, if the upper atmosphere warms faster than the surface, more radiation is lost to space—this is a negative feedback to global warming, referred to as a negative lapse-rate feedback. This feedback is nonlocal in nature: For example, atmospheric heat transport propagates tropical surface warming toward the poles, causing increased midtropospheric temperatures in the high latitudes. This increases outgoing longwave radiation in that region, which affects the local feedback magnitude even in the absence of local surface warming (Po-Chedley et al. 2018; Dong et al. 2019).

To investigate this effect in LOVECLIM, we look at the deviation from vertically uniform warming, normalized by global mean surface warming:

$$\frac{\Delta T(z, \mathbf{r}, t) - \Delta T_{\text{surface}}(\mathbf{r}, t)}{\Delta T_{\text{surface}}(t)}. \quad (3)$$

The zonal mean of this quantity is shown in Fig. 5, for the running means around year 100 (transient; Fig. 5a) and year 985 (near-equilibrium; Fig. 5b). Regions where this quantity is positive (negative) contribute to a negative (positive) lapse-rate feedback, as described in the previous paragraph. We cannot provide a quantitative estimate for the lapse-rate

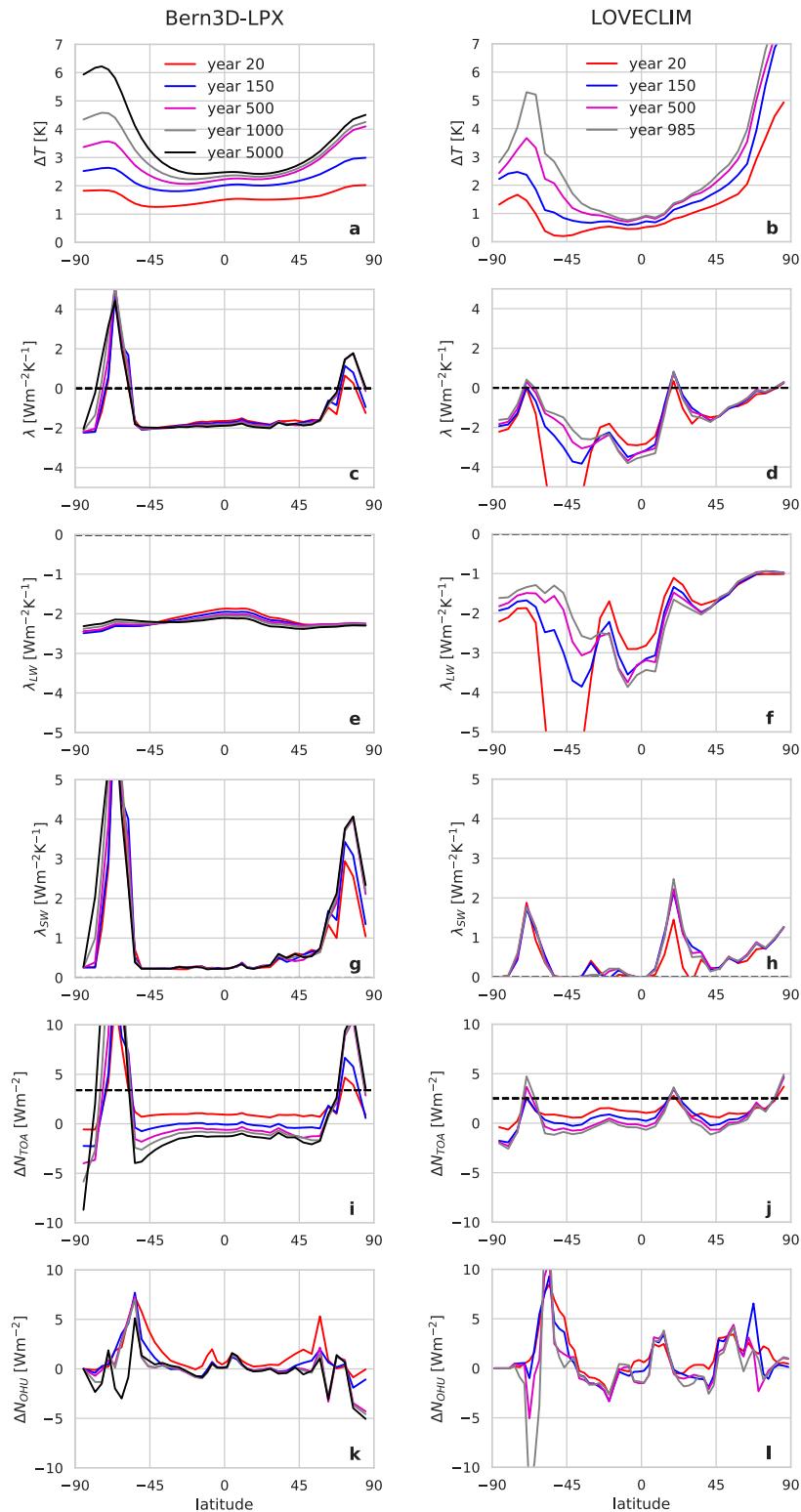


FIG. 4. Zonal averages of warming, feedbacks, TOA radiation change, and OHU change in the  $2 \times \text{CO}_2$  simulation of the (left) Bern3D-LPX and (right) LOVECLIM models. Colors denote different simulation years (see legend).

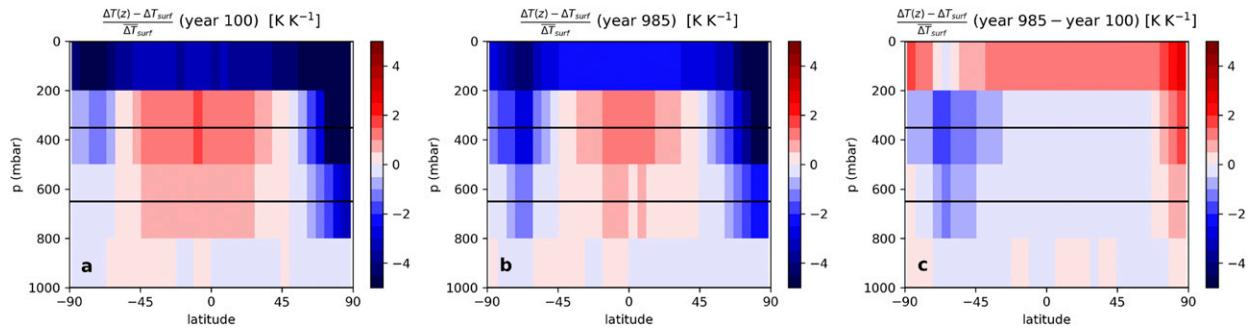


FIG. 5. Zonal-mean profiles of the deviation from vertically uniform warming, normalized by global-mean surface temperature change [Eq. (3)], in the  $2 \times \text{CO}_2$  simulation of LOVECLIM: (a) transient, (b) near-equilibrium, and (c) the difference [(b) - (a)] (as in Fig. 3). Black lines show LOVECLIM's four temperature levels at 0, 350, 650, and 1000 mb. Negative (positive) regions in (c) contribute to an increasing (decreasing) lapse-rate feedback.

feedback for LOVECLIM. Such an estimate could be obtained based on the radiative kernel method (Soden et al. 2008), by multiplying the data shown in Fig. 5 by a radiative kernel matrix, which quantifies the change in radiation per change in temperature. However, radiative kernels are not available for LOVECLIM, and interpolating radiative kernels from high-resolution models onto LOVECLIM's four atmospheric layers may lead to an erroneous quantification of the global mean feedback. Therefore, we do not further decompose the longwave feedback into Planck, lapse rate, and water vapor contributions, but discuss the simulated lapse-rate processes qualitatively in the following.

