Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2018-18-AC1, 2018

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Application of ensemble transform data assimilation methods for parameter estimation in nonlinear problems" *by* Sangeetika Ruchi and Svetlana Dubinkina

Sangeetika Ruchi and Svetlana Dubinkina

s.dubinkina@cwi.nl

Received and published: 12 June 2018

Point by point comment.

This manuscript aims to compare performance of ensemble transform Kalman smoother (ETKS) and ensemble transform particle smoother (ETPS) in nonlinear parameter estimation problem. The authors conducted observing system simulation experiments and obtained reasonable results. The scope discussed in this manuscript suits well to Nonlinear Processes in Geophysics. I do not have any major concerns for the experiments presented in the manuscript. However, some discussions and

descriptions are difficult to follow due to insufficient explanation. Here I list the concerns, which would be beneficial to improve the manuscript further.

[General Comments]
1. Scientific Significance: The authors addressed that they applied the ETPS for estimating a large number of uncertain parameters (P2L34). It can be a good motivation; however I could not understand the scientific significance that can be achieved by applying the ETPS and ETKS for the large-dimensional problem. Please address this point clearly in abstract and conclusion.

Reply: The large number of uncertain parameters is of a particular interest for subsurface reservoir modelling as it allows to parameterise permeability on the grid. The most reliable methods of MCMC are computationally expensive and sequential ensemble methods such as ensemble Kalman filers and particle filters provide with a favourable alternative.

2. Lack of explanations: I could not follow several logics of the manuscript, therefore, my major comments includes many "whys" and "reasons". Most of the issues should be solved by adding sufficient explanations.

Reply: we have added more explanations (please see point-by-point answer).

3. Results (Figures): Some figures were discussed insufficiently. It is better to remove figure(s) if they are not needed.

Reply: we have adopted the revised version accordingly.

4. Methods: The author compared the ETKS and ETPS. I am wondering the difference between the ETPS used in this study and a nonlinear ensemble transform filter by

Tödter and Ahrens (2015) (Tödter, J., and B. Ahrens, 2015: A second-order exact ensemble square root filter for nonlinear data assimilation. Mon. Wea. Rev., 143, 1347–1367).

Reply: The paper by Tödter and Ahrens (2015) addresses an important issue of ensemble Kalman (square root) filter being biased for nonlinear models and makes a correction for that. The resulting algorithm does not attempt to estimate the full analysis pdf in contrast to the ETPF.

Also, it is better to compare the localization methodology with local particle filters (Penny and Miyoshi 2016; Poterjoy 2016). Please add more discussion on difference from existing methods.

Penny, S. G. and T. Miyoshi, 2016: A local particle filter for high-dimensional geophysical systems. Nonlin. Processes Geophys., 23, 391-405. Poterjoy, J. (2016). A localized particle filter for high-dimensional nonlinear systems. Monthly Weather Review, 144(1), 59-76.

Reply: The localization methodology considered in this manuscript was particularly developed for ETPF by S. Reich and C. Cotter (2015). The same holds for localization for ensemble transform Kalman filter by Hunt at al (2007). Therefore we keep this comparison. However, we mention Penny and Miyoshi 2016; Poterjoy 2016 in the localization section to give a flavour of an ongoing research in local particle filters.

[Major Comments]
1. P1L13: Please add reason(s) why ETPS is very sensitive w.r.t. the initial ensemble.

Reply: ETKF is very robust while ETPF is very sensitive with respect to the initial ensemble due to a sampling error.

2. P1L15: Please add reason(s) why the localization deteriorated the posterior estimation.

Reply: An issue of an increase in the root mean square error after data assimilation is performed in ETPF for a high-dimensional test problem is resolved by applying distance-based localization, which however deteriorated the posterior estimation of the first mode by largely increasing the variance, which is due to a combination of less varying localized weights, not keeping the imposed bounds on the modes via the Karhunen-Loeve expansion and the main variability explained by the first mode.

3. P7L15: Isn't it possible to apply the localization between variables?

Reply: Localization can be defined for variables that depend on space. Thus it cannot be applied to geometrical parameters, for example.

4. P8L1: I could not understand the sentence " is made such that $y_{\text{obs}} = 48$". Please rephrase this sentence.

Reply: The true parameter u^{true} gives $h(u^{\text{true}}) = 48$.

5. P10L14: Please explain more about reason(s).

Reply: It is interesting to note that ETKF overestimates the tails of the pdfs while ETPF underestimates them, which indicates that there is not enough spread in the ensemble.

6. P11L5: Why? Does it relate to the resampling issue discussed later?

Reply: ETPF provides ensemble members that stay within the original bounds, while ETKF—outside the bounds. Moreover the optimal transport problem solved by ETPF results in some particles being almost identical. Therefore ETPF gives smaller spread than ETKF.

