**Climate
of the Past**

Discussions

Open Access

EGU

Interactive comment on "Comparison of observed borehole temperatures in Antarctica with simulations using a forward model driven by climate model outputs covering the past millennium" by Zhiqiang Lyu et al.

Anonymous Referee #1

Received and published: 13 April 2020

The article provides the first systematic model-data comparison based on borehole temperature-depth profiles in Antarctica. They elaborate two techniques (depth and time domains) to compare these profiles and their reconstructions from four sites with climate model output. They conclude by outlining some useful metrics for future model-data comparison and highlight the importance of internal variability on the observed tendencies.

Below are some points that need clarifying or addressing in the manuscript.

1. L51: "Since the variable measured in the borehole is the temperature itself,..." In most cases, resistivity is measured, which is easily converted to temperature. Is this true of your measurement techniques?

2. L53: "the surface temperature history makes the reconstruction mathematically undetermined." The equation of temperature at depth usually results in a system of linear equations which is mathematically under and overdetermined. Varying mathematical inversion techniques are then utilized to reconstruct the ground surface temperature history. Please clarify. This should also be clarified on L.267.

3. L.92: "Previous studies using forward models driven by climate model outputs were focused on ground temperature and not to borehole..." Please provide a couple examples (references) here.

4. In Section 2.1, the borehole measurements and reconstructions are briefly explained. Since they come from four different publications, how can differences in inversion/reconstruction techniques affect the results presented in Figure 1?

5. L. 136: "CESM1-CAM5 and MPI-ESM-P are not continuous in 1850." What is meant by this? Please clarify.

6. In Section 2.2, for all models, excluding the CESM ensemble, which realization (r1i1p1) is used? Is it the only realization available? If not, why was this one selected? Please clarify.

7. L. 160: "...for Mill Island, the heat flux is set to zero..." How realistic is this? Furthermore, this is a different technique than the other sites. Could this influence the results? If the heat flux is set to zero, how is the steady-state temperature calculated?

8. L. 161: "For WAIS, a vertical step of 1 m for the upper 500 m and up to 25 m for the deepest part, and for other sites where the depth of borehole is close or less than 500 m, the step is set to 1 m for overall depth." Why are various techniques used again? What is the benefit of this?

9. L. 183: "At WAIS-Divide, the spread of the sensitivity tests is lower than the spread if the different scenarios." What is meant by scenarios? Is it the different models being analyzed?

10. L.196: "...but the deviation in the top 100 m show that there is climate information stored in the upper part of the profile, and that this profile cannot be fully determined by boundary conditions." Climate is not the sole reason why the top 100 m would show deviation. How can you be sure it is climatic information?

11. At the start of the paragraph at L.198, it is stated that internal climate variability and the different characteristics of the climate models are the main sources of differences. The results from the CESM ensemble have not been discussed in this section. To strengthen this point, I recommend adding in a discussion of it. Furthermore, the statement that internal climate variability and different characteristics of the climate models being the main source of differences does not hold true for Mill Island. In Figure 1, only different depths of the zero heat flux are considered. More tests must be added to conclude the importance of the influence of internal variability and different model characteristics to the differences at this site.

12. In Figure 3, why are different smoothing techniques used? Can it influence the results?

13. The reconstructions from the climate models presented in Figure 3 are calculated using what technique? Their errors bounds are also not presented. How does this influence the results? Does the reconstruction from the climate model always lie within the error bounds? Please clarify in the manuscript.

14. L.215: "In order to remove the bias on the mean state for each climate model, anomalies are shown using the total period covered by each reconstruction as reference." Which figure is being referred to? The paragraph starts discussing Figure 3 but these are not anomalies.

15. Since the temperature variability increases as you go back in time, there is less confidence with respect to the timing of events. Timing of events varies within climate models. Could this further explain any discrepancies of the timing of events? Please discuss.

16. What causes the decrease in temperature at ∼1980 and ∼2000 in Styx and Larissa (Figure 3)? Is it climatic in origin or an artifact of the reconstruction technique?