LOVECLIM's zonal mean profile of the deviation from vertically uniform warming at year 100 (Fig. 5a) is in broad agreement with more comprehensive models (Andrews and Webb 2018; Po-Chedley et al. 2018; Dong et al. 2019), showing enhanced upper tropospheric warming from the equator up to around  $60^\circ\text{S}$  and  $50^\circ\text{N}$ . At higher latitudes, the upper troposphere warms less than the surface. The transition to the colder highest layer (spanning from 200 to 0 mb; 1 mb = 1 hPa) is abrupt, because this layer represents the stratospheric layer at LOVECLIM (close to 0 mb).

Toward equilibrium, the region of enhanced tropospheric warming diminishes. Most notably, the difference between equilibrium and transient warming deviation (Fig. 5c) is negative in the mid- to upper troposphere between roughly  $30^\circ$  and  $85^\circ\text{S}$ , with a minimum around  $60^\circ\text{S}$ . This region thus contributes to an increasing lapse-rate feedback, and coincides with the region where the total longwave feedback becomes substantially less negative (Fig. 3i). This is also agreement with more comprehensive models, which show an increasing lapse-rate feedback in the same region in the multimodel mean (Ceppi and Gregory 2017). We conclude that the increasing lapse-rate feedback is a plausible mechanism to explain, or at least contribute to, LOVECLIM's longwave feedback time dependence above the Southern Ocean.

In the stratospheric layer and in the midtroposphere over the northern high latitudes, Fig. 5c shows opposite lapse-rate changes that contribute to a decrease in lapse-rate feedback. Although the radiative impact of the stratospheric layer is limited by its lower mass compared to the troposphere, it may

thus contribute to the longwave feedback becoming more negative in the lower latitudes (Fig. 3i). However, no such change in total longwave feedback is diagnosed in the northern high latitudes. The weak influence of the lapse-rate changes on the local longwave feedback in this region is in line with the spatial pattern of the radiative kernel (Soden et al. 2008), which shows a weaker radiative response to temperature changes in the highest latitudes, compared to the midlatitudes. Nevertheless, this may also indicate that there are other counteracting longwave feedbacks (Planck or water vapor). On the other hand, the lack of local longwave feedback change could also be a consequence of our assumption of a globally uniform  $R_{\text{eff}}$ . This is investigated in the following.

The  $\text{CO}_2$  radiative forcing is not globally uniform in LOVECLIM (Goosse et al. 2010) or in comprehensive models (e.g., Huang et al. 2017; Nalam et al. 2018), but LOVECLIM's forcing pattern is not available to us (section 2). As a substitute, we use the zonal mean forcing pattern from a  $2 \times \text{CO}_2$  simulation of a GCM, the Community Earth System Model (CESM) version 1.2 (Hurrell et al. 2013). We have obtained this zonal mean feedback by reading Fig. 1a) of Huang et al. (2017), and approximating their values by a forcing function that is piecewise linear with latitude  $\varphi$ . We have scaled this to match LOVECLIM's global mean  $R_{\text{eff}}$ , yielding the zonal mean forcing pattern  $R_{\text{scaled}}(\varphi)$  shown in Fig. 6a.

To assess the influence of a nonuniform forcing on the feedback pattern, we have diagnosed LOVECLIM's longwave feedback pattern again, using  $R_{\text{scaled}}(\varphi)$  instead of  $R_{\text{eff}}$  in Eq. (2) (Figs. 6b–e). The difference between this feedback pattern  $\lambda(R_{\text{scaled}})(\mathbf{r}, t)$  and the feedback pattern with uniform forcing  $\lambda(\mathbf{r}, t)$  is not purely zonal, as it is scaled by  $1/[\Delta T(\mathbf{r}, t)]$  (Fig. 6c). The shortwave feedback is not affected by the forcing pattern.

The longwave feedback change from transient to equilibrium looks very similar when diagnosed with  $R_{\text{scaled}}$  (Fig. 6d) as with  $R_{\text{eff}}$  (Fig. 3i), with a few notable differences (Fig. 6e). First, the feedback change over the Arctic Ocean is now slightly negative, compatible with the lapse-rate feedback decrease expected from Fig. 5c. We cannot compare the magnitude of the decrease, as explained above. Second, the diagnosed feedback changes with  $R_{\text{scaled}}$  are less pronounced above the

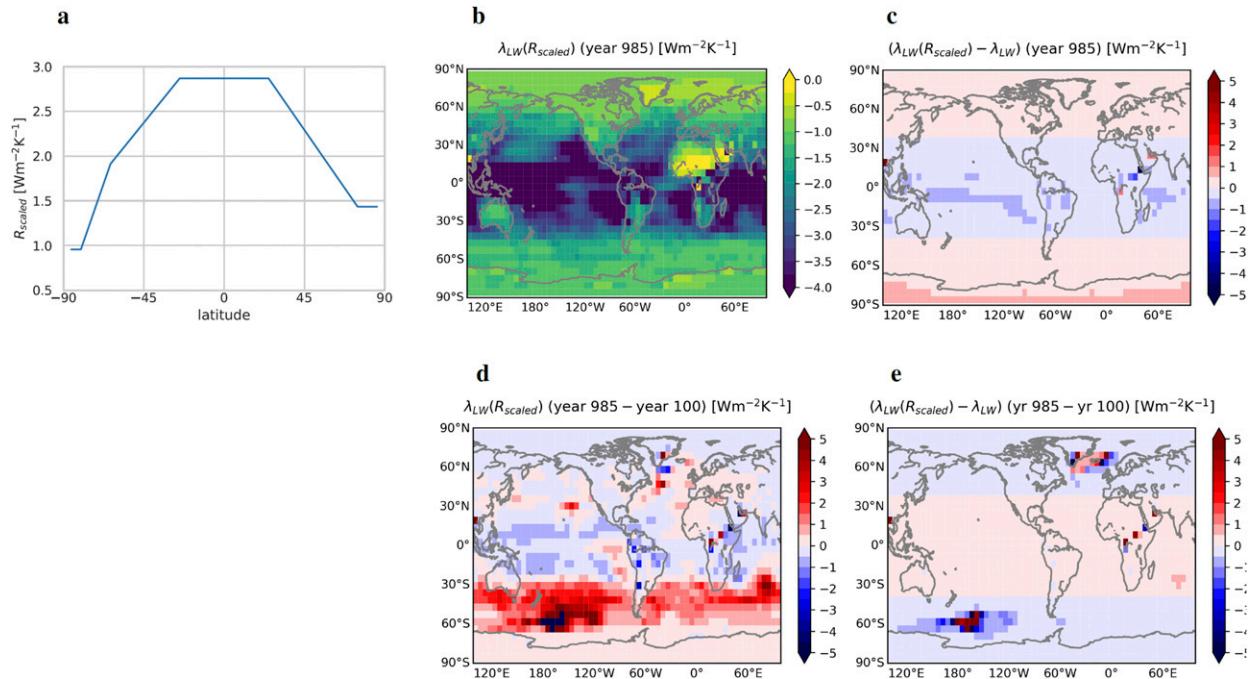


FIG. 6. Influence of a diagnostic zonal forcing profile on LOVECLIM's longwave feedback pattern. (a) Diagnostic forcing  $R_{\text{scaled}}$ , obtained from scaling LOVECLIM's global-mean  $R_{\text{eff}}$  by the zonal  $\text{CO}_2$  forcing pattern from a GCM (see text). (b) Near-equilibrium longwave feedback pattern diagnosed using  $R_{\text{scaled}}$ . (c) Difference between (b) and the pattern diagnosed using uniform  $R_{\text{eff}}$ . (d) As in Fig. 3i, but using  $R_{\text{scaled}}$  instead of  $R_{\text{eff}}$ . (e) Difference in the evolution of the longwave feedback pattern when  $R_{\text{scaled}}$  is used instead of  $R_{\text{eff}}$  [i.e., (d) minus Fig. 3i].

