



EGUsphere, author comment AC3
<https://doi.org/10.5194/egusphere-2022-184-AC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on CC1

Michael P. Erb et al.

Author comment on "Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation" by Michael P. Erb et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-184-AC3>, 2022

Thank you for these comments. We've replied to your comments in order below.

1) Proxy records in the Temp12k database have a large range of temporal resolutions. Some ice core, speleothem, and two wood records have high resolution, although these archive types are in the minority. Additionally, a considerable amount of marine sediment, lake sediment, and peat records have mean resolutions lower than 200 years. While bioturbation may smooth some of this signal, our approach attempts to retain higher resolution information rather than smoothing the proxy signals up front. This results in a climate reconstruction which, while nominally decadal, simply represents the information contained in the proxy database. We will clarify this in the paper and stress that the decadal nature of the reconstruction does not imply that it contains robust decadal information. Instead, we want to use the data present in the database. This is similar to the approach used in Marcott et al, 2013 ("A Reconstruction of Regional and Global Temperature for the Past 11,300 Years"), which presents a temperature composite with a resolution of 20 years despite a median resolution of 120 years in their proxy data.

2) In our approach, the data assimilation is computed on a decadal timescale, so we need to have assimilated proxy values on the same timescale. To regrid proxy data to a decadal resolution, nearest neighbor interpolation was used to regrid data to annual resolution (e.g., the value for a given year is equal to the closest value), then binned to decadal. The intermediate step of interpolating to annual is meant to better account for sub-decadal data and decades mid-way between data points. A similar approach is used in some of the composites in Kaufman et al., 2020 ("Holocene global mean surface temperature, a multi-method reconstruction approach"). The approach is simply meant to represent multi-decadal data on a decadal timescale. With this said, we may take a different approach that also addresses age uncertainty concerns; see #3 below for more explanation.

To investigate how our interpolation approach affects the reconstruction, two reconstructions were compared. Both use 200 year bins. In one case, proxy data was interpolated between bins using nearest neighbor interpolation (this reconstruction is shown in Fig. B2h of the paper). In the other reconstruction, no interpolation was done. Both approaches lead to similar global-mean temperatures. The uncertainty range of the reconstruction using no interpolation was larger, but only by a small amount. The main difference between the reconstructions occurs near 11 ka, where many of the proxy records are cut off. The non-interpolated reconstruction has a rapid change in temperature here, which is likely not a climate signal, while the interpolated version is smoother.

3) Currently, age model uncertainty is not accounted for in the reconstruction, as mentioned in lines 542-543 & 586 of the paper. To account for age model uncertainties, we plan to rerun the reconstructions using a new methodology, which will work as follows: for each proxy record, we will generate a set of possible proxy realizations using the geoChronR R package (<https://nickmckay.github.io/GeoChronR/>) that sample the proxy's age and magnitude uncertainties. Each proxy's ensemble will be used to quantify the median proxy estimate and the joint age/magnitude uncertainty at a decadal timescale. (This approach will be used instead of the nearest neighbor interpolation approach described in the last comment.) The median proxy estimate will be assimilated and the joint age/magnitude uncertainty will be used to define a time-varying uncertainty term at each decade for each proxy (i.e., the R term in Eq. 2 of the paper). By combining age uncertainty and magnitude uncertainty into the R term, this method will account for age uncertainty without having to run a large collection of reconstructions. To our knowledge, this implementation hasn't been used yet for paleoclimate DA, but should account for proxy uncertainties with a reduced computational expense. If there is a good reason why this is not a good solution, please let us know.

4) In the multi-timescale DA approach, the model data is used to compute covariances between different temporal resolutions. To use an example, let's say that a proxy data point represents a 50 year mean. To relate 50 year means to the decade being reconstructed, the model-based proxy estimate is computed as a set of 50 year means, with each 50 year mean centered (to the extent possible) on the decades in the prior using a running average. In theory, this lets us quantify how 50 year means relate to decadal climate, and this information is used in the data assimilation. Additionally, since a data point representing a 50 year mean spans 5 decades, the process is repeated for each of the decades, with the only difference being our use of slightly different model values due to the moving prior window. The use of a running mean for multi-decadal averages is not perfect, but difficult to avoid without having a very large number of modeled years with which to construct the prior. In the paper, we will clarify the description of the multi-timescale data assimilation approach.

This multi-timescale approach is theoretically useful, as it is designed to allow each proxy to inform the reconstruction on its own timescale. While the final result is not drastically different from a reconstruction made using a single-timescale binned approach, this paper is partly an exploration of new methodological choices. We feel like that is a good justification for using the method and alternate experimental designs are shown in Appendix B. Regarding model covariances on different timescales, it's possible that the models do not properly quantify climate covariances on different timescales; however, that would be a shortcoming of the data rather than the method. In the released code, if users do not want to use the multi-timescale approach, it can be disabled by setting two variables (`time_resolution` and `maximum_resolution`) equal to each other in the config file.

