Interactive comment on "Technical Note: The analogue method for millennial-scale, spatiotemporal climate reconstructions" *by* Oliver Bothe and Eduardo Zorita

Oliver Bothe and Eduardo Zorita

oliver.bothe@googlemail.com

Dear editor, dear referees,

Hereby we want to thank the editor and the referees for their evaluation of our manuscript and their helpful comments. Below we provide a response to their remarks.

We note that a couple of our replies ask for guidance by the editor.

On behalf of the authors

Yours sincerely

C1

Oliver Bothe

Referee 1

I think the paper addresses an important topic and provides a useful extension of the Analogue Method, combining the data and dates uncertainties. However, I think the paper lacks clarity in the description of the method (some fundamental steps as the generation of the confidence ellipses are not properly explained) and I have concerns about the pseudo-proxy setup.

Response: We thank the referee for their generous evaluation of our manuscript. Below we address these points in more detail. In our revisions, we will particularly take care to clarify our method. Below we also address the concerns about the pseudo-proxy setup.

In particular, I think the assessment of the method' skill (of course possible under pseudo-proxy conditions) is flawed: the same run used as "truth" is used inside the Analogue Pool leading, therefore, to a potential overestimation of the skill. In addition to that, I can not comprehend why the authors selected a pseudo-proxy network design (number of proxies, locations of proxies, period covered, uncertainties, etc.) that do not resemble at all the real-world case they later try to reconstruct. I recommend the authors to re-do the exercise generating a pseudo-proxy environment as close to the real-case as possible. Of course, later the here presented pseudo-proxy setup could be considered informative as how would the method perform if more proxies were considered, etc. but the generation of a closer to real situation is nonetheless essential and I suggest for it not to be bypassed.

Response: There are two points to address here:

1. We will do/redo one experiment with a pseudoproxy setup as close to the real-world cases as possible.

2. We understand the concern of the referee, and agree that the setup is suboptimal. We are confident that the submitted manuscript already was careful not to overestimate

the skill of the method. However, we are going to be even more clear in stating how good the method may be.

Here, we want to point out why we use the setup as criticised by the referee: For one, at the time of our study, the Trace-21ka simulation was the only available simulation providing a continuous deglacial climate trajectory in annual resolution. Tests showed that in our chosen setup the method does only find analogues from the Trace-21ka simulation from which we also constructed the pseudo-proxies, as we wrote in the manuscript.

Thus, if we exclude Trace-21ka from the candidate pool the method fails completely, and if we don't use Trace-21ka for the pseudoproxy construction, we cannot use a simulation in interannual resolution. We will discuss this in our revisions. For this reason, for the time being, and unless the editor advises us to change this, we do not plan to change our setup. We allow that we could have used the QUEST FAMOUS simulations to provide pseudo-proxies in, presumably, centennial resolution.

General Comments

- The description of the method is not clear enough. How do you define the uncertainty ellipses? I don't see anywhere in the text the methodology followed to find such ellipses. Also, related to that, what's the difference between 90%, 99% or 99.99% uncertainty ellipses? Please, provide a clear methodology to follow to find them. What are the confidence intervals? Please, define.

Response: We will clarify this in our revisions, possibly including a figure if the text becomes too cluttered with technical descriptions.

- Why in the pseudoproxy setup you don't mimic the real-world conditions your are trying to reconstruct? I find it confusing that the pseudoproxy and real-world proxy locations and time-spans are not the same. As it is now, because the pseudoproxy

and proxy cases differ, the results of the pseudoproxy analysis are not completely transferable to the real-world case. I suggest to generate a pseudoproxy network that is exactly the same as in the real-world case, and show the results in that case.

Response: We will do/redo one pseudo-proxy setup with a setup as close to one of the real-world cases as possible.

- The title of the manuscript talks about climate reconstructions. However, the manuscript deals only with surface temperature reconstructions. I suggest to modify the manuscript title to reflect this and to add some discussion on how the method could/could not be applied to reconstruct some other climate variables (particularly, how do you expect the results to change when reconstructing a more challenging variable as precipitation?).