North Atlantic and in the most strongly changing region above the Southern Ocean. This suggests that these strong changes, including the transient positive feedback anomaly in the region of slight transient cooling (section 4b), may be overestimated by the assumption of a uniform forcing. However, the finding that the overall feedback above the Southern Ocean becomes less negative is robust.

### 5. Testing the approximation of constant local feedbacks

We have shown that the local feedback patterns of Bern3D-LPX and LOVECLIM are time dependent. How does this time dependence affect the global mean feedback? In this section, we test how well the time dependence of the global feedback is captured by the approximation of constant local feedbacks suggested by Armour et al. (2013).

By definition, the global mean feedback  $\lambda(t)$  can be calculated from the local feedbacks using a warming-weighted averaging (Armour et al. 2013):

$$\lambda(t) = \overline{\lambda(\mathbf{r}, t) \frac{\Delta T(\mathbf{r}, t)}{\Delta T(t)}}, \quad (4)$$

where overbars denote global averages. If most of the time dependence of  $\lambda(t)$  stems from the warming pattern  $\Delta T(\mathbf{r}, t)$  rather than changes in the feedback pattern  $\lambda(\mathbf{r}, t)$ , it may be reasonable to make the approximation of constant local feedbacks  $\lambda(\mathbf{r}, t) \approx \lambda_{\text{eq}}(\mathbf{r})$ :

$$\lambda(t) \approx \overline{\lambda_{\text{eq}}(\mathbf{r}) \frac{\Delta T(\mathbf{r}, t)}{\Delta T(t)}}. \quad (5)$$

Figure 7a shows the global  $\lambda(t)$  time series for both models. Solid lines are calculated from the exact Eq. (4), dashed lines from its approximation given by Eq. (5). The magnitude of the global-mean feedback change is roughly comparable among the two models. However, the constant local feedback approximation closely captures the feedback time dependence for Bern3D-LPX, but not at all for LOVECLIM.

The global-mean longwave and shortwave feedbacks (Figs. 7b,c) confirm that LOVECLIM's feedback time dependence is mainly due to longwave effects. This can be attributed to an increasing lapse-rate feedback in the Southern Ocean region (section 4c). As described there, this feedback is nonlocal in nature, therefore the assumption of constant local feedbacks does not capture the feedback time dependence in the global mean. On a side note, our assumption of a globally uniform forcing pattern (section 4c) does not bias the global-mean feedback estimate.

Increasing shortwave feedbacks contribute to LOVECLIM's feedback time dependence only in the first 50 years, due to rapid initial melting and continental albedo shifts in response to the abrupt  $\text{CO}_2$  forcing. The initial melting is stronger in the Northern Hemisphere (NH; dashed line in Fig. 7d) than in the Southern Hemisphere (SH; solid line), due to the rapid warming over the Arctic Ocean (section 4b). However, like in Bern3D-LPX, the long-term melting is stronger in the Southern Hemisphere.

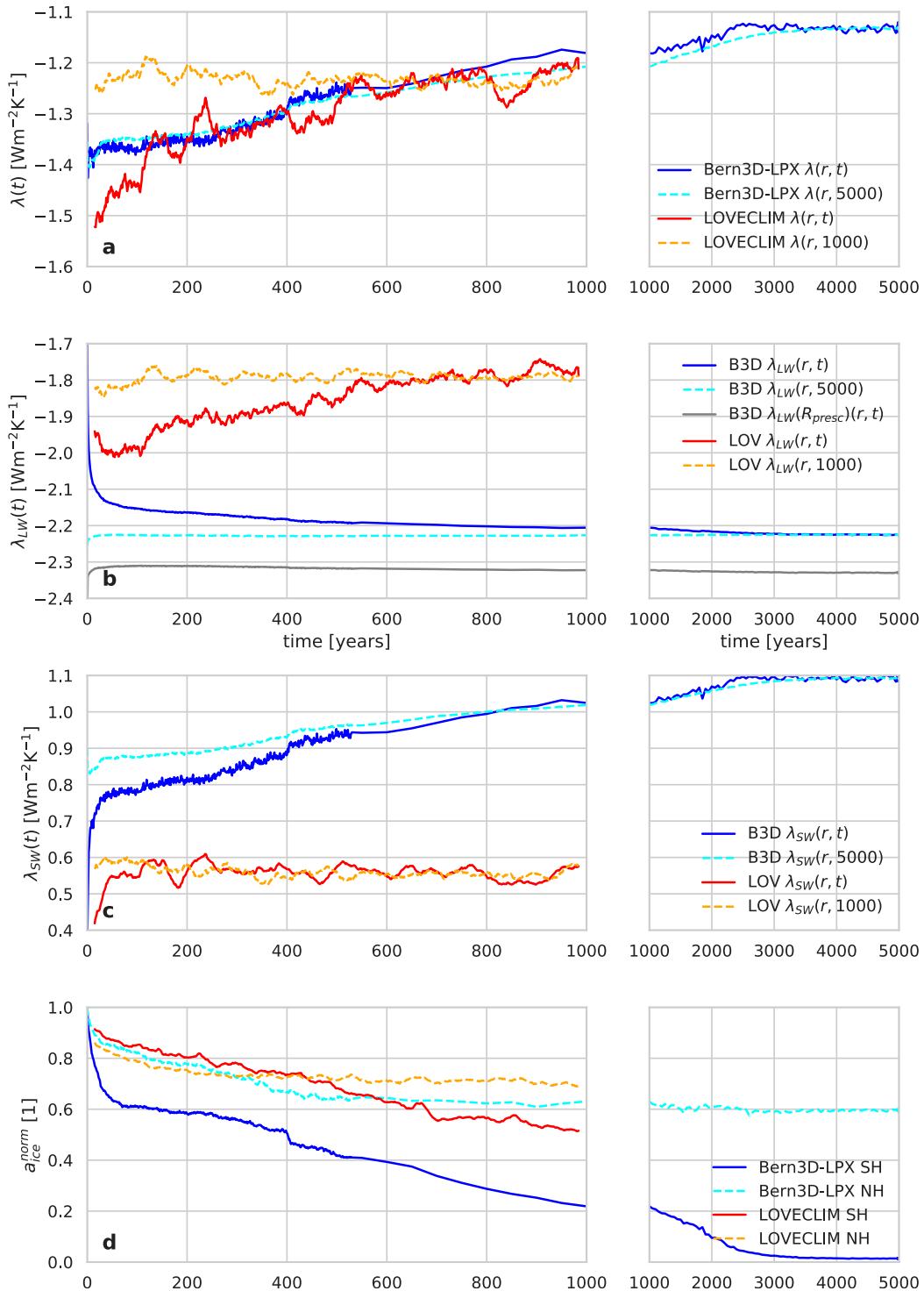


FIG. 7. Global mean feedbacks: (a) net, (b) longwave, and (c) shortwave, as well as (d) normalized hemispheric sea ice extent, in the  $2 \times \text{CO}_2$  simulation of Bern3D-LPX (blue/cyan) and LOVECLIM (red/orange). In (a)–(c), solid lines are calculated from the exact Eq. (4) and dashed lines from the approximation of constant local feedbacks, Eq. (5). The gray line in (c) is calculated with Bern3D-LPX’s prescribed feedback  $R_{\text{presc}}$  instead of  $R_{\text{eff}}$  (see text). In (d),  $a_{\text{ice}}^{\text{norm}}$  is the remaining fraction of the preindustrial sea ice extent in the Southern (solid) and Northern (dashed) Hemisphere.

In Bern3D-LPX, we have inferred from the feedback pattern (Fig. 2f) that the feedback changes are dominated by shortwave (albedo) feedback changes. Also in the global-mean shortwave feedback (Fig. 7c), we see a strong increase up until simulation year 3000.

