Clim. Past Discuss.,
doi:10.5194/cp-2016-98-AC3, 2017

Climate of the Past Discussions
Open Access
EGU

Interactive comment on "Testing the analog method in reconstructing the global mean annual temperature during the Common Era" *by* Juan José Gómez-Navarro et al.

Juan José Gómez-Navarro et al.

jjgomeznavarro@um.es

Received and published: 13 January 2017

This paper introduces analog method and uses it to reconstruct spatial fields of temperature covering the whole globe. The paper looks at several variants of the analog technique and evaluates which performs best. Although not the first time analog method has been used to reconstruct past climate this study is a useful addition to the field and I found that it is interesting, provides a good review of the field, is well written and is logical and clear. As such I would definitely like to see this paper published. I do however first have several comments which I hope will improve the paper – I've split the comments into major and minor, although not all the major comments are that major.

Thank you very much for the time devoted to carefully read the manuscript and the

Printer-friendly version

Discussion paper

positive view expressed about it. We will try to address all major concerns pointed out by the reviewer.

Major comments (in no particular order).

I find the title and to a lesser extent the abstract to be slightly misleading. The title suggests that the paper focuses on reconstructing global mean temperature – however this is only briefly mentioned in the main paper (fig 9 and one paragraph at the end of section 5) with the focus instead on spatial reconstructions. While I have no problem with the emphasis of the paper, I do think that the title should be changed accordingly or alternatively more emphasis should be placed in the analysis of the global mean temperature to make for a more consistent paper.

We will re-think the title according to this suggestions and the changes carried out in the reviewed manuscript.

Equally I find the use of the acronym "MAT" slightly confusing. It is introduced as "global near-surface mean annual temperature (MAT)" - l139, but is used frequently to refer to local temperatures and line 501 mentions the "annual series of MAT". I'm therefore unsure what the M in MAT actually refers to (is it a global mean or an annual mean?). Personally I would prefer the less ambiguous SAT (surface air temperature) to be used. But MAT would also be OK if properly defined and consistently used.

We chose that acronym following the naming convention used some references we use. We will carefully reconsider what naming convention leads to the more natural reading for the expected target audience of the paper.

One interesting finding of the paper is that the simulations and reconstruction of the Arctic has reduced variance compared to the observations. This is however a region where there is no or little coverage in the HadCRUT4 dataset. How sensitive is the result therefore to the infilling technique used? Would the results be changed if a different infilling was used e.g. that of Cowtan and Way 2014?

Using a target infilled differently would have a minor role in the output of the reconstruction, as the analogs themselves are independent on this. The only effect would come from their choice, which can be affected by having obtained slightly different calibrations in the arctic area. But we believe this is a second order effect that is anyway masked by the weight of the rest of proxies worldwide.

However, what is more important is the fact that using a different target would have an important and direct effect on the metrics we obtain from the evaluation of the reconstruction. In particular, using the Cowtan and Way infilling as target as suggested by the reviewer will certainly reduce the underestimation of variance we report here. Therefore this is an important point we will develop in the discussion.

Since you discuss the sensitivity to individual models I think it would be useful to mention explicitly which models correspond to which bar in fig. 4. I think I can more or less piece it together based on the text but would like this stated clearly either in the text, figure caption or a separate table. When doing this the GISS model used as the target could be number 16 (or 1) so that it can be left off of figure 4 to avoid confusion. I would also be tempted to group colours by model e.g. make all MPI-ESM models different shades of green – as I think this would improve clarity. I think figure 4 would also benefit from panels b and d being square with a straight line through x=y, to highlight the important point that the results are not scattered around this line, except for the end of the 20th century. Is this also the case for volcanic years?

In the first version of the manuscript we decided not to label models on purpose, since the nature of the method consist of considering all models equally as members of a pool with all models contributing equally without any consideration of their intrinsic skill, and where the method selects years blindly without other consideration than the chosen metric. Therefore we are not sure if labeling models is sensible in this context. Still, we will carefully consider these suggestions and decide how to proceed giving our reasons in the final version of the manuscript.

The majority of the results presented are obtained using only one model (GISS r1i1p121) as the target. In the text you describe this model as unusual in having low variance. I think it would therefore strengthen the paper if analysis was also carried out using all the other models (or at least some) as targets. I don't envisage all the results actually being shown in the paper, since this would be far too many, but I do think a statement saying that the results are not sensitive to the model used as a target would be very useful (if this is indeed the case).

This is a comment shared by the other reviewer, so we will indeed carry out this test, although we will surely not include the figures to keep the paper as concise as possible.

In the theory or discussion section no mention is made of how this method could be used to provide uncertainties to the reconstructions. Would it be possible to add some information regarding how the analyses presented here could be used to provide an uncertainty estimate on the reconstruction, for future use?

We will enlarge the paper to provide a method to estimate uncertainties (see also our response to the above comment of E. Boucher regarding uncertainties). We envisage a method similar to those proposed in regression-based methods.

No mention is also made regarding the drop in coverage of the proxy network back through time. This surprised me as I thought this would be one of the key considerations when reconstructing the climate of the past millennium and beyond. Given that some of your PPE experiments have changing coverage through time could you show the performance of your method as a function of time. Equally I think it would be very informative to include a case in figure 8 (and section 6) with a sparse proxy coverage reflecting, for example, long proxies which cover the whole period 1000-2000 (PAGES-1000?), as this would better reflect the performance of the method during these earlier periods when less data is available.

We believe these are interesting suggestion. We will consider additional tests to explore the sensitivity of the performance to the variable number of proxies.

In section 5 you comment that the difference between figure 7 and figure 5 6 suggests that the level of noise employed in the first PPE is an underestimation. Could it not also be due to errors in the model fields?

This could be an alternative explanation we will consider in the discussion.

If one of the goals of this paper is to lay the ground work for a future analysis which reconstructs the global mean climate of the last 1000 or 2000 years then I think that a comparison of your global annual mean reconstruction (fig 9) with simpler commonly used reconstructions methods which only use the proxies scaled to the observationsto make a composited reconstruction, as well as a comparison to just the raw model results would be very useful. This would then allow the reader some sense of how much the method presented here adds to simply using the proxies or the models on their own. I would consider that at least some improvement over both of these would be the minimum requirement for applying this to the climate of the past to produce a global mean reconstruction, although I appreciate, as made clear in the paper, that this method does add much more valuable spatial and multi-variable information and is not focused on just producing a global mean reconstruction. If this is not a goal of the analysis than this should be clearly stated.

We wanted to keep the paper as short as possible by 1) focusing on the performance of the method to reproduce the spatial patterns, rather than the global average and 2) not including a full range of alternative methodologies, as this work is indeed under preparation to be submitted as a separate publication. However we agree with the reviewer that such a comparison with CFR would not be very complex and could enrich the paper. Therefore we will include a comparison with a global mean reconstruction carried out with Composite plus Scaling.

Minor comments:

- *M in eq 4 is not defined.*

- *It would be good to state what value of L is used i.e. how many EOFs are retained.*
- *L3 – explain acronym AM-CFR*
- *L4 -> As a test bed*
- *L6 –> simulations from PMIP3 are used*
- *L10-> provided by the PMIP3 ensemble*
- *L99-> with respect to*
- *L395 -> especially*
- *L556 – tree -> three?*
- *Figs2 and others – what period is the correlation calculated over?*

We will carefully review these comments in the final version of the manuscript.

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2016-98, 2016.

Printer-friendly version

Discussion paper