Geophysical Research Letters®

Key Points:

- A multivariate climate feedback framework is introduced that takes into account the transient state dependency of climate feedbacks
- Using the new framework, changes in feedback processes can be analyzed per time scale and temporal evolution can be tracked
- Within the framework, it is possible to create transient and equilibrium projections of (the spatial patterns of) climate feedbacks

Correspondence to:
R. Bastiaansen,
r.bastiaansen@uu.nl

Projections of the Transient State-Dependency of Climate Feedbacks

Robbin Bastiaansen[1] [ID], **Henk A. Dijkstra[1,2]** [ID], and **Anna S. von der Heydt[1,2]** [ID]

[1]Department of Physics, Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, The Netherlands, [2]Department of Physics, Centre for Complex System Studies, Utrecht University, Utrecht, The Netherlands

Abstract When the climate system is forced, for example, by the emission of greenhouse gases, it responds on multiple time scales. As temperatures rise, feedback processes might intensify or weaken. Such state dependencies cannot be fully captured with common linear regression techniques that relate feedback strengths linearly to changes in the global mean temperature. Hence, transient changes are difficult to track and it becomes easy to underestimate future warming this way. Here, we present a multivariate and spatial framework that facilitates the dissection of climate feedbacks over time scales. Using this framework, information on the composition of projected transient future climates and feedback strengths can be obtained. The new framework is illustrated using the Community Earth System Model version 2.

Plain Language Summary When the Earth warms, the internal processes of the climate system change. This can lead to additional warming, forming a feedback loop. For instance, as the ice melts due to increased temperatures, less solar radiation gets reflected back to outer space, causing temperatures to rise even more. To properly understand and assess (future) climatic changes, it is, therefore, necessary to quantify these so-called climate feedbacks and track how they change over time. However, with current techniques, it is not easy to explicitly track these temporal changes, which hampers the interpretation of long-term projections of the future climate state and easily leads to systematic underestimation of future (committed) warming. Here, we present a new feedback framework that can capture the temporal evolution of climate feedbacks. Consequently, a better insight into the development of a (projected) future climate is obtained, because not only global mean temperatures can be tracked, but also the temporal change in individual climate feedbacks, including their spatial distribution.

1. Introduction

Were the climate system free of feedback processes, it would be easy to predict and control the future climate (Arrhenius, 1896). Unfortunately, the climate system contains many feedback loops (Heinze et al., 2019; Von der Heydt et al., 2020). Because of this, climate change can get suppressed or enhanced, making future projections hard (Sherwood et al., 2020). Therefore, detailed knowledge of all relevant feedback processes is required to accurately assess potential future climates. However, knowledge of the current strengths of climate feedbacks is not enough. Over time, as the climate state changes, the strengths of climate feedbacks also change (Armour et al., 2013; Gregory & Andrews, 2016; Marvel et al., 2018); for instance, the albedo-increasing effect of ice is only relevant when there still is ice.

As the Earth warms, the strengths of feedbacks change (Bony et al., 2006). Therefore, projections based only on current knowledge of climate feedbacks misestimate future climate change–especially the committed warming that is to come even if zero-emission is achieved (Goodwin, 2018; Marvel et al., 2018; Senior & Mitchell, 2000). To properly assess different emission scenarios, it is crucial to identify all relevant feedback mechanisms and, additionally, to quantify how their strengths change over time.

For future temperature projections with global climate models, the focus lies with the following feedback processes (Cess, 1975; Klocke et al., 2013; Zelinka et al., 2020). First, the Planck radiation feedback suppresses warming due to increased outgoing radiation. Second, the lapse rate feedback also suppresses warming due to an increase in long-wave radiation escaping to space (Sinha, 1995). The third is the ice-albedo feedback that enhances warming through a less reflective Earth surface (Curry et al., 1995). Fourth is the water vapor feedback which boosts warming because of increased atmospheric water vapor content (Manabe &

Wetherald, 1967). Finally, there is the cloud feedback, describing changes in cloud type and distribution, which could enhance or suppress warming (Cess & Potter, 1987).

The effect of a feedback process is quantified by its contribution to the climate system's radiative response to a certain experienced radiative forcing—in the current Anthropocene primarily caused by greenhouse gas emissions. Specifically, the top-of-atmosphere radiative imbalance ΔN (that equals zero when the climate system is in equilibrium) is given by the sum of the radiative forcing F and the radiative response ΔR, that is,

$$\Delta N(t) = F(t) + \Delta R(t). \tag{1}$$

This response ΔR is the sum of the *feedback contributions* of all relevant feedbacks. So, writing ΔR_j for the feedback contribution of feedback j,

$$\Delta R(t) = \sum_{j \in \mathcal{F}} \Delta R_j(t), \tag{2}$$

where the sum is over all feedback processes. Classically, the *feedback strength* λ_j (also called *feedback parameter*) is given by the feedback contribution per unit warming, that is,

$$\lambda_j := \frac{\Delta R_j(t)}{\Delta T(t)}. \tag{3}$$

From this, the top-of-atmosphere radiative imbalance ΔN can be related to the warming ΔT as

$$\Delta N(t) = F(t) + \left(\sum_{j \in \mathcal{F}} \lambda_j \right) \Delta T(t), \tag{4}$$

which can be used for climate projections (Gregory et al., 2004).

However, the climate system responds on many, vastly different time scales (Von der Heydt et al., 2020). In particular, not all feedback processes react similarly on all of these time scales. Consequently, there is no linear relationship between all feedback contributions ΔR_j and the global warming ΔT (Andrews et al., 2015; Armour, 2017; Knutti & Rugenstein, 2015). Hence, computations of feedback strengths λ_j often find these to change over time (Boer et al., 2005; Klocke et al., 2013; Meraner et al., 2013; Senior & Mitchell, 2000; Zelinka et al., 2020); they are certainly not constant, as seems implied by Equation 3.

We propose here to extend the classical framework to include dynamics on multiple time scales. To this end, feedback strengths should not be computed for the total time series, as in Equation 3, but they should be separated in modal contributions instead. This makes the time-dependency of the various feedback processes explicit. Moreover, when combined with linear response theory (Lucarini & Sarno, 2011; Ragone et al., 2016; Ruelle, 2009), it opens the possibility of projections for individual feedback contributions under any forcing scenario without having to perform the model simulations. This has been used for temperature projections already in many models, including global climate models (Aengenheyster et al., 2018; Hasselmann et al., 1993; Lembo et al., 2020; Maier-Reimer & Hasselmann, 1987).

