Answer to anonymous referee #3's commnets

Fallah et al.

February 27, 2018

We wish to thank the reviewer for his/her critics and positive feedback. We will answer the comments(*italic*) point by point in the following (**Bold**) :

1- The paper describes an approach for high-resolution climate reconstruction using an off-line assimilation of proxies into a set of regional climate model simulations. The set-up is tested with the COSMO-CLM model and a number of sensitivity studies are carried out. While the study is interesting, it somehow stops at a point that is still too distant from applications, and it is not clear what sort of applications the authors have in mind. I think the authors should better demonstrate how high-resolution climate reconstruction actually will be obtained and how they will be applied. Furthermore, I found the methodology not very well explained. However, because this is the first paper I am aware of that applies paleo DA to regional climate, I think we can learn a lot and therefore, in my view, the paper is potentially worth being published after major revisions.

We wish to thank you for your suggestions. According to your and reviewer 2's concerns on the real application of the methodology and how such high resolution information will be obtained, we set up new experiments using pre-computed time-slice COSMO-CLM simulations over Europe during the Holocene and pollen-based temperature reconstructions. Please refer to our answer to reviewer 2's 5th question. We hope that the new manuscript is closer to real application and will be accepted for climate of the past. The appli-

cation of such high resolution climate data for the paleo-modeling community, might be of interest for example for evaluation of coupled simulations against high resolution climate maps over the target region (here Europe).

1) The design of the study is still far from a real world sparse proxy network. Just one example: They assimilate 500 observations, they even go up to 2700 stations and consider 100 "a small number of stations". I would be far more curious to see how the approach works with fewer observations, and what the author's view is concerning other variables (precipitation). Will this eventually work for tree rings? The results section is only 2 pages (part of which, i.e., the localisation, should actually be in the methods section).

This comment is similar to reviewer 2's comment number 1. We briefly answer it again here: In very recent studies focusing on reconstruction of climatic variables the number of records used are similar to the number we have chosen: for example [Mauri et al., 2015] have used "879 selected pollen sites representing nearly 60,000 pollen counts" (see Figure 1 of their paper `https://ars.els-cdn.com/content/image/1-s2.0-S0277379115000372-gr1.jpg`**). Or [Franke et al., 2017] have used a proxy network which is very dense over Europe (Figure 3 of their paper :**`https://www.nature.com/articles/sdata201776/figures/3`**). Or the location of sites in the study of [Marlon et al., 2017] in North America is also a dense one. Or [Cook et al., 2010b] where they used 1,854 annual tree-ring chronologies over North America (Figure 5 of their study :** `http://onlinelibrary.wiley.com/doi/10.1002/jqs.1303/pdf`**). Or the study of [Cook et al., 2010a] where they used a 327-series tree-ring chronology network to reconstruct the Palmer Drought index over Asian mansoon area (Figure 1 of their paper :** `https://d2ufo47lrtsv5s.cloudfront.net/content/sci/328/5977/486/F1.large.jpg`**). Therefore, we believe 500 stations over Europe looks realistic with new advances in paleo-data collection, synthesis and stewardship (for example** `https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets`**). Following your first comment we also assimilated a proxy**

number of ~ 300 in the real application.

On your comment about precipitation or usage of tree rings, we should add that here we use an inverse model's outcome (temperature reconstructions) and one can also use a proxy forward model of any kind and repeat the assimilation on proxy space instead of the model space. We will add more detail on that in the discussion. We have used such an approach in a recent study ([Acevedo et al., 2017]) with an extended version of Vaganov–Shashkin Lite (VSL) process-based tree-ring-width forward model ([Tolwinski-Ward et al., 2011]) and the SPEEDY climate model. The error reduction for the precipitation was shown to be not significant using the Enkf. However, the model was of intermediate complexity in that study.

2) The paper places itself in the sequence of recent work on paleo DA - it does not mention existing 0.5 -resolved statistical reconstructions. The motivation of many of the global paleo DA studies is to obtain a physically consistent global climate for time periods with spatially very heterogeneous coverage. There are good reasons for suspecting the same on the regional level, particularly for Europe (given the orography and land-sea contrast), but I think this needs to be better justified.

We hope that by showing the real application with RCM simulations throughout the Holocene and adding the references of the pollen data, especially the work of [Mauri et al., 2015], we will cover this comment. We will add a paragraph of previous statistical reconstruction efforts on Europe in the introduction of the new manuscript. On the comment on the usage of RCM instead of GCM in paleo-data assimilation, we should mention that one of our main motivations is to resolve the gap of regional to local scale climate change, which might be of interest in the paleo-community. For example the uncertainties of proxy data are bound at regional scales. However, as you mentioned in question 1, the performance of the RCM might vary for different variables. Assuming that in an RCM there are more realistic

physical processes implemented than the GCM (especially complex topography) which otherwise had to be parameterised in GCMs, the resolution of such models are of advantage when comparing with the proxy data. One problem of using coarse GCM data might be for example the process of selecting the best observation within a grid cell for DA scheme ("data thinning") or sampling with averaging for the observations within a grid cell ("super-obing"). By using the RCMs, we are reducing such problems. For comparison of proxy and GCM, one might use classical approaches (statistical downscaling, upscaling, forward model), however, such methods for a coarse resolution of GCM might be very challenging as one has to evaluate or train such models with very short observation time-window. The proxy-data relation might also change over time (not stationary). On the other hand, the feedback of regional climatic changes on the global scale is ignored using one-way nesting approach applied in our simulations. Finally, we should mention that with our domain set up the RCM is constrained by the GCM at the lateral boundaries and therefore its internal variability is similar to the driving GCM at large scale. This behavior is detectable by the maps of ensemble spread shown in the manuscript and answer to reviewer 1. By changing the domain or initialization time, the RCM simulations do not vary dramatically for the averaged seasonal maps.

