REVIEW 2

Dear reviewer, first of all we would like to thank you for your effort. We were happy to address the major and minor comments. We hope the paper is publishable according to your opinion.

We believe it does not make sense to list all improvements as we enhanced the contextual knowledge at various occasions. Instead, we created a file that gives the difference between the latest draft and the final manuscript which is submitted with this report. Please review this file to get an overview of the changes made.

We use the following fonts to mark the comment by the reviewer, the reply by the author, and the citation from the source code of the manuscript:

COMMENT

REPLY

CITITATION

COMMENT

Summary:

This paper presents interesting and useful new results on the timescales of the climate response to CO2 forcing, exploiting 1000-year long step forcing AOGCM experiments. While the results are novel, I found the presentation rather complex and hard to follow, so I am requesting major revisions to make the paper more accessible.

Main comments:

Presentation:

Overall I found the text difficult to read, despite it being well polished and free of typos – to the point that I didn't understand everything despite a careful read. I ended up becoming frustrated and skipped most of section 5. The issues start with the abstract, where things should be kept simpler in my opinion. In particular, I struggled with the sentence L11–13, which I'm still not sure I fully understand after reading the paper. Can this be explained more simply, or perhaps omitted?

REPLY

The abstract now reads: line 1-15

CITITATION

\abstract{We explore to which degree the temperature dependence of the climate radiative feedback influences the slow mode of the surface temperature response, which describes the surface air temperature adjustment to forcing on a centennial timescale. We question whether long-term climate change is described by a single e-

folding mode with a constant timescale which is commonly assumed to be independent of temperature or forcing and the evolution of time. To do so, we analyze Atmosphere-Ocean General Circulation model (AOGCM) simulations which have an integration time of 1000 years and are forced by atmospheric CO$_2$ concentrations ranging from two times (2X) to eight times (8X) the preindustrial level. Our findings suggest that feedback temperature dependence strongly influences the equilibrium temperature response and adjustment timescale of the slow mode. The timescale of the slow mode is thus state-dependent. In addition, the effective heat capacity of the slow mode increases over time, which makes the adjustment timescale also time-dependent. The state-dependence and time-dependence of the adjustment timescale of long-term climate change call into question common eigenmode decomposition with a fast and a slow timescale, in the sense that the slow mode is not well described by a single linear e-folding mode with a constant timescale. Instead, we find that any eigenmode decomposition will depend on the forcing level, and that an additional mode or a multiple mode and timescale structure of the slow adjustment is necessary to reproduce the details of AOGCM simulated long-term climate change even at a single forcing level.
}

We changed the overall presentation of the manuscript.

COMMENT

The introduction begins rather abruptly, and assumes a fairly high level of background knowledge – for example, that it is commonly understood that the response to CO2 forcing can be decomposed into fast and slow components. The notion that climate feedbacks are temperature dependent is also assumed. I think these concepts should be introduced more slowly, with references to the relevant prior literature:

REPLY

We changed the introduction accordingly. We introduce the two-timescale approach first and mention our hypothesis. We then provide reason for our hypothesis. Finally, we illustrate the study content by a simple energy balance model.
line 15-125

CITITATION

\section{Introduction}
Studies of climate change have long found that the response of surface warming to radiative forcing occurs at multiple timescales. These studies typically differentiate between a fast timescale of response, that occurs within the first decade and is associated with the thermal inertia of the ocean's mixed layer and land, and a slow timescale of response, that occurs over centuries and is associated with the thermal inertia of the intermediate and deeper ocean \citep[e.g.][]{dickinson1998}. The latter is denoted as slow mode. Many studies have sought to understand how this slow mode of warming will unfold in time. This understanding is critical to predicting long-term warming. Commonly the slow mode has been modeled as exponential decay to equilibrium, and the decay has been assumed to be constant relative to the forcing level. In this paper we analyze the slow mode of the climate response in light of feedback temperature dependence and inconstant global heat capacity using abrupt CO$_2$ experiments with Atmosphere-Ocean General Circulation models (AOGCMs). More precisely, we analyze the timing of long-term climate change and therefore the adjustment timescale of the slow mode, which varies with temperature or forcing and time according to our study. We define the variation of the adjustment timescale with temperature or forcing as being state-dependent, and we define the variation of the adjustment timescale with the evolution of time as being time-dependent. We use this phrasing throughout the study. That is, we put forward the idea that the adjustment timescale of long-term climate change on a centennial timescale is state- and time-dependent, with the surface air temperature being the state-variable.
\newline

