



On simple representations of the climate response to external radiative forcing

Tim Rohrschneider¹ · Bjorn Stevens¹ · Thorsten Mauritsen²

Received: 23 August 2018 / Accepted: 14 February 2019 / Published online: 27 February 2019
© The Author(s) 2019

Abstract

Global warming in response to external radiative forcing is determined by the feedback of the climate system. Recent studies have suggested that simple mathematical models incorporating a radiative response which is related to upper- and deep-ocean disequilibrium (ocean heat uptake efficacy), inhomogeneous patterns of surface warming and radiative feedbacks (pattern effect), or an explicit dependence of the strength of radiative feedbacks on surface temperature change (feedback temperature dependence) may explain the climate response in atmosphere-ocean coupled general circulation models (AOGCMs) or can be useful for interpreting the instrumental record. We analyze a two-layer model with an ocean heat transport efficacy, a two-region model with region specific heat capacities and radiative responses; a one-layer model with a temperature dependent feedback; and a model which combines elements of the two-layer/region models and the state-dependent feedback parameter. We show that, from the perspective of the globally averaged surface temperature and radiative imbalance, the two-region and two-layer models are equivalent. State-dependence of the feedback parameter introduces a nonlinearity in the system which makes the adjustment timescales forcing-dependent. Neither the linear two-region/layer models, nor the state-dependent feedback model adequately describes the behavior of complex climate models. The model which combines elements of both can adequately describe the response of more comprehensive models but may require more experimental input than is available from single forcing realizations.

Keywords Nonconstant global feedback · Energy balance models · Timescales

1 Introduction

The need to understand factors influencing climate feedback and sensitivity make simple models indispensable as means to investigate the climate response to changes in the Earth's energy balance. In this paper we discuss the relationship between the energy balance of the climate system and its temperature response. We focus on the temporal behavior using the conceptual framework of four commonly used conceptual models. We reveal similarity and essential differences among these models, and we provide insight into the climate response in the Max Planck Institute Earth System Model 1.2 (MPI-ESM1.2) as a counterpart to our theoretical

considerations. On mathematical grounds, we question the difference between the concept of ocean heat uptake efficacy and a pattern effect due to regional feedbacks weighted by a geographic pattern of surface warming. Further, we challenge the common assumption of these concepts that the characteristic timescales on which the climate system adjusts are independent of the strength of the external perturbation.

A popular framework to analyze the radiative response to external radiative forcing is the linear forcing-feedback framework (e.g. Gregory et al. 2004). It is based on a linear relationship between global net top-of-the-atmosphere (TOA) radiative flux N , radiative forcing F and surface temperature perturbation T which in turn actuates the radiative feedback parameter λ ;

$$N = F + \lambda T. \quad (1)$$

The linear forcing-feedback framework Eq. (1) is an appropriate first-order description for the relationship between N and T in experiments of state-of-the-art atmosphere-ocean coupled general circulation models (AOGCMs) in which

✉ Tim Rohrschneider
tim.rohrschnieder@mpimet.mpg.de

¹ Max Planck Institute for Meteorology, Hamburg, Germany

² Department of Meteorology, Stockholm University, Stockholm, Sweden

the preindustrial atmospheric CO₂ concentration is abruptly doubled and then held constant. Details, such as the possibility of a nonlinear dependence on T are omitted from this model by construction. Numerical experiments with more comprehensive models suggest that these details may matter, and that Eq. (1) becomes an increasingly poor description as the forcing is increased to 4 or 8 times preindustrial CO₂. In these cases, there is a tendency for the radiative response in several coupled climate models to decrease over time and with higher temperatures (e.g. Andrews et al. 2015; Bloch-Johnson et al. 2015). If one assumes that Eq. (1) holds at each instance of time, this can be interpreted as nonconstant global feedback. This is modeled by allowing λ to vary over time and rearrange Eq. (1) to obtain

$$\lambda(t) = \frac{N(t) - F(t)}{T(t)}, \quad (2)$$

which is sometimes paraphrased as effective feedback (Senior and Mitchell 2000; Williams et al. 2008; Armour et al. 2013). To avoid ambiguities in terminology, we refer to $\lambda(t)$ as the radiative response of the climate system relative to the global mean surface temperature perturbation T . The outcome of Eq. (2) may, however, also be influenced by additional state-variables. It is itself a specific approach to analyze the climate response to external perturbations and provides a simple and vivid description of the radiative response. Unfortunately, to the extent that λ changes with t it does not alone provide a framework for understanding the origin of such changes which becomes evident during the course of this study.

Currently, there are two different branches of literature on conceptual models to explain the more complex climate response in AOGCM experiments. The first branch suggests that on decadal timescales the variation of $\lambda(t)$ arises from different radiative responses associated with a fast and a slow component of the climate system (two-timescale approaches). The second branch accounts for the possibility that the strength of radiative feedbacks depends explicitly on surface temperature change (feedback temperature dependence).

Hasselmann (1979) already recognized the multiple timescale structure of the climate response, but at this time the global feedback was assumed to depend on the global mean surface temperature only. Recently, these characteristic adjustment timescales have been explicitly linked to nonconstant global feedback associating them with different state-variables each of which influences the radiative response (e.g. Held et al. 2010; Winton et al. 2010; Geoffroy et al. 2013a, b; Armour et al. 2013; Andrews et al. 2015; Hedemann et al. 2019). Held et al. (2010) and Geoffroy et al. (2013a, b) describe a model based on two distinct timescales to reproduce the transient global mean

temperature response of a wide range of AOGCMs. They introduce a globally averaged two-layer ocean model with an efficacy parameter modifying the upper ocean response to heat uptake. Hereafter it is referred to as a two-layer model with efficacy, or just the two-layer model. Two or more adjustment timescales can also be explicitly linked to a one-layer model with two regions. In contrast to Held et al. (2010) and Geoffroy et al. (2013a), Armour et al. (2013) suggest a geographic approach in which constant regional feedbacks are weighted by an evolving pattern of surface warming. This approach is motivated by the idea that the adjustment over parts of the Earth's surface which are weakly coupled to the state of the deep-ocean is more rapid than over regions which are more strongly coupled to the state of the deep-ocean. Consequently, this can be called as pattern effect (Stevens et al. 2016). In addition to time-varying sensitivity, two-region models have also been used to analyze the effect of perturbation heat transport on stability properties and equilibrium sensitivity (e.g. Bates 2012, 2016) or to explore polar amplification (e.g. Langen and Alexeev 2007). In all of these models, the time-variation of $\lambda(t)$ does not depend on the strength of the forcing—in this sense these models are linear in forcing. The tendency of such models to have a value of $\lambda(t)$ that varies with time is thus often paraphrased as a time-dependent feedback though it arises from a dependence of the radiative response on an additional state-variable with its own distinct temporal evolution. Using AOGCM output, e.g. Senior and Mitchell (2000) and Hedemann et al. (2019) discuss time-dependent feedback as a way to understand the radiative response in more comprehensive climate models.

A different approach to representing variation in the radiative response has been to relax the assumption that a temperature perturbation would be proportional to the associated radiative forcing F (e.g. Roe and Baker 2007; Zaliapin and Ghil 2010; Roe and Armour 2011; Meraner et al. 2013; Bloch-Johnson et al. 2015). Studies adopting this approach investigate the long-term response on centennial to millennial timescales as well as the nonlinear change in equilibrium sensitivity between different forcing strengths. Conceptual models on nonlinear feedbacks of the climate system can be traced back at least to Budyko (1969) and Sellers (1969). In a recent study, Bloch-Johnson et al. (2015) extend the linear forcing-feedback framework Eq. (1) by a quadratic coefficient a representing second-order feedback temperature dependence and solve for stationary solutions, $-F = \lambda T + aT^2$. They estimate the range of likely feedback temperature dependencies for various GCM simulations and find $-0.04 \leq a \leq 0.06 \text{ W m}^{-2} \text{ K}^{-2}$. This range is in line with Roe and Armour (2011) who find $-0.058 \leq a \leq 0.06 \text{ W m}^{-2} \text{ K}^{-2}$. Positive feedback temperature dependence makes the feedback of the climate

system continuously less effective as the surface temperature increases, while the opposite is true for negative feedback temperature dependence. In this connection, the radiative response depends explicitly on the climate state; i.e., feedback temperature dependence gives rise to state-dependent feedback such that the variation of $\lambda(t)$ is changed as the strength of the forcing is altered.

We formulate the different conceptual models discussed above as simple ordinary differential equations and study their dynamics. To provide a basis, in Sect. 2 we describe the climate response found in an AOGCM. Then, in Sect. 3, we discuss dynamical systems of the two-timescale approaches and demonstrate that the two-layer model with efficacy can be projected onto a two-region model. In Sect. 4 we analyze dynamical systems which incorporate feedback temperature dependence to point out that the temporal behavior of the climate system, or more precisely the temporal behavior of the comprehensive model developed at our institute, is indeed forcing-dependent.