Over the first 500 simulation years, only part of the  $\lambda_{\text{SW}}$  increase is captured by the assumption of constant local feedbacks (Fig. 7c). This is because this increase is mostly due to shifts in sea ice melting regions toward higher latitudes, which changes the magnitude of the local albedo feedback (section 4). This includes a rapid increase during the first 10–50 years, where the initial melting regions become established on both hemispheres, but mainly in the Southern Ocean (Fig. 7d).

During the further increase between years 500 and 3000, melting regions are fully established, and the global-mean feedback is well described by the approximation that the increasingly polar-amplified warming pattern acts on a constant local feedback.

Also in the Arctic Ocean of Bern3D-LPX, the ice–albedo feedback increases, mainly due to enhanced melting north of Bering Strait (Figs. 2f, l–n). This increase takes place in the first 500 years (Fig. 4g shows that there is no further increase thereafter). This is consistent with the fact that Northern Hemispheric sea ice melting progresses only very little after year 500, while about half of the Southern Ocean sea ice melting occurs on a longer time scale (Fig. 7d).

To investigate how much this Arctic melting contributes to the increase in global mean shortwave feedback (Fig. 7c), we have performed simulations in which the albedo was fixed in one hemisphere, and allowed to evolve freely in the other hemisphere [as detailed in Pfister and Stocker (2017)]. In this way, we have found that about  $0.1 \text{ W m}^{-2} \text{ K}^{-1}$  of the  $\lambda_{\text{SW}}$  increase during the first 500 years can be attributed to Arctic contributions, while the remaining increase (including the full increase after 500 years) is due to Southern Ocean contributions. Therefore, we focus on southern sea ice changes in section 7.

Why does the approximation of constant local feedbacks capture the total feedback evolution so well in Bern3D-LPX (Fig. 7a), even though it does not capture the shifting sea ice melting regions as described above? This is because global-mean longwave feedback becomes more negative (Fig. 7b) mainly during the first 500 years, which counteracts the initial shortwave feedback increase. However, this longwave feedback time dependence is a purely diagnostic feature due to the assumption that fast adjustments are reflected in the effective radiative forcing  $R_{\text{eff}}$  rather than in the feedbacks.

If Bern3D's longwave feedback is diagnosed using the prescribed forcing  $R_{\text{presc}} = 3.71 \text{ W m}^{-2} \text{ K}^{-1}$  (gray line in Fig. 7b) rather than the effective forcing  $R_{\text{eff}} = 3.44 \text{ W m}^{-2} \text{ K}^{-1}$  (blue line), the global-mean longwave feedback remains roughly constant. This indicates that, in the global-mean feedback calculation using  $R_{\text{eff}}$ , some of the initial  $\lambda_{\text{SW}}$  increase is falsely accounted for as a rapid adjustment rather than a shortwave feedback increase. Sea ice, as well as snow and land ice, rapidly absorbs some of the initial energy imbalance, lowering  $R_{\text{eff}}$  to a smaller value than  $R_{\text{presc}}$ .

Only because this initial perturbation is accounted for in  $R_{\text{eff}}$  rather than  $\lambda(t)$ , the approximation of constant local feedbacks works well in Bern3D-LPX. This calls for caution when diagnosing global-mean longwave feedbacks using an effective radiative forcing, also in more comprehensive models. Although the fast adjustments in Bern3D-LPX are physically better described as an increase in  $\lambda_{\text{SW}}$ , the processes causing fast adjustments in comprehensive models may be distinct (Rugenstein et al. 2016b), justifying the use of an effective radiative forcing—apart from the fact that the actual forcing before rapid adjustments is more difficult to diagnose, as it is not directly prescribed like in Bern3D-LPX. Nevertheless, rapid ice melting may also influence the  $R_{\text{eff}}$  diagnosis in comprehensive models.

We have used  $R_{\text{eff}}$  for Bern3D-LPX throughout this study mainly to be consistent with our LOVECLIM analysis, and with comparable analyses of comprehensive models. Note that the results discussed in section 4, including the change in longwave feedback patterns (Fig. 2i), are robust to the choice of forcing ( $R_{\text{presc}}$  or  $R_{\text{eff}}$ ).

## 6. The ocean heat uptake efficacy perspective

Alternatively to a time-dependent  $\lambda$ , the deviation from linearity in Fig. 1 can also be explained by introducing an additional factor  $\varepsilon$  in the global EBM, accounting for the efficacy of OHU (Winton et al. 2010):

$$-\lambda_{\text{eq}} \Delta T(t) = R_{\text{eff}} - \varepsilon N(t), \quad (6)$$

where  $\lambda_{\text{eq}}$  is constant. An OHU efficacy  $\varepsilon > 1$  accounts for the fact that the cooling caused by a global mean OHU of  $1 \text{ W m}^{-2}$  is stronger than the warming caused by a global mean forcing of  $1 \text{ W m}^{-2}$ . The first-order explanation of this difference is that the OHU takes place mainly in the high latitudes where local feedbacks are strongest, while the forcing is more uniform (Winton et al. 2010; Rose et al. 2014). An efficacy of  $\varepsilon > 1$  is diagnosed for most comprehensive models (e.g., Winton et al. 2010), while  $\varepsilon$  is close to 1 for most EMICs except for the two studied here, which are more consistent with the comprehensive models (Pfister and Stocker 2018).

Assuming a constant efficacy  $\varepsilon$  exceeding 1 can account for the fact that the model output lies below the dashed line in Fig. 1 (i.e., for a less negative but constant slope of the  $N/\Delta T$  relation). However, to explain the changing slope with a constant  $\lambda_{\text{eq}}$ , we need a time-dependent  $\varepsilon(t)$  (Armour et al. 2013; Paynter and Frölicher 2015). To better understand this time dependence,  $\varepsilon$  can be decomposed into a contribution by changing warming patterns and a contribution by changing local feedbacks, following Rose et al. (2014):

$$\varepsilon_{\text{Rose}}^{-1}(t) = 1 + \underbrace{\frac{\lambda_{\text{eq}}(\mathbf{r})}{\bar{\lambda}_{\text{eq}}} \left[ \frac{\Delta T_{\text{ohu}}(\mathbf{r}, t)}{\Delta T_{\text{ohu}}} - \frac{\Delta T_{\text{eq}}(\mathbf{r}, t)}{\Delta T_{\text{eq}}} \right]}_{R_1} + \underbrace{\frac{\lambda_{\text{ohu}}(\mathbf{r}, t) - \lambda_{\text{eq}}(\mathbf{r})}{\bar{\lambda}_{\text{eq}}} \frac{\Delta T_{\text{ohu}}(\mathbf{r}, t)}{\Delta T_{\text{ohu}}}}_{R_2}. \quad (7)$$

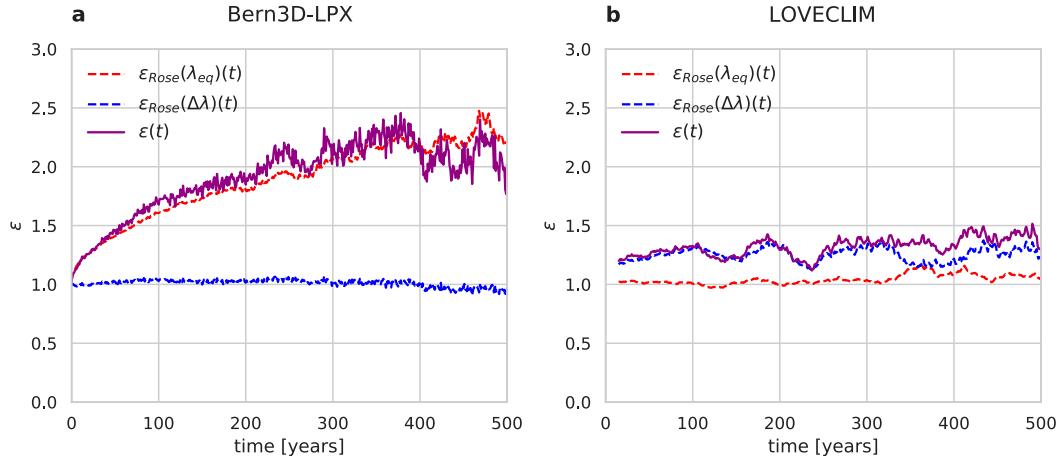


FIG. 8. Time-dependent OHU efficacy in the  $2 \times \text{CO}_2$  simulation Bern3D-LPX and LOVECLIM (purple line). The contributions of changing warming patterns (red) and changing local feedbacks (blue) are also shown; see Eqs. (7), (11), and (12) for the definition of  $\varepsilon_{\text{Rose}}$ .