The rest of this paper is structured as follows. In Section 2, we describe the theoretical framework, define new feedback strength metrics, and explain the computations of the feedback contributions. In Sections 3.1 and 3.2, the framework is applied in a study of climate feedbacks in a 1,000 year long experiment in the "Community Earth System Model version 2" (CESM2), in which atmospheric CO_2 was quadrupled from the start (a so-called "abrupt4×CO_2" experiment). In Section 3.3, as illustrative example, projections of feedback contributions are made for and compared against output of an experiment in which atmospheric CO_2 was gradually increased with 1% yearly until quadrupling and kept constant thereafter (a "1pctCO_2" experiment). Finally, Section 4 contains a summary and discussion.

2. Method and Model

2.1. Method: The New Climate Feedback Framework

In the linear regime of a climate model, the change (response) in an observable O due to a specific forcing g is given by

$$\Delta O(t) = \left(G^{[O]} * g \right)(t) = \int_0^t G^{[O]}(t - s) \, g(s) ds, \tag{5}$$

where $G^{[O]}$ is the Green function for observable O (Aengenheyster et al., 2018; Hasselmann et al., 1993; Maier-Reimer & Hasselmann, 1987; Proistosescu & Huybers, 2017). Many Green functions in the climate system seem well-approximated by a sum of decaying exponentials, with only the contribution of the different modes depending on the observable O. That is,

$$G^{[O]}(t) = \sum_{m=1}^M \beta_m^{[O]} \, e^{-t/\tau_m}, \tag{6}$$

where the sum is over all M modes, τ_m is the m-th mode's time scale and $\beta_m^{[O]}$ is the contribution of mode m to the observable's Green function. Because of this, following for example, Proistosescu and Huybers (2017), the response of observable O can also be written as

$$\Delta O(t) = \sum_{m=1}^M \beta_m^{[O]} \, \mathcal{M}_m^g(t), \tag{7}$$

where $\mathcal{M}_m^g(t) := \int_0^t e^{-(t-s)/\tau_m} g(s) ds$ contains all time-dependencies and $\beta_m^{[O]}$ all observable-dependencies. Since all feedback contributions ΔR_j and the global warming ΔT adhere to this formulation, it is possible to define the feedback strength of feedback j per mode m as

$$\lambda_j^m := \frac{\beta_m^{[R_j]}}{\beta_m^{[T]}}. \tag{8}$$

Additionally, the instantaneous feedback strength at time t can be defined as the incremental change in feedback contribution ΔR_j per incremental change in temperature ΔT (i.e., the local slope of the graph $(\Delta T, \Delta R_j)$ at time t):

$$\lambda_j^{\text{inst}}(t) := \frac{\dfrac{d}{dt}\Delta R_j(t)}{\dfrac{d}{dt}\Delta T(t)} \tag{9}$$

Using this framework, it is also possible to estimate properties of the eventual equilibrium state—if such state exists. For instance, if $g(t) \equiv \overline{g}$, $\mathcal{M}_m^g(t) = \overline{g}\tau_m \left[1 - e^{-t/\tau_m} \right]$, the equilibrium ΔO^* is given by

$$\Delta O^* = \sum_{m=1}^M \beta_m^{[O]} \, \overline{g} \, \tau_m. \tag{10}$$

The equilibrium feedback strength λ_j^* for feedback j is then

$$\lambda_j^* := \frac{\Delta R_j^*}{\Delta T^*} = \frac{\sum_{m=1}^M \tau_m \, \beta_m^{[R_j]}}{\sum_{m=1}^M \tau_m \, \beta_m^{[T]}}. \tag{11}$$

A more elaborate treatment of the mathematics of this framework is given in Text S1 in Supporting Information S1. Below, we use this framework on two idealized experiments in CESM2 to understand the temporal evolution of the strength of climate feedbacks in these scenarios.

The above equations can be used empirically to analyze a single forcing experiment (Proistosescu & Huybers, 2017). However, if an ensemble of runs are used and the expressions are interpreted as ensemble averages, rigorous results from linear response theory become available (Lucarini & Sarno, 2011; Ragone

et al., 2016; Ruelle, 2009). In particular, then Green functions obtained from one forcing experiment can be used to make projections for the ensemble average in another forcing scenario by changing the forcing in Equation 7. If such procedure is applied on a single run—like is done in Section 3.3 for illustrative purpose—results may not be correct quantitatively.

2.2. Model: CESM2 Abrupt4×CO$_2$ and 1pctCO$_2$ Experiments

We focus on forcing experiments conducted with the model CESM2 (Danabasoglu et al., 2020), specifically on the model's abrupt4×CO$_2$ (1,000y long) and 1pctCO$_2$ (150y long) experiments in CMIP6 (Eyring et al., 2016). The abrupt4×CO$_2$ experiment is used for fitting of the parameters and computations of feedback strengths per mode; the 1pctCO$_2$ experiment is used to illustrate the possibilities of projections made within the presented framework.

As feedback contributions are not outputted directly by any climate model, these have been determined using additional methodology. For this, we have used the radiative kernel approach (Shell et al., 2008; Soden et al., 2008), in which feedback contributions are computed (per grid point at location **x**) as

$$\Delta R_j(\mathbf{x},t) = \frac{\partial \Delta N}{\partial y_j}(\mathbf{x}) \, \Delta y_j(\mathbf{x},t),\tag{12}$$

where y_j is the relevant (derived) state variable for the j-th feedback and $\partial N/\partial y_j(\mathbf{x})$ is the pre-computed radiative kernel. We have used the publicly available CESM-CAM5 kernels from Pendergrass et al. (2018); Pendergrass (2017a, 2017b). There are others available, but differences between kernels are typically small. The procedure for the cloud feedback contribution is somewhat different because of its more nonlinear nature, but is nevertheless computed following common practice (Klocke et al., 2013; Shell et al., 2008; Soden et al., 2008)—see Text S4 in Supporting Information S1 for more details.