2 Nonconstant global feedback in MPI-ESM1.2

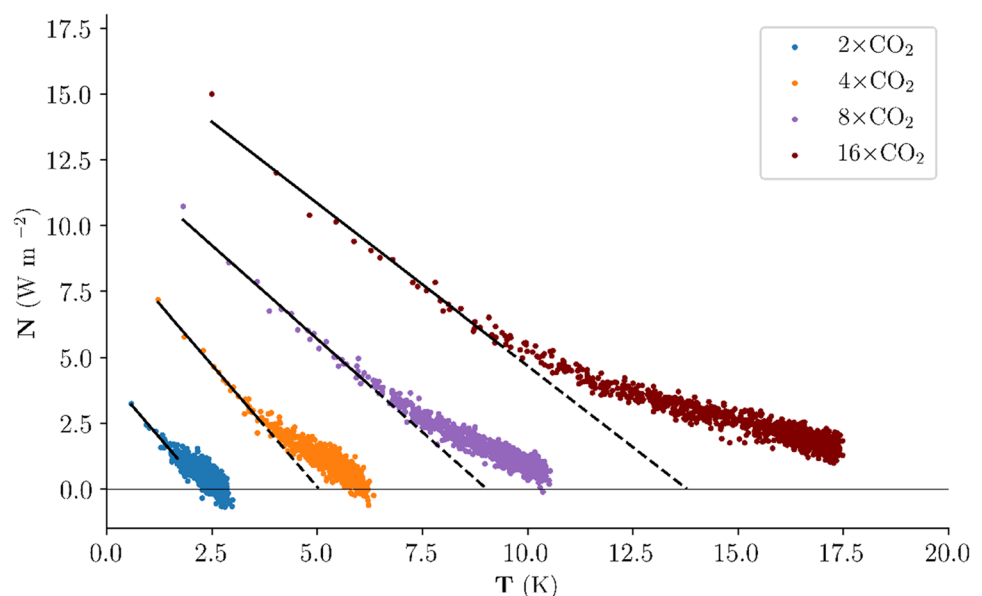
To understand the radiative response of the climate system, state-of-the-art climate models such as the Max Planck Institute Earth System Model 1.2 (MPI-ESM1.2), thoroughly described in Mauritsen et al. (2018), can be forced by step forcing—usually an abrupt increase in atmospheric CO₂ concentration. Such experiments reveal changes or common features in the relationship between N and T (e.g. Gregory et al. 2004). We briefly illustrate the MPI-ESM1.2 response to an abrupt increase in atmospheric CO₂. A detailed analysis of

radiative feedbacks in MPI-ESM1.2 or in its atmospheric component ECHAM6.3 is provided by Mauritsen et al. (2018) and Hedemann et al. (2019). MPI-ESM1.2 is forced by abrupt 2, 4, 8 and 16 times preindustrial CO₂ (284.7 ppm) and is integrated forward in time to 1000 years. In this case, the radiative response is related to radiative feedbacks which emerge on decadal and centennial timescales such as the water vapor, lapse rate, clouds and snow/sea ice albedo feedbacks.

Figure 1 shows the MPI-ESM1.2 relationship between global and annual mean N and T . A common feature of these experiments is that there is an abrupt change in this relationship on decadal timescales. The model output for each CO₂ doubling deviates significantly from the linear least-square regression (solid black line) which is computed using only output from year 1 to 25. Extrapolating the initial response (dashed lines), as would be appropriate if the radiative response to temperature were constant, leads to an underestimation of the equilibrium temperature response, increasingly so for larger forcing. In this connection, the initial variation of $\lambda(t)$ can be explained by distinct timescales on which the climate system adjusts. This is at the heart of the theory of the two-timescale approaches (Sect. 3).

At the same time, the change in the long-term temperature response between the consecutive doublings substantially increases with warmer climates. State-of-the-art models reveal a nearly logarithmic dependence of the radiative forcing F on atmospheric CO₂ concentration as predicted by theory which can be traced back to Arrhenius (1896). That is, the change in the associated radiative forcing F imposed by each doubling cannot explain the nonlinear rise in the equilibrium temperature response: extrapolating the relationship between N and T in MPI-ESM1.2 linearly from year

Fig. 1 The relationship between the TOA imbalance N and surface temperature change T in MPI-ESM1.2 abrupt CO₂ experiments using annual means. The black line represents a linear least-square regression from year 1–25. The dashed line is the extrapolation of this regression to $N = 0$

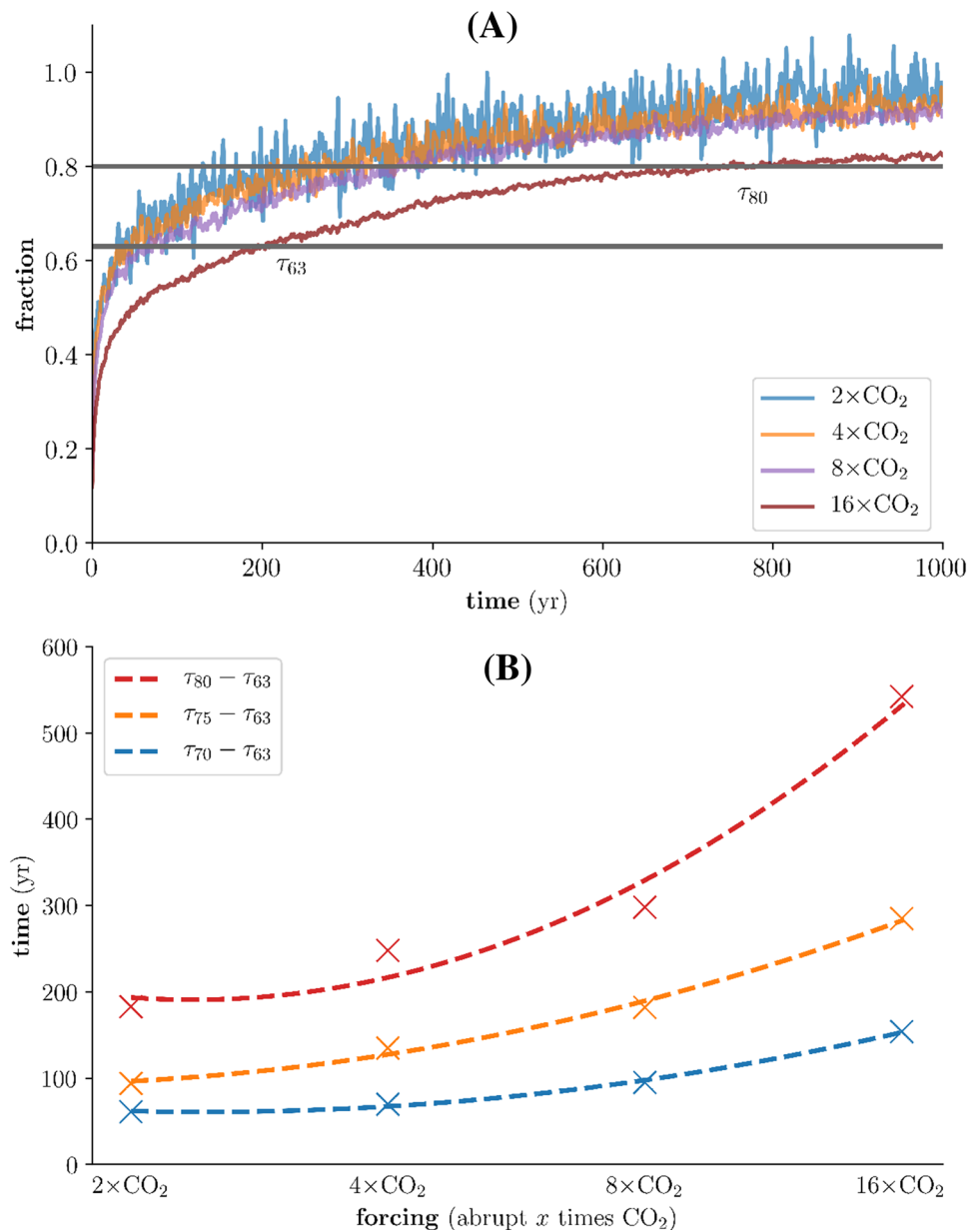


100 to 1000 gives a change in equilibrium sensitivity from $2 \times \text{CO}_2$ to $4 \times \text{CO}_2$ of about 3.6 K and from the $4 \times \text{CO}_2$ experiment to the $8 \times \text{CO}_2$ of about 4.9 K, while the change from the $8 \times \text{CO}_2$ experiment to the $16 \times \text{CO}_2$ is roughly 9.8 K. As discussed in the introduction, this type of nonlinearity has been attributed to feedback temperature dependence (e.g. Meraner et al. 2013; Bloch-Johnson et al. 2015).

After 1000 years the $8 \times \text{CO}_2$ and $16 \times \text{CO}_2$ experiment are still far from steady state, whereas the $2 \times \text{CO}_2$ experiment is nearly equilibrated. After a forcing is applied, the temperature relaxes to a new equilibrium at specific timescales and the temperature increase actuates radiative feedbacks that amplify or attenuate surface warming. That the adjustment time-scale depends on the forcing is demonstrated by

comparing at what time a given fraction of the equilibrium warming is realized as a function of forcing (Fig. 2a). The e-folding time τ_{63} and τ_{80} , defined as the time when a fraction of 63 and 80 percent of the equilibrium response is reached, are compared. Evidently, the response times change as the radiative forcing is altered, and these changes occur in a nonlinear way. The change in response times acts on all timescales while the absolute change between different CO_2 doublings increases with higher forcing and higher fractions of the equilibrium response. To highlight this behavior, we plot the difference between specific response times (τ_{80} , τ_{75} , τ_{70}) to the e-folding time τ_{63} (Fig. 2b). To compute specific response times, the MPI-ESM1.2 temperature output has been smoothed by running mean with a window of thirty

Fig. 2 The MPI-ESM1.2 temporal behavior in abrupt CO_2 experiments using annual means. **a** The fraction of the equilibrium response realized at a given time. For a linear response the lines should lie atop one another. **b** The time difference of specific response times (τ_{80} , τ_{75} , τ_{70}) to the e-folding time τ_{63} . Dashed lines are quadratic fits to the data points



years. The difference between these response times increases approximately quadratically as atmospheric CO₂ is progressively doubled. Section 4 explains this phenomenon.

3 Two-timescale approaches

In the linear forcing-feedback framework Eq. (1) the imbalance is caused by the radiative forcing F which is equivalent to a radiative perturbation imposed by a radiative forcing agent such as CO₂. In the following we explore the temporal dynamics of the energy budget and temperature response of the conceptual models subject to a step (Heaviside) radiative forcing, F . This is conceptually equivalent to the forcing imposed by an abrupt increase in atmospheric CO₂ in a more comprehensive model. The time-dependent unit-area perturbation equation for a zero-dimensional model like Eq. (1) can be generalized by

$$\frac{dE}{dt} = N \tag{3}$$

where N is the downward energy flux perturbation per unit area at TOA, and E is the system’s enthalpy per unit area. We assume that the atmosphere is always in energetic balance because its heat capacity is much smaller than the oceanic heat capacity. This way, we can parameterize the change in the climate system’s enthalpy by the surface temperature perturbation T multiplied by the effective oceanic heat capacity C such that

$$\frac{dE}{dt} \equiv C \frac{dT}{dt} = N. \tag{4}$$

We obtain an ordinary differential equation in which surface temperature is the state-variable of an initial value problem. Different models in the literature propose different parametrizations for N . The ocean heat uptake efficacy model is a globally averaged two-layer model (Held et al. 2010; Geoffroy et al. 2013a, b) and based on two state-variables which represent an upper ocean- or mixed-layer temperature (T , comprising atmosphere, land and ocean) and a deep-ocean temperature (T_D). As a counterpart, we consider a configuration of the pattern effect which comprises two distinct regions. We demonstrate that, on mathematical grounds, these models can describe the same global temperature evolution and radiative response.

