

Interactive comment on “A systematic comparison of bias correction methods for paleoclimate simulations” by Robert Beyer et al.

Anonymous Referee #1

Received and published: 1 March 2019

The manuscript is focused on the evaluation of three statistical methods to correct the biases of the output of paleoclimate simulations. This evaluation is carried out by comparing the corrected simulated values of annual temperature and annual precipitation with the corresponding reconstructions based on proxy record for three periods in the past (Mid-Holocene, Last Glacial Maximum and Last Interglacial Period). The three statistical methods are the Delta Method, Quantile Mapping and Generalized Additive Models that include additional geographical predictors to correct the simulated climatologies. The main conclusion is that, in general, the Delta Method outperforms the other two.

My general impression of the manuscript is positive, but I have two main concerns that I think should be addressed in a revised version. In addition, I have a few more

C1

suggestions to particular points in the manuscript that would benefit from a further clarification.

My evaluation is that the manuscript would need some (not drastic) revisions, but I would like to evaluate the revised version.

Main concerns

1) The evaluation criterion is essentially the difference between the corrected and reconstructed climatology. However, the Delta Method has been specifically constructed to eliminate this difference between simulated climatology and present-day climatology. Quantile Mapping pursues a more general correction, namely to correct the whole probability distribution of annual temperature (or precipitation). The GAM method is a statistical model that incorporates (in my understanding) simulated and observed grid-point climatologies as predictors and predictands, and additionally some other factors like distance to the ocean, etc. The GAM method is therefore also not specifically tailored to eliminate the bias. I wonder if the main result of the manuscript, namely the best performance of the Delta method, is not an artifact. The Delta Method is precisely tailored to maximise the evaluation criterion and thus, it is for me not surprising that it outperforms the other two methods. I am not sure which other, fairer, evaluation criterion could be introduced, but I think that this issue should be addressed or at least thoroughly discussed.

2) The difference between the corrected simulated climatologies and the reconstructed climatologies does not take into account the presumably large uncertainty in the reconstructions and in the corrected simulated climatologies (the former being presumably much larger?). This needs to be incorporated in the evaluation of the three methods. If the inter-methodological differences are much smaller than the uncertainties in the estimated paleo-bias, it would be difficult to claim that one particular method is superior to other two. I think that the manuscript should include also these uncertainty estimations, or at least place the inter-methodological differences in the frame of the

C2

reconstruction uncertainties.

3) The readability of the illustrations is poor. It is, for instance, very difficult to discern anything in Figure 2 and Figure 3. The lettering, axis labels, etc, in most figures is too small (e.g. Figure 4)

Particular points

4) What is the original spatial resolution of the climate reconstructions? Were they regridded, and how?

5) The text refers sometimes to bias, other times to 'error', whereas in my understanding very often both terms carry the same meaning. This can be confusing for some readers. I would recommend to stick to one of those terms when possible.

6) The text also refers to the climate reconstructions as 'the observations', e.g. in equation 5. This can also be confusing. It would be clearer to use 'climate reconstructions' when referring to the proxy-reconstructed climatologies and 'observations' when referring to present-day climatologies.

7) The main conclusion is derived from the analysis of only one model. Perhaps I missed it but I think this a caveat that should be mentioned.

Interactive comment on Clim. Past Discuss., <https://doi.org/10.5194/cp-2019-11>, 2019.