Clim. Past Discuss., 7, C2352–C2357, 2012
www.clim-past-discuss.net/7/C2352/2012/

Climate
of the Past
Discussions

Interactive comment on "A novel approach to climate reconstructions using Ensemble Kalman Filtering" *by* J. Bhend et al.

J. Bhend et al.

jonas.bhend@csiro.au

Received and published: 16 January 2012

We thank Pavel Sakov for valuable comments to the manuscript. The responses to the specific comments follow below. Paragraphs from the original review are marked with '$$'; our comments follow immediately after the respective paragraphs.

$$ Major issues

$$ 1. The DA method used in the manuscript can not be characterised as a filter because assimilation of the past data does not affect the current state of the DA system. Perhaps, it could be characterised as an ensemble based data fitting method.

The authors are perfectly aware that our implementation of the EnSRF algorithm is different from standard data assimilation (as the procedure is not cycled in our setup).

We agree with the reviewer that this has to be made clearer in the manuscript and adjusted paragraphs in the introduction and discussion accordingly.

$$ 2. Because of the design of the method, it is impossible to use the standard methods of assessment of the skill of the system, such as comparing the forecast versus persistence. The authors use a "reduction of error" metric that characterises the relative reduction of the distance between the estimation and the truth. It is difficult to see the importance of this metric for assessing the system performance in the context of climate reconstruction. If, for example, one simply replaces the observed model state elements with observations, this metric will normally show a positive skill, while the overall quality of the analysis remains unknown.

Simply replacing the reference simulation (i.e. the pseudo-proxy series derived from the reference simulation) by observations would result in generally more positive skill using the proposed metric as not only internal variability, but also differences in the simulated and observed forced response would be used to constrain the ensemble (thus making it more likely that we find positive skill). However, we strongly disagree that the proposed metric is unimportant for assessing system performance. We explicitly show in the manuscript (with the proposed metric) that the data assimilation without localisation leads to negative skill in regions far away from the assimilated information.

Following comments from reviewer 4, we also discuss the effect of assimilating data on ensemble spread. By analysis rank histograms, we investigate, whether filter divergence would become a problem were the procedure to be cycled. The rank histogram complements the analysis using correlation and RE and provides a simple check of whether the procedure suffers from overfitting.

$$ 3. In this context, a demonstrated improvement of composite indices observed in the manuscript could be considered, generally, as a good indicator of the skill of the method. Unfortunately, the design of the experiment has little to do with assimilating real observations. Namely, the true field is represented by one of the members of

the unconstrained ensemble of model runs; consequentially, it uses the same model and the same forcing as the rest of the ensemble. These conditions of perfect model and perfect forcing can not be satisfied with real observations; therefore, the positive correlations between the analysis and the truth observed in the course of the study are unlikely to take place in practice.

We are aware of and try to be candid about the difference between this explorative study into technical aspects of assimilating proxy data using the EnSRF and a full-fledged assimilation (which needs to include cycling with coupled ocean-atmosphere model and an explicit and more complex proxy forward model). The present setup allows us to explore certain aspects of the assimilation procedure – obviously at the cost of generality and maybe at the cost of relevance for real-world applications. As the perfect-model framework we operate on is by design optimistic, we expect skill in this framework a necessary (but not sufficient) condition for skill in real-world applications.

$$ Minor issues

$$ 1. P. 2839, l. 14: "In order to keep computations tractable, we thin out the initial model grid..." The computational complexity of the EnKF in regard to the state vector dimension is linear; modern EnKF based DA systems routinely function with the state vectors of 10**8-10**9 elements.

We agree that an EnKF algorithm with a much larger state vector is technically feasible. However, as we only have desktop computers available to do the computations, we keep the state vector small to allow for quick computations and to explore various aspects of the algorithm.

$$ 2. I can not see the relevance of section 2.3 "Ensemble Square Root Filtering" for the rest of the manuscript – see major issue 1. In particular, the discussion of filter divergence is completely irrelevant for the method involved.

Following the suggestions of reviewers 3, we revise the introduction to Ensemble

Square Root Filtering and expand slightly on it, however, we do feel that the issue of filter divergence is an important aspect of an Ensemble Kalman Filter data assimilation scheme. Overfitting (leading to filter divergence) is an important issue even in our setting, since the analysis ensemble spread no longer captures the uncertainty in the case of overfitting. Being able to quantify hindcasting uncertainty (through the ensemble spread) in a natural and useful way is a major advantage of ensemble based methods over other methods for reconstruction, thus we consider the problem of overfitting or filter divergence to be crucial.

$$ It is a good idea to give credit to the authors when describing methods. Equation (4) represents (a parallel) square root filter solution by Andrews (1968). As such, it is not used in the EnSRF. The EnSRF uses the serial solution by Potter. We thank the reviewer for bringing the above publications to our attention.

$$ 3. P. 2843, l. 21-22. "With localisation, skill is less confined to the regions where we assimilate data". But it is zero outside these regions?

Skill outside the regions where proxy information is added is not zero. To clarify, we added in: "For example, we find positive skill throughout Eurasia in boreal winter with localisation, whereas without localisation, positive skill is confined to western Europe, northern Siberia, and central Asia, where proxy information is assimilated. "

$$ 4. P. 2844, l. 4-6. "The spread of the ensemble - here expressed as the intra-ensemble standard deviation - indicates hindcasting uncertainty." Once again – only for the twin experiment involved. It will not represent hindcasting uncertainty when assimilating real observations.

We agree that it will be very difficult to make sense of the ensemble in a real-world application. In the hypothetical case in which we have a sufficiently accurate model (that isn't biased in it's representation of the mean climate, forced response, and internal variability), the ensemble would represent hindcasting (or reconstruction) uncertainty. In a more realistic case for which the above conditions are not met, ensemble spread will only provide weak indication of hindcasting uncertainty of the true climate. We rephrase the corresponding paragraphs to better discuss the additional problems expected in making sense of ensemble spread in real-world applications.

$$ 5. P. 2850, l. 20-22. "This approach extends previous suggestions for data assimilation in paleoclimatology to a high-resolution GCM with data assimilation as used in weather forecasting applications." Once again – the proposed approach has little in common with data assimilation methods used in weather forecasting applications.

We rephrased this to highlight the difference from data assimilation frameworks in weather forecasting.

$$ Conclusion

$$ Data assimilation into climate models represents a major challenge due to the sheer complexity of the physical system. Reconstruction of the past climate through data assimilation of paleo observations seems almost unthinkable to me, and any attempt in this direction must be admired. There is little (I tend to say "no") hope of constraining dynamic models to such a degree that the DA system could have a positive forecasting skill. It less obvious though whether the variational or ensemble methods used in atmospheric or ocean forecasting systems can be useful for extracting some (or any) information about the state of the system from paleo observations. The manuscript gives a positive answer; however in my view this conclusion is not substantiated due to fundamental difference in properties of the system used in the twin experiment (perfect model, perfect forcing) and practice.

While we agree with the reviewer in that our approach by design is overly optimistic, we are far less pessimistic about the potential merit of data assimilation schemes in the paleoclimatological context as a few first approaches have been successfully applied (see introduction). Data assimilation in a paleoclimatology context adds value in that it complements existing approaches for climate reconstruction by making full use of the available information about the physical system (through climate models) and about

the past climate states (through observations). Hence, we cannot see why ensemble methods should be less efficient in extracting information from paleo observations than traditional empirical methods for reconstruction.

Interactive comment on Clim. Past Discuss., 7, 2835, 2011.