3.1 Two-layer model

The two-layer model with ocean heat uptake efficacy is given by

$$C \frac{dT}{dt} = F + \lambda_{eq} T - \varepsilon \eta (T - T_D) \tag{5}$$

$$C_D \frac{dT_D}{dt} = \eta (T - T_D) \tag{6}$$

where $C \ll C_D$ are the heat capacities of the upper- and deep-ocean respectively, λ_{eq} is the equilibrium feedback parameter, η is the heat transport efficiency and ε the efficacy factor for ocean heat uptake. The concept of ocean heat uptake efficacy has been originally introduced by Winton et al. (2010) in analogy to the efficacy of radiative forcing agents (e.g. Hansen 2005). Following Stevens et al. (2016) we prefer to think instead of the radiative response depending on the disequilibrium of the deep-ocean because a certain surface temperature pattern may accompany this disequilibrium. This can be more easily seen by writing the TOA imbalance $N = C \frac{dT}{dt} + C_D \frac{dT_D}{dt}$ as

$$N = F + \lambda_{eq} T - (\varepsilon - 1) \eta (T - T_D). \tag{7}$$

Hence, $(\varepsilon - 1) \eta (T - T_D)$ is an additional radiative flux which is lost to space, while the heat transport into the deep-ocean corresponds to $\eta (T - T_D)$.

To begin with, we briefly characterize basic features of the two-layer model. Subsequently, we project the two-layer model onto a two-region model. The initial change of N with respect to T is

$$\left. \frac{dN}{dT} \right|_{t=0} = \lambda_{eq} - (\varepsilon - 1) \eta, \tag{8}$$

while it approaches λ_{eq} as the mixed-layer and deep-ocean temperatures converge. Thus, equilibrium sensitivity is given by $-F/\lambda_{eq}$ and independent of the coupling between the mixed-layer and deep-ocean, whereas the coupling determines the transient response. More precisely, nonunitary efficacy causes a temporal change in the radiative response that attenuates warming; i.e., $\varepsilon > 1$ leads to a decrease in the magnitude of the radiative response over time as the deep ocean heat uptake changes. Applying Eq. (2), the evolution of $\lambda(t)$ is solely determined by the ratio between T and T_D and it is a monotonically increasing function of time starting from a strong negative feedback. Strictly speaking, the time-variation of $\lambda(t)$, or time-dependent feedback, does not depend explicitly on time but implicitly because it emerges from an additional state-variable which also adjusts on forcing-invariant timescales. In the case of the two-layer ocean model, $\lambda(t)$ increases or decreases gradually over time because it is relative to the surface temperature perturbation T . This gradual increase or decrease, however, leads to an apparent abrupt change in the relationship between N and T , and the strength of this change is determined by the efficacy factor. The abrupt change in MPI-ESM1.2 on decadal timescales implies an efficacy factor above unity.

The analytical solution for each layer, thoroughly described in Geoffroy et al. (2013a, b) but presented here independently to highlight important features, consists of the sum of the

equilibrium response and two modes which increase exponentially towards zero with their characteristic timescales. These modes are characterized by the fast and slow adjustment timescale τ_f and τ_s , their amplitudes ψ_f and ψ_s which sum up to the initial ocean heat uptake temperature, and the mode parameters ζ_f and ζ_s which scale the adjustment of the deep-ocean layer (parameters are listed in Table 1);

$$T(t) = \frac{F}{-\lambda_{eq}} + \psi_f e^{-t/\tau_f} + \psi_s e^{-t/\tau_s} \tag{9}$$

$$T_D(t) = \frac{F}{-\lambda_{eq}} + \zeta_f \psi_f e^{-t/\tau_f} + \zeta_s \psi_s e^{-t/\tau_s}. \tag{10}$$

The amplitudes ψ_f , ψ_s and $\zeta_s \psi_s$ correspond to the negative heat uptake temperatures in the mixed-layer and the deep-ocean, and $\zeta_f \psi_f$ is the positive contribution of the mixed-layer to the deep-ocean adjustment. The fast timescale τ_f is related to the small heat capacity of the mixed-layer and decreases as the coupling $\epsilon\eta$ or the strength of the feedback parameters λ_{eq} increases. The slow timescale τ_s is related to the high heat capacity of the deep-ocean and likewise decreases as the strength of η or λ_{eq} increases. However, the opposite is true for ϵ , because the efficacy factor causes a temporary enhancement of the radiative response such that the heat exchange between the layers is prolonged.

3.2 Two-region model

Regional feedbacks can also be formulated as linearization about a regional energy budget. A common approach is a first-order linearization which gives

$$N(r, t) = F(r, t) + \lambda(r)T(r, t) - Q(r) \tag{11}$$

where r denotes (not necessarily contiguous) regions and $Q(r)$ denotes heat transport between these regions (e.g. Armour et al. 2013; Rose and Rayborn 2016). Even though regional feedbacks $\lambda(r)$ are taken as constant this system may introduce nonconstant global feedback, because regional feedbacks are weighted by a pattern of surface warming. In the following we assume a two-region model in which the climate response in each region is characterized by a specific

heat capacity $C(r)$ and a feedback parameter $\lambda(r)$. In analogy to the two-layer or two-region model we assume Heaviside step forcing input in both regions and five model parameters. The global mean TOA imbalance is given by

$$\bar{N} = F + \chi \lambda_1 T_1 + (1 - \chi) \lambda_2 T_2 \tag{12}$$

where the bar denotes the spatial mean and χ and $(1 - \chi)$ are the corresponding area-fractions. The evolution of $\lambda(t)$ results from the weighting of the regional feedback parameters by the regional temperatures normalized by the global mean temperature perturbation (Armour et al. 2013). Further, we assume that the two regions are not coupled. Regions such as latitudinal bands can hardly be taken as decoupled, since oceanic and atmospheric energy transport alter the regional energy budget significantly. That is, dynamical aspects of climate sensitivity become important and the regional radiative response is related to remote forcing. Adding a parametrization for $-Q(r)$ which scales linearly with the temperature gradient, the characteristic timescales τ_1 and τ_2 are relatively shortened while the nonlinearity between \bar{N} and \bar{T} is partially compensated for (“Appendix”). Langen and Alexeev (2007) and Bates (2012, 2016) focus on implications of perturbation heat transport for the temperature response in tropical-extratropical two-region models. Here, instead, we assume that an AOGCM can be aggregated on two different regions. The analytical solution for regional temperature change in the two-region model is characterized by the sum of the equilibrium response $-F/\lambda(r)$ and a single mode which converges exponentially towards zero on its own (regional) timescale;

$$T(t, r) = \frac{-F}{\lambda(r)} (1 - e^{-t/\tau(r)}). \tag{13}$$

The previously discussed two-layer model with ocean heat uptake efficacy is a coupled model. However, it is a linear system, and this allows us to find analytical relationships between it and the uncoupled two-region model. These relationships are based on the condition that the globally averaged surface temperature T and TOA imbalance N of the two-layer model are equal to the global mean temperature perturbation \bar{T} and the global mean TOA imbalance \bar{N} of the two-region model. We assume that the radiative response of the climate system can be aggregated on these different regions and link the parameters for (regional) feedbacks to the fast component of the climate response and to the slow component of the climate response. The feedback parameters are then given by

$$\lambda_1 = \lambda_{eq} + (\epsilon\eta - \eta)(\zeta_f - 1) \tag{14}$$

$$\lambda_2 = \lambda_{eq} + (\epsilon\eta - \eta)(\zeta_s - 1) \tag{15}$$

where difference in the radiative response coefficients is set by ζ_f and ζ_s . The term ζ_f is negative and scales the

Table 1 Parameters of the analytical solution for the two-layer model

$(b_1, b_2) = (\frac{\lambda_{eq} - \epsilon\eta}{C}, \frac{\epsilon\eta}{C})$
$(b_3, b_4) = (\frac{\eta}{C_D}, -\frac{\eta}{C_D})$
$(\alpha_f, \alpha_s) = \frac{b_1 + b_4}{2} (1 \pm \sqrt{1 - 4 \frac{b_1 b_4 - b_2 b_3}{(b_1 + b_4)^2}})$
$(\tau_f, \tau_s) = (-\alpha_f^{-1}, -\alpha_s^{-1})$
$(\zeta_f, \zeta_s) = (\frac{\alpha_f - b_1}{b_2}, \frac{\alpha_s - b_1}{b_2})$
$(\psi_f, \psi_s) = (\frac{F/\lambda_{eq}(1 - \zeta_s)}{\zeta_f - \zeta_s}, \frac{F/\lambda_{eq}(1 - \zeta_f)}{\zeta_s - \zeta_f})$

contribution of the mixed-layer to the deep-ocean adjustment. The term ζ_s is positive and scales the slow mode or heat uptake temperature of the deep-ocean component itself. Thus, the regional feedback parameters can be decomposed into a background value $\lambda_{eq} - \varepsilon\eta + \eta$ and a contribution from the fast or slow mode. The regional timescales must correspond to τ_f and τ_s and the regional heat capacity is thus given by