In this study feedback temperature dependence describes how the radiative feedback of the climate system depends on the global mean surface temperature change. The global feedback is given by the derivative of the global mean energy budget N with respect to the global mean surface air temperature pertubation T, $\frac{d N}{d T}$. Commonly, the global feedback is assumed to be constant. However, it may change with temperature and time. We find that the dependence of the radiative feedback on temperature makes the timing of long-term climate change state-dependent. We also find a systematic time-dependent component of the adjustment timescale of the slow mode. Time-dependence of the adjustment timescale is possible due to changes in ocean heat uptake efficiency and horizontal heat transport, or an inconstant global ocean heat capacity. Oceanic timescales or heat capacities are commonly assumed to be constant. However, the circulation of the deep ocean may change \citep[e.g.][]{knutti2015} and diffusive ocean heat uptake progresses over time \citep{hansen1985,wigley1985}, which can make the deep ocean effective heat capacity inconstant. Another approach for a time-dependent timescale from a global perspective is that there exists a geographic timescale pattern that emerges from the ocean circulation or a spatial pattern of local heat capacities.
\newline

To illustrate the study content, we make use of a simple energy balance model. The simple energy balance model for the global mean surface temperature response to forcing is given by
$C \frac{\mathrm{d}T}{\mathrm{d}t}= F + \lambda T$
where C is the effective heat capacity of the global system, T is the surface temperature perturbation relative to a reference state, F is the radiative forcing which is e.g. the forcing by atmospheric CO$_2$ or aerosols, and λ is the feedback parameter. In equilibrium, it is called the linear forcing-feedback framework. It is a first-order differential equation which analytical solution is given by $T(t) = \frac{F}{-\lambda} (1-e^{-t/\tau})$. The equilibrium warming is

$T(\infty)=\frac{F}{-\lambda}$ and thus linear in forcing. The temporal adjustment is described by the e-folding timescale $\tau=\frac{C}{-\lambda}$ at which the single e-folding mode unfolds. That is to say, the response times of the temperature adjustment depend on C and λ. The stronger the feedback parameter or the smaller the effective heat capacity, the more rapidly the system adjusts to forcing. In fact, λ itself may depend on temperature which gives rise to temperature or forcing-dependent adjustment timescales. At the same time, C may not be constant but time-dependent due to the changes in the ocean circulation and diffusive ocean heat uptake in response to global warming, or a spatial pattern of local heat capacities that prescribes the thermal inertia geographically.
\newline

…

COMMENT

What do we know about feedback temperature dependence? Is this commonly simulated by GCMs? Do we know the sign of this dependence, or is this still a subject of ongoing research? The text asserts that feedbacks become more amplifying with warming (L25), yet this is inconsistent with two out of four GCMs used in this study (Table 1). 4

REPLY

We add the following: line 81-85

CITITATION

State-of-the-art AOGCMs exhibit both positive and negative feedback temperature dependence in warming experiments under modern-day boundary conditions. There, are, however, only a few studies which quantify the degree to which the global feedback depends on temperature. \citet{bloch2015} and \citet{roe2011} have quantified feedback temperature dependence for various AOGCMs. In a recent study, \citet{bloch2020} show that feedback temperature dependence is positive for 10 out of 14 state-of-the-art GCMs.
\newline

COMMENT

Another confusing aspect for me was the introduction of the two conceptual models (Eqs. 4–6) -

What physics underlie the 1st model (based on two regions, Eqs. 4–5)? Presumably this is meant to reflect the SST pattern effect, but I don't think this was explained.