5) You're correct that 0.09C may be too small to be considered "warmth," as suggested by the uncertainty bands in Fig. 12, which include 0 from ~8 ka to the present. Using a t-test of two related samples of 1002 values each (the size of the prior), the mid-Holocene period (5.5 - 6.5 ka) is significantly different from the recent period (0-1 ka). However, it's unclear if this is the right significance test to use. If ensemble members are analyzed individually, 88% of them have a positive anomaly at the mid-Holocene and 12% have a negative anomaly. We welcome better suggestions for a significance test. If a clear significance test cannot be established, the paper text can be modified to remove the assertion of mid-Holocene warmth, instead stating that our reconstruction does not support a cooler mid-Holocene.

6) For this first paper, we use the calibrated data from the Temperature 12k database, which is readily available for a wide range of proxy types. PSMs provide many potential advantages for data assimilation and represent an exciting avenue for exploration, but are

not available for all of the proxies we use (e.g. pollon). Additionally, some PSMs are best used with isotope-enabled model simulations, which limits the choices of prior for the Holocene. The use of PSMs will be a focus of future study.

Line 65: We will be more specific about the inclusion of land records. However, it is difficult to describe the full difference in the databases, since Osman also includes many marine sediment records which are not currently in the Temp12k database.

Line 80: We will clarify the time-resolution and age control proxy criteria.

Line 98: We will include a figure of proxy temporal resolution.

Line 99: It would be useful to have a better sense of the temporal coverage of data points. However, that data is not widely available in the Temp12k database. As such, we've decided to treat data as continuous and mentioned our justification in the paper. While unrealistic, this assumption has the benefit of avoiding possible non-climatic signals generated by the alternation between proxy data and data gaps. To explore the effect of interpolating proxy data, we did the comparison described in point #2 above; however, we plan to replace the interpolation approach with the ensemble approach described in point #3 above.

Line 126: The minor checkerboard pattern is present in the results of the HadCM3 simulation. It seems to result from processes involving the ocean streamfunction and the model orography, although the details of this are not completely clear. In time-slice simulations, rerunning a simulation with smoothed orography causes the checkerboard pattern to disappear but the general climate remains the same. Because of this, we decided to simply apply a smoothing filter to the model results, which is mentioned in line 126 of the paper. We do not expect this to have much effect on the reconstruction.

Line 137: Yes, decadal resolution is higher than the mean resolution of most of the records. This is done by design, because we want to utilize information from our high-resolution records. Making use of data at a variety of temporal resolutions is one of the primary goals of the multi-timescale data assimilation methodology. As mentioned in point #1 above, the widely-cited Marcott 2013 proxy composite also uses a temporal resolution at the high end of the utilized proxy data. Regardless, we will add a note that our reconstruction is "nominally" decadal, but actually represents the information content in the proxies themselves.

Line 158: We agree this wording was unfortunate, since of course there are many PSMs that have been published by both you and by Sylvia and others. We will rephrase this statement. This first study uses the calibrated temperature records and we plan to move forward with exploration of proxy system models in the future. One complication to the use of proxy system models is that many use isotopes as inputs, and there is a limited availability of isotope-enabled simulations relevant to the full Holocene. Osman et al. 2021 used newly-run isotope-enabled simulations for their prior, but running new simulations is a considerable undertaking. Instead, we plan to explore the use of existing simulations, such as iTRACE. Additionally, by using calibrated temperatures for this first reconstruction, we can use it as a baseline to explore the added value of proxy system models in the future.

Line 515: This calculation uses the time windows of 5.5 - 6.5 ka vs. 0 - 1 ka. This is specified on line 515.

Line 518: This is discussed in comment #5 above.

Line 537: We will cite Osman et al., 2021 in this discussion.

Line 576: Unfortunately, metadata about sample thickness is not present in the Temp12k database for the vast majority of records. It would be good to have this information standardized and included in future databases.

Regarding localization: Covariance localization presents some benefits and some drawbacks, which are discussed in Appendix B.2. Among the drawbacks: 1) it is not clear which long-distance relationships are valid, so covariance localization is largely arbitrary and 2) covariance localization diminishes the overall impact of proxies on the reconstruction, allowing the prior to have greater influence on the reconstruction. However, the approach also has benefits, so we show some experiments using covariance localization in Fig. B2 and localization is an option in our released code. We will continue to explore its use in future reconstructions, and will clarify the discussion in Appendix B.2 if needed.

Line 835: Interesting! I didn't realize that the covariance localization was possible with the simultaneous data assimilation approach. We will fix this oversight in the text.