**Climate
of the Past
Discussions**

Interactive
Comment

Interactive comment on "An assessment of climate state reconstructions obtained using particle filtering methods" *by* S. Dubinkina and H. Goosse

S. Dubinkina and H. Goosse

svetlana.dubinkina@uclouvain.be

We would like to thank the Anonymous Referee 2 for his helpful remarks and suggestions, which were taken into account in the revised version.

The title announces a comparison of climate state reconstructions from different particle filtering methods, however, the study focusses on a comparison between reconstructions based on paricle filtering and nudging. I suggest adjusting the title to better reflect the contents of the title.

Response: In the revised version it is changed to "An assessment of particle filtering methods and nudging for climate state reconstructions".

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

The setup of the experiment complicates the comparison between the different approaches and the interpretation of the results. Also, the motivation to chose this particular setup is not clear to me. Especially the choice of nudging with near surface temperature over oceans seems particular in the light of recent studies. Keenlyside et al. 2008 have used this simple approach, but Swingedouw et al. 2012 note that this choice (with strong nudging) is problematic and 'optimising' the strength of the nudging parameter is crucial. You don't motivate your setup further, but the results indicate that either the nudging parameter used is too strong, or nudging doesn't work for the SH. More discussion on the motivation for the current setup and the implications is needed (e.g. should nudging be discarded for reconstructions?).

Response: This is taken into account in the revised version. We add a new sensitivity experiment for the nudging and more discussion on the issue: "The nudging is performed over the global ocean by introducing a term into the computation of heat fluxes coming from the atmosphere to the ocean. The nudging parameter α is chosen such that a corresponding relaxing timescale for a mixed layer of $50\mathrm{m}$ depth is 20 days. Moreover, a maximum of $50\mathrm{Wm}^{-2}$ is imposed on the heat flux adjustment due to the nudging. This nudging is 3 times stronger than the nudging used in [Swingedouw et al. 2012] and 5 times weaker than the nudging used in [Keenlyside et al. 2008]. In our standard experiment, sea surface salinity obtained by the nudging has always negative correlations with the pseudo-observations. This could be due to the strong nudging used in this study, since [Swingedouw et al. 2012] argued that a short relaxing timescale of 4 days used in [Keenlyside et al. 2008] resulted in non-physical heat fluxes, while a relaxing timescale of 60 days used in [Swingedouw et al. 2012] did not impose such a spurious behavior of ocean heat content. Therefore, we perform experiments with a weaker nudging with a relaxing timescale of 60 days, as in [Swingedouw et al. 2012]. This nudging is 3 times weaker than the nudging used in the standard experiment. Correlations between sea surface salinity obtained by the weaker nudging and the pseudo-observations are still negative (not shown). Hence, our results do not appear to depend crucially on the value of a relaxing timescale."..."The nudging term

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

strongly modifies the mixing (not shown) leading to a wrong vertical ocean temperature profile and to wrong vertical salinity. Thus in climate models, nudging has to be used with caution at least when applied to a region like the Southern Ocean where a small modification of the vertical stratification of the ocean variables has a strong impact on the surface."

You chose correlation and RMSE to assess the performance of the different assimilation methods. These metrics are simple and easy to understand, however, they do not allow the reader to quickly grasp the added information through assimilation. A measure of skill that relates the performance of the simulations with data assimilation to the performance of the unconstrained ensemble would be much more informative. At the very least, the correlation and RMSE of the unconstrained ensemble should be mentioned in the text and shown in the figures along with the correlation and RMSE of the data assimilation results.

Response: Comparison with simulations without data assimilation is added to the revised version.

Please rewrite the abstract to better summarise the most significant findings and their implications.

Response: Done.

p44l7: 'twin experiment' has not been introduced previously and may be difficult to understand. Consider rephrasing.

Response: This sentence is removed.

p44l8ff: "The net of..." is this sentence needed here?

Response: This sentence is removed.

p45l7ff: Consider rephrasing to "... are biased in the sense that the analysis is linearized and they thus assume gaussian distributions." as it is not only the prior that is

gaussian. Whether the gaussian assumption represents a serious limitation in paleo-climatological applications with generally fairly aggregated data, however, remains to be seen.

Response: This is taken into account in the revised version: "These methods, however successful, are limited in the sense that the analysis is linearized and thus the methods assume Gaussian distributions."

p45l21ff: Similarly, EnKF suffers from spurious off-diagonal covariances when the ensemble is small, with the concesquence of filter divergence and unreliable (underdispersive) probabilistic estimates.

Response: The text in the revised version is modified in order to be more accurate: "Particle filtering has no assumption of gaussianity, uses a full nonlinear model to propagate the particles, but unfortunately, suffers from the "curse of dimensionality" [Snyder et al. 2008], meaning that for a high-dimensional system particles (ensemble members) tend to drift apart during their forward evolution leading, consequently, to large variance in the corresponding importance weights. If the ensemble size is small, after a few data-assimilation cycles all but one of the particles have importance weights close to zero, and an ensemble that has collapsed to a single particle can no longer approximate the probability distribution function of the state."

p47l8: not 'used by'. Better 'used in' or 'used with'

Response: Done.

p47l16: '... which allows us to ...'