REPLY

The two-region model without the coefficient for feedback temperature dependence (\$a\$) mimics the pattern effect only. The pattern effect is associated with different state-variables which are the temperatures in different regions, which in turn actuate a regional radiative feedback. The pattern effect is a time-dependent radiative response and emerges from the interplay of at least two state-variables. Feedback temperature dependence introduces an additional state-dependent radiative response, since the feedback in each region now depends on temperature. In this connection, the two-region model with regional feedback temperature dependencies combines time-dependent and state-dependent feedback. We focus on the response in one effective region only. One can imagine that the surface air temperature response and radiative feedbacks are aggregated onto different regions which represent the fast mode and the slow mode. Conceptually, we analyze the slow adjustment only and therefore neglect the time-dependent radiative response associated with different state-variables, having state-dependent feedback in one region only. We add some more explanations. line 127-160

COMMENT

It would help to discuss the commonalities and differences between the two models. My understanding would be that using an efficacy term (epsilon) in the 2nd model could be mathematically equivalent to using spatially-varying feedbacks in the 1st model – is this correct? The 2nd model additionally includes a heat transport efficiency term – what physics does this involve and does it make the 2nd model different from the first?

REPLY

In the case of zero feedback temperature dependence the two-region model and the two-layer model are mathematical equivalent as demonstrated by Rohrschneider et al. (2019). They provide a thorough discussion of the two-layer model and the two-region model, which shouldn't be repeated. The efficacy factor in the two-layer model makes the fast mode and the slow mode having different radiative feedbacks. Heat transport efficiency is an inertia parameter in the two-layer model and changes the timing of climate change and the magnitude of the fast mode and the slow mode. Heat transport efficiency is associated with the two-region model by changing the heat capacity that influences the timing of climate change as well as the regional feedback parameter that gives the magnitude of the surface air temperature response. We add some more explanations. line 161-205

COMMENT

The authors ultimately choose to focus on the two-region model (Eqs. 4–5), as stated L182. Why this choice, and how does it affect the interpretation of the results? Do we even need both models in the paper? I feel like it might help to use an appendix to discuss some of the more technical aspects of the two conceptual models and/or the methodological choices, so as to keep the main text simpler and more focused on the key results and their interpretation.

REPLY

We wouldn't like to have an appendix because the conceptual models are needed to provide the theoretical background to understand the paper. We choose the two-region model because it provides a simple framework to understand the slow mode, having a simple expression for the slow mode in order to analyze the different parameters. By its nature, the two-layer model is more complicated, having more than one inertia parameter. However, it is necessary to mention the two-layer model because the slow mode can also be understood as a function of the deep ocean component, and changes in ocean heat uptake (heat uptake efficiency) changes the behavior of the slow mode. We wouldn't like to miss these conceptual insights.

The simple model section now reads: line 127-205

CITITATION

\section{Conceptual insights}
Before exploring the slow mode's behavior in AOGCMs, we provide conceptual insights about the slow mode using simple climate models. We explain conceptually the slow mode as well as demonstrate conceptually the state- and time-dependence of the adjustment timescale of the slow mode. The simple models considered here are energy balance models and outlined in detail in \citet{geoffroy2013}, \citet{geoffroy2013b}, \citet{armour2013}, \citet{rohrschneider2019}, among others. We bring together these existing concepts to lay out the parameter dependencies of the slow mode in order to provide a solid basis and motivation for our experimental analysis. With this section we provide insight how the fast e-folding mode and the slow e-folding mode (Eq. 3) emerge from simple assumptions using energy balance models. We present two recent concepts: the two-region framework which is used in this study to analyze the slow mode; and a two-layer model in which the slow mode is a function of the Earth's deep ocean component. The two-region model is much

simpler than the two-layer model, while the two-layer model accounts explicitly for changes in the ocean circulation. However, the two-layer model can be expressed mathematically by the two-region model.
\newline