Response: We will consider the recommendations of both reviewers and change the title. We will also discuss the extension and the transferability of the method. However, already here we want to stress, one needs to remain sceptical that additional variables can be reconstructed. This would amount to assuming that precipitation is tightly constrained by temperature. While this may be true on a global scale, at regional scales precipitation is also modulated by shifts in storm tracks and changes in the atmospheric circulation in general. This can be spatially very heterogeneous.

- It would be interesting to compare with the results of having a fixed number of Analogues, for example 1 Analogue.

Response: Indeed this would be an interesting experiment. One may, however, ask whether it is really meaningful considering the large uncertainties of the proxies, as one would only test relative to the best estimate for each proxy. We consider to provide such a test, but so far do not see it as essential for this manuscript. Analyses for shorter time scales (compare recent works from Juan José Gómez-Navarro and colleagues in Climate Dynamics and Climate of the Past) have found that fewer analogues lead to higher variability in the reconstructions but also to lower skill. This is plausibly also the

case here, as these effects can be explained by statistical sampling. Thus, it would not be a novel contribution in this regard.

- How do results change if using a less years for the sliding window-mean? For example 50-years-means? There is a mention to interannual data in the discussion session, but not comparison plots are shown.

Response: We consider to provide a test with shorter or longer sliding-window means. We note that Appendix Figure A5a already provides a comparison to interannual data for experiment E01.

- For the pseudo-proxy setting: The selected reality is the simulation Trace21k. Most of the Analogues come also for this simulation. It would be fairer for assessing the method's skill if the chosen reality is excluded from the analogues pool. How do the results change if done so? When allowing the same simulation as reality to enter the pool, results might be overly optimistic. For some of the simulations listed there are several runs available, in that case one run could be selected as reality and the other pool of Analogues.

Response: See our response to a previous point: We understand the concern of the referee, and agree that the setup is suboptimal. We are confident that the submitted manuscript already was careful not to overestimate the skill of the method. However, we are going to be even more clear in stating how good the method may be.

Here, we want to point out, why we use the setup as criticised by the referee: For one, at the time of our study, the Trace-21ka simulation was the only available simulation providing a continuous deglacial climate trajectory in annual resolution. Tests showed that in our chosen setup the method does only find analogues from the Trace-21ka simulation from which we also constructed the pseudo-proxies as we write in the manuscript.

Thus, if we exclude Trace-21ka from the candidate pool the method fails completely,

and if we don't use Trace-21ka for the pseudoproxy construction, we cannot use a simulation in interannual resolution. We will discuss this in our revisions. For the time being, and unless the editor advises us to change this, we are not going to change our setup. We allow that we could have used the QUEST FAMOUS simulations to provide pseudo-proxies in, presumably, centennial resolution.

Specific Comments

Abstract:

- In the first paragraph authors talk about the last 21 kyr. However, in the second paragraph the target is reduced to the last 15kyr. Please, rephrase or explain failure in the target.

Response: We will be more clear about the temporal scope of the manuscript. The discrepancy is solely related to the different temporal extent of the pseudoproxy and real-world applications.

- The authors could emphasize that in the present for the reconstruction method seems to be no better than a long-term mean.

Response: We are not fully clear to what "in the present" refers here, but we will put this information into a revised manuscript.

- These fields reveal that uncertainty are also large locally. Please, change for . . . uncertainty is also. . .

Response: We will correct this.

Introduction:

- Please clarify the definition of nonillion

Response: We will do so.

Section 2:

- Here you sometimes use the word Analog instead of Analogue. Please, unify throughout the manuscript

Response: We are sorry for this oversight and we will do so.

- Figure 1: Please, add latitude and longitude. Also introduce the acronyms P01 and E01, as so far they have not been introduced in the text.

Response: We will do both.

