Reply on

Interactive comment on **"On the Future Role of the most Parsimonious Climate Module in Integrated Assessment"** ***by*** **Mohammad M. Khabbazan and Hermann Held**

Anonymous Referee #1

Received and published: 6 June 2017

Original comments by the referee are highlighted in italic font.

The manuscript by Khabbazan and Held assesses the performance of a very simplified climate module currently in use in some IAMs. In particular, the study is motivated by the need to adjust the existing tools to the capability of this module in the light of assessments of below 2°C scenarios. To that end, it is fitted to different CMIP5 RCP2.6 AOGCMs.

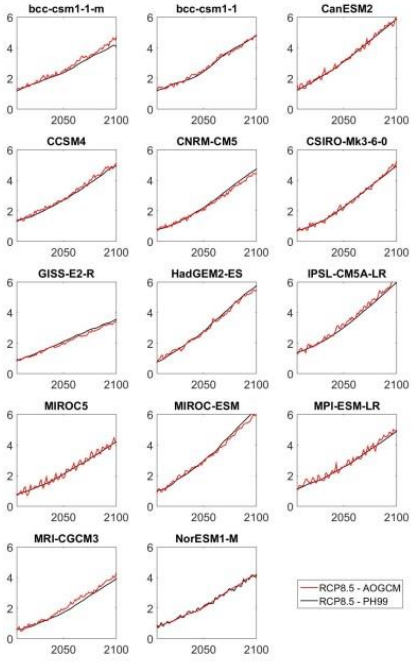
This is the very point of the first version of our ms indeed.

The manuscript contains no fundamental flaws although a re-read is in order and the literature list should be checked.

This will be done for the next version (that we call ‘V2’ for ‘version 2’ thereafter).

One key paper (Foster et al. 2013) is for example missing from the literature list. I presume it’s Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models.
J. Geophys. Res. Atmos., 118, 1139– 1150

The referee is right. This is the very reference that represents the data basis for our work. We sincerely apologize for this flaw. We will make sure that flaws like this one will not occur in V2.



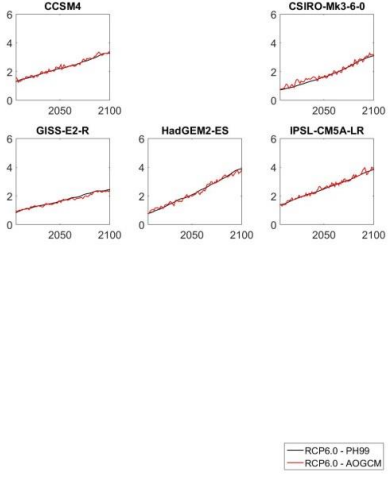
Reply-Figure 1: Intercomparison of global mean temperature from various AOGCMs and PH99 for RCP8.5. PH99 was fitted to RCP2.6.

More fundamentally, however, the scientific advancement presented of this study in my assessment is rather poor and it neglects important recent literature in this context (in fact, the literature list is rather short and at least 3 years old).

V2 will state clearer what the point of our ms is. We recommend recalibrating PH99 before using it and thereby we also allow for a re-interpretation of existing literature. Furthermore, we will add a Section that demonstrates a mechanism that might be the key cause of the PH99-AOGCM discrepancy. We will show that already when moving from a 2-box model as used by the most influential work of W. Nordhaus in the integrated assessment literature (see e.g. Nordhaus and Sztorc) to PH99 the so far reported discrepancies can occur (for a mild detuning of Nordhaus' model). From this we will strive at explaining why a recalibration of PH99 is able to solve the reported problem and would do so not only for RCP2.6, but would then do so with identical calibration for *all* four RCPs, including RCP8.5 (see Reply-Fig. 1 for illustration). Thereby we will extend the scope of our article. Regarding omitted literature, to our impression the referee mainly refers to the effects outlined in the following paragraph. As outlined below the next paragraph, our ms is *not* about the issue raised by the referee in that paragraph.

On the methodological approach: What's the justification of using the PH99 model (apart from it 'being there')? The authors argue that it's computational efficiency, …

PH99 can be interpreted as an energy balance model (for details of the justification see Petschel-Held et al., 1999, and references therein; the authors' list contains Klaus Hasselmann! Furthermore, the model was validated in Kriegler and Bruckner, 2004 – however the validation was done in a different way and also did not have the forcing reconstructions by Forster et al., 2013, at hand). Even today, computational efficiency is key, e.g. when it comes to decision-making under endogenous learning (see e.g. Webster et al., 2012).



Reply-Figure 2: Start of an intercomparison of global mean temperature from various AOGCMs and PH99 for RCP6.0. PH99 was fitted to RCP2.6.

