Clim. Past Discuss., 9, C276–C280, 2013
www.clim-past-discuss.net/9/C276/2013/

Climate of the Past Discussions

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

Received and published: 29 March 2013

We would like to thank Prof. Van Leeuwen for his helpful remarks and suggestions, which were taken into account in the revised version.

1. p46, l 24: The wording is inaccurate on the curse of dimensionality. It is mentioned that 'it leads to large variance in the particles'. I assume the authors mean that the particles tend to drift apart during their forward evolution. Note that this is not related to the ensemble size, just to the model dynamics.

Response: We took this into account in the revised version: "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."

2. The naming 'Extremely efficient particle filter' is not ideal, firstly because it is too general, and secondly because the crucial step that makes that method really efficient is not implemented here. My suggestion is to rename the method as something like 'Particle filter with nudging proposal', or 'Nudging proposal particle filter'.

Response: It is changed to the nudging proposal particle filter in the revised version.

3.On the normalisation factor K. Perhaps it is good to mention that K=M in eq (2), just to help the reader a bit, and the crucial point that K does not depend on the state psi, so it is a constant, the same for each particle. Because the particle filter works with only with the relative weights of the particles, so only compares the weights between the particles, this constant is irrelevant.

Response: This is now taken into account in the revised version: "K (equal to M here) is a normalization factor. (Hereinafter any normalization factor will be denoted by K. Since K is the same for every particle, it is irrelevant for the weight comparison.)"

4. In the discussion of the SIR just above section 2.2 the authors mention that 'the remaining particles are duplicated and perturbed'. How is this perturbation done?

Response: Description of the perturbation is added to the revised version: "The perturbation is computed from the empirical orthogonal function analysis of the model error, which is the difference between a control model run and the instrumental records of surface temperature HADCRUT3 over the last 150 years. The perturbation, which is

a sum of the first ten modes each multiplied by a random scalar, is added to initial conditions for surface temperature."

5. The nudging in section 2.2 is very primitive and much more complex scheme's have been proposed and used in which alpha is not just a scalar, but a matrix that incorporated information of error correlations between variables. It would be good to mention that, since a proper choice here could improve the results drastically.

Response: This is now taken into account in the revised version: "In complex nudging schemes, the parameter α is a matrix that incorporates information of error correlation between variables. We, however, consider α to be a scalar matrix for simplicity and also because a scalar α is still used in many recent studies using climate models, e.g. [Swingedouw et al 2012]."

6. To be consistent with the other methods the SIR needs to have model errors (the zeta term) too. Is that indeed the case?

Response: The SIR does not have the zeta term since we want to compare to the nudging proposal particle filter to the SIR used in our previous publications, e.g. Goosse, H., Crespin, E., Dubinkina, S., Loutre, M., Mann, M., Renssen, H., Sallaz-Damaz, Y., and Shindell, D.: The role of forcing and internal dynamics in explaining the Medieval Climate Anomaly", Clim. Dynam., doi:10.1007/s00382-012-1297-0, 2012.

7. End of section 3, the variance in C is assumed to be $0.5\mathrm{C}^2$. What is used exactly? $(0.5\mathrm{C})^2$ or $0.5(\mathrm{C})^2$?

Response: $(0.5\mathrm{C})^2$. The text is changed accordingly.

8. The discussion of the results needs some sharpening. On p54, line 15 in mentioned that the EEPF outperforms the SIR. It should be mentioned that this is due to the small ensemble size. Note that both methods try to solve for the posterior probability density function (pdf), given the prior assumptions on initial conditions and covariance matrices C and R. Both are constructed such that in the limit of an infinitely large ensemble they

converge to the same posterior pdf. The EEPF seems to be more efficient than the SIR for small ensemble sizes.

Response: This is now taken into account in the revised version: "This 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]."

9. It would be really interesting to see a few histograms as representations of the pdf's. Note that the particle filters try to represent the full posterior pdf, not just the mean.

Response: Indeed, the assessment of approximations of the pdf obtained by different data-assimilation methods would be interesting. However, we decided to concentrate on the means.

10. p56, line 11: It is mentions that the correlations are not significant. Please specify how 'significant' is defined.

Response: It is changed in the revised version to low correlations.

11. In the conclusions, last section, it is mentioned that 'some developments are still needed...' It should be good to realsie, and add to the discussion, that the data assimilation system can only do so much. If the observations do not constrain the geopotential height the data-assimilation system should not influence geopotential height. That information should come from other observations. The data-assimilation system should only affect those variables that are dynamically connected to the observations. The system evolution should take care of further spreading of the information. The only way to improve the posterior pdf when we do not have more observations is to improve on the prior, so on the initial conditions, or on the covariance matrices C and R. There is some discussion on that in Van Leeuwen 2010.

Response: This is now 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. Therefore, 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]."

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