The terms  $\Delta T_{\text{eq}}$  and  $\lambda_{\text{eq}}$  are the equilibrium warming and feedback, that is, the warming and feedback caused by the forcing once the global mean OHU becomes zero. For LOVECLIM, we used  $\Delta T_{\text{eq}} = \text{ECS}$ , and therefore also had to scale the equilibrium warming pattern to the extrapolated ECS for consistency:

$$\Delta T_{\text{eq}}(\mathbf{r}, t) = \frac{\text{ECS}}{\Delta T(t_{\text{end}})} \Delta T(\mathbf{r}, t_{\text{end}}), \quad (8)$$

where ( $t_{\text{end}}$ ) is shorthand for the average over the end of the simulation (i.e., years 970–1000). Note that this scaling is minor, as the factor  $\text{ECS}/\Delta T(t_{\text{end}}) \approx 1.05$  is close to 1, but it is required to obtain a global mean  $\varepsilon$  consistent with Eq. (6).

The terms  $\Delta T_{\text{ohu}}$  and  $\lambda_{\text{ohu}}$  are the cooling and feedback caused by the OHU. In contrast to Rose et al. (2014), we do not perform separate prescribed-OHU simulations, and therefore  $\Delta T_{\text{ohu}}$  and  $\lambda_{\text{ohu}}$  have to be diagnosed as the differences between transient and equilibrium values:

$$\Delta T_{\text{ohu}}(\mathbf{r}, t) = \Delta T(\mathbf{r}, t) - \Delta T_{\text{eq}}(\mathbf{r}), \quad (9)$$

$$\lambda_{\text{ohu}}(\mathbf{r}, t) = \frac{\lambda(\mathbf{r}, t) \Delta T(\mathbf{r}, t) - \lambda_{\text{eq}}(\mathbf{r}) \Delta T_{\text{eq}}(\mathbf{r})}{\Delta T_{\text{ohu}}(\mathbf{r}, t)}. \quad (10)$$

The separate contributions of changing warming patterns [ $\varepsilon_{\text{Rose}}(\lambda_{\text{eq}})(t)$ ] and changing local feedbacks [ $\varepsilon_{\text{Rose}}(\Delta\lambda)(t)$ ] to the time dependence of  $\varepsilon(t)$  can be determined by omitting either term  $R_1$  or  $R_2$  in Eq. (7):

$$\varepsilon_{\text{Rose}}(\lambda_{\text{eq}})(t) = \frac{1}{1 + R_1}, \quad (11)$$

$$\varepsilon_{\text{Rose}}(\Delta\lambda)(t) = \frac{1}{1 + R_2}. \quad (12)$$

Figure 8 shows the total OHU efficacy, as well as the two approximations  $\varepsilon_{\text{Rose}}(\lambda_{\text{eq}})(t)$  and  $\varepsilon_{\text{Rose}}(\Delta\lambda)(t)$ . The total OHU efficacy calculated from Eq. (7) is identical to the OHU efficacy calculated from the global EBM, Eq. (6).

The OHU efficacy is larger than 1 for both models and increases with time, reflecting the increasing deviation from linearity in Fig. 1. We do not show OHU efficacy after 500 simulation years, because this quantity becomes noisy and hard to interpret once the planetary heat imbalance  $N_{\text{TOA}}$  becomes small.

The processes in Bern3D-LPX are consistent with the basic understanding of an OHU efficacy that is greater than one (Winton et al. 2010). The transient change in OHU is strongest in the Southern Ocean (Figs. 4k,l), in agreement with GCMs (Frölicher et al. 2015). Due to the sea ice–albedo feedback in this region, the local feedback  $\lambda_{\text{eq}}(\mathbf{r})$  is stronger than anywhere else in Bern3D-LPX (Figs. 4c,g). Thus, the region of strongest OHU change corresponds with the region of anomalously strong feedbacks. This explains the stronger impact of OHU on temperature, compared to the uniform forcing.

The processes in LOVECLIM are comparable to a more recent mechanistic interpretation of an OHU efficacy larger than one, related to changing local feedbacks. In comprehensive models, changes in both cloud and lapse-rate feedbacks are linked to changes in atmospheric stability (Rose and Rayborn 2016; Rose and Rencurrel 2016; Ceppi and Gregory 2017, 2019). Interactive cloud effects are missing in LOVECLIM. But we have shown that the lapse-rate feedback can explain, or at least contribute to, the feedback change toward less negative values in the Southern Ocean (section 4c), which is the cause of the nonunitary  $\varepsilon$ . In the first 100 simulation years, even in the absence of cloud feedbacks, LOVECLIM’s OHU efficacy amounts to roughly  $\varepsilon = 1.3$ , which is comparable to comprehensive GCMs (e.g., Winton et al. 2010; Pfister and Stocker 2018).

For the first 500 years, the OHU efficacy is closely matched by the contribution of changing warming patterns for Bern3D-LPX (red line in Fig. 8a) and by the contribution of changing local feedbacks for LOVECLIM (blue line in Fig. 8b). The OHU efficacy decomposition thus yields consistent results with the direct feedback analyses shown in sections 4 and 5. The contributions from changing warming patterns become more important for LOVECLIM after 300 years. This indicates that

