Nonlin. Processes Geophys. Discuss., 2, C283–C291, 2015
www.nonlin-processes-geophys-discuss.net/2/C283/2015/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on “Hybrid Levenberg–Marquardt and weak constraint ensemble Kalman smoother method” *by* J. Mandel et al.

Anonymous Referee #2

Received and published: 21 July 2015

1 General impression

I believe the paper is interesting. In particular, the use of the EnKS to solve the inner loop problem is the real novelty of the paper worth investigating.

I am less pleased with the treatment of the literature. Some contributions need to be mentioned. Others are discussed and mentioned but not properly described, or part of the results relevant to this paper omitted. Grey literature is mentioned. In theory it should not. I personally don't mind but then you should also mention other non peer-reviewed contributions of other colleagues.

Moreover, there are a few unjustified statements. For instance, the standard EnKS as presented as if it was a novelty. Also, the paper does not truly deliver on the promise, especially at the end of Section 5. The numerics is technically fine, but not entirely convincing.

Overall I would ultimately recommend the publication of this paper, but on the condition that the following remarks are properly addressed.

2 Main comments

1. (a) page 869, l.11-41: This passage has wrong statements, and uses gray and peer-reviewed literature in a biased way.
   First of all, let me say that the IEnKF/IEnKS is quite complementary to your idea of using the EnKS to solve the inner loop problem. It has always been claimed (Bocquet and Sakov, 2012, 2013, 2014) that the IEnKS/IEnKS could use a different optimizer (on the shelf, Quasi-Newton, Levenberg-Marquart, etc.). Quasi-Newton and Levenberg-Marquart methods have indeed also been used in those papers. The IEnKS could easily incorporate your idea and use the EnKS to to solve the inner loop problem, which would make a nice blending!

   (b) "Additional work appeared after the first version of this paper was written (Mandel et al., 2013). Bocquet and Sakov (2014) extend the method of Bocquet and Sakov (2012) to 4DVAR..." :
   This chronology is biased and incorrect for these reasons:
   - If you use gray literature then you should mention: http://www.meteo.fr/cic/meetings/2012/ensemble.conference/presentations/session04/1.pdf.

- Bocquet and Sakov (2014) appeared online in final form with a doi number in 2013.
- Please also cite Bocquet and Sakov (2013), which additionally offers a comparison with a (fully cycled) 4D-Var.

(c) Sakov et al. (2012); Bocquet and Sakov (2012, 2013, 2014) not only use finite-differences but also an ensemble transform approach without rescaling which proved to lead to very similar performances. Finite-difference/bundle is interesting in that it mimics the tangent linear, although the ensemble transform is more elegant. This is of direct relevance to your discussion of τ in Section 4. Please mention it.

(d) "However, Bocquet and Sakov (2014) nest the minimization loop for the 4DVAR objective function inside a square root version of the EnKS and minimize over the span of the ensemble, rather than nesting EnKS as a linear solver inside the 4DVAR minimization loop over the full state space as here." This sentence seems nice but it is partially misleading in at least two ways: (i) the IEnKS is more than what is implicit here as it incorporates cycling which is one of the main results of Bocquet and Sakov (2014). So the sentence should start with something like "Focusing only on the variational analysis..." (ii) Bocquet and Sakov (2012, 2013, 2014) emphasized that the minimization can be performed differently opening the way to many consistent variant in the variational analysis. Using your idea of the EnKS for solving would actually be a nice addition to the IEnKS.

(e) "Their method is tied to the use of the sample covariance matrix of the state without localization of the covariance and to strong-constraint 4DVAR": This is partially incorrect for the second statement and plain wrong for the first. Please remove entirely this sentence. I agree that Bocquet and Sakov (2014) strongly rely on the strong-constraint hypothesis (which is not the case for Bocquet and Sakov (2013)). As for localization, it seems that it was

not used in Bocquet and Sakov (2014) on purpose. But it was not claimed it is not possible to use it, only that this is not as simple as with the EnKF. Actually, localization can be used in the IEnKS. Preliminary results were reported early in 2013 http://das6.umd.edu/program/das6_program.html in the largest international data assimilation conference. Please mention clearly that localization has been shown to be possible with the IEnKS.

(f) "However, limiting the EnKF to linear combinations only does not allow common approaches to localization (Sakov and Bertino, 2011)." This is wrong. Please remove the sentence. Local analysis/domain analysis which limits the EnKF to local linear combinations, is extensively used in data assimilation, notably, but not only, via the popular LETKF (Ott et al., 2004). Please read Sakov and Bertino (2011); Nerger et al. (2012). That is why it is rather straightforward to implement localization in the IEnKS. It seems to me that you try to create an opposition that does not exist.

