Clim. Past Discuss., 9, C45–C47, 2013
www.clim-past-discuss.net/9/C45/2013/

Climate of the Past Discussions

CPD
9, C45–C47, 2013

Interactive Comment

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

prof van Leeuwen (Referee)
p.j.vanleeuwen@reading.ac.uk

Received and published: 20 February 2013

This is an interesting paper that describes the performance of three data-assimilation methods on state estimation within the climate model LOVECLIM1.2. The methods are described in sufficient detail with proper referencing, as is the climate model. The results are interesting and potentially of great value for the community. There are, however a few more technical points that I suggest to address before publication, detailed below:

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.

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'.

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.

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?

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.

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

7. End of section 3, the variance in C is assumed to be 0.5 degrees Cˆ2. What is used exactly? (0.5 degrees C)ˆ2 or 0.5 (degrees C)ˆ2?

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

CPD
9, C45–C47, 2013
Interactive Comment

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

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.

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

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.

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