Clim. Past Discuss.,
https://doi.org/10.5194/cp-2018-177-RC1, 2019

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

Anonymous Referee #1

Received and published: 20 February 2019

In this work, the authors make a first attempt to investigate the linearity of forced Holocene variability in dependence on the temporal and spatial scale – by using global surface temperature fields obtained from the TraCE-21ka paleo-climate CGCM simulation. The topic is interesting and the results seem to indicate that further research into this direction may provide further knowledge, that will be useful for both, (a) the interpretation of paleo records and (b) the attribution and detection field of research. Nonetheless, I have a number of concerns regarding the conceptual approach (see 'general comments') which, I think, need clarification before the conclusions, drawn by the authors, can be thoroughly evaluated. A few specific questions are listed at the end (see 'Specific comments').

1 General comments

(1) Throughout this work, it seems that the following two *different* questions are mixed up, which makes it basically impossible to evaluate the conclusions drawn from the results:

Q-1. How linear is the response to external forcing? If we denote the temperature resonse to the full external forcing, $F_{all}(t) = F_1(t)+F_2(t)+F_3(t)+F_4(t)$, by $T_R(F_{all}(t))$, the response to the individual forcings by $T_R(F_i(t))$ (with $i = 1, \ldots, 4$), and the internal temperature variability of the five model simulations by $T_{I,all}$, $T_{I,1}$, $T_{I,2}$, $T_{I,3}$ and $T_{I,4}$, respectively, then the linearity of the response could be defined by the extent to which the statement

$$T_R(F_{all}(t)) = \sum_{i=1}^{4} T_R(F_i(t)) \quad (1)$$

holds, and the linearity could be measured by the correlation between the forced response on the left and that on the right hand side of the above equation. In the manuscript, however, the correlation is computed (see Section 2.2) from the full 'forced plus internal' temperature variability, i.e. between $T_R(F_{all}(t)) + T_{I,all}(t)$ and $\sum_{i=1}^{4} T_R(F_i(t)) + T_{I,i}(t)$.

Since this latter correlation is influenced by the signal-to-noise ratio, Var(T_R)/Var(T_I), a small correlation does *not* necessarily indicate the absence of linearity, because it could be that simply the signal-to-noise ratio is small, although the response is still perfectly linear. (One would need ensembles of model simulations for each of the five forcing scenarios, and then use the ensemble average in order to suppress the internal variability.) Hence, Q-1 cannot be answered by this approach (without additional information), unless one would always obtain correlations close to unity, which would indicate strong linearity.

Q-2. What is the relative importance of externally forced vs. internal variability, assuming the response were linear? To answer this question one could use the correlation computed from the full temperature variability, as done in the manuscript, but one had to assume the linearity which, however, is to be proven by this work, in particular, for different temporal and spatial scales.

Hence, the authors should clarify the above issues, and make explicit which of their results contributes to which one of the above two questions. This will also help to clarify the implications of the conclusions for various research fields.

(2) Even if we had ensembles available for each of the forcing scenarios, it would still be possible to obtain a large correlation coefficient although the response is only weakly linear (i.e., mostly non-linear), if the individual response to, for example, one of the forcings F_i is much larger than the responses to the remaining forcings, because in this case the full temperature variability might still be dominated by the response to the strong forcing (the non-linear interactions might still be relatively small). Thus, one would need to know the strength (e.g., in terms of variance) of the responses to the various individual forcings.

(3) In the Conclusions section it should be mentioned that, even if strong linearity for the given model simulations were proven, then this conclusion is valid only for the given range of forcing amplitudes as non-linearities may appear for stronger forcings.

(4) How is it justified to estimate the variance of the internal variability at orbital and millennial time scales by the full variance at centennial and decadal variability (page 7, last paragraph)? That is, why should we have

$$\text{Var}_{orb,mill}(T_{I,all}) \approx \text{Var}_{cent+dec}(T_{I,all} + T_R(F_{all}))? \tag{2}$$

Even if we assume that $\text{Var}_{cent+dec}(T_R(F_{all}))$ is small compared to $\text{Var}_{cent+dec}(T_{I,all})$, this does not imply anything about the relation between $\text{Var}_{cent+dec}(T_{I,all})$ and $\text{Var}_{orb,mill}(T_{I,all})$. Maybe it could be helpful to investigate the power spectra of the temperature variability under the various forcing scenarios?

2 Specific comments

(5) It would be nice if the reasons for showing the linear error index L_e were explicitly discussed, and what the implications of this index are for the linearity. And what is the added value of this index over the correlation coefficient?

(6) Please, be a bit more explicit how the significance levels are computed. For example, how is the AR(1) fit done in case of the correlation, and what is the bootstrap design for the error index?

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