

**Dr Charles JR Williams, BA DPhil FRGS**

**Research Fellow**

1.2n, School of Geographical Sciences  
University Road, Bristol, BS8 1SS

Wednesday 21 July 2021

To whom it may concern,

**Re: Response letter to Reviewer 1, after submission of manuscript “*Simulation of the mid-Pliocene Warm Period using HadGEM3: Experimental design and results from model-model and model-data comparison*” by CJR Williams *et al.* to *Climate of the Past*.**

I extend my sincere appreciation to Reviewer 1 for their thorough examination of my manuscript, and their detailed and highly constructive comments. I propose to address all of your concerns.

Here, I address the your suggestions, comment-by-comment. In the following, your comments are blue and in a smaller font, and my corresponding response follows in a standard font.

I very much hope that my responses will satisfy you and meet your expectations.

Yours faithfully,

A handwritten signature in blue ink that reads "C. R. Williams". The signature is written in a cursive style with a horizontal line underneath.

Dr Charles JR Williams, and co-authors

**REVIEWER 1**

Williams et al. described the experimental setup and model results from the HadGEM3 mid-Pliocene simulation. Modeled temperature responses are compared with proxy, previous generations of HadGEM models, and the other climate models that participated in the PlioMIP2. Consistent with the high climate sensitivity of HadGEM3, its Pliocene simulation is warmer than the other CMIP6/PMIP4 models and the previous generations of HadGEM models. The simulated Pliocene warmth is also larger than proxy reconstructions.

The paper is in general well written and easy to read. The authors did an excellent job describing details of the simulation setup. Simulation results are also presented clearly. In my view, major drawbacks of the manuscript are its short simulation length and the lack of in-depth analysis of simulation results. The simulation length is certainly limited by computing resources but the implications on results presented in the manuscript should be more clearly stated. The lack of in-depth mechanistic understanding of model results may be due to the descriptive nature of the manuscript but clearly makes the manuscript scientifically less interesting. Please see my detailed comments and suggestions below.

1. The short length of the simulation. The fully coupled simulation was carried out for a total of 576 years. At the end of the simulation, model temperatures and radiation still exhibit a large trend (Table 1). In particular, Figure S5 shows that the net energy imbalance at top-of-the-atmosphere (TOA) is close to  $1.0 \text{ W m}^{-2}$ , which is not a small number. This is an important piece of information and should be better presented in the manuscript. (1) I suggest the authors add a “Gregory-like plot” by simply plotting the TOA energy imbalance against the global mean surface temperature for the production simulation. This new figure will provide information on the equilibrium state of the production run and perhaps provide some estimates of the magnitude of the final warming if the simulation is run to equilibrium state. (2) When discussing results on model-model and model-data comparison, this caveat should be more clearly stated. For example, the global mean SAT/SST anomaly in the model and the root mean squared error from proxy data are very likely greater than values listed in the manuscript. In that sense, the performance of HadGEM3 in the simulation of mid-Pliocene could be poorer than indicated in the manuscript. (3) If possible, please also document the equilibrium state (e.g. TOA energy imbalance) of the other HadGEM model versions and take it into consideration in model-model comparison. I think it is critically important to state that this simulation is far from equilibrium in terms of the surface climate, as the simulation data has been uploaded to the ESGF and future users may easily overlook this information.

We thank the reviewer for this comment, and agree entirely that with unlimited computational resources a longer simulation would be preferable; using this version of the model, however, it is simply not feasible to run for the thousands of years required to reach full equilibrium. This point will be made more clear throughout the manuscript and in particular we will add a new row into Table 1 to include mean TOA radiation, and appropriate discussion. To respond to the comments individually:

- (1) A Gregory plot will be added into the Supplementary Material and will be discussed in the manuscript.
  - (2) The caveat of the relatively short run length will be made more clear, and will be added into the abstract, results and conclusions.
  - (3) The equilibrium state (i.e. the TOA radiation balance) will be calculated and presented for all of the Hadley Centre models used here, in a new table in the Supplementary Material, as well as appropriate discussion in the main manuscript.
2. The lack of in-depth analysis. The current manuscript is highly descriptive and lacks in-depth analysis to provide any mechanistic understanding of simulation results. The Abstract basically states that the HadGEM3 Pliocene simulation is warmer than the previous generation of models, the other PlioMIP2 models, and likely the proxy reconstruction. In the main text, the authors have not provided sufficient analysis to probe why the simulation is so warm and which aspects of the model simulation may be responsible for the excessive warmth. I suggest the authors add scatter plots of model ECS against the simulated Pliocene warming for both PlioMIP2 models and also the HadGEM family models. Moreover, the authors can show multi-model results on the ratio of Earth system sensitivity over the equilibrium climate sensitivity. These simple plots are not as sophisticated as a comprehensive forcing-feedback analysis but they can provide important insights on mechanisms for the warmth in the Pliocene simulations.

A scatterplot of model ECS against simulated Pliocene warming, as well as against ESS, for all available models will be included, with appropriate discussion in the manuscript.

3. The “Short title” could be improved. The current one describes “research activities” but not “findings”. I think it is better to describe the most important findings, especially when the main title is already on “research activities”. A revised one could be “HadGEM3 simulates a warmer Pliocene than proxy data and other models”.

We thank the reviewer for this suggestion, and will incorporate the revised short title.

4. Line 340: Which boundary layers are you referring to, atmosphere or ocean?

This will be clarified, to insert the word “atmospheric”.