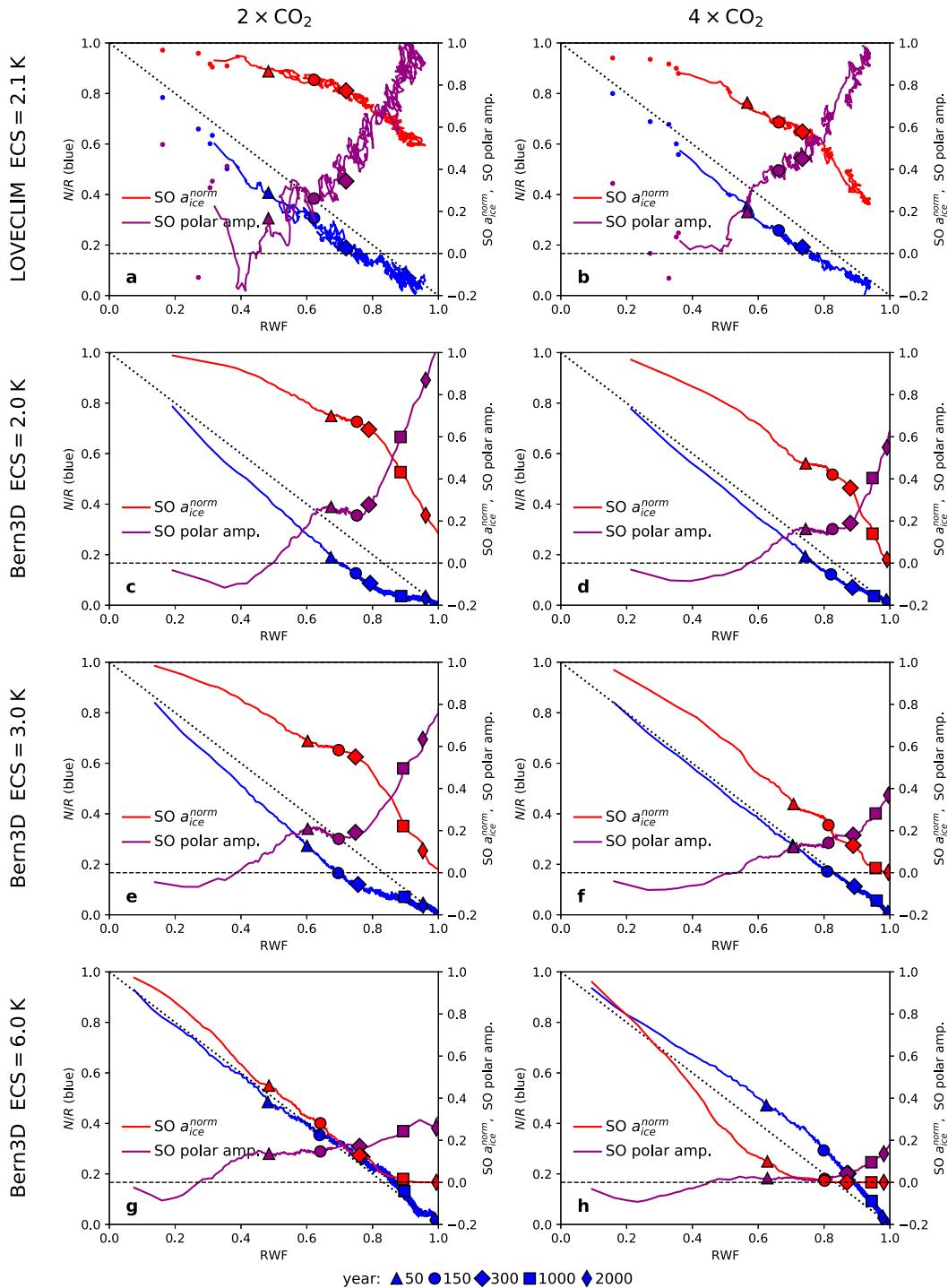


FIG. 9. Normalized global heat uptake (blue; left y axis), Southern Ocean sea ice area and Southern Ocean polar amplification (red and purple; right y axis) compared to the realized warming fraction (RWF). Time series (lines) and specific years (symbols; legend below the figure) are shown. Shown are the (left)  $2 \times \text{CO}_2$  and (right)  $4 \times \text{CO}_2$  simulations. (a), (b) LOVECLIM, and (c)–(h) three Bern3D-LPX versions with increasing ECS. The dashed grid line marks zero on the right y axis.

the long-term change in global mean feedback toward less negative values may be a combination of two factors: 1) the increasing local lapse-rate feedback and 2) the intensifying polar amplification in the Southern Ocean, acting on a net local feedback that is stronger than the global average (due to a combination of the near-constant sea ice–albedo feedback, and the increased lapse-rate feedback).

### 7. The influence of forcing magnitude, climate sensitivity, and sea ice availability

So far, we have focused on the  $2 \times \text{CO}_2$  scenario in the standard versions of both models. In this section, we compare this to the  $4 \times \text{CO}_2$  scenario, and also to a low- and high ECS version (section 2) of the Bern3D-LPX model.

To enable this comparison, each panel of Fig. 9 combines the information from Fig. 1 with two quantities that were shown to be influential in the sections above. These quantities are the normalized Southern Ocean sea ice area  $a_{\text{ice}}^{\text{norm}}$  (Fig. 7b) and the Southern Ocean polar amplification SPA, defined as

$$\text{SPA} = \frac{\Delta T(52^\circ - 90^\circ S)}{\Delta T} - 1, \quad (13)$$

where  $\Delta T(52^\circ - 90^\circ S)$  is the mean atmospheric warming over the southern polar cap (i.e., averaged both zonally and over latitudes  $52^\circ - 90^\circ$ ). SPA values above zero imply that warming is stronger over the southern high latitudes than in the global average. Values below zero indicate the opposite, which occurs in the early years of all model versions due to faster polar amplification in the Northern Hemisphere.

In contrast to Fig. 1, the axes in Fig. 9 are normalized: The  $x$  axis corresponds to the realized warming fraction  $\text{RWF}(t) = \Delta T(t)/\text{ECS}$  (Frölicher and Paynter 2015; Pfister and Stocker 2018) and the left  $y$  axis corresponds to  $N/R_{\text{eff}}$ . This is done to enable direct comparison of the different forcings and climate sensitivities. The right  $y$  axis is shared between  $a_{\text{ice}}^{\text{norm}}$  and SPA. The time information is contained in the different symbols, which correspond to simulation years 50, 150, 300, 1000, and 2000.

For LOVECLIM, both forcings produce qualitatively similar results (Figs. 9a,b). Although the global heat imbalance decays to about 20% by year 300, half or more of the southern sea ice melting and polar amplification takes place between years 300 and 1000. As discussed in the previous section, this is not the main cause of the global-mean feedback time dependence in LOVECLIM, but may act to enhance it. In LOVECLIM's  $4 \times \text{CO}_2$  simulation, about 25% of the preindustrial sea ice remains in the Southern Ocean after 1000 simulation years. Due to the doubled forcing, the plotted time series appears less noisy, as the natural variability is relatively smaller compared to the forcing response. Otherwise, the simulation closely corresponds to the  $2 \times \text{CO}_2$  simulation.

In Bern3D-LPX, the different forcings and different ECS produce qualitatively different results. We first compare the  $2 \times \text{CO}_2$  and  $4 \times \text{CO}_2$  simulations in the standard model version (Figs. 9e,f). In the  $2 \times \text{CO}_2$  simulation, about 60% of the southern sea ice melting and 70% of the polar amplification take place after year 300, although only about 10% of the global heat imbalance remains. This delayed amplification

causes the increased climate sensitivity (shallower slope of the blue line) after year 300.

In contrast, the climate sensitivity of the  $4 \times \text{CO}_2$  simulation is almost constant over time (Fig. 9f, the blue line follows the dashed black line). This is due to lack of sea ice availability in the Southern Ocean. About 85% of the Southern Ocean sea ice has already vanished by year 300 due to the stronger, near-uniform warming in response to the stronger forcing. Therefore, the delayed warming amplification in the SO is less than half compared to the  $2 \times \text{CO}_2$  simulation, because the ice–albedo feedback is much weaker with so little ice remaining.

Sea ice availability thus limits the delayed polar amplification, and thereby the time dependence of the global mean feedback, in Bern3D-LPX. In the low-ECS model version (Figs. 9c,d), both forcings yield a qualitatively similar evolution to the standard  $2 \times \text{CO}_2$  simulation (Fig. 9e). The global mean feedback is smaller due to the low-ECS tuning (section 2), and therefore enough sea ice remains by year 300 to enable the late polar amplification. Compared to the standard  $2 \times \text{CO}_2$  simulation, the late polar amplification is less pronounced in the low-ECS  $4 \times \text{CO}_2$  simulation, and more pronounced in the low-ECS  $2 \times \text{CO}_2$  simulation, which is the only simulation with no complete meltdown of the Southern Ocean sea ice within 5000 years.