$$C(r) = -\lambda(r)\tau(r). \tag{16}$$

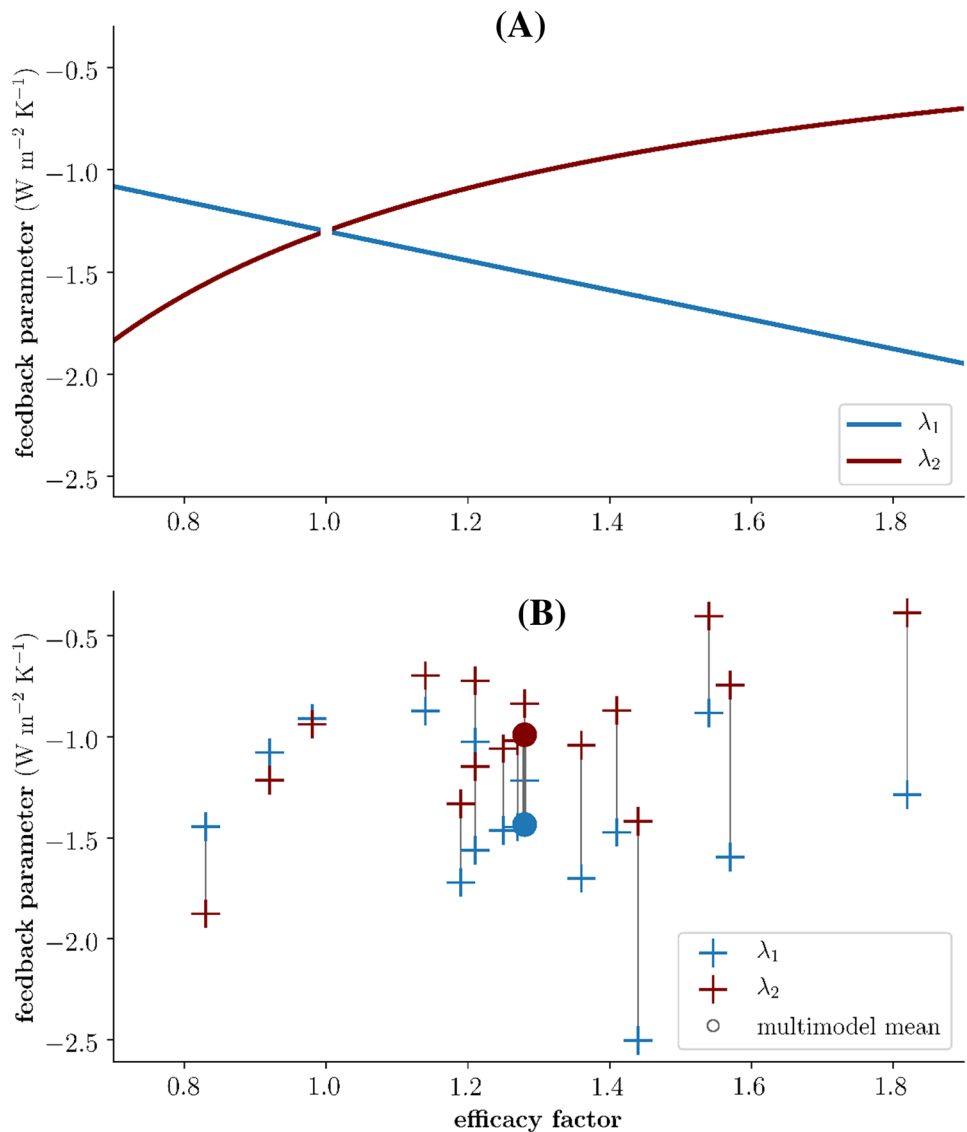
Finally, the regional sensitivity or radiative response is weighted by the corresponding area-fraction χ and $(1 - \chi)$ to result in \bar{T} and \bar{N} . This weighting is done by

$$\chi = \left(\frac{\lambda_2}{\lambda_{eq}} - 1 \right) / \left(\frac{\lambda_2}{\lambda_1} - 1 \right). \tag{17}$$

The projection of the two-layer model with efficacy onto two regions is not limited to Heaviside step forcing input as it is independent of the form of the forcing.

Based on an idealized configuration, in Fig. 3 the feedback parameters λ_1 and λ_2 are plotted as a function of ocean heat uptake efficacy ε . An efficacy factor below unity implies a feedback parameter λ_1 of smaller magnitude compared to a strong feedback parameter λ_2 in the second region. In this case, λ_1 is actuated by fast temperature adjustment, while λ_2 is actuated by slow temperature adjustment on a slow timescale which originally characterized the adjustment due to the deep-ocean component. An efficacy factor above unity, indicated by the MPI-ESM1.2 abrupt CO₂ experiments, implies a strong feedback parameter λ_1 compared to a feedback parameter λ_2 of smaller magnitude in the second region. Likewise, λ_1 is actuated by fast temperature adjustment, while λ_2 is actuated by slow temperature adjustment.

Fig. 3 Ocean heat uptake efficacy and pattern effect. **a** The regional feedback parameters λ_1 and λ_2 as a function of the efficacy factor ε for $\lambda_{eq} = -1.3 \text{ W m}^{-2} \text{ K}^{-1}$, $\eta = 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ while $C = 10 \text{ W year m}^{-2} \text{ K}^{-1}$ and $C_D = 100 \text{ W year m}^{-2} \text{ K}^{-1}$. **b** The projection of CMIP5 model output onto the regional feedback parameters λ_1 and λ_2



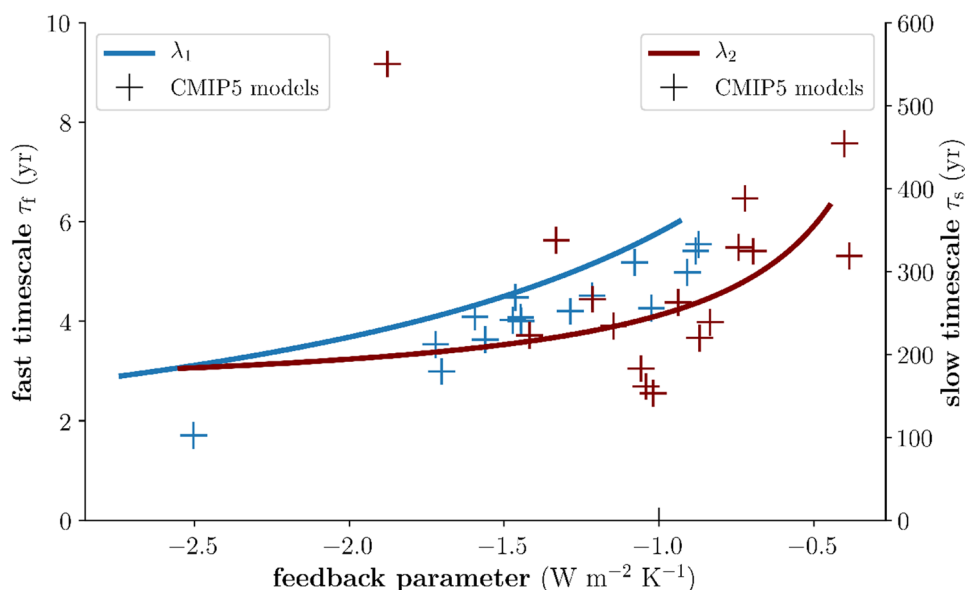
In the case of unitary efficacy, the analytical solution is still characterized by a fast timescale τ_f and a slow timescale τ_s . However, the relationship between T and N in the two-layer model would be linear with slope λ_{eq} and no pattern of surface warming and feedback could be resolved. That is, the fast and the slow adjustment timescales must be directly coupled to radiative feedbacks to cause time-variation of $\lambda(t)$. In Fig. 3b we plot the range of λ_1 and λ_2 for output from the Coupled Model Intercomparison Project (CMIP5) ($4 \times \text{CO}_2$) using parameter estimates from Geoffroy et al. (2013b) who fit the two-layer model to 16 CMIP5 AOGCMs. In line with the MPI-ESM1.2 experiments, most of these models exhibit an efficacy factor above unity ($C_1 \ll C_2$ and $\lambda_1 < \lambda_2$). However, these models differ in their temporal adjustment, which is discussed in the following.

According to the projection of the two-layer model onto the two-region model, we plot the fast timescale τ_f and the slow timescale τ_s against the magnitude of λ_1 and λ_2 (Fig. 4). Again, we use an idealized configuration of the two-layer model (curves) and vary the efficacy factor to compute regional feedbacks. Furthermore, we plot the pairs of regional feedbacks and timescales for the output from CMIP5 models. The regional approach establishes a simple relationship between regional feedbacks and regional timescales, $\tau(r) = -C(r)/\lambda(r)$. That is, the regional timescales are inversely proportional to the feedback parameters. The question arises whether CMIP5 models describe a unique relationship or whether the relationship is appropriate to characterize complex climate model behavior. For a given configuration, the relationship between $\lambda(r)$ and $\tau(r)$ is independent of the global mean equilibrium feedback parameter: using the idealized case, changing the equilibrium feedback parameter λ_{eq} changes the regional feedback parameters

and timescales in such a way that we move along the same curves to higher or lower values. Using CMIP5 output, the fast timescales are approximately linearly related to the regional feedback parameters in the region of fast adjustment. The effective heat capacity of this region is small and τ_f is determined by the magnitude of λ_1 . This is commonly observed in regions which are weakly coupled to the state of the deep-ocean as found in low latitudes. As far as the slow timescale is concerned, the idealized configuration shows an exponential increase of τ_s as λ_2 gets more positive. The CMIP5 models deviate from the idealized configuration because they differ in C , C_D and η , and these model discrepancies induce changes in the temporal behavior of the climate response. The variation of these inertia parameters would change the relationship between $\tau(r)$ and $\lambda(r)$, and this way we would move along a different curve. The model parameters fitted to CMIP5 models, however, do support our general understanding that a more sensitive region adjusts on a longer timescale, and that this timescale changes in a nonlinear way as the magnitude of the radiative feedback is reduced.

