

Earth Syst. Sci. Data Discuss., referee comment RC1  
<https://doi.org/10.5194/essd-2022-38-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on **essd-2022-38**

Anonymous Referee #1

---

Referee comment on "LegacyClimate 1.0: a dataset of pollen-based climate reconstructions from 2594 Northern Hemisphere sites covering the last 30 kyr and beyond" by Ulrike Herzschuh et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-38-RC1>, 2022

---

Review of LegacyClimate 1.0: A dataset of pollen-based climate reconstructions from 2594 Northern Hemisphere sites covering the late Quaternary by Herzschuh et al.

The authors provide temperature and precipitation reconstructions based on pollen assemblage time series. They provide three different types of reconstructions and provide a clear description of the methods. The dataset is highly valuable and the manuscript is clearly written and the figures are of high quality (if sometimes a bit small). The manuscript seems to be part of a set of articles (a trilogy?): a manuscript describing the raw pollen data, a manuscript dedicated exclusively to the chronology and the present manuscript about the pollen-derived climate reconstructions. I can to some degree follow the rationale of the sequence, but I think this (last?) article would benefit from a closer integration with the article describing the chronology. The chronology, and importantly its uncertainty, is an integral part of the climate reconstruction that the authors present here. In addition, I have some further recommendations and points that require clarification in a revised manuscript.

### Major issues

#### Integration with chronology:

this manuscript focuses entirely on the reconstruction of temperature and precipitation, yet the time series also have a chronology with associated uncertainty. By separating these two aspects into two manuscripts it becomes unclear how the full uncertainty of the

paleoclimate time series can be derived. Looking at the data (on pangaea.de) it seems that the provided error only accounts for the reconstruction, not for the chronology. This is not the full story and the manuscript would be tremendously improved if the authors made this third manuscript of the sequence a true integration of the papers on the chronology and the climate reconstruction. In L341-343 the authors even touch on this possibility, but they refrain from taking the logical next step that would make the data product more useful for other researchers.

This means that the first order analysis of the time series as shown in figures 5 and 6 should include some combined error resulting from the reconstruction and the chronology and a clear description of the methodology to combine these errors. The provided data sets should also contain uncertainties that reflect both the chronological and the reconstructions errors. This is not a complicated step, but would massively improve the value of the data product.

### **Meaning of reconstruction differences:**

The authors also mention other reconstruction methods (L372), which begs the question why MAT and WA-PLS were chosen. Only because they are widely used, or because they yield superior results?

In addition, the authors provide three different reconstructions for each time series. What I miss is a discussion of how these different reconstructions can be used. Does the difference between them represent additional uncertainty on the reconstruction? How should the user include or use this information? Are certain reconstruction methods better than others? If so, which is to be preferred? If not, how can the (information from the) reconstructions be combined?

### **Reconstruction quality:**

The CCA suggests that only some part of the variance in the training sets is explained by T and precip and the significance testing indicates that a shocking 60-70 % of the reconstructions are basically noise. Whilst the authors go some way and filter out the time series that do not pass the significance test, I feel that the authors hardly mention this, let alone discuss. I also realise that this manuscript should not analyse the data, but perhaps some discussion in place and the different ways in which (pollen) assemblages could be used in paleoclimate science, including forward modelling, could be highlighted.

### **Land use issues/human influence:**

Some of the time series must bear an imprint of human influence. Can the authors briefly discuss to what degree and if and how this influences the reconstructions?

### **Insufficient explanation and detail in the methods:**

- 2,000 km radius for training set. Please explain why this was done and why the distance is (globally) appropriate.
- Why were seven analogues used for MAT? Are the reconstructions weighted to analogue quality, or simply the arithmetic mean of the seven closest analogues?
- How is the calibration error determined? Was spatial autocorrelation taken into account? From the code it seems that this is not the case, why?
- What is the sample-specific error based on? Why is this provided and not the calibration error?
- If I am correct, the tailoring approach serves the purpose of reducing the effect of co-variation between T and P. Please mention this earlier in the methods. I understand the point and that this goes some way to alleviating the problem. But what is done in cases where the correlation is not reduced? After all, there still is a large proportion of the sites for which there is a marked correlation in the training set. Some discussion would be appropriate here.
- Please provide more detail on the significance test. How were the random environmental fields generated? Simple permutation, or taking spatial correlation into account. Why?
- Why were the tailoring and the significance testing not applied to the MAT reconstructions?
- The CCA seems to be the first step in the development of the transfer function model to demonstrate that T and Precip really explain the variance in the assemblages. Would it not be better placed earlier in the description? And why are the implications barely discussed?
- How are poor analogues treated? Do they occur at all after the lumping? There is some discussion in L327-332, but it is unclear what the user of the data can do with this information.

### **Minor issues:**

L3: reconsider the use of "late quaternary" in the title. The meaning is actually rather vague and something along the lines of 30,000 years would be more informative.

L108: not sure what the policy is to refer to submitted manuscripts.

L131: please provide a bit more detail on WorldClim 2. For instance, what are the data based on, over what period are the data integrated, etc.

L385: crucially, this manuscript does not describe a fossil pollen data set, but a data set of temperature and precip

L402-404: this seems a somewhat dangerous statement. Are the two reconstructions really independent?

Why is the x axis of figure 6 on a log scale?

Whilst glancing through the code I missed the significance testing and the CCA. (But thumbs up for sharing the code.)