

Interactive comment on “Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling” by Robert Tardif et al.

Robert Tardif et al.

rtardif@atmos.washington.edu

Received and published: 18 January 2019

Authors' Responses to Anonymous Reviewer 2

We thank referee 2 for comments on the manuscript.

This reviewer believes the methodology, analysis and the final LMR product presented herein are too premature to be acceptable for formal publication, let alone for its stated purpose to serve as the basis for the first publicly released NOAA last millennium reanalysis.

C1

The data from the first release described in Hakim et al (2016) has been publically available for over two years. The basic method on which Hakim et al (2016) and the present paper are based has been evaluated and tested extensively in the literature (e.g. Bhend et al, 2012; Steiger et al , 2014; Matsikaris et al, 2015; Acevedo et al, 2017; Franke et al, 2017; Okazaki and Yoshimura, 2017; Steiger et al , 2018).

It is misleading for this study (and its prototype in H16) to call the DA method used ... as an ensemble Kalman filter (EnKF). As in Evensen (1994) and subsequent studies, the primary promise of the EnKF is the use of flow dependent background error covariance represented by the forecasting ensemble. The current so-called “offline” DA method has none of that: the ensemble perturbations are randomly sampled from a past-millennium climate simulation that has no relation to the prior estimate, and the same set of sampled perturbations were used at all analysis times.

The reviewer appears to have a narrow view of data assimilation limited to operational weather forecasting. In fact, the prior may come from a wide variety of sources, and Monte Carlo sampling of that distribution using ensembles has proved to be a powerful solution method. The “offline” EnKF approach was originally described by Oke et al (2002, 2005, 2007), and Evensen himself described it in his 2003 review of the ensemble Kalman filter (Evensen, 2003). In revision we plan to add a few more references to basic data assimilation theory for readers not familiar with the applicability of the technique in general, and to the aforementioned literature for the offline method in particular.

This method used in this study is similar to the commonly used 3D-Var method for numerical weather prediction with static background error covariance, and is arguably less advanced than 3D-Var since 3D-Var in NWP used the dynamic

C2

model to propagate the previous cycle's analysis as the prior before the analysis. The current so-called "offline" DA method neither cycles the analysis nor the ensemble perturbations, with the stated reason that the forecast model is not good enough to do either.

We regret the reviewer's interpretation that the motivation for offline DA is because "the forecast model is not good enough," which is not the case. The choice is a result of a cost-benefit analysis: predictive skill of Earth System models on proxy timescales is small, but the cost of ensemble forecasts with these models is high. We plan to make that point clearer in the revised manuscript. For an extension of the LMR method to online DA, and comparison to the offline method, please see Perkins and Hakim (2017).

If the forecast model is not good enough to cycle the mean analysis or the analysis uncertainties to provide the best estimate of the prior estimate and related prior uncertainties, why would this model(s) be good at all for use as the prior estimate that the LMR reanalysis depends critically on? In this regards, it is premature to state (line 10) that the "LMR employs the ensemble data assimilation to optimally blend the information from the proxies and the climate model data". The current method is more like an objective analysis method.

Yes, the method is a form of "OI," although we believe that using such jargon is not helpful to the readership of this journal.

It is not clear whether the authors are aware that the traditional static 3D-Var methods also derive the background covariance from an ensemble of perturbations, as is traditionally called "the NMC method" using the sampled forecast

C3

divergence between different lead times from many realizations. The Kalman filter update in this case is equivalent to the variational update using the 3D-Var algorithm, though again the 3DVar in NWP cycles the analysis and forecast during data assimilation, which is the most basic function in combining the model and data.

Yes, we are aware of the NMC method, which samples forecast differences on the timescale of the DA cycle. In our case that is one year, and the random sampling method we employ assumes that forecast differences on that timescale have converged on the climatological distribution; we lack analyses and forecasts over the Common Era to formally apply the NMC method.

The validation performed in this study for the prototype and updated LMR "reanalysis" with several existing 20th-century reanalysis is misleading at best. The quality of the LMR reanalysis for the 20th century is the least issue given the availability of the modern much more advanced reanalysis and given the exponentially increased number of proxies or model instrumental observations. The validation currently focuses exclusively on the 20th century says little on the quality and performance of the LMR products, in particular over the early period when the proxy data are scarce. A more appropriate validation can potentially be done in two objective methods: (1) perform the 20th century "reanalysis" through thinning the observation density and maybe also degrading the observation accuracy to those representation of different periods of the past millennium; and/or (2) performing observing system experiments in which a certain number of observations are not assimilated but reserved for independent validation (or all of them in cross validation).

Part of the method described in this paper involves withholding 25% of the proxies for

C4

independent validation, which we do both before and during the instrumental period. Furthermore, we perform 51 realizations over each experiment, randomly sampling the proxies, so that all proxies participate in validation. In revision, we plan to move those results from the supplementary material into the main body of the paper. Although not described here, we have done experiments consistent with suggestion (1), where we vary the percentage of withheld proxies and there is little sensitivity to the 25% value.