- Page 4: 'Our interest is in temperature', please clarify if it is surface, annual mean, etc. What is a temperature calibration?

Response: We will clarify these questions.

-Page 5: Please explain better the meaning of "at best centennial" Does this mean that there are no proxy records with resolution finer that 100 years?

Response: We will clarify this.

- Page 6: why not consider the same period for real and pseudo proxy setups?, how are ellipses of confidence constructed? Please, provide the appropriate ellipse equation for its construction.

Response: We are able to extend the reconstruction period for the pseudoproxy approach back to the Last Glacial Maximum. We regard this an interesting exercise.

Therefore, we will keep the different periods. We acknowledge that there are formal arguments for synchronising the setups.

We will add a clearer explanation of how we construct the ellipses.

- Page 8: What is a credible interval? Please, define.

Response: We will clarify our terminology.

- Page 10: The authors say: "randomly chosen pseudo age uncertainties". How are those selected? Is the random process a Gaussian distribution? Which mean and variance? This needs more clarification.

Response: We will clarify this briefly.

Figure 3: Isn't it easier to show the plots in the form of line-plots? Specially plot a is difficult to read, as it looks like a huge black block, differences are hard to distinguish.

Response: We think the vertical lines better represent the discrete character of the approach but we will reconsider how to visualize the data.

Figure 4: Please, put all the plots in the same scale

Response: The reviewer's suggestion would make it close to impossible to identify changes in individual series. However, we will synchronise the absolute range of the temperature-axes for all panels.

Section 3:

- Page 14: The authors indicate very little variability in the reconstruction median over certain periods. This probably arises due to too many Analogues are selected in those periods. How could you constrain the Analogue selection?

Response: There is a trade-off between considering the uncertainty of the proxies

and constraining the number of analogues. We consider providing a discussion of this aspect.

Figure 5: Panel c: please add name like "warmer case", "colder case" and the respective locations (lon, lat). Panel d: add the subtitle "Regional average" Panel e: add the subtitle "Grid point: (lon, lat)" In panels d and e: I can't understand what the authors mean by "examples". Why some of the examples look like dots and some as dashed lines? Are the dots (dashes) associated to the warmer (colder) case shown in panel c? It would be interesting to discuss the moments when the Target is outside the envelope (Figure 5a)

Response: We will try to improve the visualisation of our results, and we will provide a clearer description of the results. We are not going to discuss in detail the specific cases when the target is outside the envelope but we will thoroughly discuss reasons why the target may generally fall outside the envelope.

Figure 6: Please, add the units directly above the colorbar. Also, indicate the year that is being shown as Example.

Response: We will clarify the Figure.

Figure 7: Please put all the plots in the same scale.

Response: As for Figure 4, the reviewer's suggestion would make it hard to identify changes in individual series. However, we will synchronise the absolute range of the temperature-axes for all panels.

Figure 8: Similar considerations as in Figure 5. Figure 9: similar considerations as in Figure 6.

Response: We will clarify all four Figures.

Page 24: In the summary the authors say that the method succeeds in the pseudo-proxy setup. I think that sentence might be overestimating the skill of the method, as

the authors used one model run (Trace21k) both as truth and as proxy pool. Please, remove the truth from the possible pool of Analogues to be able to properly analyse the method's skill.

Response: We cannot remove the truth from the pool of analogues as we then would not be able to apply the method due to the lack of suitable simulations as we wrote in the manuscript. We will more clearly discuss this.

For the real-case the authors say the reconstructions fail. How can you assure failure when you don't know the truth? I think the sentence should be re-phrased and the only thing that can be known for sure is the failure to find Analogues within the selected pool. I think that it needs to be made clearer that not knowing the truth in the real-case is exactly the reason for making pseudo-proxy analysis. Which leads, again, for the importance of the pseudo-proxy setup (design of the network, period covered, etc.) to be as similar as possible to the real-case.

Response: We will rephrase this to make it clear that a failure of the method is equivalent to a failure of finding analogues in the candidate pool.