Clim. Past Discuss.,
https://doi.org/10.5194/cp-2019-81-RC3, 2019

Interactive comment on "Proxy surrogate reconstructions for Europe and the estimation of their uncertainties" *by* Oliver Bothe and Eduardo Zorita

Anonymous Referee #3

Received and published: 23 August 2019

Abstract

This study proposes a new climate reconstructions for Europe for nearly the full last millennium. The approach is based on the Analog Method, also known in the literature as Proxy Surrogate Reconstruction. One of the main novelties of this manuscript is how the authors extend the methodology to explicitly account for uncertainties. The authors present several reconstructions and compare them to the Euro 2k reconstruction, as well as independent data from the BEST project. Similarities, differences, advantages and caveats are discussed through the manuscript.

General comment

Most classical reconstruction methods produce a single reconstruction which does not explicitly account for uncertainty, although it is acknowledged that it populates this type of data-sets. This is problematic because uncertainty is not only ubiquitous, but it is heterogeneous both in time and space. This is an important limitation that precludes the proper assessment of the limitations of the knowledge we can gather from climate reconstructions. In this sense, I think this study is important and necessary to improve one prominent tool to produce such reconstructions, the Analog Method.

The design of the study is sensible, and I have mostly minor comments regarding details I could not fully understand and therefore might deserve clarification. Should not be for the issue I discuss below, I would recommend publication after minor revision.

There is however and important aspect that has to be improved in the manuscript under the light of very recent bibliography published even after this discussion was started. There exists a published extension to the Analog Method that allows to estimate uncertainties. This is part of a recent publication with a more general aim (Neukom et al., 2019). There, authors briefly introduce and apply a methodology which largely differs from the one presented here, but that aims at the same purpose: explicitly assess uncertainties in climate field reconstructions with the Analog Method. I think this work should somehow account for the existence of this already published method. The level of modification applied to the manuscript depends on the authors. At the minimum, the differences between approaches should be discussed (for example, the approach opted by Neukom et al. (2019) does not produce missing values, being in principle an important advantage). At best, the approach adopted by Neukom et al. (2019) could be implemented here as well, and a comparison could be done between both methods. In my opinion, the latter would greatly improve the interest of this manuscript, but it is perhaps a major modification of the work that falls beyond its original scope. I leave it up to the authors and I would not be disappointed if they decide not to tackle this task.

Minor comments

1. Page 2, Line 20: I think the correct citation is Gómez-Navarro et al. (2014)
2. Fig 1: Maybe excluded locations could be shown with grey symbols, as well as the are representative for Central Europe. The location of these proxies is relevant for example to understand Figure 5.
3. Page 4, Lines 28-29. I think it is more correct to say that, only when Var$_{res}$ and Var$_{sig}$ are uncorrelated, the total variance is the sum of both (because in that case the covariance term vanishes).
4. Page 5, Line 6: typo (modfied)
5. Page 5, Lines 17–21: I do not understand where the 2.57 comes from. How it is related to the minimum number of 39 proxies? Please clarify.
6. Page 5, Lines 22-23: why is it the only one? why 2105 is special? why not 1.5 SD$_{noi}$?
7. Overall, in the two paragraphs aforementioned, it lies the core of the two reconstructions carried out. I think this is important, and it should be made more explicit that the two approaches represent different method used for real below. Perhaps this can be made more explicit with some structural element, such as an un-ordered list or similar.
8. Page 6, Table 1: I assume this is exactly the correlation used to define the SD$_{noi}$ in each proxy location, right? If so, this could be clarified in the main text (especially in Section 2.1.2).
9. Page 6, lines 6–8: The criterion to exclude two proxies is not very clear. What is meant by "relevant portion of variance"? In Fig. 5 we learn that the reconstruction

in these sites is poor. Would it be better if these sites were part of the network. Surely the answer is yes. I understand that the amount of climate information we get is poorer than in the other locations, but still we could benefit for having *some* information. At worst, if the proxies were pure noise, it would not be necessarily worse than not having information at all. In other words, I think having poor information is better than having none, and it's not fully obvious to me why proxies should be excluded from the analysis based on relatively low correlation alone.