The use of a 2,5000-km covariance localization is highly questionable for the use of a 100 sets of fixed ensemble perturbations. At midlatitudes, this is amount to the observation impacts across the entire global latitude belt. The use of a fixed set of 100 sample perturbations also means a high rank deficiency over such a large area with this large localization distance.

This reasoning is consistent with covariance lengthscales on weather timescales. Covariance lengthscales on annual timescales are much longer, and the effective dimension of the covariance matrix is comparatively smaller.

The current final LMR reanalysis derives from the mean of 51 such 100-member analyses, should it be the same if the 5100 samples of perturbations are used simultaneously in the Kalman filter update given the Kalman filter used is largely a linear operation?

The fact that we get better results by averaging over Monte Carlo realizations (multiple analyses) as compared to larger ensembles is not completely understood. We believe that this is an artifact of poorly estimated analysis errors for a subset of the proxies, but fully exploring this issue is beyond the scope of the present paper. In revision we will highlight the issue and the hypothesis we have for it.

C5

How much is the result sensitive to the choice of this arbitrary number of sample perturbations? It is also worth noting the the NMC method used for 3D-Var uses singular value decomposition to make it full rank. Such a approach is different from (and likely more advantageous over) the current Kalman filter update using purely non-involving static ensemble covariances.

We have found little sensitivity to the ensemble size, provided it is at least 100 members and that covariance localization is used (effectively increases the rank of the covariance matrix). We find larger improvements by randomly subsampling the proxies through many Monte Carlo realizations.

It is unclear what is the purpose of such as hastily done LMR reanalysis products with such ad-hoc DA approaches and the not-good-enough forecast models? The so derived climate trend is almost certainly depending too much on the climate models used as a prior and ensemble sampled perturbations (and maybe the assumed climate forcings used in these models), as well as the density of observations over different periods. It could do more harm if such a premature reanalysis product is used or misused and if it were publicly released through NOAA, unfortunately. A more careful vetting of the products, and a more concerned effort in refined DA methodology are warranted before NOAA sanctioned such a product as reanalysis, in this reviewer's opinion.

If you wish to see more sensitivity analysis with respect to these issues, please carefully read Hakim et al (2016), where we not only considered the performance statistics of analyses using different priors and calibrations of proxy forward models, but also examples of the differences that result in the spatial fields for an individual year.

C6

References

- Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N. J. and Perkins, W. A., (2016) The Last Millennium Climate Reanalysis Project: Framework and First Results, *JGR: Atmospheres*, 121, 6745–6764, doi:10.1002/2016JD024751, 2016
- Bhend, J., Franke, J., Folini, D., Wild, M. and Brönnimann, S., An ensemble-based approach to climate reconstruction, *Climate of the Past*, 8, 963–976, 2012.
- Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S. and Roe, G. H., Assimilation of time-averaged pseudoproxies for climate reconstruction, *J. of Climate*, 27, 426–441, doi: 10.1175/JCLI-D-12-00693.1, 2014
- Matsikaris, A., Widmann, M. and Jungclaus, J., On-line and off-line data assimilation in palaeoclimatology: a case study, *Climate of the Past*, 11, 81–93, doi:10.5194/cp-11-81-2015, 2015
- Acevedo, W., Fallah, B., Reich, S. and Cubasch, U., Assimilation of pseudo-tree-ring-width observations into an atmospheric general circulation model, *Climate of the Past*, 13, 545–557, doi:10.5194/cp-13-545-2017, 2017
- Franke, J., Brönnimann, S., Bhend, J. and Brugnara, Y., A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations, *Sci. Data*, 4, doi: 10.1038/sdata.2017.76, 2017
- Okazaki, A. and Yoshimura, K., Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction, *Climate of the Past*, 13, 379–393, doi:10.5194/cp-13-379-2017, 2017
- Steiger, N. J., Smerdon, J. E., Cook, E. R. and Cook, B. I., A reconstruction of global hydroclimate and dynamical variables over the Common Era, *Sci. Data*, 5, doi: 10.1086/sdata.2018.86, 2018
- Oke, P. R., Allen, J. S., Miller, R. N., Egbert, G. D. and Kosro, P. M., Assimilation of surface velocity data into a primitive equation coastal ocean model, *JGR:Oceans*, 107, doi:10.1029/2000JC000511, 2002
- Oke, P.R., Schiller, A. Griffin, D. A. and Brassington, G. B., Ensemble data assimilation for an eddy-resolving ocean model of the Australian region, *Quart. J. of the Royal Meteor. Soc.*, 131, 3301–3311, 2005
- Oke, P. R., Sakov, P. and Corney, S. P., Impacts of localisation in the EnKF and EnOI: experiments with a small model, *Ocean Dyn.*, 57, 32–45, 2007
- Evensen, G., The ensemble Kalman filter: Theoretical formulation and practical implementa-

C7

tion, *Ocean Dyn.*, 53, 343–367, 2003

- Perkins, W. A. and Hakim, G. J., Reconstructing paleoclimate fields using online data assimilation with a linear inverse model, *Climate of the Past*, 13, 421–436, 2017

Interactive comment on Clim. Past Discuss., <https://doi.org/10.5194/cp-2018-120>, 2018.