3. Results

3.1. Temporal Evolution, Abrupt4×CO$_2$

From CESM2 model output, feedback contributions $\Delta R_j(\mathbf{x},t)$, top-of-atmosphere radiative imbalance $\Delta N(\mathbf{x},t)$ and near-surface temperature increase $\Delta T(\mathbf{x},t)$ were computed; changes in observables are determined with respect to the temporal average of a 1,200 year control experiment with pre-industrial CO$_2$ levels. From these, globally averaged datasets $\Delta R_j(t)$, $\Delta N(t)$ and $\Delta T(t)$ were derived. These global time series have been fitted to Equation 7, with exception of $\Delta N(t)$ that also includes the radiative forcing and is fitted to

$$\Delta N(t) = F_{\text{abr}} + \Delta R(t) = F_{\text{abr}} + \sum_{m=1}^{M} \beta_m^{[R]} \mathcal{M}_m^{g_{\text{abr}}}(t),\tag{13}$$

where the subscript "abr" indicates that the forcings are from the abrupt4×CO$_2$ experiment. More details of the procedure can be found in Text S2 in Supporting Information S1.

First, the amount of modes M needs to be determined. Tests with various values for M have indicated $M = 3$ is the best choice here: $M < 3$ does not capture all dynamics and $M > 3$ has no significant benefits; see Figure S1 in Supporting Information S1. This is also in agreement with other reports that typically find two or three relevant modes (Bastiaansen et al., 2020; Caldeira & Myhrvold, 2013; Tsutsui, 2017; Proistosescu & Huybers, 2017). With $M = 3$ fixed, the other parameters are fitted with nonlinear regression. The values for the fitted parameters are given in Table S1 in Supporting Information S1. In Figure 1, the datasets are shown along with the fits and projections, showing a very good resemblance.

From the fitted parameters, the feedback strengths per mode are computed using Equation 8. In Table 1, the results are shown per mode and for the equilibrium values. A plot of the change in instantaneous feedback strength over time is given in Figure 2a. This shows behavior on three distinct time scales: (a) a short time scale $\tau_1 = 4.5y$, (b) an intermediate time scale $\tau_2 = 127y$, and (c) a long time scale $\tau_3 = 889y$. The shorter time scale is in agreement with other studies (Zelinka et al., 2020), and long-term behavior seems very plausible. The Planck feedback does not change much over time. The lapse rate feedback becomes weaker over
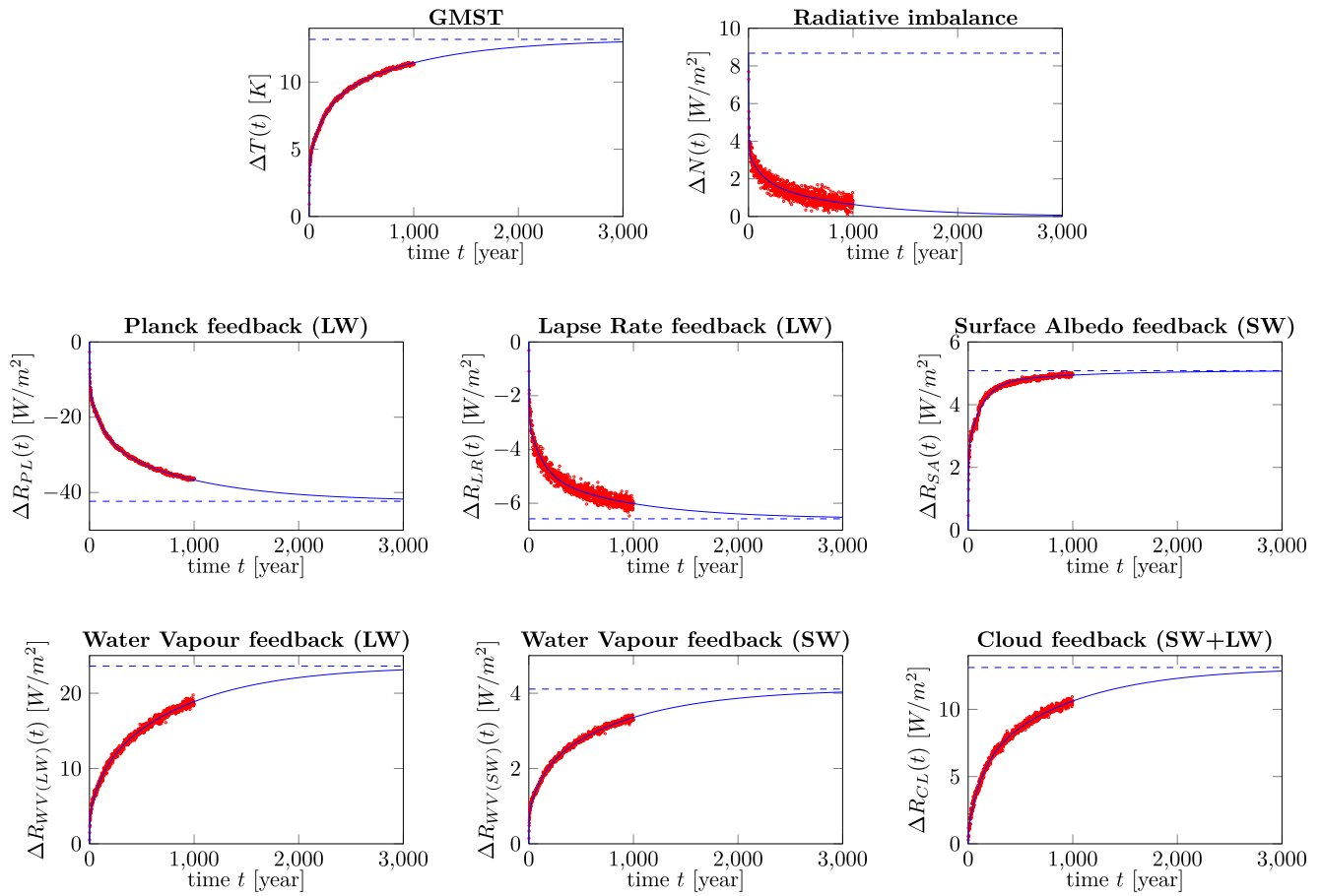
Figure 1. Evolution of temperature increase ΔT, radiative imbalance ΔN and feedback contributions ΔR_j for the Community Earth System Model version 2 abrupt4$\times$CO$_2$ experiment. Red circles denote data points; blue lines the obtained fits. The dashed lines indicate the estimated equilibrium values or initial forcing in the case of the radiative imbalance plot. Table S1 in Supporting Information S1 contains the values for the fitted parameters.