For high ECS and high forcing (Fig. 9h), most of the Southern Ocean sea ice melts within 150 years. For the remaining duration of the simulation, warming continues at a lower total feedback because the southern sea ice–albedo feedback is disabled by the lack of sea ice. This leads to  $\varepsilon < 1$  if the curvature in the  $N/\Delta T$  relation is attributed to  $\varepsilon$ . A similar but less pronounced curvature is simulated by other climate models, mainly under the abrupt  $4 \times \text{CO}_2$  forcing. This includes all six EMIC-AR5 models apart from Bern3D-LPX and LOVECLIM that do not feature interactive clouds, as well as one EMIC (the CLIMBER-3 $\alpha$  model) with interactive clouds (Pfister and Stocker 2018, their Fig. S2). It also includes the comprehensive ECHAM5 model (Li et al. 2013, their Fig. 11). Our results suggest that this could be due to a weakened albedo feedback after the depletion of sea ice, although the depleted region (Southern Ocean or Arctic Ocean) may differ between models, or other effects could be prevalent. In the case of ECHAM5, a strong Southern Ocean polar amplification (Li et al. 2013, their Fig. 10) points toward a similar sea ice depletion in the south; the simulated sea ice extent should be investigated to confirm or reject this hypothesis.

## 8. Conclusions

We have investigated the causes of the time dependence of the transient climate sensitivity in two EMICs: the Bern3D-LPX and the LOVECLIM model. This time dependence can be understood from two different perspectives: First, the global mean feedback becomes less negative due to a combination of increasingly polar-amplified warming and changing local feedback patterns; second, OHU efficacy increases due to strong OHU in the Southern Ocean, where an anomalously strong local feedback dominates. This is not a contradiction: Haugstad et al. (2017) have demonstrated in their aquaplanet simulations that the two perspectives are equivalent.

These effects are in agreement with comprehensive models (e.g., Armour et al. 2013; Andrews et al. 2015; Proistosescu and Huybers 2017), and are in principle independent of the mechanisms causing the polar-amplified feedback pattern. However, mechanistic analyses of comprehensive models have pointed toward a dominant influence of cloud and lapse-rate feedbacks for the time dependence of  $\lambda(t)$ , as discussed in the introduction. Here we have shown that a time dependence of  $\lambda(t)$  is also found in two models of intermediate complexity that do not simulate cloud feedbacks.

In both models, the time dependence of the global-mean feedback is dominantly driven by processes in the Southern Ocean region. There are two main differences between the two models. First, both the local and global-mean feedback changes are dominated by longwave feedbacks in LOVECLIM, and by shortwave feedbacks in Bern3D-LPX, as detailed below. Second, the relative importance of changing warming patterns and changing local feedbacks is opposite in the two models. This was illustrated both in a regional feedback analysis and in an OHU efficacy decomposition following Rose et al. (2014).

In LOVECLIM, the global-mean feedback becomes less negative mostly because the local feedback in the Southern Ocean region becomes less negative. In agreement with comprehensive models (e.g., Po-Chedley et al. 2018), LOVECLIM's atmospheric warming profile suggests that this can be attributed to a lapse-rate feedback increase. Only on the multicentury time scale, the change in the global-mean feedback toward less negative values is enhanced by the delayed warming over the Southern Ocean. That is because this regional warming acts on the local feedback that is higher than the global mean, due to the near-constant sea ice–albedo feedback and the increased lapse-rate feedback.

In Bern3D-LPX, the global-mean feedback becomes less negative mainly due to delayed polar-amplified warming acting on near-constant local feedbacks (Armour et al. 2013). The local feedback pattern is mainly determined by the sea ice–albedo feedback, which is strongest in the Southern Ocean. The delayed warming in this region acting on this strong local feedback causes most of the global-mean feedback change. However, the assumption of constant local feedbacks only holds well if fast adjustments in sea ice melting are accounted for in the effective radiative forcing  $R_{\text{eff}}$  rather than the feedback. Otherwise, local shortwave feedback pattern changes within the first 500 years (due to poleward shifts in sea ice melting regions) contribute to a less-negative global-mean feedback, in addition to the shifting warming pattern.

In summary, these results show that the assumption of constant local feedbacks can hold well in models with a one-layer atmosphere (like Bern3D-LPX). There, the only non-constant feedback contributions are nonlinearities in the Planck and water vapor feedbacks, as well as shifts in melting regions. However, these shifts may strongly affect initial local albedo feedback changes, such that the assumption holds true only on multicentury to millennial time scales. In models with a three-dimensional atmosphere, the assumption no longer holds due to the nonlocal lapse-rate feedback changes (like in LOVECLIM), and even less so if nonlocal cloud feedbacks are resolved in addition (Ceppi and Gregory 2017).

Furthermore, we have shown that the feedback change caused by delayed Southern Ocean warming is limited by sea ice availability in Bern3D-LPX. In model simulations with low global mean warming, due to low ECS tuning or low forcing or both, the feedback becomes less negative over time due to the effects described above. Contrarily, in simulations with a higher global mean warming, the global mean feedback becomes time independent, because the delayed polar amplification is inhibited by the faster loss of sea ice. In the simulation with the highest ECS (6.0°C) and forcing ( $4 \times \text{CO}_2$ ), the feedback even becomes more negative over time.

There are some caveats to this finding. First, there is a warm bias of around 3°C in the Southern Ocean surface air temperature of Bern3D-LPX compared to reanalysis data. Second, the ice–albedo feedback of Bern3D-LPX amounts to  $0.75 \text{ W m}^{-2} \text{ K}^{-1}$  in the  $2 \times \text{CO}_2$  simulation (estimated by the feedback difference to a fixed-albedo simulation; Pfister and Stocker 2017), which is an overestimation compared to the multimodel mean of  $0.51 \text{ W m}^{-2} \text{ K}^{-1}$  from 15 comprehensive models (with a minimum-to-maximum range of  $0.29\text{--}0.69 \text{ W m}^{-2} \text{ K}^{-1}$ ; Ceppi and Gregory 2017). Finally, in the thermodynamic sea ice component of Bern3D-LPX, sea ice is only advected following prescribed wind fields; wind changes and ice dynamics are not simulated. More comprehensive models are thus required to confirm this finding.

The global-mean feedback also becomes less negative over time in some other EMICs (Pfister and Stocker 2018) and in the  $4 \times \text{CO}_2$  simulation of the comprehensive ECHAM5 model (Li et al. 2013; Rugenstein et al. 2020). We suggest that sea ice depletion is a possible mechanism for the feedback time dependence of these model simulations. In an intercomparison of millennial simulations, the global-mean feedback becomes less negative over time in most comprehensive models, but the multimodel mean local feedback shows a robust decrease in the Southern Ocean sea ice region (Rugenstein et al. 2020). This indicates that the regional shift and/or depletion of sea ice simulated in Bern3D-LPX are also relevant in comprehensive models. In the Community Climate System Model version 3 (CCSM3), the Southern Ocean becomes ice-free under a  $4 \times \text{CO}_2$  forcing (Goosse et al. 2018). Recent observations by Parkinson (2019) show that the Southern Ocean sea ice extent has decreased rapidly over the last 5 years, to levels unprecedented in at least 40 years. This underlines the importance of understanding the climate feedback and warming response to strong sea ice melting in the Southern Ocean.

*Acknowledgments.* This study is supported by the Swiss National Science Foundation through projects 200020\_159563 and 200020\_172745. This project is TiPES contribution 16. This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement 820970. François Klein performed the LOVECLIM simulations and provided us with regridded model output and support for its analysis. Hugues Goosse and Michael Eby provided helpful information on the EMIC-AR5 model simulations. The constructive comments of Kyle Armour and two anonymous reviewers have substantially improved this paper.

*Data availability statement.* The model output analyzed for this study is publicly available online (<https://doi.org/10.5281/zenodo.4066686>). Information on structure and usage is provided in the description of this repository dataset.

