

Editor comments (author responses in red)

We appreciate the in-depth comments and suggestions from the editor to improve the manuscript, from general comments to specific line-item comments. We have taken most of these into account, and detail our changes and responses to other suggestions below.

The term “paleotemperature measurements” comes up frequently. In some contexts I think this is appropriate, for example where it’s clear that you’re referring to measurements undertaken to derive a paleotemperature estimate. In other cases, for example line 23 in the abstract, my sense is that “paleotemperature estimates” is a more appropriate because the datum of information (e.g. MAT at X location Y thousand years ago) was not measured but rather estimated from measurements.

In the abstract and introduction, we do mean to refer to the measurement (act) rather than estimate (measurement output). However, we recognize that the description did not make that clear. We have modified the text accordingly (lines 23-25, 65-67) to make this more clear.

The point about uncertainty and geological time is an important one and it’s elegant to turn to Urey’s pioneering work to make the point. But phrasing of the point is a bit awkward and I recommend some wordsmithing.

We agree that the wording was awkward there – we have modified this to improve clarity.

Line 186: You may not agree, but it seems optimistic to suggest that calibration can “effectively remove” the influence of CCFs in even the most well-characterized proxy system. Is it more realistic to instead say “account for” or “mitigate against”?

We agree – while control measurements do “effectively remove” the influence of CCFs, proxies (the subject of the paragraph in question) do, at best, account for CCFs. This has been changed accordingly.

The section from lines 305 to 320 is interesting, and I’m not sure if I agree with the assessment that sample processing, human error, etc give rise to “unquantifiable uncertainty”. Tools such as secondary reference standards (ie separate from instrument calibration standards), replicate analyses, and process blanks are all ways to account for and, in some cases, quantify the impacts of sample handling. I suggest a long chat with someone who runs a mass spectrometer for a living... they would likely have some good thoughts on this.

We recognize that the structure and wording in that paragraph was not clear. We have modified that section to better make clear how unquantified and quantified uncertainties relate to analytical uncertainty, making a more straightforward description of how researchers can indeed quantify such uncertainty.

Re: transfer functions, I was surprised to see the discussion come up relatively late. Since transfer functions are the sets of equations used in “correlation constrained proxies”, it would seem natural to bring up the term there. This would also provide an opportunity to give some more examples of correlation constrained proxies (e.g. chironomids, pollen, ice core/permafrost water isotopes, etc). This would contribute to addressing, in part, the next point...

We don't quite agree that transfer functions are universally used by correlation-constrained proxies to the point that they should be brought up immediately when introducing those proxies. We see them as more of a tool that is employed to push correlation-constrained proxies closer to the observation/inference-constrained proxy side (by identifying CCFs and their relative influence). Where the discussion of transfer functions is situated currently represents that idea of transfer functions as tools used to improve our understanding of proxies, ultimately to push them away from correlation-constrained proxy.

Reading the manuscript again, there's a perception of beating up a bit on the TEX proxy. I'm not sure if this is intended or not (or maybe I'm reading more into it because I'm also handling a manuscript that is critical of current TEX calibration schemes). I can't point to anything specific, but if that wasn't your intention I'd urge you to take a close read of the sections relating to TEX and see if the tone can be subtly shifted to avoid that perception. You could also consider adding a few sentences of text to summarize some of the major contributions made by workers applying TEX to paleoclimate problems.

We, like many others, do think that too much confidence has been ascribed to the TEX86 proxy, given its relationships with CCFs we outline in our manuscript. In many ways, our manuscript provides an epistemological critique of the proxy that is meant to highlight shortcomings and room for improvement. We have changed the language of some sections (e.g., lines 317-330) to focus more on how the data is represented (rather than how researchers represent the data), and feel that our discussions of the ongoing work to improve TEX leave the paper with an optimistic tone.

Fig 1: the numbers on the X axis seems impossibly low to me. Surely there are (at least!) 90 papers a month published in various journals that discuss paleotemperature. You don't need this figure to support the point that there has been an increasing scholarly interest in paleoclimate over the last several decades, so I suggest deleting it or re-doing the analysis with more inclusive search criteria.

We agree that the figure adds very little to the manuscript, and have removed it.

Re: “first quantitative paleotemperature proxy” (line 152), we are not aware of any truly quantitative paleothermometer prior to 1950 – neither tree rings nor varved sediments, to our knowledge, provided/provide quantitative temperature estimates. If the editor is aware of such applications, we are happy to change this language, but currently feel that the $\delta^{18}\text{O}$ calcite proxy is well known to be the first.

We have also made small edits throughout the manuscript to improve clarity elsewhere.