How valid is such a regional approach or two-region model conceptually? In their experimental model study based on aquaplanet simulations Haugstad et al. (2017) show that local TOA radiative feedbacks depend on the pattern of the climate forcing modified by ocean heat uptake but the same radiative response arises when the surface temperature pattern induced by that forcing is prescribed. In that respect, ocean heat uptake induces a surface temperature pattern but it is secondary how this surface temperature pattern is induced because the same radiative feedbacks govern the relationship between N and T . We have shown theoretically that, from the perspective of global N and global T , there

Fig. 4 The fast timescale τ_f and the slow timescale τ_s as a function of the magnitude of the regional feedback parameter using the projection of the two-layer model onto the two-region model as well as CMIP5 output. For the idealized case (curves) we set $\lambda_{\text{eq}} = -1.3 \text{ W m}^{-2} \text{ K}^{-1}$, $\eta = 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ while $C = 10 \text{ W year m}^{-2} \text{ K}^{-1}$ and $C_D = 100 \text{ W year m}^{-2} \text{ K}^{-1}$



is no difference in the radiative response and temperature evolution between the two-layer model and the two-region model. Whereas the two-layer and two-region approaches are mathematically equivalent, the former may be attractive in some cases simply because it does not predicate a fixed spatial distribution of regional feedbacks.

4 Feedback temperature dependence

The concepts discussed in the previous section describe a temporal adjustment on forcing-invariant timescales. In this section, following Bloch-Johnson et al. (2015), we add an explicitly state-dependent component and analyze conceptually the implications of feedback temperature dependence for the transient response of the climate system. We first provide insight into the analytical solution of a single mixed-layer to demonstrate that feedback temperature dependence introduces a timescale that depends on the strength of the forcing. Finally, we analyze the interplay between the two-timescale approach and feedback temperature dependence in the case of the two-layer ocean model, since the relationship between N and T in MPI-ESM1.2 suggests both a time-dependent feedback due to the interaction of different state-variables and state-dependent feedback due to feedback temperature dependence.

4.1 Forcing-dependent timescale

Radiative feedbacks may change in strength as the surface temperature T increases, and this temperature dependence can cause a nonlinear temperature response. The first-order feedback framework $\Delta N = \frac{\partial N}{\partial T} \Delta T$ can be extended by retaining terms up to second-order in the Taylor Series expansion of N . This results in nonlinear relationships as

$$\Delta N = \frac{\partial N}{\partial T} \Delta T + \frac{1}{2} \frac{\partial^2 N}{\partial T^2} \Delta T^2 \tag{18}$$

gives a quadratic relationship; i.e., the quadratic coefficient $\frac{1}{2} \frac{\partial^2 N}{\partial T^2} |_T$ relates the change in the radiative feedback to T . We call the quadratic coefficient *feedback temperature dependence* and denote it by a . Using the zero-dimensional model Eq. (1) extended by second-order feedback temperature dependence and solving for the MPI-ESM1.2 equilibrium response for $2 \times \text{CO}_2$, $4 \times \text{CO}_2$ and $8 \times \text{CO}_2$ gives a positive quadratic coefficient a of roughly $0.04 \text{ W m}^{-2} \text{ K}^{-2}$. In the following we assume positive feedback temperature dependence.

We add second-order feedback temperature dependence to a time-dependent mixed-layer approximation to explore associated changes in the temporal behavior. In general, a mixed-layer approximation can be obtained by combining Eqs. (1) and (4). Alternatively, one can assume that $C_D \rightarrow \infty$

such that the two-layer ocean model Eq. (5, 6) reduces to the globally-averaged one-layer model with deep ocean heat uptake.

$$C \frac{dT}{dt} = F + \lambda T - \epsilon \eta T. \tag{19}$$

In Eq. (19) heat is removed from the mixed layer at constant rate $\epsilon \eta$, and mathematically there is no difference to the time-dependent formulation of the linear forcing-feedback framework which describes each zone in the two-region model ($\epsilon \eta$ acts as an additional radiative response parameter). To avoid confusion for later analysis, we aggregate first-order radiative response coefficients in $\Lambda = \lambda - \epsilon \eta$. Again, we assume Heaviside step function input for the radiative forcing F . Adding the quadratic coefficient representing feedback temperature dependence, our problem is described by a first-order nonlinear differential equation with constant coefficients;

$$C \frac{dT}{dt} = F + \Lambda T + aT^2. \tag{20}$$

Substituting

$$T = -\frac{C}{a} \frac{dS/dt}{S} \tag{21}$$

this equation can be transformed into the second-order linear differential equation

$$\frac{d^2 S}{dt^2} = \frac{1}{C} \left(\Lambda \frac{dS}{dt} - \frac{aF}{C} S \right). \tag{22}$$

After re-transformation and applying the initial condition $T(0) = 0$ the analytical solution reads

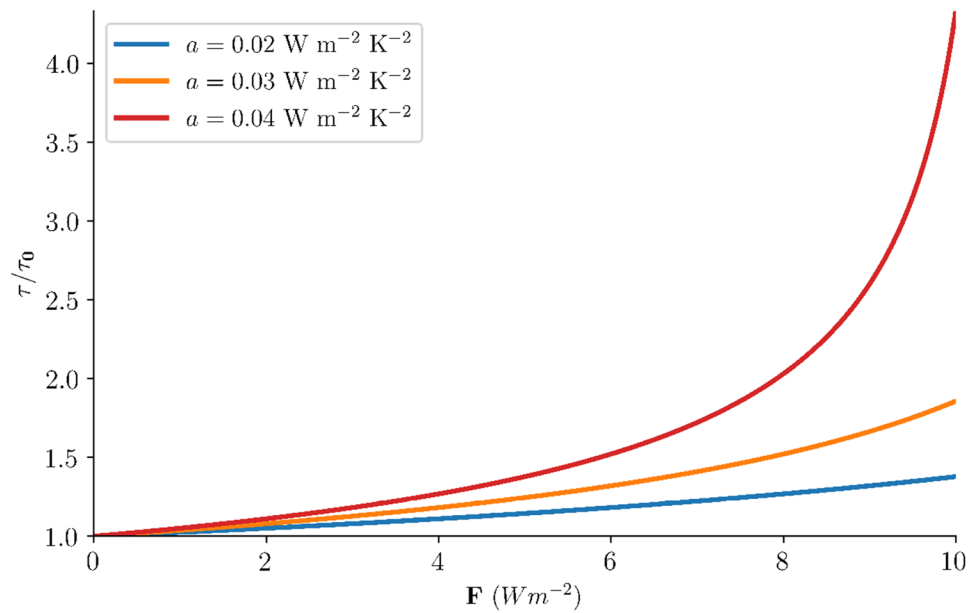
$$T(t) = C \frac{\alpha_2}{a} \frac{e^{\alpha_1 t} - e^{\alpha_2 t}}{\alpha_1} \text{ or } T(t) = C \frac{\alpha_2}{a} \frac{1 - e^{t/\tau}}{\alpha_1 e^{t/\tau} - 1}$$

where $\alpha_{1,2} = \frac{\Lambda}{2C} \pm \frac{\sqrt{\Lambda^2 - 4aF}}{2C}$ and $\tau = \frac{C}{\sqrt{\Lambda^2 - 4aF}}$. (23)

Stability depends additionally on the radiative forcing F and the quadratic coefficient a , and we get a quadratic runaway in the relationship between N and T in the case that $F > \frac{\Lambda^2}{4a}$. Likewise, the characteristic timescale τ on which temperature adjusts depends additionally on F and a , and changes in F or a lead to temporal changes in the adjustment of the system. In this connection, the temporal behavior of the climate response is determined by the strength of the forcing.

A stronger step forcing or a more positive a causes the equilibrium response to be reached far later in time compared to a configuration without feedback temperature dependence. In MPI-ESM1.2 the $8 \times \text{CO}_2$ and $16 \times \text{CO}_2$ experiments are still far from stationarity after 1000 years of integration, while the

Fig. 5 The normalized change of the characteristic timescale τ in the mixed-layer model as a function of F . The radiative response coefficient Λ is set to $\Lambda = -1.3 \text{ W m}^{-2} \text{ K}^{-1}$



$2 \times \text{CO}_2$ experiment is nearly equilibrated. Figure 5 shows an example of forcing-induced changes in the normalized timescale of the mixed-layer model with feedback temperature dependence against $a \rightarrow 0$; i.e.,

$$\tau/\tau_0 = \frac{-\Lambda}{\sqrt{\Lambda^2 - 4aF}} \tag{24}$$

In this case, the choice of the effective heat capacity is arbitrary because the timescale τ is proportional to C . The characteristic timescale increases in a nonlinear way as the forcing is increased, and this nonlinearity is reinforced by a more positive quadratic coefficient. The positive contribution from feedback temperature dependence to the temperature or radiative response is counteracted by the first-order radiative response coefficients comprised in Λ .

4.2 Two timescales and feedback temperature dependence

The one-layer model with feedback temperature dependence adjusts on a single timescale and the relationship between N and T is characterized by a continuous function with smoothly changing slope which can describe the long-term response in MPI-ESM1.2 on centennial timescales. The change in the relationship between N and T on a decadal timescale, however, cannot be reproduced due to the missing multiple timescale structure. We combine the two-timescale approach and state-dependent feedback by adding feedback temperature dependence to the mixed-layer energy budget of the two-layer ocean model Eqs. (5, 6). The TOA imbalance N is now given by

$$N_{\epsilon \neq 1} = F + \lambda_b T + aT^2 - (\epsilon - 1)\eta(T - T_D) \tag{25}$$

with λ_b the background or initial feedback parameter: in Eq. (5) we replace λ_{eq} by $\lambda_T(T) = \lambda_b + aT$. Solving for stationary solutions $-F = \lambda_b T + aT^2$ gives equilibrium sensitivity which is still independent of the coupling between the layers.

We split the surface adjustment T into a fast response T_f when the deep-ocean is not yet significantly perturbed and a slow component T_s when the deep-ocean responds; that is, $T = T_f + T_s$.