## 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>.
- Armour, K. C., 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>.
- 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**, 13 126–13 131, <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>.
- 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>.
- 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>.
- Eby, M., and Coauthors, 2013: Historical and idealized climate model experiments: An intercomparison of Earth system models of intermediate complexity. *Climate Past*, **9**, 1111–1140, <https://doi.org/10.5194/cp-9-1111-2013>.
- Feldl, N., and G. H. Roe, 2013a: Four perspectives on climate feedbacks. *Geophys. Res. Lett.*, **40**, 4007–4011, <https://doi.org/10.1002/grl.50711>.
- , and —, 2013b: The nonlinear and nonlocal nature of climate feedbacks. *J. Climate*, **26**, 8289–8304, <https://doi.org/10.1175/JCLI-D-12-00631.1>.
- Frölicher, T. L., and D. J. Paynter, 2015: Extending the relationship between global warming and cumulative carbon emissions to multi-millennial timescales. *Environ. Res. Lett.*, **10**, 075002, <https://doi.org/10.1088/1748-9326/10/7/075002>.
- , J. L. Sarmiento, D. J. Paynter, J. P. Dunne, J. P. Krasting, and M. Winton, 2015: Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *J. Climate*, **28**, 862–886, <https://doi.org/10.1175/JCLI-D-14-00117.1>.
- Fyfe, J. C., and O. A. Saenko, 2006: Simulated changes in the extratropical Southern Hemisphere winds and currents. *Geophys. Res. Lett.*, **33**, L06701, <https://doi.org/10.1029/2005GL025332>.
- Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. J. L. Olivie, and S. Tyteca, 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>.
- Goosse, H., and Coauthors, 2010: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2. *Geosci. Model Dev.*, **3**, 603–633, <https://doi.org/10.5194/gmd-3-603-2010>.
- , and Coauthors, 2018: Quantifying climate feedbacks in polar regions. *Nat. Commun.*, **9**, 1919, <https://doi.org/10.1038/s41467-018-04173-0>.
- Gregory, J. M., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, **31**, L03205, <https://doi.org/10.1029/2003GL018747>.
- , T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response parameter under increasing CO<sub>2</sub>. *Philos. Trans. Roy. Soc.*, **373A**, 20140417, <https://doi.org/10.1098/rsta.2014.0417>.
- Haugstad, A. D., K. C. Armour, D. S. Battisti, and B. E. J. 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>.
- Huang, Y., Y. Xia, and X. Tan, 2017: On the pattern of CO<sub>2</sub> radiative forcing and poleward energy transport. *J. Geophys. Res. Atmos.*, **122**, 10 578–10 593, <https://doi.org/10.1002/2017JD027221>.
- Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94**, 1339–1360, <https://doi.org/10.1175/BAMS-D-12-00121.1>.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., <https://doi.org/10.1017/CBO9781107415324>.
- Kushner, P. J., I. M. Held, and T. L. Delworth, 2001: Southern Hemisphere atmospheric circulation response to global warming. *J. Climate*, **14**, 2238–2249, [https://doi.org/10.1175/1520-0442\(2001\)014<0001:SHACRT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0001:SHACRT>2.0.CO;2).
- Li, C., J.-S. von 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>.
- Myhre, G., E. J. Highwood, K. P. Shine, and F. Stordal, 1998: New estimates of radiative forcing due to well-mixed greenhouse gases. *Geophys. Res. Lett.*, **25**, 2715–2718, <https://doi.org/10.1029/98GL01908>.
- Nalam, A., G. Bala, and A. Modak, 2018: Effects of Arctic geo-engineering on precipitation in the tropical monsoon regions. *Climate Dyn.*, **50**, 3375–3395, <https://doi.org/10.1007/s00382-017-3810-y>.
- Parkinson, C. L., 2019: A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic. *Proc. Natl. Acad. Sci. USA*, **116**, 14 414–14 423, <https://doi.org/10.1073/pnas.1906556116>.
- Paynter, D., and T. L. Frölicher, 2015: Sensitivity of radiative forcing, ocean heat uptake, and climate feedback to changes in anthropogenic greenhouse gases and aerosols. *J. Geophys. Res. Atmos.*, **120**, 9837–9854, <https://doi.org/10.1002/2015JD023364>.
- Pfister, P. L., and T. F. Stocker, 2017: State-dependence of the climate sensitivity in Earth system models of intermediate

- complexity. *Geophys. Res. Lett.*, **44**, 10643–10653, <https://doi.org/10.1002/2017GL075457>.
- , and —, 2018: The realized warming fraction: A multi-model sensitivity study. *Environ. Res. Lett.*, **13**, 124024, <https://doi.org/10.1088/1748-9326/aaebae>.
- 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>.
- Proistosescu, C., and P. J. Huybers, 2017: Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Sci. Adv.*, **3**, e1602821, <https://doi.org/10.1126/sciadv.1602821>.
- Ritz, S. P., T. F. Stocker, and F. Joos, 2011: A coupled dynamical ocean–energy balance atmosphere model for paleoclimate studies. *J. Climate*, **24**, 349–375, <https://doi.org/10.1175/2010JCLI3351.1>.
- 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>.
- , and M. C. Rencurrel, 2016: The vertical structure of tropospheric water vapor: Comparing radiative and ocean-driven climate changes. *J. Climate*, **29**, 4251–4268, <https://doi.org/10.1175/JCLI-D-15-0482.1>.
- , 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>.
- Roth, R., S. P. Ritz, and F. Joos, 2014: Burial-nutrient feedbacks amplify the sensitivity of atmospheric carbon dioxide to changes in organic matter remineralisation. *Earth Syst. Dyn.*, **5**, 321–343, <https://doi.org/10.5194/esd-5-321-2014>.
- Rugenstein, M. A. A., K. Caldeira, and R. Knutti, 2016a: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophys. Res. Lett.*, **43**, 9877–9885, <https://doi.org/10.1002/2016GL070907>.
- , J. M. Gregory, N. Schaller, J. Sedláček, and R. Knutti, 2016b: Multiannual ocean–atmosphere adjustments to radiative forcing. *J. Climate*, **29**, 5643–5659, <https://doi.org/10.1175/JCLI-D-16-0312.1>.
- , and Coauthors, 2020: Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophys. Res. Lett.*, **47**, e2019GL083898, <https://doi.org/10.1029/2019GL083898>.
- Russell, J. L., K. W. Dixon, A. Gnanadesikan, R. J. Stouffer, and J. R. Toggweiler, 2006: The Southern Hemisphere westerlies in a warming world: Propping open the door to the deep ocean. *J. Climate*, **19**, 6382–6390, <https://doi.org/10.1175/JCLI3984.1>.
- Sen Gupta, A., A. Santoso, A. S. Taschetto, C. C. Ummerhofer, J. Trevena, and M. H. England, 2009: Projected changes to the Southern Hemisphere ocean and sea ice in the IPCC AR4 climate models. *J. Climate*, **22**, 3047–3078, <https://doi.org/10.1175/2008JCLI2827.1>.
- Sherwood, S., and Coauthors, 2020: An assessment of Earth’s climate sensitivity using multiple lines of evidence. *Rev. Geophys.*, e2019RG000678, <https://doi.org/10.1029/2019RG000678>.
- Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying climate feedbacks using radiative kernels. *J. Climate*, **21**, 3504–3520, <https://doi.org/10.1175/2007JCLI2110.1>.
- Stocker, B. D., and Coauthors, 2013: Multiple greenhouse-gas feedbacks from the land biosphere under future climate change scenarios. *Nat. Climate Change*, **3**, 666–672, <https://doi.org/10.1038/nclimate1864>.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, <https://doi.org/10.1256/qj.04.176>.
- 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>.
- Yoshimori, M., M. Watanabe, H. Shiogama, A. Oka, A. Abe-Ouchi, R. Ohgaito, and Y. Kamae, 2016: A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. *Prog. Earth Planet. Sci.*, **3**, 21, <https://doi.org/10.1186/s40645-016-0096-3>.
- 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>.