Response: Done.

p47l26ff: You perform two sets of experiments, one with information for assimilation available everywhere, which may be thought of indicating the upper limit of skill through assimilation of near-surface temperature, and the second experiment with information where observed temperature is available. Arguably, the spatial coverage of proxy infor-

mation would be even coarser. Some discussion as to how the skill of the assimilation might be affected in a more 'paleo' and thus data sparser context would be interesting.

Response: This is taken into account in the revised version: "For a more distant past, the number of proxies is substantially smaller. Therefore, the performance of data assimilation is expected to be weaker but still satisfactory, if the signal recored in sparse proxies is strong and the aim is to reconstruct large-scale features, e.g. [Annan Hargreaves 2012, Mathiot et al. 2013]."

p47l26ff: A figure illustrating the two different cases (spatial locations used for assimilation) would be very helpful. As the sparser case (using HadCRUT3 locations) has varying temporal density, you may want to show average coverage in an early and late period (say 1850-1900 vs 1950-2000).

Response: New figure is added to the revised version.

p49l2: omit 'following'

Response: Done.

p53l3: How does the nudging parameter compare to nudging in Keenlyside et al., 2008 and Swingedouw et al. 2012?

Response: See the answer above.

p53l4ff: I do not understand what 'taking into account the instrumental surface temperature records HADCRUT3 ...' means in this context. Do you project HadCRUT3 on model-derived EOFs to construct the stochastic error? Please clarify.

Response: We clarify this in the revised version: "The stochastic error ξ^n is constructed as following: we perform empirical orthogonal function (EOF) analysis of the model error, which is the difference between the output of a control model run and instrumental surface temperature records HADCRUT3 [Brohan et al. 2006] over the last 150 years. Then, the noise is a sum of the first ten modes each multiplied by a random coefficient,

and this noise together with the nudging term is added to the equation of heat fluxes."

p53l12: The model error covariance is assumed to be diagonal. This almost certainly overestimates the degrees of freedom in model errors considerably as there is significant spatial correlation to be expected. Please justify your choice and discuss potential biases resulting from this.

Response: This is taken into account in the revised version: "The error covariance of the observations R is computed using the instrumental error and the error of representativeness, as in [Dubinkina et al. 2011], and the model error covariance C is assumed to be a scalar matrix with $(0.5C)^2$ on the diagonal. The latter assumption is relatively crude and it would be more adequate to build the covariance matrix C by taking into account spatial correlation of the model error. We, however, consider a scalar matrix for representing the covariance matrix C and a scalar matrix for the nudging in order to assess the performance of the nudging proposal particle filter when it was obtained by a transition density of a simple form." In future, a more complex nudging (non-scalar matrix α) and better approximations of covariance matrices C and R should be considered, as it was discussed in [van Leeuwen 2010].

p53l26ff: How well do the different data assimilation methods compare with the unconstrained ensemble. That is, how much of the correlation is due to external forcing and how much is due to internal variability. Such a comparison would facilitate the interpretation of differences between the different assimilation methods. Especially for the case with nudging south of 66S, such a comparison might be interesting as it would highlight the importance (or lack) of correlation between mid- and high-latitude weather.

Response: The results of simulations without data assimilation are added in the plots and in the discussion as suggested.

p54l26ff: Consider rephrasing "We obtain that ... smallest mean RMS error." to remove redundancy and increase readability.

Response: Done.

p55l14: From Fig. 2, ocean heat content may need more than 15 years to adjust. What about spinning up with perpetual 1850 conditions (and for the different assimilation methods) to overcome the potential bias?

Response: Indeed, ocean heat content needs more than 15 years to adjust. However, the initial conditions for the year 1850 were derived using a long spin up: We took an equilibrium run with LOVECLIM1.2 in the year 850 and perturbed it four times in order to have four different initial conditions. Then, starting from these initial conditions four transient simulations were performed over 850-1850. These gave us four 1850 conditions for the data-assimilation experiments. We didn't mention how we obtained the initial conditions for 1850 in the previous version, but we do now in the revised version.

p55l19f: replace 'providing with' with 'resulting in'

Response: Done.

p55l20ff: 'Even a free model run...' please provide quantification (see comment above).

Response: The results of simulations without data assimilation are added in the plots and in the discussion as suggested.

p56l13: What about the other variables and what about the unconstrained ensemble in this case?

Response: Other variables are added to the plot. The results of simulations without data assimilation are added in the plots and in the discussion as suggested.

p56l16: Do you mean " ...close to degeneracy for the larger domain."

Response: Yes, but this sentence is now removed.

p56l16: I don't agree with your statement on how degeneracy affects correlation. If the

ensemble collapses to the pseudo-observations, we would expect higher correlations with degeneracy, but very small ensemble spread. Please clarify.