This allows us to characterize the interplay between the two-timescale approach and feedback temperature dependence. The fast timescale, on the order of a few years for reasonable parameters, and the related fast response T_f can be approximated by

$$\begin{aligned} \tau_f &\approx C/\sqrt{\Lambda^2 - 4aF} \text{ and} \\ T_f &\approx \frac{-\Lambda - \sqrt{\Lambda^2 - 4aF}}{2a} \text{ with } \Lambda = \lambda_b - \epsilon\eta. \end{aligned} \tag{26}$$

We assume that T_f is constant due to the fast adjustment of the mixed-layer; making the fast adjustment time-dependent leads to Eq. (21) with the analytical solution described by Eq. (23). On timescales longer than τ_f , the slow component is approximated by

$$T_s(t) \approx \frac{\sqrt{\Lambda^2 - 4aF} - \sqrt{\Lambda^2 - 4aF - 4a\epsilon\eta T_D(t)}}{2a} \tag{27}$$

where the term $a\epsilon\eta T_D(t)$ is the time-dependent control. Equations (26) and (27) indicate that there is no linear superposition between the two-timescale approach and feedback

temperature dependence. State-dependent feedback and time-dependent feedback are entangled, and thus indistinguishable in a single forcing realization. If time-dependent feedback and state-dependent feedback were not coupled, we could configure a purely time-dependent model and a purely state-dependent model to obtain the same output.

The choice of parameters, however, determines to what extent the nonlinear behavior of the climate response is related to the fast mode T_f and to the slow mode T_s . We use an idealized case and plot T_f and T_s for different feedback temperature dependencies (Fig. 6a) and for different forcing input (Fig. 6b). For simplicity we choose $F_{2\times\text{CO}_2} = 4 \text{ W m}^{-2}$ and take multiples thereof. Furthermore, we choose a high ocean heat uptake efficacy $\varepsilon = 1.5$. In Fig. 6a the effect of the different feedback temperature dependencies on the fast response is counteracted by $\Lambda = \lambda - \varepsilon\eta$. The first-order

coefficient Λ comprises the high ocean heat uptake efficacy ε and T_f becomes thus approximately independent of a . The effect of feedback temperature dependence on the radiative response is reinforced as the slow component evolves. In contrast, both the fast component and the slow component change significantly as the radiative forcing F is linearly increased (Fig. 6b). The fast response depends sensitively on the magnitude of Λ , and time-dependent feedback dictates to what extent T_f and T_s behave nonlinearly as we change F . In the case of an efficacy factor below unity, the magnitude of the slow mode T_s is reduced compared to an efficacy factor above unity. Likewise, the degree to which the slow mode T_s behaves nonlinearly diminishes, because the radiative response is initially suppressed and subsequently enhanced.

Nonconstant global feedback in MPI-ESM1.2 can be explained by the combined effect of time-dependent

Fig. 6 The fast (dashed) and the slow (solid) mode of the two-layer ocean model with feedback temperature dependence, for **a** different feedback temperature dependencies a and for **b** different step function input for the radiative forcing F . In **(A)** the forcing is set to $F = 12 \text{ W m}^{-2}$ ($\sim 3 \times F_{2\times\text{CO}_2}$). In **b** the quadratic coefficient is set to $a = 0.03 \text{ W m}^{-2} \text{ K}^{-2}$. The background feedback parameter is $\lambda_b = -1.3 \text{ W m}^{-2} \text{ K}^{-1}$, $\eta = 0.7 \text{ W m}^{-2} \text{ K}^{-1}$, $\varepsilon = 1.5$, $C = 10 \text{ W yr m}^{-2} \text{ K}^{-1}$ and $C_D = 100 \text{ W yr m}^{-2} \text{ K}^{-1}$

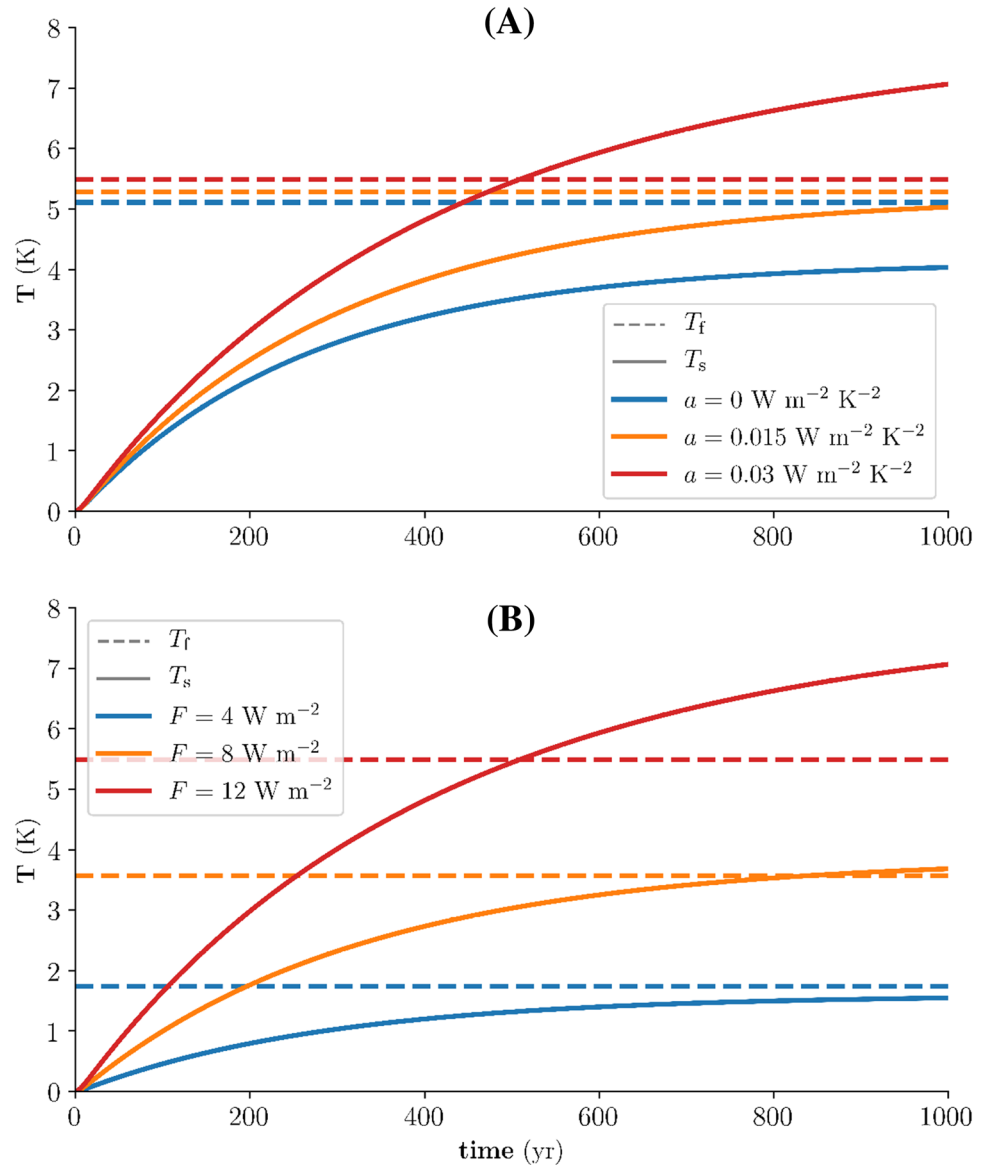
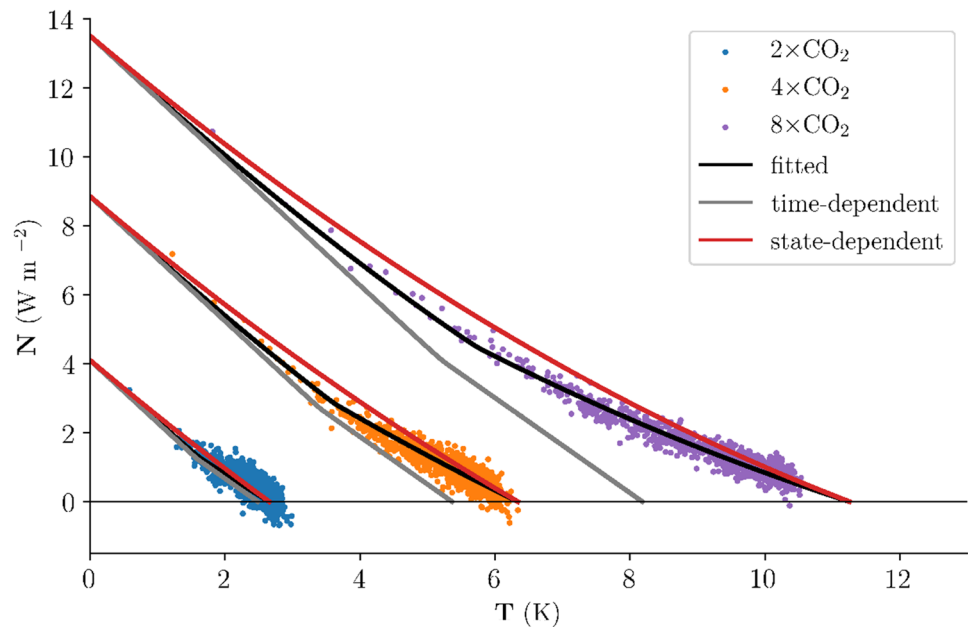


Fig. 7 Time-dependent feedback and state-dependent feedback in MPI-ESM1.2 based on the two-layer ocean model with feedback temperature dependence. The background feedback parameter is set to $\lambda_b = -1.65 \text{ W m}^{-2} \text{ K}^{-1}$, $a = 0.04 \text{ W m}^{-2} \text{ K}^{-2}$, $\eta = 0.8 \text{ W m}^{-2} \text{ K}^{-1}$, $\varepsilon = 1.2$, $C = 5 \text{ W year m}^{-2} \text{ K}^{-1}$ and $C_D = 150 \text{ W year m}^{-2} \text{ K}^{-1}$



feedback and state-dependent feedback. We can determine purely time-dependent feedback and purely state-dependent feedback based on parameter estimates for the two-layer ocean model fitted to AOGCM output. Parameter estimates are taken from Mauritsen et al. (2018) who fit the two-layer model with feedback temperature dependence to the different MPI-ESM1.2 abrupt CO_2 experiments. In Fig. 7 we plot time-dependent feedback ($a = 0$ and $\varepsilon \neq 1$) and state-dependent feedback ($a \neq 0$ and $\varepsilon = 1$) in these experiments. Time-dependent feedback can explain the abrupt change in the relationship between N and T in the MPI-ESM1.2 on decadal timescales as well as the variation in the $2 \times \text{CO}_2$ experiment. For larger forcing such as $4 \times \text{CO}_2$ and $8 \times \text{CO}_2$, the effect of feedback temperature dependence on the radiative response becomes increasingly important and dictates the long-term response. We could quantify contributions from a and ε to the change of N or the variation of $\lambda(t)$ in the combined run ($a \neq 0$ and $\varepsilon \neq 1$). However, these contributions are relative to the temperature adjustments of the system while the characteristic timescales on which the system adjust themselves depend on F ; i.e., these contributions are related to the combined effect of time-dependent and state-dependent feedback. Interpreting time-dependent feedback in AOGCM experiments when both time-dependent feedback and state-dependent feedback are present thus depends sensitively on the conceptual framework chosen and the assumptions made on the temperature-dependent feedback parameter.