Table 1
Values for the Climate Feedback Parameters λ_j^m per Mode

	Mode 1	Mode 2	Mode 3	Equilibrium
τ_m	4.5 ($\pm$ 0.1)	127 ($\pm$ 3.8)	889 ($\pm$ 50)	−
λ_m	− 1.28 ($\pm$ 0.08)	− 0.38 ($\pm$ 0.03)	− 0.37 ($\pm$ 0.02)	− 0.66 ($\pm$ 0.03)
Planck (LW)	− 3.16 ($\pm$ 0.02)	− 3.24 ($\pm$ 0.02)	− 3.23 ($\pm$ 0.01)	− 3.21 ($\pm$ 0.05)
Lapse rate (LW)	− 0.73 ($\pm$ 0.03)	− 0.50 ($\pm$ 0.03)	− 0.32 ($\pm$ 0.03)	− 0.50 ($\pm$ 0.01)
Surface albedo (SW)	+ 0.62 ($\pm$ 0.04)	+ 0.56 ($\pm$ 0.02)	+ 0.08 ($\pm$ 0.10)	+ 0.39 ($\pm$ 0.01)
Water vapor (LW)	+ 0.97 ($\pm$ 0.03)	+ 1.38 ($\pm$ 0.02)	+ 2.71 ($\pm$ 0.01)	+ 1.79 ($\pm$ 0.04)
Water vapor (SW)	+ 0.21 ($\pm$ 0.09)	+ 0.26 ($\pm$ 0.05)	+ 0.43 ($\pm$ 0.02)	+ 0.31 ($\pm$ 0.01)
Clouds (SW + LW)	+ 0.27 ($\pm$ 0.36)	+ 1.19 ($\pm$ 0.02)	+ 1.43 ($\pm$ 0.01)	+ 1.00 ($\pm$ 0.03)
Sum	− 1.82 ($\pm$ 0.37)	− 0.36 ($\pm$ 0.07)	+ 1.09 ($\pm$ 0.11)	− 0.22 ($\pm$ 0.08)
Residue	+ 0.54 ($\pm$ 0.38)	− 0.02 ($\pm$ 0.08)	− 1.46 ($\pm$ 0.11)	− 0.43 ($\pm$ 0.08)

Note. These have been computed from the fitted parameters (Table S1 in Supporting Information S1) via Equation 8. Time scales τ_m have units "year"; feedback parameters λ_j^m have units "$W/m^2/K$." plusminus values indicate 95% confidence intervals, which were propagated from the 95% confidence intervals for the fitted parameters assuming a normal distribution for all errors and no existing correlations between parameters.
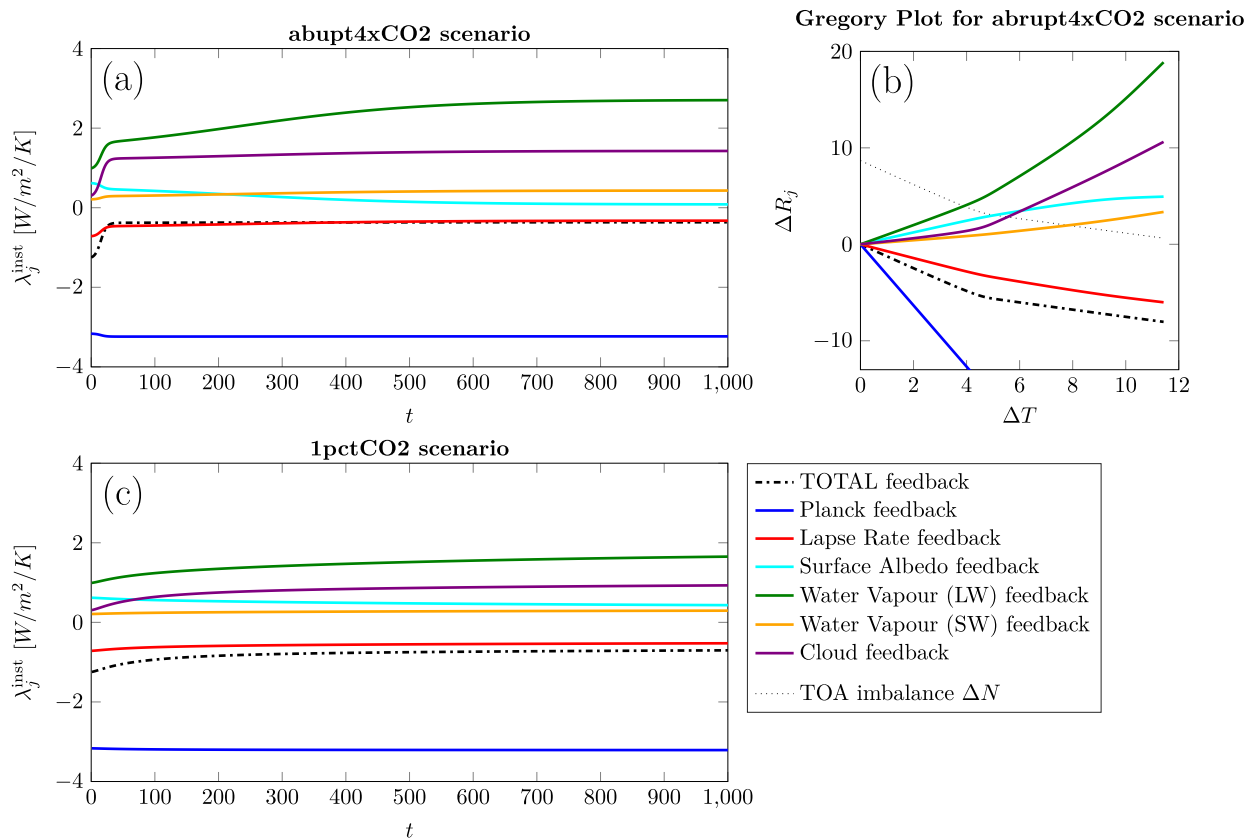
Figure 2. The evolution of climate feedback strengths over time in (a) The abrupt4×CO₂ experiment and (c) The 1pctCO₂ experiment, and (b) a Gregory plot for climate feedback contributions ΔR_j against warming ΔT in the abrupt4×CO₂ experiment. Instantaneous feedback strengths $\lambda_j^{\text{inst}}(t)$ are computed via Equation 9 from the fits derived from the abrupt4×CO₂ experiment and show projected feedback strengths under a continued 1% yearly CO₂ increase.

time since temperatures increase. The surface albedo feedback diminishes as sea-ice is melting. The water vapor feedback, however, strengthens as warmer air can contain more water, which—as demonstrated here—does lead to an almost tripling in this feedback's strength over time. Finally, cloud feedback becomes more dominant over time. The estimated equilibrium values are a bit higher generally compared to other studies (e.g., Zelinka et al., 2020) since longer time scales are incorporated in the present study, which leads to higher temperature values and more positive total feedback strength (Goodwin, 2018; Marvel et al., 2018; Proistosescu & Huybers, 2017; Rugenstein et al., 2020).