Response: Yes, we agree, but correlations are high if the ensemble collapses to the pseudo-observation not to something else, if it collapses to something else correlations would be smaller. This, however, was not clearly explained in the previous version. Therefore, we modify the text in the revised version: "Keeping the total number of particles the same (96) but reducing the assimilation area, thus reducing the number of degrees of freedom, results in more particles with relatively high importance weights. Consequently, the mean obtained by these particles has higher correlation and the smaller RMS error with the pseudo-observations."

p57l14ff: I do not understand what the prior distribution is. Please rephrase or clarify what you mean by prior distribution in Sec. 2.

Response: In the revised version it is changed to "We assimilate the sparse pseudo-observations over the area southward of 60S in order to decrease the number of degrees of freedom and avoid degeneracy."

p57l20: Is the nudging applied at HADCRUT3 locations only? I assume so from comparing Fig. 2 and 6, but please clarify.

Response: This is taken into account in the revised version: "The nudging is still applied over the global ocean but at the HADCRUT3 locations only."

p57l25: Annan and Hargreaves (2012) have also . . .

Response: Done.

p58l11: Do you suggest here that you can estimate the forced change with data assimilation without changing the forcing of the model over time? I.e. are you arguing that the assimilated information provides constraints that are strong enough to override changes in forcing? Please clarify.

Response: The text in the discussion paper was not clear enough. We do need changes in the forcing. The text is modified in the revised version: "When the forcing is unknown and a random one is applied, the trend as well as the forcing can be still estimated due to data assimilation, see [Dubinkina et al. 2011]."

p58l13: please show the five different initial condition experiments in Fig. 2. This may clarify your point.

Response: Done.

p59l10ff: What is the correlation of the unconstrained ensemble?

Response: The results of simulations without data assimilation are added in the plots and in the discussion as suggested.

p59l10ff: The statement that the correlation is different between the nudging and the particle filters for the end of the 19th century is not backed by the plot. Nudging performs (significantly?) worse from 1907-1948, but not earlier.

Response: This is taken into account in the revised version: "The correlations given by the nudging for the periods 1907–1927 and 1928–1948 are significantly worse than the particle filters. For 1907–1927 the mean correlation of the nudging is even smaller than the mean correlation of simulations without data assimilation, unlike in the case of assimilating the dense pseudo-observations, when for every time period the nudging provides with the mean correlation higher than the mean correlation of simulations without data assimilation (not shown). Therefore, when the pseudo-observations have low density the nudging may not be able to propagate the sparse signal. "

p60l6ff: consider replacing with " ...linked to the pseudo-observations such as surface air temperature and sea ice concentration, but also variables such as geopotential height and sea surface salinity."

Response: Done.

p61l1: This statement is interesting. The assessment of skill for variables that are less closely related to the assimilated information provides a stricter test in that the use of physically ill-conditioned levers are exposed (as in the case with nudging and ocean temperatures at various depths). The extremely efficient particle filter does not seem to suffer from such severe deficiencies, but it is also not able to outperform the sequential importance resampling for variables that are less closely related to the assimilated information. This is somewhat unexpected and it would be worth discussing the strengths and limitations of the extremely efficient particle filter and the sequential importance resampling in more detail.

Response: The efficiency of the nudging proposal particle filter is due to the small ensemble size, since with an infinitely large ensemble both of the particle filters will converge to the same posterior probability density function [van Leeuwen 2010]. To improve reconstructions of variables that are not directly linked to the observations better approximations of covariance matrices C and R should be considered. This is now added to the revised version.

p61l13: estimates

Response: Done.

p61l13: 'reliable' might be misunderstood to mean not over- or underdispersed. You do not discuss these issues here, therefore I suggest to rephrase the sentence. Also, you need to clarify why you think the reconstructions of geopotential height and salinity are not 'reliable'.

Response: This is taken into account in the revised version: "Some developments, however, are still needed in order to get better estimations of variables that are not strongly linked through the model dynamics to the assimilated surface air temperature such as geopotential height and salinity." Additionally, "not reliable" are replaced by "having low correlation".

p61l20: "Past4Future contribution no. X" placeholder?

Response: In the revised version it is changed to "Past4Future contribution no 44".

Fig. 1: Consider reorganising the plots with only 1 panel per location with the different assimilation methods and the truth superimposed for better comparison. Furthermore, you may increase the readability of the plots by freeing up space for the main plot through a reduction of redundant axis labels where possible (e.g. one common axis across multi-panel plot).

Response: We decided to leave 1 panel per method for better readability.

Fig. 3,4,7-10: Please add the 'No data assimilation' case for all figures.

Response: Figures are adjusted in the revised version.

Fig. 7-10: Please clarify that these plots relate to the sparse pseudo-observations.

Response: Figures are adjusted in the revised version.

Interactive comment on Clim. Past Discuss., 9, 43, 2013.