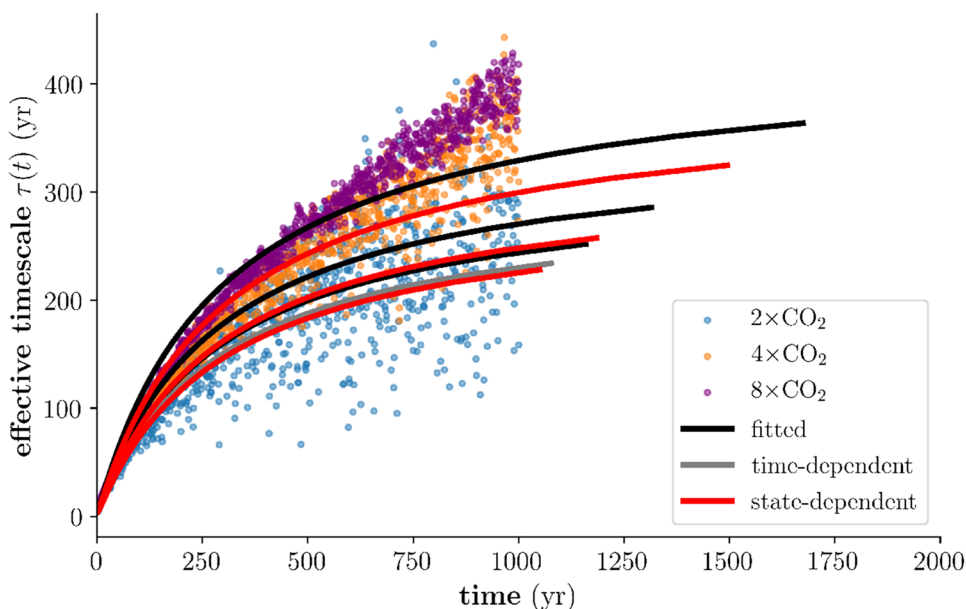
To highlight the differences in the temporal adjustment between state-dependent and time-dependent models as well as their combination, we define a timescale $\tau(t)$ which is allowed to vary over time. We define this timescale as

$\tau(t) = -t / \ln\left(\frac{T(t)-T_{\text{eq}}}{-T_{\text{eq}}}\right)$ where $T(t)$ is the globally averaged surface temperature perturbation and T_{eq} is the equilibrium response. That is, we assume a single mode. Figure 8 shows $\tau(t)$ for purely time-dependent feedback, purely state-dependent feedback and their fitted combination. In general, $\tau(t)$ increases monotonically from relatively small values which are related to the fast mode of the climate response to relatively high values which are related to amplified warming due to the slow component on a centennial timescale. In the case of purely time-dependent feedback, the variation of $\tau(t)$ over time is independent of the strength of the radiative forcing F , since the characteristic timescales of the two-timescale approaches are constant and the amplitudes of the temperature adjustment are linear in forcing. In the case of purely state-dependent feedback due to feedback temperature dependence and unitary efficacy, the variation of $\tau(t)$ depends additionally on the radiative forcing F . Due to the interaction of time-dependent feedback and state-dependent feedback the temporal adjustment is even prolonged and the effective timescales deviate from purely state-dependent feedback despite the same equilibrium response. Understanding the forcing-dependency of the characteristic adjustment and its interaction with time-dependent feedbacks is therefore important to narrow the timing of transient warming on a centennial timescale.

5 Conclusion

The two-timescale approaches give rise to time-dependent feedback but we paraphrase time-dependence as implicit because it emerges from the interplay of at least two state-variables and relates to the evolution of the radiative

Fig. 8 The timescale $\tau(t)$ for time-dependent feedback and state-dependent feedback in MPI-ESM1.2 based on the two-layer ocean model with feedback temperature dependence. The background feedback parameter is set to $\lambda_b = -1.65 \text{ W m}^{-2} \text{ K}^{-1}$, $a = 0.04 \text{ W m}^{-2} \text{ K}^{-2}$, $\eta = 0.8 \text{ W m}^{-2} \text{ K}^{-1}$, $\epsilon = 1.2$, $C = 5 \text{ W yr m}^{-2} \text{ K}^{-1}$ and $C_D = 150 \text{ W yr m}^{-2} \text{ K}^{-1}$. In the case of time-dependent feedback, the lines simply lie atop one another



response over time rather than to an explicit time-dependence. The recent conceptual frameworks of the two-timescale approaches, the two-layer model with efficacy and the two-region model, can describe the same global temperature evolution and radiative response. On mathematical grounds, the two-layer ocean model with nonunitary efficacy can be projected onto a two-region model, and we can use the regional model feedbacks and timescales to characterize complex climate model behavior. The mathematical equivalence raises the question which specific mechanism causes time-dependent feedback. We cannot reject one or the other framework given only global mean output for the TOA imbalance N and surface temperature change T . We suggest that the radiative response of the climate system can be aggregated on different regions such as regions directly coupled to the state of the deep-ocean and not directly coupled to the state of the deep-ocean.

In contrast to time-dependent feedback, state-dependent feedback due to feedback temperature dependence introduces a nonlinearity which makes the timescale on which the climate system adjusts forcing-dependent. The finding that the timescales on which the climate system adjusts depend on the strength of the forcing has not been addressed to date. As we alter the forcing, we change the system's response times and a certain fraction of the equilibrium response or temperature level is reached earlier or later in time. Further, this forcing-dependency and the combined effect or entanglement of time-dependent feedback and state-dependent feedback poses challenges for determining contributions from different sources such as feedback temperature dependence or a pattern effect. Single forcing realizations may not constrain the model parameters, and this complicates attempts to understand temperature change especially

in more strongly forced climates on a centennial timescale but also the details of the instrumental record.

Acknowledgements Open access funding provided by Max Planck Society. This study was supported by the Max-Planck-Gesellschaft (MPG) and TM received funding from the European Research Council (ERC) Consolidator grant 770765. Computational resources were made available by Deutsches Klimarechenzentrum (DKRZ) through support from Bundesministerium für Bildung und Forschung (BMBF).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Solution to a linear two-dimensional first-order differential system

A linear two-dimensional first-order ordinary differential system as the two-layer ocean model or the two-region model with heat exchange between these regions is described by

$$\frac{d}{dt} \mathbf{X} = \mathbf{Z} \mathbf{X} + \mathbf{F},$$

$$\mathbf{X}(t) = \begin{pmatrix} T_1(t) \\ T_2(t) \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}, \mathbf{F}(t) = \begin{pmatrix} \phi_1(t) \\ \phi_2(t) \end{pmatrix} \quad (28)$$

where ϕ_i are the time-dependent forcing functions. The general solution of the homogenous system is described by

$$\mathbf{X}_H(t) = \psi_1 \mathbf{K}_1 e^{\alpha_1 t} + \psi_2 \mathbf{K}_2 e^{\alpha_2 t} \quad (29)$$

where α_1 and α_2 are the eigenvalues of the system and \mathbf{K}_1 and \mathbf{K}_2 are the corresponding eigenvectors. The eigenvalues α_1 and α_2 satisfy

$$\det(\mathbf{Z} - \alpha\mathbf{E}) = 0, \quad \alpha_{1,2} = \frac{z_1 + z_4}{2} \left(1 \pm \sqrt{1 - 4 \frac{z_1 z_4 - z_2 z_3}{(z_1 + z_4)^2}} \right). \tag{30}$$

The eigenvectors are given by

$$\mathbf{K}_1 = \begin{pmatrix} 1 \\ \kappa_1 \end{pmatrix} \text{ and } \mathbf{K}_2 = \begin{pmatrix} 1 \\ \kappa_2 \end{pmatrix} \text{ such that} \tag{31}$$

$$\kappa_1 = \frac{\alpha_1 - z_1}{z_2} \text{ and } \kappa_2 = \frac{\alpha_2 - z_1}{z_2}$$

Given a nonzero forcing function ϕ , the analytical solution is the sum of the general solution of the homogenous system $\mathbf{X}_H(t)$ and a particular solution $\mathbf{X}_P(t)$ to the non-homogeneous part. The particular solution $\mathbf{X}_P(t)$ depends on the forcing functions comprised by $\mathbf{F}(t)$. Heaviside step function input in both dimensions reads

$$\mathbf{F}(t) = \begin{pmatrix} \phi_1(t) \\ \phi_2(t) \end{pmatrix} = \begin{pmatrix} \Delta F_1 \times 1(t) \\ \Delta F_2 \times 1(t) \end{pmatrix}, \tag{32}$$

and the particular solution is given by the equilibrium response

$$\mathbf{X}_P = \begin{pmatrix} \Delta T_1 \\ \Delta T_2 \end{pmatrix}, \quad \Delta T_1 = \frac{z_2 C_1 \Delta F_2 - z_4 C_2 \Delta F_1}{z_1 C_1 z_4 C_2 - z_2 C_1 z_3 C_2} \text{ and} \tag{33}$$

$$\Delta T_2 = \frac{z_3 C_2 \Delta F_1 - z_1 C_1 \Delta F_2}{z_1 C_1 z_4 C_2 - z_2 C_1 z_3 C_2}.$$

