Clim. Past Discuss.,
https://doi.org/10.5194/cp-2018-177-RC3, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Climate of the Past Discussions

Interactive comment on "On the linearity of the temperature response in Holocene: the spatial and temporal dependence" *by* Lingfeng Wan et al.

Anonymous Referee #3

Received and published: 26 February 2019

The linearity of the externally forced temperature evolution during the Holocene is investigated using climate model simulations forced by the total or by individual external forcing factors. In particular, it is tested whether the total forced Holocene temperature variability is a superposition/sum of the individual externally forced temperature responses. Moreover the linearity of the forced temperature response is tested on different spatial and temporal scales. The addressed topic is interesting and important.

Major comments:

- please revise the method section. Sometimes it is not clear what was done and why it was done. Please see specific comments below.

- the discussion should be more extensive, in particular the limitations of the study

(please see the following remarks)

- only a single simulation for each forcing is available. Therefore, a correct definition of external and internal variability is not possible. The internal variability likely differs between the individual simulations and the internal variability is likely not constant during the individual simulations. By summing up the four individual simulations it is not certain that the internal variability cancels out. Moreover, the internal variability might depend on the time and spatial scale. In addition, the ALL-forcing experiment still includes the internal variability. Please make this more clear in the text and discuss.

- an ensemble of Holocene simulations with that model is not available. Therefore, although incorrect, because the internal variability might depend on the forcing, it might be useful to get an estimate of the internal variability of the different time and spatial scales from a long control simulation with the same model.

- I am wondering if it makes sense to investigate the shorter time and also partly the regional scales if only one ensemble member is available. The signal to noise ratio on the shorter time and regional scales might require a larger ensemble size to make a robust statement? Using a control simulation - please see previous point - an estimate of the signal to noise rate might be possible.

- I am wondering if the following definition is useful: "Since our study above shows that the linear response is largely valid for orbital and millennial variability, but not for centennial and decadal variability, we define the variance of the orbital and millennial variability crudely as the linear signals, while define the variance of the sum of the centennial and decadal variability, which is dominated by internal variability, as the linear noise." Please comment.

- Laepple and Huybers (2014) have shown that "a multiproxy estimate of sea surface temperature variability that is consistent between proxy types and with instrumental estimates but strongly diverges from climate model simulations toward longer timescales. At millennial timescales, model−data discrepancies reach two orders of magnitude in

the tropics, indicating substantial problems with models or proxies". Please discuss the implications in the context of the findings

- please describe the filtering method in more detail. It is not clear to me what kind of polynomial was used for the LOESS. Moreover, it is not clear whether the authors used several iteration to get more 'robust' estimates. More important, what is the influence of the LOESS-filtering method on the result, in particular on the linearity of the response.

- please describe the method - used to compute the significance of the correlation - in more detail. If I understand the authors correctly, an AR1 process is only fitted to the ALL-forcing simulation on the different time scales. The Monte-Carlo method is then used to produce an ensemble (PDF) of fitted curves. Then the correlations between the fitted curves and the ALL forcing run are computed and the 95% confidence level is determined afterwards. If I understood the authors correctly, I am wondering if this method is sufficient. I would think that an AR1 process has to be fitted to the ALL forcing run and the superposition (sum of the response of the four individual simulations). Then two ensembles - one for the ALL forcing and one ensemble for the superposition - have to be computed using the Monte-Carlo method. The correlations between these two ensembles have to be used to determine the confidence level. Please make also more clear why you choose the AR1 as a benchmark and how robust the parameter of the AR1 process is, in particular for the orbital time scale.

- it is not clear to me why the authors did not do a spectral analysis of the runs like e.g. wavelet analysis, power spectrum, cross power spectrum ...

- why was the analysis based on the model grid and not on climate modes using e.g. EOF analysis?

Minor comments:

- please be more precise (whole text): please rewrite sentences like 'the linear response is strong' => the response is almost linear; the response is similar to that of a

linear system

- whole text: I would prefer: forcings => forcing factors

- page 3, line 8-9: Please rewrite the sentence

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2018-177, 2019.