…but how exactly are they convinced that their treatment of non-CO2 GHGs is appropriate. For strong mitigation pathways, these 'minor' differences may become very important, last but not least to determine net-zero global GHG forcing etc. I'd think they would need validate their fit using other strong mitigation scenarios with different non-GHG trajectories (if no others are available then from the GeoMIP experiment) rather than RCP4.5. In particular, it appears that non-CO2 gases obscure our assessments of ECS (see e.g. Myre et al. 2016)
As it stands, I'm not convinced that the simplified model is capable of including non- GHG forcing in a sufficient fashion for the question at hand (i.e. staying below 2°C or 1.5°C).

Apparently here we provoked a key misunderstanding what our article is about. It is *not* about how to generate a global total forcing out of regional forcings. Our ms is about how to get from a global total forcing to global mean temperature and whether PH99 does a good enough job in that, i.e. could substitute for AOGCMs or more complex integrated assessment climate modules in that regard. We will make this point very clear in V2. We nevertheless started assembling validation data on RCP6.0 which contains a strong ozone component. This will be completed in V2. Reply-Fig. 2 shows the success of PH99 also in this case. We are not aware of other scenarios rather than the RCP2.6 for which the total forcings had been reconstructed.
Furthermore, we are grateful for the hint that non-CO2 gases might obscure an assessment of ECS. This provides an alternative explanation for the discrepancies

that we report prior re-calibration of PH99. While above we report a failure of PH99 to reproduce certain features of even as simple model as a 2-box model, in fact ECS of various AOGCMs might come with some reconstruction errors. So part of the problem might lie outside the 'responsibility of PH99'.
Finally, any author within integrated assessment of the coupled climate-economy-problem has to deal with how to construct a meaningful global forcing out of regionally disaggregated AOGCM forcings, as AOGCMs cannot directly be utilized in economic optimizations. This is a difficult discussion indeed, but *any* climate module within integrated assessment would face this issue, not only PH99. This discussion, however, is not the subject of our ms.

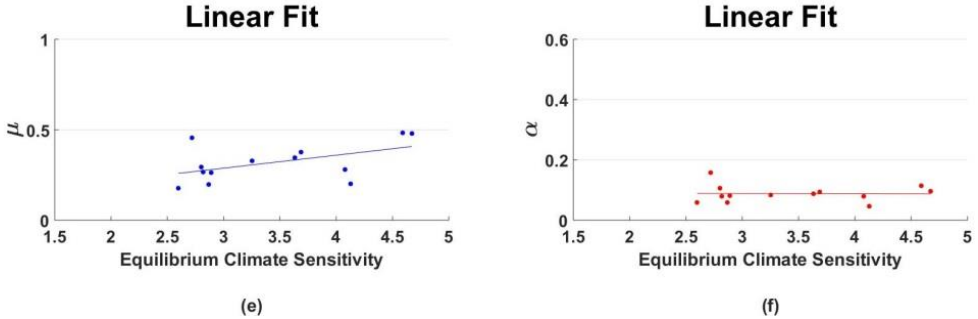
On the application: I didn't fully the motivation for step 1. What was the reasoning for the authors to assume that their PH99 model would work with AOGCM diagnostics from Forsters et al. directly? Obviously, the derived feedback response time parameter 1/alpha of 34.5 years in the multi-model mean is quite unphysical. It seems that the PH99 model is not equipped to be used in that context.

This is the very point. We want to highlight that *current practice* in integrated assessment of directly prescribing ECS and other parameters from AOGCMs leads to biases. As an AOGCM also contains time scales faster than 35 years, it is not immediately clear that an average time scale of smaller and larger time scales would

be meaningless.

In a next step, the authors find that with two free parameters they are capable of achieving better fits. That's not particularly surprising, but a physical interpretation of these differences is virtually absent? ECS is substantial decreased by almost 1°C. Can this be understood? The authors continue with fitting derived and fitted ECS and TCR, but I would rather like to see a physical interpretation or an extension of the PH99 model that would correct for this. The authors should also consider their approach in the light of alternative simplified approaches out there i.e. based on a response function approach as in Ragone et al. (2016).

V2 will deliver on this. By having found that the discrepancy we report can be explained by the move from a 2-box to a 1-box model, we have generated an anchor for explaining the reported ECS effect. A key problem of the 1-box model is that it replaces the slow-component response by the averaged faster one and hence would lead to an overreaction for peak-and-decline forcing scenarios such as RCP2.6. In that sense we expect RCP2.6 as particularly difficult to emulate. Kriegler and Bruckner, 2004, could not do this validation, as mitigation scenario forcing reconstructions were not available at that time.