3) The methodology could be explained better. I already stumbled over p.4/l.5, which I first read as impliyng that X_TRUE and X_A is the same (is an "and" missing?). The terms X_NATURE and "free ensemble run" appear before they are introduced. There are some other instances (listed below).

Thanks. We will go through the formulation of mathematical terms and describe them as they appear in the text. We take care of all your minor comments in the new version of the manuscript.

1 Minor comments

P. 1, l.12: How can the selection of proxies reduce the background error? **We change the text accordingly. The analysis error is reduced compared to the background.**

p. 2, l.20: states **Changed.**

p. 3, l.13: Since the sentence cites DA approaches that were actually "applied", it might be good to cite Franke et al. (Scientific Data, 2017). **Done.**

p. 3, l.20: The sentence is somehow odd: "optimum" in the first part implies a choice, C2"truly" implies an esimation. **Changed.**

p. 5, eq. 8: X_NATURE is not introduced yet p. 5, l. 5: X_Analysis was called X_A before p. 5, l. 9: X_Analsis -¿ X_Analysis **Done.**

p. 5, eq. 11 and 12 are both said to represent "the error covariance of the analysis" **Changed. The trace of P^a is the total error variance of the analysis and P^a is the error covariance of the analysis.**

p. 6, l. 9: Is the added noise spatially uncorrelated? **They could be correlated. It is common practice to assumed that they are uncorrelated for the sake of simplicity and affordability.**

p. 7, l. 14: When describing the shift, the state vector should be defined (because it can no longer include the entire model domain - is it the "evaluation domain", which on my first reading I interpreted as the domain in which evaluations are done). **We will describe it clearer. It was previously described in the caption of the figure 1.**

p. 7, l. 19: The analysis skill should be in the title, and mentioned in the text upfront. Some measure of dispersiveness might be interesting. **Done.**

p. 7, l. 20: What is a "free ensemble run"? This term is not introduced. **A run without data assimilation. We will describe it clearer.**

P. 9, l. 13: Please explain the "universal behaviour of fluctuations of terres-

trial near-surface temperature" **Done. There exist a correlation between the temperature and the topography. The power-law behavior seen in topography also exists for the near surface temperatures. There exist a universal persistent role in the static geometry of the Earth which controls the dynamics of atmosphere.**

References

[Acevedo et al., 2017] Acevedo, W., Fallah, B., Reich, S., and Cubasch, U. (2017). Assimilation of pseudo−tree−ring−width observations into an atmospheric general circulation model. *Clim. Past*, 13(5):545–557.

[Cook et al., 2010a] Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., and Wright, W. E. (2010a). Asian monsoon failure and megadrought during the last millennium. *Science*, 328(5977):486–.

[Cook et al., 2010b] Cook, E. R., Seager, R., Heim, R. R., Vose, R. S., Herweijer, C., and Woodhouse, C. (2010b). Megadroughts in north america: placing ipcc projections of hydroclimatic change in a long-term palaeoclimate context. *J. Quaternary Sci.*, 25(1):48–61.

[Franke et al., 2017] Franke, J., Brönnimann, S., Bhend, J., and Brugnara, Y. (2017). A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations. *Scientific Data*, 4:170076–.

[Marlon et al., 2017] Marlon, J. R., Pederson, N., Nolan, C., Goring, S., Shuman, B., Robertson, A., Booth, R., Bartlein, P. J., Berke, M. A., Clifford, M., Cook, E., Dieffenbacher-Krall, A., Dietze, M. C., Hessl, A., Hubeny, J. B., Jackson, S. T., Marsicek, J., McLachlan, J., Mock, C. J., Moore, D. J. P., Nichols, J., Peteet, D., Schaefer, K., Trouet, V., Umbanhowar, C., Williams, J. W., and Yu, Z. (2017). Climatic history of the northeastern united states during the past 3000 years. *Clim. Past*, 13(10):1355–1379.

[Mauri et al., 2015] Mauri, A., Davis, B., Collins, P., and Kaplan, J. (2015). The climate of europe during the holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation. *Quaternary Science Reviews*, 112(Supplement C):109–127.

[Tolwinski-Ward et al., 2011] Tolwinski-Ward, S. E., Evans, M. N., Hughes, M., and Anchukaitis, K. J. (2011). An efficient forward model of the climate controls on interannual variation in tree-ring width. *Climate Dynamics*, 36(11-12):2419–2439.