A way to represent the global mean surface temperature response to forcing is to assume two effective regions, $T=(\chi -1) T_\mathrm{F} + \chi T_\mathrm{S}$, where χ is the effective fractional area:
\begin{equation}
{C_{\mathrm{F}}} \frac{\mathrm{ d} T_{\mathrm{F}}}{\mathrm{d} t}= F +
(\lambda_\mathrm{F} + a_\mathrm{F} T_\mathrm{F}
) T_\mathrm{F}.
\end{equation}
and
\begin{equation}
{C_{\mathrm{S}}} \frac{\mathrm{ d} T_{\mathrm{S}}}{\mathrm{d} t}= F +
(\lambda_\mathrm{S} + a_\mathrm{S} T_\mathrm{S}
) T_\mathrm{S}.
\end{equation} F is the radiative forcing, C is the constant effective heat capacity, λ is the background feedback parameter, and a is the coefficient for feedback temperature dependence. Each region behaves similarly to Eq. (2), and according to this framework, the climate response is characterized by a fast mode T_{F} and a slow mode T_{S}. The two-region model without the coefficient for feedback temperature dependence (a) mimics the pattern effect only. The pattern effect is associated with different state-variables which are the temperatures in different regions, which in turn actuate a regional radiative feedback. The pattern effect is a time-dependent radiative response and emerges from the interplay of at least two state-variables. Feedback temperature dependence introduces an additional state-dependent radiative response, since the feedback in each region now depends on temperature. In this connection, the two-region model with regional feedback temperature dependencies combines time-dependent and state-dependent feedback. We focus on the response in the slow effective region only. One can imagine that the surface air temperature response and radiative feedbacks are aggregated onto different regions which represent the fast mode and the slow mode. Conceptually, we analyze the slow adjustment and therefore neglect the time-dependent radiative response associated with different state-variables, having state-dependent feedback in one region only. Positive feedback temperature dependence causes the equilibrium response of the slow mode to increase. Furthermore, feedback temperature dependence introduces a timescale that depends on the strength of the forcing. Considering the temporal behavior, the thermal inertia of the slow mode is represented by a single effective heat capacity which is much higher than the heat capacity of the fast mode ($C_{\mathrm{F}} \ll C_{\mathrm{S}}$). At this point, C_{S} is constant over time and does not change with the climate state. .
\newline

Another conceptual framework with a fast mode T_{F} and a slow mode T_{S} is the two-layer ocean model with ocean heat uptake efficacy and feedback temperature dependence \citep{held2010,winton2010}. We extend this model by introducing a coefficient for feedback temperature dependence. This model then also combines time-dependent feedback due to the evolution of two different state-variables and state-dependent feedback due to temperature-dependent feedback. The model configuration with ocean heat uptake efficacy and feedback temperature dependence is given by
\begin{equation}
{C}\frac{\mathrm{d} T}{\mathrm{d} t}= F + (\lambda_{\mathrm{b}} + a T) T - \epsilon \eta (T-T_{\mathrm{D}})
\end{equation}
\begin{equation}

{C_{\mathrm{D}}} \frac{\mathrm{ d} T_{\mathrm{D}}}{\mathrm{d} t}= \eta (T-T_{\mathrm{D}})
\end{equation}
where $C \ll C_{\mathrm{D}}$ are the heat capacities of the upper- and deep-ocean, λ_{b} is the background feedback parameter and a the coefficient for feedback temperature dependence. The parameter η is the heat transport efficiency and ϵ the efficacy factor for ocean heat uptake. The slow component is approximated by
\begin{equation}
T_\mathrm{s} (t) \approx \frac{\sqrt{\Lambda^2 - 4 a F}-{\sqrt{\Lambda^2- 4 a F - 4 a \epsilon \eta T_\mathrm{D} (t)}}}{2 a} \: \:   \text{with} \: \: \Lambda=\lambda_\mathrm{b}- \epsilon \eta
\end{equation}
after the fast contribution from the surface, as derived in \citet{rohrschneider2019}. Following this conceptual framework, the slow mode is a function of the deep ocean component T_D because the slow mode emerges from the heat transport into the deep ocean and the convergence of the state-variables over time towards the same equilibrium temperature perturbation.
\newline