10. Page 6 (but relevant for the whole study): why do you restrict the reconstruction to the period 1260 to 2003? The reconstruction could have been applied further back in time. The number of proxies varies in time, but this could be even beneficial for this study, focused on the validation of new methodologies. It would show how the estimates of the uncertainty presented here are sensible to a varying number of proxies. I feel that this choice has unnecessarily limited the scope of the manuscript.
11. Page 7, Table 2: it could be interesting to write the total number of analogues, i.e. the pool size. It would make more meaningful the number of proxies used to produce ensembles. For example, having 817 analogues (as in Fig. 8) has a clearer meaning when you add that they are 817 out of, let's say, 25000. It shows that you are still selecting a relatively minor number of relatively good analogues.
12. Page 7, Lines 6–7: I think having a consistent bias through the pool is not necessarily good, as it seems to be implied by the wording. It ensures that the bias are translated into the reconstruction. This is partly avoided using structurally different models to build the pool. I do not mean that the authors should necessarily rebuild the reconstruction with a larger set of models, but I think that at least they should not imply that using a single model is somehow beneficial.
13. Page 8, Fig. 2: I think a line marking the 0 K anomalies would help to read the series. This pertains mostly panels b and c, where the sign of the anomaly is

important, but difficult to appreciate without such a line. This argument applies to Figs 6 and 7 as well.

14. Page 8, Fig. 2: It's not fully clear to me what this figure (as well as Figs 6 and 7) show. Does "summary" mean spatial average?

15. Page 9, Line 10: please change "degree Kelvin" to "Kelvin". Please review it, as there are other locations where I saw this in the manuscript.

16. Page 9, Lines 9–16. The order of these two paragraphs can be exchanged. It's a bit unusual and therefore confusing to discuss Fig. 2c before Fig. 2b.

17. Page 10, Lines 10–15: I think the fact that the reconstruction underestimate the intra-location variability is a problem of the pool, not the Analog Method itself. Do the authors think that this could be improved if higher resolution models were used to build the pool?

18. Page 11, Fig. 3: The list of locations in the caption is misleading (the name and the ID are written all together). It seems a detail, but it puzzled me for a while until I realised that Tor92 and Torneträsk are not two proxies, but the ID and the name of the same one. You could easily remove this by using for instance parenthesis to separate name from ID or vice versa.

19. Page 11, Line 26: "The general agreement between the Euro 2k and the analogue..." this reads odd at this point, as the reader does not know where to find the information the authors are referring to. It turns out that this comparison is introduced later, in Figure 6 in Page 14.

20. Page 15: Lines 17–23: The reduced variance could be quantified (how much is notably smaller variance in Line 19?). Further, the lost of variance when more analogues are considered is common in this approach, and generally in any statistical approach, i.e. there is a bias-variance trade off. It could be noted here

that this has been comprehensively discussed in the bibliography of the Analog Method.

21. Page 15, Line 32: do the authors have a theory on what could be the reason for such systematic differences? Are they meaningful, can they be used to discuss merits or problems in the reconstructions? Or are they rather low-frequency random fluctuations highly sensitive to method parameters?

22. Page 16: Lines 25–26: The presence of missing values in years with volcanic eruptions is a major caveat of the method, as those are typically the years most interesting in climate studies. Here it would be specially relevant my comment about a comparison with the method presented by Neukom et al. (2019).

23. Page 16, Lines 27–31: I do not see why it is "unsurprising" this lack of analogues for the recent period. The pool contains this warming as well, so the search should not present more problems for this period than in any other.

24. Page 19, Lines 18–20: Maybe I'm miss-evaluating this, but I think that anchoring the reconstruction within a range of 8 K is a poor result. It shows that the 800 analogues are indeed poorly constrained in this region, so we have little idea of how the actual climate was in that period and region. More generally, I have the concern that the spread shown for example in Fig. 7 might provide an optimistic measure of the actual uncertainty. Fig 7e for instance shows the range in the spatial average, which is about 2 K. But this is after spatial average, where regional differences can cancel out! I wonder how large is the range in each location. This might perhaps be illustrated with a map of (temporally averaged) ranges? Eventually, my guess is that using as many as 800 analogues or more, really far away from "the best" is, as outlined by the authors, too much.

References

Neukom, R., Steiger, N., Gómez-Navarro, J. J., Wang, J., and Werner, J. P.: No evidence for globally coherent warm and cold periods over the preindustrial Common Era, Nature, 571, 550, https://doi.org/10.1038/s41586-019-1401-2, https://www.nature.com/articles/s41586-019-1401-2, 2019.

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2019-81, 2019.