Since the individual feedback contributions ΔR_j should sum to the total radiative response ΔR, feedback strengths should satisfy $\lambda^m = \sum_{j \in \mathcal{F}} \lambda_j^m$. However, because of ignored feedback interactions, imperfect radiative kernels and other limitations of the feedback contribution computations, a mismatch of up to 10% is deemed inconsequential. Looking at the feedback parameters per mode in Table 1, we see a good balance for the intermediate time scale ($-0.02\ Wm^{-2}K^{-1}$ imbalance), but not for the short ($+0.54\ Wm^{-2}K^{-1}$ imbalance) or long time scales ($-1.46\ Wm^{-2}K^{-1}$ imbalance). Taking the large uncertainty into account for the shortest time scale, this imbalance might be insignificant still. However, for the long time scale, the imbalance is large and even the signs are not the same. This might point to a missing negative feedback on this time scale (or inaccuracies in the linear computational methods). Because the equilibrium values show a smaller imbalance ($-0.43\ Wm^{-2}K^{-1}$), the alleged missing feedback (or feedbacks) seem to be especially important during the transient state.
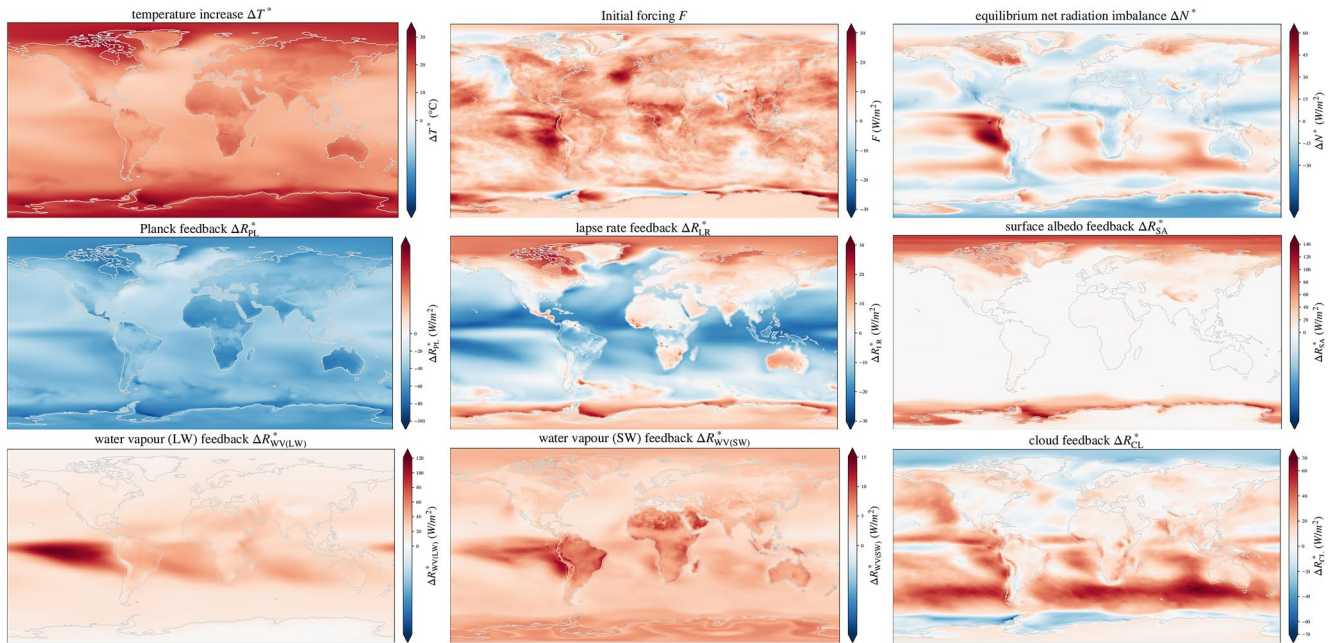
Figure 3. Estimated spatial distribution of warming $\Delta T_*(\mathbf{x})$, radiative imbalance $\Delta N_*(\mathbf{x})$ and feedback contributions $\Delta R_j^*(\mathbf{x})$ in equilibrium, and initial effective radiative forcing $F(\mathbf{x})$ for the CESM2 abrupt4×CO$_2$ experiment. Equilibrium distributions are derived from the fitted spatial modes; initial forcing $F(\mathbf{x})$ is obtained directly from fits.

3.2. Spatial Distribution, Abrupt4×CO$_2$

With a slight modification of Equation 8, the spatial structure of the modes can be included. Following a similar procedure in Proistosescu and Huybers (2017), we take

$$\Delta R_j(\mathbf{x}, t) = \sum_{m=1}^{M} \beta_m^{[R_j]}(\mathbf{x}) \, \mathcal{M}_m^{g_{\text{abr}}}(t),\tag{14}$$

and similar formulas for $\Delta N(\mathbf{x}, t)$ and $\Delta T(\mathbf{x}, t)$. More details can be found in Text S2.2 in Supporting Information S1.

The space-dependent $\beta_m^{[R_j]}(\mathbf{x})$ encompasses the spatial structure of the observable ΔR_j in the m-th mode. Again, it is also possible to estimate equilibrium properties, leading to an estimation of the spatial distribution of feedback contributions and warming in equilibrium. Using the time scales τ_m from the fits of the global observables, the spatial coefficients $\bar{\beta}_m^{[R_j]}(\mathbf{x})$ can be computed efficiently using linear regression methods.