17. L.390: "Fig. 8 shows the spatial correlation in the Antarctica Peninsula (AP)." Do you mean the spatial correlation of the gridcells? Please clarify.

18. L390-393: "Despite the correlation coefficient decreasing as the grid getting far away from the Larissa, the values, at least around Larissa for each model, are higher than 0.6, showing that this metric is representative of the whole peninsula region, and not extremely site-specific." A correlation coefficient of 0.6 means that it only explains ∼36% of the variance. How can you conclude that it is representative of the entire peninsula?

19. Why are the CESM ensemble members not presented in Figure 8? How is this metric influenced by internal variability?

20. L.426-428: "A model that responds clearly to the Ozone forcing, and has a strong SAM signature should exhibit this dipole pattern, and it is interesting that some models do not show it, indicating that the Ozone forcing is not dominating over internal variability." The CESM ensemble members are not seen in Figure 10 and 11 nor discussed. How can this be concluded?

21. Borehole temperature profiles and their ground surface temperature histories are compared with those from climate models. They ability of the climate models to reconstruct the ground surface temperature was evaluated and three distinct metrics were created. From all of this, how do you think climate models could improve? From your analyses, what are their areas of weaknesses? It would be beneficial to add a section

outlining this to the conclusions.

Technical Points:

There are many grammatical errors throughout the article impeding the reader's comprehension. They are not all outlined below but should be addressed in the revised manuscript.

1. L36: Please define acronym AP.

2. It would facilitate comprehension if the depths of each borehole were added to Table 1. This would help explain the various time periods for the reconstructions found in Figure 1.

3. Please add the units of elevation to the map in Figure 1.

4. In Figure 2, in the boxes below Figures 2a,b and 2c,d, there is a typo in the word accumulation. The thermal diffusion used along with its units should be included. In the caption, " 2) sensitivity tests using the temperature history of once CESM member...", do you mean one CESM member? Also "The shade area represents the simulated subsurface temperature ensemble driven by CESM" should read "The shaded area..."

5. L.192: "At Mill island,..." Should read Mill Island to be consistent throughout the text.

6. L.192: "..the ice thickness is much deeper..." Ice thickness cannot be deeper. Should read thicker.

7. In Figure 4, the y-axis of 4a,e and f are crowded. Either decrease the amount to y-ticks or increase the figure size. Some of the symbols, in particular the yellow triangle of CCSM4, are difficult to see. I would recommend increase the size of the markers for the climate models and the reconstruction. Also, the labels on 4c and d are cut off by the below figures. Please fix.

8. L.254: "Larissa shows a temperature minimum in 1940's..." should read ...1940s.

9. L.270 a period is missing at the end of the sentence.

10. In Figure 5, please use a different colour for the observations.

11. For consistency, use CCSM or CCSM4.

12. For the techniques/metrics elaborated in Section 4, please be consistent with the use of grid, grid-point, and gridcell. Since you are comparing with data from the gridcell, I'd recommend the use of that word to facilitate the reader's comprehension.

13. L.372: "For most of the models,..." It would be best to include a number or percentage of models to really illustrate your point.

14. In Figures 6,9,12, please add the units to the colour bar as these are surface temperatures tendencies.

15. In Figures 6,7,9, the circle illustrating the location of the observations is not clear. Maybe make it bolder or another colour.

16. L.394: "Figure 9 shows the same temperature trend (1825-1925) for all models." Do you mean surface temperature since Figure 9 shows varying trends.

17. L.395-396: "A majority of the CESM members(CESM1, 7, 8, and 9)..." Do you mean minority? 4/12 is not a majority.

18. L.403: There is a typo in the word overestimation.

19. In Figure 8, the dotted contour line is not clear to the reader. Also, indicate that the colour bar represents the correlation coefficient.

20. L.427 please define acronym SAM.

21. Figure 10, some of the numbers in the colour bar appear to be cut-off. The red-dashed line is not visible to the reader. Please correct.

22. Figure 11, the y-label of d is overlapping with c. Please clarify that it is the linear trends of surface temperature in the caption.

23. Figure 12 is not referenced in the text.

C7