In the case of the two-layer ocean model Eqs. (5, 6), $\phi_2(t) = 0$, or $\Delta F_2 = 0$, which simplifies the preceding equations to $\Delta T_{1,2} = -\Delta F_1 / \lambda$. Finally, the constants ψ_1 and ψ_2 are obtained from the initial conditions $T_1(0) = 0$ and $T_2(0) = 0$ such that

$$\psi_1 = \frac{\kappa_2 \Delta T_1 - \Delta T_2}{\kappa_1 - \kappa_2} \text{ and } \psi_2 = \frac{\kappa_1 \Delta T_1 - \Delta T_2}{\kappa_2 - \kappa_1}. \tag{34}$$

Perturbation heat transport

The regional approach can be extended by adding a parametrization for perturbation heat transport $-Q(r)$ between the regions. Bates (2016) proposed a parametrization for baroclinic eddy transport; Langen and Alexeev (2007) proposed a parametrization which also accounts for greater latent heat transport with increased low- and midlatitude temperatures;

$$-Q(r) = -\eta_1(T_1 - T_2) - \eta_2 T_1. \tag{35}$$

Table 2 Parameters of the analytical solution for the two-region model with perturbation heat transport

$(z_1, z_2) = \left(\frac{\lambda_1 - \eta_1 - \eta_2}{C_1}, \frac{\eta_1}{C_1} \right)$
$(z_3, z_4) = \left(\frac{\eta_1 + \eta_2}{C_2}, \frac{\lambda_2 - \eta_1}{C_2} \right)$
$(v_1, v_2) = \frac{z_1 + z_4}{2} \left(1 \pm \sqrt{1 - 4 \frac{z_1 z_4 - z_2 z_3}{(z_1 + z_4)^2}} \right)$
$(\tau_1, \tau_2) = (-\alpha_1^{-1}, -\alpha_2^{-1})$
$(\kappa_1, \kappa_2) = \left(\frac{\alpha_1 - z_1}{z_2}, \frac{\alpha_2 - z_1}{z_2} \right)$
$\Delta T_1 = \frac{\Delta F(z_2 C_1 - z_4 C_2)}{z_1 C_1 z_4 C_2 - z_2 C_1 z_3 C_2}$
$\Delta T_2 = \frac{\Delta F(z_3 C_2 - z_1 C_1)}{z_1 C_1 z_4 C_2 - z_2 C_1 z_3 C_2}$
$(\psi_1, \psi_2) = \left(\frac{\zeta_2 \Delta T_1 - \Delta T_2}{\zeta_1 - \zeta_2}, \frac{\zeta_1 \Delta T_1 - \Delta T_2}{\zeta_2 - \zeta_1} \right)$

The first term mimics the increase in transport with increasing baroclinicity, while the second term is only proportional to the tropical temperature to mimic the change in latent heat transport. The related two-region model is described by

$$C_1 \frac{dT_1}{dt} = F + \lambda_1 T_1 - \eta_1(T_1 - T_2) - \eta_2 T_1 \tag{36}$$

and

$$C_2 \frac{dT_2}{dt} = F + \lambda_2 T_2 + \eta_1(T_1 - T_2) + \eta_2 T_1. \tag{37}$$

The analytical solution to the same Heaviside step forcing $\Delta F \times 1(t)$ in both region reads

$$T_1(t) = \Delta T_1 + \psi_1 e^{-t/\tau_1} + \psi_2 e^{-t/\tau_2} \tag{38}$$

and

$$T_2(t) = \Delta T_2 + \kappa_1 \psi_1 e^{-t/\tau_1} + \kappa_2 \psi_2 e^{-t/\tau_2}. \tag{39}$$

At first glance, this solution looks similar to the two-layer ocean model but each region converges towards the local equilibrium temperature ΔT_1 and ΔT_2 (Table 2). Taking the spatial mean results in global sensitivity while local instability in one region does not imply global instability. Compared to the uncoupled two-region model, global sensitivity is decreased in the case that $\eta_1 > \eta_2$. Further, the characteristic timescales τ_1 and τ_2 are relatively shortened and the nonlinearity between \bar{N} and \bar{T} is partially compensated for. More precisely, positive η_1 decreases τ_1 and τ_2 , while positive η_2 decreases τ_1 but partially compensates for the decrease in τ_2 . To find a simple expression describing the temporal behavior, the fast and slow timescale can be approximated using the diagnostic limit such that

$$\tau_f \approx \frac{C_1}{\eta_1 + \eta_2 - \lambda_1} \text{ and } \tau_s \approx \frac{C_2}{\eta_1 - \lambda_2 - \frac{\eta_1^2 + \eta_1 \eta_2}{\eta_1 + \eta_2 - \lambda_1}} \tag{40}$$

while we assume that $C_1 \ll C_2$ and $\lambda_1 < \lambda_2$.

References

- Andrews T, Gregory J, Webb M (2015) The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J Clim* 28:1630–1648. <https://doi.org/10.1175/JCLI-D-14-00545.1>
- Armour KC, Bits CM, Roe GH (2013) Time-varying climate sensitivity from regional feedbacks. *J Clim* 26:4518–4534. <https://doi.org/10.1175/JCLI-D-12-00544.1>
- Arrhenius S (1896) On the influence of carbonic acid in the air upon temperature of the ground. *Philos Mag J Sci* 41:237–276
- Bates JR (2012) Climate stability and sensitivity in some simple conceptual climate models. *Clim Dyn* 38:455–473. <https://doi.org/10.1007/s00382-010-0966-0>
- Bates JR (2016) Estimating climate sensitivity using two-zone energy balance models. *ESS* 3:207–225. <https://doi.org/10.1002/2015EA000154>
- Bloch-Johnson J, Pierrehumbert RT, Abbot DS (2015) Feedback temperature dependence determines the risk of high warming. *Geophys Res Lett* 42:4973–4980. <https://doi.org/10.1002/2015GL064240>
- Budyko M (1969) The effect of solar radiation variations on the climate of the earth. *Tellus* 21:611–619. <https://doi.org/10.3402/tellusa.v21i5.10109>
- Geoffroy O, Saint-Martin D, Olivie D, Voldoire A, Bellon G, Tyteca S (2013a) Transient climate response in a two-layer energy-balance model. Part I: analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J Clim* 26:1841–1856. <https://doi.org/10.1175/JCLI-D-12-00195.1>
- Geoffroy O, Saint-Martin D, Olivie D, Voldoire A, Bellon G, Tyteca S (2013b) 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 Clim* 26:1859–1875. <https://doi.org/10.1175/JCLI-D-12-00196.1>
- Gregory JM, Ingram WJ, Palmer MA, Jones GS, Stott PA, Thorpe RB, Lowe JA, Johns TC, Williams KD (2004) A new method for diagnosing radiative forcing and climate sensitivity. *Geophys Res Lett* 31. <https://doi.org/10.1029/2003GL018747>
- Hansen J et al (2005) Efficacy of climate forcings. *J Geophys Res* 110. <https://doi.org/10.1029/2005JD005776>
- Hasselmann K (1979) On the problem of multiple time scales in climate modeling. *Dev Atmos Sci* 10:43–55. <https://doi.org/10.1016/b978-0-444-41766-4.50011-4>
- Haugstad AD, Armour KC, Battisti DS, Rose BEJ (2017) Relative roles of surface temperature and climate forcing patterns in the inconstancy of radiative feedbacks. *Geophys Res Lett* 44. <https://doi.org/10.1002/2017GL074372>
- Hedemann C, Mauritsen T, Jungclaus J, Marotzke J (2019) The evolution of the pattern effect in response to CO₂ forcing. manuscript submitted, *J Clim*
- Held IM, Winton M, Takahashi K, Delworth T, Zeng F, Vallis GK (2010) Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J Clim* 23:2418–2427. <https://doi.org/10.1175/2009JCLI3466.1>
- Langen PL, Alexeev VA (2007) Polar amplification as a preferred response in an idealized aquaplanet GCM. *Clim Dyn* 29:305–317. <https://doi.org/10.1007/s00382-006-0221-x>
- Mauritsen T et al (2018) Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂. manuscript submitted, JAMES
- Meraner K, Mauritsen T, Voigt A (2013) Robust increase in equilibrium climate sensitivity under global warming. *Geophys Res Lett* 40:5944–5948. <https://doi.org/10.1002/2013GL058118>
- Roe GH, Armour KC (2011) How sensitive is climate sensitivity? *Geophys Res Lett* 38. <https://doi.org/10.1029/2011GL047913>
- Roe GH, Baker MB (2007) Why is climate sensitivity so unpredictable. *Science* 318(5850):629–632. <https://doi.org/10.1126/science.1144735>
- Rose B, Rayborn L (2016) The effects of ocean heat uptake on transient climate sensitivity. *Curr Clim Change Rep* 2:190–202. <https://doi.org/10.1007/s40641-016-0048-4>
- Sellers W (1969) A global climatic model based on the energy balance of the earth-atmosphere system. *J Appl Meteorol* 8:392–400
- Senior CA, Mitchell JFB (2000) The time-dependence of climate sensitivity. *Geophys Res Lett* 27(17):2685–2688. <https://doi.org/10.1029/2000GL011373>
- Stevens B, Sherwood SC, Bony S, Webb MJ (2016) Prospects for narrowing earth's equilibrium climate sensitivity. *Earths Future* 4:512–522. <https://doi.org/10.1002/2016EF000376>
- Williams KD, Ingram WJ, Gregory JM (2008) Time variation of effective climate sensitivity in gcms. *J Clim* 21. <https://doi.org/10.1175/2008JCLI2371.1>
- Winton M, Takahashi K, Held IM (2010) Importance of ocean heat uptake efficacy to transient climate change. *J Clim* 23:2333–2344. <https://doi.org/10.1175/2009JCLI3139.1>
- Zaliapin I, Ghil M (2010) Another look at climate sensitivity. *Nonlinear Proc Geophys* 17:113–122. <https://doi.org/10.5194/npg-17-113-2010>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.