In Figures S3–S10 in Supporting Information S1, the spatial modes of all feedback contributions, warming, and radiative imbalance are given. The estimated equilibrium distributions are shown in Figure 3. The patterns of feedback contributions are in agreement with previous studies (Andrews et al., 2015; Andrews & Webb, 2018; Armour et al., 2013; Boeke et al., 2020; Dessler, 2013; Dong et al., 2020; Proistosescu & Huybers, 2017; Soden et al., 2008), although they seem more pronounced here—probably due to a better incorporation of long-term behavior. Near the poles, temperature increase gets close to + 30°C (compared to + 12°C near the equator) and large changes (up to + 135 Wm^{-2}) in surface albedo are found related to sea-ice melting. Moreover, the lapse rate feedback contribution increases by about 10 Wm^{-2} in these polar regions. Near the equator, the water vapor feedback gets much stronger (more than + 50 Wm^{-2}) over the oceans—especially in the Pacific cold tongue where changes up to + 120 Wm^{-2} are found. Cloud feedback changes mostly over southern hemisphere oceans (up to + 69 Wm^{-2}), and northern hemisphere changes are smaller (on average only + 2.7 Wm^{-2}) compared to previous studies (Andrews & Webb, 2018; Armour et al., 2013; Dong et al., 2020).
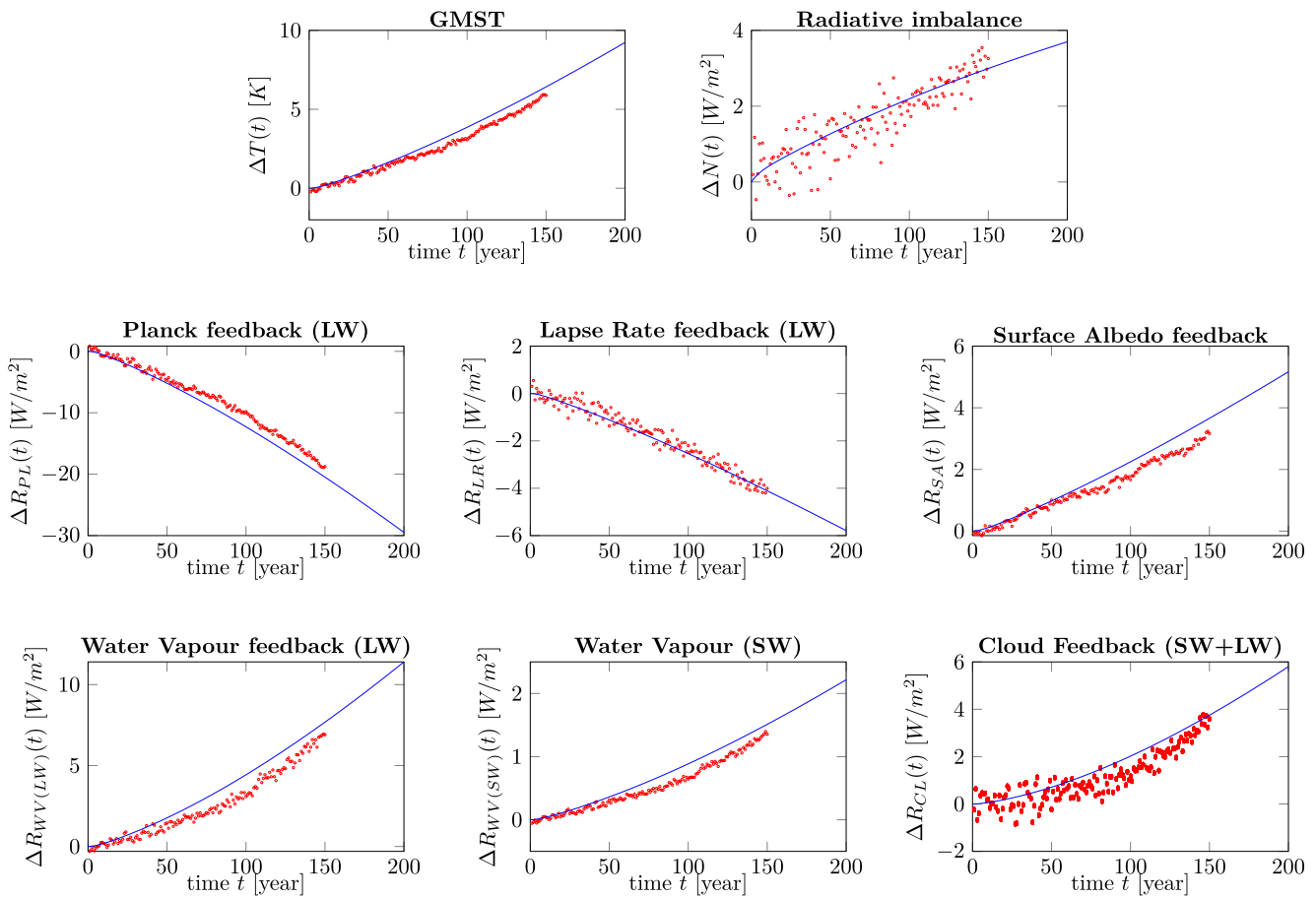
Figure 4. Projections and actual values for globally averaged feedback contributions ΔR_j, warming ΔT and top-of-atmosphere radiative imblance ΔN in the Community Earth System Model version 2 1pctCO$_2$ experiment. Red circles denote data points obtained from the model output. The lines show projections based on data from the abrupt4×CO$_2$ experiment using the introduced framework.

3.3. Projections for 1pctCO$_2$ Experiment

In the 1pctCO$_2$ experiment, the forcing (g_{grad}) is given by a yearly 1% increase in atmospheric CO$_2$. Following linear response theory, replacing $\mathcal{M}_m^{g_{abr}}$ by $\mathcal{M}_m^{g_{grad}}$ in Equation 8, while using the values for τ_m and β_m as fitted with an ensemble of abrupt4×CO$_2$ experiment runs, gives a projection for (the ensemble average of) feedback contributions and warming in the 1pctCO$_2$ experiment. Since we did not have access to a suitable ensemble, we apply the idea here on a single run to illustrate its potency qualitatively (see Section 2.2). More details can be found in Text S3 in Supporting Information S1.

In Figure 4, projections for global feedback contributions are shown along with data coming directly from the CESM2 1pctCO$_2$ experiment. Qualitatively, the projections are able to reproduce the actual trends in the globally averaged observables. The quantitative values overestimate feedback contributions by 10% − 20% (less for the lapse rate feedback; more for cloud and long-wave water vapor feedbacks up to about $t = 120\,y$). Projections for the temporal evolution of instantaneous feedback strengths in this scenario are given in Figure 2b. These show lasting changes over centuries—primarily in long-wave water vapor feedback, surface albedo feedback, and cloud feedback—and a less negative total feedback parameter, indicating the climate system becomes more sensitive to radiative forcing over time.