Using linear model versions without feedback temperature dependence, the two-region model and the two-layer model are mathematically equivalent. There is no difference in the fast e-folding mode and the slow e-folding mode (Eq. 3) as well as in the global radiative response between these models. Although no analytical solution of the coupled two-layer model with feedback temperature dependence exists to date, we can approximate the temperature and radiative response associated with the slow mode by a single effective region (Eq. 5), having a single heat capacity. However, the parameters of the two-layer model modify the inertia of the slow mode. For instance, the parameter for the efficiency of ocean heat uptake η is an inertia parameter, and changes in ocean heat uptake cause C_{S} to increase or decrease. Commonly, we assume that the parameters which describe these simple models are constant. In that respect, we emphasize that the slow mode's response is described by
\begin{equation}
{C_{\mathrm{S}}} \frac{\mathrm{ d} T_{\mathrm{S}}}{\mathrm{d} t}=N_\mathrm{S}
\end{equation}
where N_S is the TOA imbalance associated with the slow mode. After having explored the imprint of feedback temperature dependence on the slow mode, we analyze the interplay of state-varying and time-varying adjustment timescales. The former arises from the presence of feedback temperature dependence while the latter arises from the inconstancy of C_{S} according to Eq. (5,9).
\newline

For the experimental analysis, we choose the two-region model because it provides a simple framework to understand the slow mode, allowing us to analyze the different parameters. By its nature, the two-layer model is more complicated, having more than one inertia parameter and an efficacy term. However, it is necessary to mention the two-layer model because the slow mode can also be understood as a function of the deep ocean component, and changes in ocean heat uptake or in the heat uptake efficiency do change the behavior of the slow mode. Heat uptake efficiency is an inertia parameter in the two-layer model and it changes the timing of climate change and the magnitude of the fast mode and the slow mode while the magnitude of the global equilibrium response remains unchanged. Heat transport efficiency is associated with the two-region model by changing the heat capacity of the region that influences the timing of climate change as well as the regional feedback parameter that gives the magnitude of the surface air temperature response. In the following we analyze briefly the parameter dependencies of the equilibrium response and timescale of the slow mode using the two-layer ocean model. We use the more complicated two-layer ocean model to show the dependence of

the slow mode on the ocean circulation besides feedback temperature dependence. We focus on the heat uptake efficiency in the two-layer model to provide an explanation for changes in C$_S$. It, C$_S$, may change with model parameter, and this model parameter may also vary with time. It is not straightforward to find a simple analytical expression for the dependence of C$_S$ on η.
\newline

...

COMMENT

I would like the authors to clarify and make explicit their definition of temperature-dependent feedbacks. It seems to me that there are two quite distinct types of temperature dependence: (a) a temperature-dependent SST pattern effect, versus (b) temperature-dependent feedback processes (independent of the SST pattern). The latter could be quantified for example using uniform SST warming or cooling experiments. My understanding is that the temperature dependence discussed in the present paper includes both processes (a) and (b), but it would be good to clarify this. Do the authors know which type of temperature dependence is more important for their findings? If we want to understand and perhaps observationally constrain the temperature dependence of climate feedbacks, it seems to me that different approaches would be needed for (a) versus (b).

REPLY

We do not focus on the pattern effect because we analyze the response of one effective region only onto which the surface air temperature response and radiative feedback of the slow adjustment is aggregated. Feedback temperature dependence is the second-order temperature dependence of the first-order radiative feedback, using Taylor-series. We hope it is clear now with the changes made in the introduction and the section on the conceptual models. We look forward to have your opinion.

SPECIFIC COMMENTS

`We meet the specific comments listed below.`

L24: “As a result” – of what?
L112: Should clarify that this isn’t the formulation used by Held et al. and Winton et al. (who didn’t consider feedback temperature-dependence, as far as I’m aware?) Agreed
L185: Shouldn’t it be N_F(t=0)?
L186–188: I wasn’t able to follow this, can you explain in more detail or illustrate this graphically? (After further reading, I see this is explained more clearly L225–227. This needs to be reorganised.)
L193–196: Again I wasn’t able to fully follow. I’d recommend explaining this in more detail in an appendix.
L248: remove extra “between”
L283–284: I didn’t follow this reasoning.
Answerd
L441: “publicly *available* experiments”
L443: The reference to year 2100 is odd, considering that the results are based on idealised step forcing experiments, rather than realistic RCP-style scenarios. (On this timescale the slow mode arises, and it behaves like in the case of step function input)