

Interactive comment on “Methodological and physical biases in global to sub-continental borehole temperature reconstructions: an assessment from a pseudo-proxy perspective” by Camilo Melo-Aguilar et al.

Anonymous Referee #1

Received and published: 6 November 2019

Review of the manuscript “Methodological and physical biases in global to sub-continental borehole temperature reconstructions: an assessment from a pseudo-proxy perspective” by Camilo Melo-Aguilar, J. Fidel González-Rouco, Elena García-Bustamante, Norman Steinert, Johann H. Jungclauss, Jorge Navarro, and Pedro J. Roldán-Gómez.

This manuscript assesses the spatial and temporal limitations of ground surface temperature histories reconstructed from borehole temperature profiles, as well as the effect of different forcings on the interpretation of those temperature histories. Authors

C1

used a large ensemble of millennial simulations, which includes experiments to evaluate the role of individual forcings in the simulated climate evolution, for this assessment. I find this work relevant for the study of the Earth’s system dynamics, paleoclimate, and climate modeling. The evaluation of limitations on the borehole database is of particular relevance. Nevertheless, I think some issues regarding the inversions and trend analysis should be addressed before the manuscript is ready for publication.

Inversions:

1- Authors limit the depth of the synthetic boreholes that they create mimicking the depth distribution of the measured borehole database. This limitation seems to be only applied to the bottom depth in the synthetic profiles, as the upper depth is not mentioned on the text. I assume, therefore, that the upper depth in the synthetic profiles is the surface. However, measured borehole temperature profiles rarely include data at the surface, and I wonder if authors have studied the case in which the upper depth of the synthetic profiles is configured as the upper measured depth in the corresponding borehole temperature profile. Would this have an effect on the results?

2- The authors aggregate inversions performed using profiles with different bottom depths. Previous studies [1,2] analyzing measured temperature profiles have shown that this practice biases the retrieved surface temperature histories, since the period of reference for each subsurface anomaly profile is different. I realize that authors are not using measured temperature profiles and that the synthetic profiles may not share this bias with the real case, but I wonder if authors have assessed the case in which all synthetic boreholes are truncated to a common bottom depth. Additionally, authors should include a brief note in the conclusions stating that although the depth masking do not affect their results, this is not the case when analyzing real borehole profiles.

3- Related to the previous comment, I have noticed that each simulated GST anomaly used to generate synthetic profiles have a different period of reference (lines 240-243). Therefore, the inversions of those synthetic profiles are relative to different climatolo-

C2

gies. An easy solution to that would be to define a common period to compute all GST anomalies before generating the synthetic boreholes, and maintain such reference in the comparison with temperatures from the model.

4- It is very surprising that global mean temperatures from inversions computed using the B_mask configuration (green lines in Figs. 2, 3, 6, 7 and 8) are perfectly aligned to start in the year 2000 of the common era. Nonetheless, authors clearly state at several points on the text that the B_mask configuration considers the logging dates of the real borehole database. Does this mean that such different dates were not considered when aggregating the inversions? The green lines should display some kind of shift relative the ideal borehole scenario (IBS, red lines in Figs. 2, 3, 6, 7 and 8) configurations due to the different logging dates.

5- Which is the difference between GST_mask and B_mask? The authors state on line 306 that GST_mask was built by sampling GST in time, space and depth following the real borehole distribution. But the maximum simulated depth is 42m (35m is the last model node). I find this statement misleading, since borehole depths are much deeper than the simulated depth and it is not indicated which temperature is GST (although I suppose it is GST_L12, see Minor Concerns below).

Trend Analysis:

6- The comparison of trends under different masking configurations constitutes the core of the work, and yet I believe more details are needed regarding the trend analysis. There is no reference to the test applied to determine the significance of the trends. A common t-test would not be suitable for climate series due to the displayed autocorrelation, thus another test should be applied.

7- Also, it is not clear which is the method followed to generate the whisker plots comparing IBS_L12-GST and IBS_SAT-SAT trends in the linear fit case. SAT and GST temperatures are annual series, while the length of the time steps in IBS_L12 and IBS_SAT is 15 years. I guess that the trend of SAT (GST) and IBS_SAT (IBS_L12)

C3

were estimated independently and then subtracted, but if so, I do not know if this is the correct approach. Should the annual temperatures be averaged in 15yr-periods, then subtract the reconstructions and estimate the trend of the resulting series? Maybe both approaches yield similar results, but a clarification here would be very helpful.

8- Crosses in Figs. 6 and 8 should represent the median of the GHG-only and LULC-only ensembles, since the authors are comparing against the median of the rest of ensembles (horizontal lines in the boxes). Note that the median is not always equal to the mean of the distribution.

9- In line 313, it is stated that “borehole reconstructions are able to retrieve the masked or unmasked GST” based on the results of Figure 2. However, there is no analysis of the trend of the GST-B_mask case. There is an analysis of the IBS_L12-B_mask case, but the IBS_L12 configuration is not the same as the unmasked case. Why not to include the trend analysis of the GST-B_mask case? Something similar can be said about the SAT-B_mask case in Figures 6 and 8.

Minor Concerns/Mistakes:

10- Equation 2 is incorrect. Right term of the equation should display the second order partial derivative of temperature. Check Carslaw and Jaeger (1959).

11- Section 3.2: does ST_L12 means GST_L12 as in the rest of the text? If so, please be consistent through the text.

12- Related to the previous comment, the definition of GST is not clear on the text since Section 4. Is it GST_L12?

13- Line 301: Which are the trends in Hartmann et al. (2013)? You could add those numbers to the text for an easier comparison with your results.

14- Line 381: “poor sampling enhances the influence of local behavior”. What do authors mean here by local behavior? Please, expand this sentence.

C4

15- Line 451: I believe authors mean “50% of the simulated trend”.

16- Lines 308 and 461: which is the level of confidence in the statistical test applied to this trends?

17- Line 506: change “bu” by “by”.

18- It is not clear which metrics are affected by the different masking configuration is Figs. 2 and 4. Is B_mask affected by those limitations or B_mask is always defined as indicated in Section 4.1? This should be easy to clarify on the captions and on the text.

References:

1- Beltrami, H., J.E. Smerdon, G. Matharoo, and N. Nickerson (2011). Impact of maximum borehole depths on inverted temperature histories in borehole paleoclimatology. *Climate of the Past*, 7, 745-756, 2011.

2- Beltrami, H., G. Matharoo, and J.E. Smerdon (2015). Impact of Borehole Depths on Reconstructed Estimates of Ground Surface Temperature Histories and Energy Storage. *Journal of Geophysical Research - Earth Surface*. 120(5): 763-778.

Interactive comment on *Clim. Past Discuss.*, <https://doi.org/10.5194/cp-2019-120>, 2019.