Using the same approach, it is also possible to make projections for the evolution of the spatial distributions. In Figures S11–S18 in Supporting Information S1 the projections for the feedback contributions ΔR_j, the radiative imbalance ΔN and the temperature increase ΔT at $t = 140\,y$ are shown. Videos showing these results for other times are also available (Movies S1–S8). From all of these, it can be seen that the spatial projections reproduce the larger spatial patterns and even some of the smaller features, although the internal variability

is of course not present in the projections. The spatial patterns are similar to those found in the abrupt4×-CO$_2$ experiment (Figure 3) and show the same kind of response in polar and equatorial regions. In general, the projections seem to overestimate the spatial feedback contributions, similar to the overestimation of the global feedback contributions.

4. Summary and Discussion

In this study, we have introduced a new multivariate feedback framework that enables analysis and projections of individual climate feedbacks, including their spatial distribution, over different time scales. This framework deviates from the common practice of linearly regressing feedback contributions to global warming (Klocke et al., 2013; Marvel et al., 2018; Meraner et al., 2013; Zelinka et al., 2020). Instead, these contributions are here taken into account directly, and temporal evolution is explicitly considered via a nonlinear regression to obtain modes for each feedback. These modes, indicating behavior over different time scales, encapsulate the climate system's dynamics and can thus be used both for analysis and projections of climate feedbacks.

Using the presented framework, we have analyzed an abrupt4×CO$_2$ experiment in CESM2. This showed evolution of climate feedbacks over time, such as a diminishing of surface albedo feedback in the polar regions as sea-ice melts. Strikingly, the water vapor and cloud feedbacks showed a large increase over time: water vapor feedback, primarily over oceans around the equator, almost tripled in strength over a 1,000 year period and cloud feedback mainly changed over southern hemisphere oceans. Analysis also showed that the commonly considered feedback processes do not sum up to the total feedback for long time scales, which might point to a missing negative feedback on these time scales. Since estimated equilibrium feedback strengths do balance, the alleged missing feedback (or feedbacks) is expected to play a role primarily during the transient. As such, the ocean heat uptake seems a good candidate to fill this gap, which would mean tracking its effect over time is necessary to accurately understand the long-term transient behavior.

Another benefit of tracking the temporal evolution of climate feedbacks is that this makes it possible to project the (ensemble average) climatic change for all sorts of emission scenarios without the need to have dedicated model experiments for them all using linear response theory. In this paper, we have illustrated the capabilities of such projections by comparing them to a single 1pctCO$_2$ experiment with the same CESM2 model. The projections did capture the qualitative trends but also showed an overestimation for almost all global and spatial climate feedback contributions. However, due to the internal variability associated with a single run, results are expected to be better quantitatively when used on an ensemble of runs, as has been shown with applications of similar linear response techniques on for example, temperature response (Aengenheyster et al., 2018; Lembo et al., 2020).

As climate models become more and more detailed, these models also become more computationally expensive. Therefore, it is practically impossible to compute climate responses for many forcing scenarios and/or on very long time scales. As climate feedbacks seem to become more positive with rising temperatures, it is necessary to track these changes to prevent underestimation of climate change and adequately account for committed (long-term) warming. Hence, both extrapolation techniques and response theory can play a significant role to alleviate the gaps. The multivariate feedback framework in this study can contribute to this. Moreover, as these projections are multivariate, they can inform us on more than global mean warming alone, and also indicate how a potential future climate state may behave differently compared to our current one.

Data Availability Statement

Data statement Radiative kernels from Pendergrass et al. (2018); Pendergrass (2017a, 2017b) have been used; the kernels were downloaded from https://doi.org/10.5065/D6F47MT6 and accompanying software from https://doi.org/10.5281/zenodo.997899. CESM2 data has been downloaded on the fly from Google's cloud storage mirror of CMIP6 data using the "intake-esm" utility package in Python (https://doi.org/10.5281/zenodo.4243421).

References

Aengenheyster, M., Feng, Q. Y., Van Der Ploeg, F., & Dijkstra, H. A. (2018). The point of no return for climate action: Effects of climate uncertainty and risk tolerance. *Earth System Dynamics*, *9*(3), 1085–1095. https://doi.org/10.5194/esd-9-1085-2018

Andrews, T., Gregory, J. M., & Webb, M. J. (2015). The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *Journal of Climate*, *28*(4), 1630–1648. https://doi.org/10.1175/JCLI-D-14-00545.1

Andrews, T., & Webb, M. J. (2018). The dependence of global cloud and lapse rate feedbacks on the spatial structure of tropical pacific warming. *Journal of Climate*, *31*(2), 641–654. https://doi.org/10.1175/JCLI-D-17-0087.1

Armour, K. C. (2017). Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nature Climate Change*, *7*(5), 331–335. https://doi.org/10.1038/nclimate3278

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, *26*(13), 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 the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *41*(251), 237–276. https://doi.org/10.1086/121158

Bastiaansen, R., Dijkstra, H. A., & von der Heydt, A. S. (2020). Multivariate estimations of equilibrium climate sensitivity from short transient warming simulations. *Geophysical Research Letters*, *48*, e2020GL091090. https://doi.org/10.1029/2020GL091090

Boeke, R. C., Taylor, P. C., & Sejas, S. A. (2020). On the nature of the Arctic's positive lapse-rate feedback. *Geophysical Research Letters*, *48*, e2020GL091109. https://doi.org/10.1029/2020GL091109

Boer, G., Hamilton, K., & Zhu, W. (2005). Climate sensitivity and climate change under strong forcing. *Climate Dynamics*, *24*(7–8), 685–700. https://doi.org/10.1007/s00382-004-0500-3

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., et al. (2006). othersHow well do we understand and evaluate climate change feedback processes? *Journal of Climate*, *19*(15), 3445–3482. https://doi.org/10.1175/JCLI3819.1

Caldeira, K., & Myhrvold, N. (2013). Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration. *Environmental Research Letters*, *8*(3), 034039. https://doi.org/10.1088/1748-9326/8/3/034039

Cess, R. D. (1975). Global climate change: An investigation of atmospheric feedback mechanisms. *Tellus*, *27*(3), 193–198. https://doi.org/10.3402/tellusa.v27i3.9901