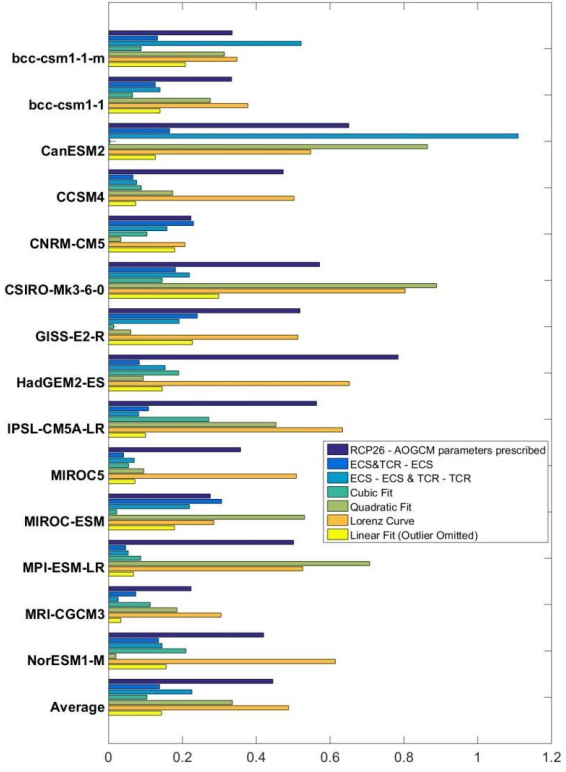


Reply-Figure 3: PH99 parameters vs. ECS.

The authors then want to apply an effective correction for their dubious model in the first place. Their results here appear to be prone to outliers. Compare e.g. the low ECS outlier in Fig. 5. When removing it, I guess even a linear fit would deliver decent results and I'm not sure I can deduce any robust trends from these graphs. . .

V2 will avoid the impression that we advertised utilizing PH99. We simply want to state how to interpret older work based on PH99, and how it could be used if someone wants to use it in the future. In addition to computational efficiency, for some applications also analytic tractability or conceptual simplicity might be

arguments for using PH99.
Furthermore we have done the sensitivity study the referee suggested by repeating some analyses w/o outliers. W/o outliers linear fits would make PH99 parameters predictable by ECS indeed (see Reply-Fig. 3). The scenario fit results deliver similar quality as before (see Reply-Fig. 4). However we would like to stress that our ms recommends direct correction of ECS along ms-Fig.6 where outliers play a less prominent role. V2 will have a discussion of the outlier issue.


bcc-csm1-1-m
bcc-csm1-1
CanESM2
CCSM4
CNRM-CM5
CSIRO-Mk3-6-0
GISS-E2-R
HadGEM2-ES
IPSL-CM5A-LR
MIROC5
MIROC-ESM
MPI-ESM-LR
MRI-CGCM3
NorESM1-M
Average
0
0.2
0.4
0.6
0.8
1
1.2
RCP26 - AOGCM parameters prescribed
ECS&TCR - ECS
ECS - ECS & TCR - TCR
Cubic Fit
Quadratic Fit
Lorenz Curve
Linear Fit (Outlier Omitted)

Reply-Figure 4: The quality of fit does not change by avoiding outliers and using a linear fit instead.

On the application of this. It seems that the model that is used in these IAMs has many flaws. The question then becomes why it is used at all? And not abandoned for a carbon budget approach that would be even more computational effective and can be determined with more complex models also for these low emissions scenarios (i.e. Rogelj 2016). That becomes in particular relevant since the mitigation challenge ahead is to define pathways that hold warming 'well below 2°C'. "Below 2°C" was interpreted as a 66% chance of non-exceedance (IPCC 2014). What's the added value of using a PH99 model in this context? Would they select an ECS at the 66% quantile and then use this as a basis for the IAM derivations? And if so, why not use carbon budgets directly?

We are grateful for this exciting suggestion. It is a fascinating question indeed when a dynamic climate module could simply be replaced by the carbon budget approach to deliver similar – or even better quality – in emulation of an AOGCM. This discussion is beyond the scope of this ms, however we would highlight this option in V2. Here we simply would like to stress that there *are* applications in climate economics where timing matters, such as cost benefit analyses (see e.g. Nordhaus and Sztorc, 2013) or cost risk analyses (see e.g. Neubersch et al., 2014).

In summary we are optimistic that a new version V2 could be acceptable for the reviewer if (i) it clarified its scope: being about the link from total forcing to temperature and (ii) if it delivered a physical interpretation of the observed effects. We are grateful for the referee's comments as they will have triggered an – in our view – considerably upgraded version of our ms.

References Review 1

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. Journal of Geophysical Research: Atmospheres, 118(3), 1139-1150, 2013.

Kriegler, E. and Bruckner, T.: Sensitivity analysis of emissions corridors for the 21st century, Climatic Change, 66, 345–387, 2004.

Neubersch, D., Held, H., & Otto, A. Operationalizing climate targets under learning: An application of cost-risk analysis. Climatic change, 126(3-4), 305-318, 2014.

Nordhaus, W., Sztorc, P., DICE 2013R: Introduction and User's Manual, dicemodel.net , 2013.

Petschel-Held, G., Schellnhuber, H.-J., Bruckner, T., Toth, F. L., and Hasselmann, K.: The tolerable windows approach: Theoretical and methodological foundations, Climatic Change, 41, 303–331, 1999.

Webster, Mort, Nidhi Santen, and Panos Parpas. An approximate dynamic programming framework for modeling global climate policy under decision-dependent uncertainty. Computational Management Science 9.3: 339-362, 2012.