7. Fig.4, Fig. 8 (b) and (c), : I did not understand why this figure is needed because they were not discussed.

Reply: In Fig. 4 we plot the true parameters $u^{\mathbf{true}}$, the mean $\bar{\bar{u}}^a$ and the spread $\bar{\bar{u}}^a \pm \bar{u}^a_{\text{std}}$ of estimated parameters averaged over 10 simulations.
In Fig. 8 (in the revised version it is Fig. 6) we plot mean, minimum and maximum over 10 simulations after data assimilation for the data misfit (a), RMSE (b), and variance (c). ETPF is shown in blue and ETKF in red. We observe that ETPF in underdispersive compared to ETKF (c). This is due to the linear transformation, as it results in some ensemble members being nearly identical. Misfit (a) given by ETPF is smaller than the one given by ETKF for almost all simulations at ensemble sizes greater than 150. The RMSE(b) on the contrary is larger.

8. P15L5, perturbation of ensemble member: In generic PF, the resampling (or inflation) method is very important to avoid the particle convergence. Could you explain why you did not need to consider this issue?

Reply: The model is time independent thus the issue of collapse does not rise here.

9. Fig. 8 (b): I was confused why the ETKS outperforms the ETPS if RMSE is used for the metric.

Reply: The first 3 modes $\mathcal{Z}$ are better estimated by ETPF than by ETKF, thus the permeability field defined only by those modes gives better resemblance to the true permeability when approximated by ETPF (please see Fig.13 in the revised version). ETKF, however, provides a better estimation of higher order modes, thus ETKF outperforms the ETPF if RMSE is used for the metric taking into account all modes.

10. Fig. 10: It is helpful to add RMSEs on the figure.

Reply: Agreed, it is added.

11. Table 1: Could you discuss why the optimal radius for the ETKS is larger than that of the ETPS?

Reply: It was also observed by Cheng and Reich (2015) that the localization radius for ETKF is larger than for ETPF. This is probably related to more noisy approximation of the posterior by ETPF than by ETKF.

Y. Chen and S. Reich, Data assimilation: a dynamical system perspective. Frontiers in Applied Dynamical Systems: Reviews and Tutorials Vol 2, 75-118, 2015

12. P16L15: Please discuss why the localization degrades the posterior estimation.

Reply: The posterior estimation of the first mode $\mathcal{Z}_1$ degraded, while of others improved. The Kullback-Leibler divergence for the first mode is 0.73 (compare to 0.21 without localization), and for second and third is 0.2 and 0.18, correspondently (compare to 0.42 and 0.6 without localization). Variance of the posteriors is larger when localization is applied for both ETPF and ETKF. The localized weights given by Eq. 11 vary less than the non-localized weights given by Eq. 3. Therefore the localized pdf is less noisy than the non-localized. However, localization applied in the form of the Karhunen-Loeve expansion given by Eq. 14 does not retain the imposed bounds on the modes $\mathcal{Z}$ as we need to invert a matrix product of eigenvalue and eigenvector matrices to obtain the modes. By increasing the localization radius to 1.2 we get the Kullback-Leibler divergence 0.65 for the first mode, and 0.14 and 0.12 for the second and third, correspondently, thus the posterior approximation improved only slightly.

13. Conclusion: It would be helpful to add findings and limitation further in this section.

Reply: MCMC methods remain the most reliable methods for estimating the posterior distributions of uncertain model parameters and states. They, however, also remain computationally expensive. Ensemble Kalman filters provide computationally affordable approximations but rely on the assumptions of Gaussian probabilities. For nonlinear models even if the prior is Gaussian the posterior is not Gaussian anymore. Particle filtering on the other hand does not have such an assumption but requires a resampling step, which is usually stochastic. Ensemble transform particle filter is a particle filtering method that deterministically resamples the particles based on their importance weights and covariance maximization among the particles.

ETPF certainly outperforms ETKF for a one parameter nonlinear test case by giving a better posterior estimation. This conclusion also holds for the five parameter test case, however demands a substantially larger ensemble size. Moreover the mean estimations obtained by ETPF are not consistently better than the ones obtained by ETKF. When the number of uncertain parameters is large (2500) a decrease of degrees of freedom is essential. This is performed by localization. At large ensemble sizes ETPF performs as well as ETKF, while at small ensemble sizes ETKF still outperforms ETPF. Even though localized ETPF overfits the data less often than non-localized, localization destroys the property of ETPF to retain the imposed bounds. This results in deterioration of the first mode posterior approximation. Another approach to improve ETPF performance is instead of applying localization to use only first modes in the approximation of log permeabilty as they are better estimated by the method. An advantage of this approach is that it is fully Bayesian. However, one needs to know at which mode to make a truncation and this is highly dependent on the covariance matrix of the log permeability.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2018-18, 2018.