Cess, R. D., & Potter, G. L. (1987). Exploratory studies of cloud radiative forcing with a general circulation model. *Tellus*, *39*(5), 460–473. https://doi.org/10.3402/tellusa.v39i5.11773

Curry, J. A., Schramm, J. L., & Ebert, E. E. (1995). Sea ice-albedo climate feedback mechanism. *Journal of Climate*, *8*(2), 240–247. https://doi.org/10.1175/1520-0442(1995)008⟨0240:SIACFM⟩2.0.CO;2

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., et al. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2). https://doi.org/10.1029/2019MS001916

Dessler, A. (2013). Observations of climate feedbacks over 2000–2010 and comparisons to climate models. *Journal of Climate*, *26*(1), 333–342. https://doi.org/10.1175/JCLI-D-11-00640.1

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., & Andrews, T. (2020). Intermodel spread in the pattern effect and its contribution to climate sensitivity in CMIP5 and CMIP6 models. *Journal of Climate*, *33*(18), 7755–7775. https://doi.org/10.1175/JCLI-D-19-1011.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Goodwin, P. (2018). On the time evolution of climate sensitivity and future warming. *Earth's Future*, *6*(9), 1336–1348. https://doi.org/10.1029/2018EF000889

Gregory, J. M., & Andrews, T. (2016). Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, *43*(8), 3911–3920. https://doi.org/10.1002/2016GL068406

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., & Williams, K. D. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3). https://doi.org/10.1029/2003GL018747

Hasselmann, K., Sausen, R., Maier-Reimer, E., & Voss, R. (1993). On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, *9*(2), 53–61. https://doi.org/10.1007/BF00210008

Heinze, C., Eyring, V., Friedlingstein, P., Jones, C., Balkanski, Y., Collins, W.,et al. (2019). ESD reviews: Climate feedbacks in the earth system and prospects for their evaluation. *Earth System Dynamics*, *10*(3), 379–452. https://doi.org/10.5194/esd-10-379-2019

Klocke, D., Quaas, J., & Stevens, B. (2013). Assessment of different metrics for physical climate feedbacks. *Climate Dynamics*, *41*(5–6), 1173–1185. https://doi.org/10.1007/s00382-013-1757-1

Knutti, R., & Rugenstein, M. A. (2015). Feedbacks, climate sensitivity and the limits of linear models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *373*, 20150146. https://doi.org/10.1098/rsta.2015.0146

Lembo, V., Lucarini, V., & Ragone, F. (2020). Beyond forcing scenarios: Predicting climate change through response operators in a coupled general circulation model. *Scientific Reports*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-65297-2

Lucarini, V., & Sarno, S. (2011). A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlinear Processes in Geophysics*, *18*(1), 7–28. https://doi.org/10.5194/npg-18-7-2011

Maier-Reimer, E., & Hasselmann, K. (1987). Transport and storage of CO_2 in the ocean——An inorganic ocean-circulation carbon cycle model. *Climate Dynamics*, *2*(2), 63–90. https://doi.org/10.1007/BF01054491

Manabe, S., & Wetherald, R. T. (1967). Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of the Atmospheric Sciences*, *24*, 241–259. https://doi.org/10.1175/1520-0469(1967)024⟨0241:TEOTAW⟩2.0.CO;2

Marvel, K., Pincus, R., Schmidt, G. A., & Miller, R. L. (2018). Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, *45*(3), 1595–1601. https://doi.org/10.1002/2017GL076468

Meraner, K., Mauritsen, T., & Voigt, A. (2013). Robust increase in equilibrium climate sensitivity under global warming. *Geophysical Research Letters*, *40*(22), 5944–5948. https://doi.org/10.1002/2013GL058118

Pendergrass, A. G. (2017a). CAM5 radiative kernels. *UCAR/NCAR Earth system grid*. https://doi.org/10.5065/D6F47MT6

Pendergrass, A. G. (2017b). CESM CAM5 kernel software. *Zenodo*. https://doi.org/10.5281/zenodo.997899

Pendergrass, A. G., Conley, A., & Vitt, F. M. (2018). Surface and top-of-atmosphere radiative feedback kernels for CESM-CAM5. *Earth System Science Data*, *10*(1), 317–324. https://doi.org/10.5194/essd-10-317-2018

Proistosescu, C., & Huybers, P. J. (2017). Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Science Advances*, *3*(7), e1602821. https://doi.org/10.1126/sciadv.1602821

Ragone, F., Lucarini, V., & Lunkeit, F. (2016). A new framework for climate sensitivity and prediction: A modelling perspective. *Climate Dynamics*, *46*(5–6), 1459–1471. https://doi.org/10.1007/s00382-015-2657-3

Ruelle, D. (2009). A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, *22*(4), 855–870. https://doi.org/10.1088/0951-7715/22/4/009

Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., et al. (2020). othersEquilibrium climate sensitivity estimated by equilibrating climate models. *Geophysical Research Letters*, *47*(4), e2019GL083898. https://doi.org/10.1029/2019GL083898

Senior, C. A., & Mitchell, J. F. (2000). The time-dependence of climate sensitivity. *Geophysical Research Letters*, *27*(17), 2685–2688. https://doi.org/10.1029/2000GL011373

Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Using the radiative kernel technique to calculate climate feedbacks in NCAR's community atmospheric model. *Journal of Climate*, *21*(10), 2269–2282. https://doi.org/10.1175/2007JCLI2044.1

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*, e2019RG000678. https://doi.org/10.1029/2019RG000678

Sinha, A. (1995). Relative influence of lapse rate and water vapor on the greenhouse effect. *Journal of Geophysical Research: Atmospheres*, *100*(D3), 5095–5103. https://doi.org/10.1029/94JD03248

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, *21*(14), 3504–3520. https://doi.org/10.1175/2007JCLI2110.1

Tsutsui, J. (2017). Quantification of temperature response to CO_2 forcing in atmosphere–ocean general circulation models. *Climatic Change*, *140*(2), 287–305. https://doi.org/10.1007/s10584-016-1832-9

Von der Heydt, A. S., Ashwin, P., Camp, C. D., Crucifix, M., Dijkstra, H. A., Ditlevsen, P., & Lenton, T. M. (2020). Quantification and interpretation of the climate variability record. *Global and Planetary Change*. https://doi.org/10.1016/j.gloplacha.2020.103399

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782