Nonlin. Processes Geophys. Discuss., 2, C722–C725, 2016
www.nonlin-processes-geophys-discuss.net/2/C722/2016/

Nonlinear Processes in Geophysics
Discussions
Open Access

NPGD
2, C722–C725, 2016

Interactive Comment

Interactive comment on "A local particle filter for high dimensional geophysical systems" *by* S. G. Penny and T. Miyoshi

C. Snyder (Referee)

chriss@ucar.edu

Received and published: 18 March 2016

Review of: A local particle filter for high dimensional geophysical systems, by S. Penny and T. Miyoshi.

Reviewed by: C. Snyder, NCAR

Recommendation: Requires major revision

This manuscript considers an approach for a spatially local update step in particle filters. This is an interesting and potentially important topic and I think the authors have an idea that is worth pursuing. On the other hand, the manuscript is overly terse and sometimes obscure in describing the method, and includes only limited analysis of results. My comments focus on those issues.

Major comments:

1. The details of the LPF implementation should be clarified:

a. How is the smoothing of weights implemented? An equation would be helpful.

b. The smoothing of weights will involve a length scale. Is that length scale related to the localization radius? If so, how and why?

c. The text mentions a transform function T (p 9, l 21; p 10, l 23). Is this the same as the ETKF transformation matrix T that appears in (11)? Is it related to the matrix E in (16), perhaps by a spatial smoothing applied to the columns of E? Please clarify.

2. Further analysis and diagnosis of the results would improve the manuscript.

a. Figure 5 shows that, for given ensemble size, the LPF has smaller error of the prior mean than LETKF only when there are more than 20 observations, despite the fact that the non-Gaussianity of the prior will even greater as the number of observations decreases below 20. This behavior is at odds with expectations and the authors' conclusions (p 16, l 4-8) that more nonlinearity or non-Gaussianity increases the advantage of the LPF relative to the LETKF. What's going on? One possibility is that there are problems with spatial continuity of members after the LPF update when observations are not dense.

b. It would be helpful to include an illustration of the problems encountered without spatial smoothing of weights, and how smoothing ameliorates those problems.

c. The manuscript has little discussion or analysis of how the weight-smoothing length scale or the localization radius might be chosen. This is a significant hole. The statements that do appear seem questionable: "For a given ensemble size, increasing the localization radius [beyond 2 grid points] degraded the accuracy of both methods." (p14, l 3) I would expect the LETKF results to improve by increasing the localization radius as the ensemble size increases, at least for dt = 0.05. d. Some examination of the behavior of the weights would also be helpful (e.g. statistics of maximum weight or

NPGD

2, C722–C725, 2016

effective sample size). Figure 5 shows that the performance of the LPF is almost independent of ensemble size beyond $\sim$75, which suggests that the localization is sufficient to keeps the weights reasonably distributed. But if that is the case, what is limiting the performance of the LPF for $\sim$20 observations or fewer?

3. Literature that should be referenced and discussed:

a. Reich (2013, SIAM J. Sci. Computing) introduces the notion of a particle filter based on transformations. Another take on the transformation view of nonlinear filtering is Metref et al. (2014, NPG); they point out the appeal of such techniques in allowing smooth spatial localization.

b. Lei and Bickel (2011) introduce the notion of computing local weights in a non-Gaussian filter, based on subset of observations that are near a given location. They apply this idea to computing the posterior mean. (Lei and Bickel are cited, but not for this contribution.)

c. Bengtsson et al. (2003, JGR) were the first to point to spatially local updating, using a local subset of observations, as a solution to difficulties of high-dimensional non-Gaussian filtering.

d. Houtekamer and Mitchell (1998) deserve citation. They introduced localization.

4. The conclusions should temper the claim that the dense-observation results from the simple model will be relevant to the atmosphere. Yes, many observing systems provide spatially dense observations, but typically only a subset of prognostic variables (or, even worse, a complicated function of a subset of variables) is observed. In that case, it will still be important to use information from the prior in updating unobserved variables, a situation that is more analogous to the few-observations portion of Fig. 5 where the LPF does not perform well.

Minor comments:

1. Section 2.4:

NPGD

2, C722–C725, 2016

Interactive Comment

a. A brief description would be welcome of how the amplitude of the additive noise was chosen and the sensitivity (or not) of the LPF results to that amplitude.

b. Many PFs apply some additional noise or "jitter" at the resampling step, or sample from a mixture of Gaussians centered at the particles, rather than directly from the empirical distribution. It's worth referencing that this is a standard, if empirical, technique.

c. I don't see that rank is a relevant concept here. The question is whether the particles are "sufficiently" different. Please re-phrase.

d. The "desirable properties" for the analysis ensemble that the authors quote from Pazo et al. are at best distracting and at worst misleading. The update step for this algorithm (i.e. the computation of weights from the observation likelihoods) assumes that the particles are a draw from the prior distribution. The forecast step achieves this (assuming it accounts appropriately for model error) as long as the analysis ensemble is a draw from the posterior distribution. That's all (!) that we require, yet of the conditions quoted only (3) is directly related to the posterior distribution.

2. Section 3.3: The authors should note that relatively simple fixes are available for the LETKF in this case. One might implement a basic quality control based on comparing the observation-minus-forecast difference to, say, 3 or 4 times its predicted standard deviation based on the ensemble and the assumed observation-error variance, or explicitly assign the observation error to one component or the other of the mixture (7) by comparing observation likelihoods.

p 4, l 7: This summary gives the impression that proposal densities for particle filters may be chosen to avoid the need exponentially large ensemble sizes as the problem size increases. Snyder et al. (2015) in fact show the opposite.

p 15, l 10: The computational requirements for the LETKF scale as the cube of the ensemble size, not exponentially with ensemble size.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 1631, 2015.