

Reply to anonymous referee 1

We thank the referee for the constructive comments.

In the following, the original referee comments are given in italics and the reply by the authors in bold.

Review of: Internal climate variability and spatial temperature correlations during the past 2000 years

I think this paper is in very good form. I only have a few minor comments.

Lines 10-12: I found this sentence confusing: 'However, combining the iLOVECLIM results with CMIP5 model results and various PAGES-2K temperature field reconstructions, we find that neither model results or reconstructions are robust.' The models and reconstructions are not robust in what sense? I think this sentence could be more clearly re-worded to be more specific or perhaps even cut entirely given the content that follows it.

Thanks for pointing this out. We have simplified the text by combining the two sentences into “However, combining the iLOVECLIM results with CMIP5 model results and various PAGES-2K temperature field reconstructions, we show overall agreement for the magnitude of continental temperature variability in models and reconstructions, but both the simulated and the reconstructed ranges are large.”

All of the main text figures have quite small font sizes on nearly all the labels and axes. Their readability would be much improved by increasing the font sizes of everything in the figures.

Thanks for pointing this out. We updated all main text figures to increase font sizes where possible.

Could you comment in the paper on how the spatial resolution of the different datasets may influence the comparisons you've done? LOVECLIM is 3 degrees, the PAGES2k reconstructions are 5 degrees, the CMIP models are ~2 degrees, and the proxy data is based on individual points, so there's a lot of spatial averaging differences that could affect the comparisons.

There seem to be two aspects to this very relevant question. Firstly, we did not find any evidence to suggest that the differences in spatial resolution between the different products that we use (2-5 degrees) impacts our findings. First averaging all fields to a common 5 degree resolution yields the same results.

There are, however, important questions related to the general comparison of point-based temperature reconstructions (like McGregor et al. 2015) and coarse resolution gridded products like model output or CFR-based reconstructions. We partially addressed this question in the discussion on lines 357 to 359 of the original manuscript. However, since this is an important question, we have rephrased it on lines 355-358 to read “More in-depth studies are need to understand and resolve the differences between reconstructed point-based temperature variability on the one hand, and temperature variability in coarse resolution products (2-5\ degree{} in this study), like model results and CFR-based reconstructions, on the other hand.”