(g) "Ensemble methods for the solution of the 4DVAR nonlinear least squares problem in the weak constraint 4DVAR, or ensemble methods for this problem which allow localization, do not seem to have been developed before.": I disagree. There are published papers (not to mention gray literature) that already discuss the issue in an ensemble variational context, some of them being difficult to ignore for the readership of Nonlinear Processes in Geophysics. For instance: Chen and Oliver (2012), Desroziers et al. (2014), Lorenc et al. (2014) to quote just a few.

2. Implementing Levenberg-Marquardt in the solution of an EnVar problem has been considered first, tested and validated in Bocquet and Sakov (2012) and Chen and Oliver (2013). Surprisingly the authors mentioned "and Bocquet and Sakov (2012), who added regularization" but not the fact that this regularized is based on the Levenberg-Marquardt scheme... Please mention those references, and make it clear. Bocquet and Sakov (2012); Chen and Oliver (2013) did not find any con-

vergence problem with their application, but rather use it as a faster convergence method, as an adaptive method between steepest descent and Gauss-Newton.

Here are quotations from Bocquet and Sakov (2013, 2014):

"One has a choice of minimization scheme: for instance, Sakov et al. (2012) used a Gauss-Newton scheme whereas Bocquet and Sakov (2012) advocated the use of the Levenberg-Marquardt scheme (Levenberg, 1944; Marquardt, 1963) for strongly nonlinear systems. In this article we shall use a Gauss-Newton scheme, because the emphasis is not specifically on strongly nonlinear systems and the number of iterations for convergence in the experiments below is rather limited for most experiments."

"The Gauss-Newton minimization scheme shown in Eq. (2) can easily be replaced by a quasi-Newton scheme that avoids the computation of the Hessian, or by a Levenberg-Marquardt algorithm that guarantees convergence of the minimization. These alternatives have been suggested and successfully tested in Bocquet and Sakov (2012)."

3. page 868, l.23-26. "Gradient methods in the span of the ensemble for one analysis cycle (i.e., 3DVAR) include Zupanski (2005); Sakov et al. (2012) (with square root EnKF as a linear solver in Newton method), and Bocquet and Sakov (2012)" This is wrong. The iterative ensemble Kalman filter in Sakov et al. (2012) and Bocquet and Sakov (2012) is already a 4D ensemble variational method as it has a temporal variational analysis. It coincides with the iterative ensemble Kalman smoother with only one batch of observations. It can be seen as a one-lag smoother. Actually your method essentially coincides with the IEnKF in the lag-one case (modulo some irrelevant details such as the use of stochastic perturbations or not)! Note that Sakov et al. (2012) actually compared two variants of the IEnKF (lag-one smoother): one with the tangent linear model and one with the nonlinear model, which is of direct relevance to your discussion of τ.

4. What you called the nonlinear EnKS (Algorithm 4) is actually the standard EnKS as implemented by the geophysical data assimilation community! You will find many variants (depending on the flavor of the EnKF, perturbed observations or not, with or without model error, with or without localization), but they strictly follow the same smoothing principle: an EnKF pass operated with the nonlinear model, and a backward smoothing pass.

   As as far as the EnKS is concerned (the question is richer in the IEnKS context, and could be in your section 5), the question of using the tangent linear model or not only appears in the EnKF pass and it has been discussed over 20 years. This is what is commonly refereed to the reduced rank Kalman filter approach (RRSQRT) versus the EnKF which differ by the use of the tangent linear or the full model in the propagation. The reason why the nonlinear model is preferred is because it is simpler and natural and capture some nonlinear effects (which turns out to be often more precise).

   Hence, what you call the *nonlinear EnKS* (which in light of the previous comment is a pleonasm) is what is actually used in Evensen (2009); Cosme et al. (2010); Nerger et al. (2014); Bocquet and Sakov (2013, 2014) and several others (see also Cosme et al. (2012)). This should be stated clearly.

5. As mentioned earlier the novel and appealing idea of this manuscript is the use of the EnKS to solve the inner loop problem of a nonlinear problem. Almost up the to end of section 5, the discussion is on the reformulation of known methods and techniques, and the expectation of the reader is great at this point. But, the final theoretical piece of the study does not seem to be given. Where do you describe the full algorithm with the regularization? It is necessary that you give it, because this should stand as the essential piece of the paper and one might think that there is nothing essentially new without it. Besides, this is where nonlinear ensemble variational methods gets trickier.

   Please, explain precisely how you solve Eq.(23) and give us the complete algorithm. This is critical for the paper.

6. The numerics is technically fine and using the OOPS QG model offers a nice illustration. But it is not entirely convincing. This seems a mere check of consistency. Some of the early claims of the paper are not supported, because, for instance, of the absence of localization and cycling (the latter being critical in ensemble methods). The use of localization could have make this paper a bit different from other contributions. I would suggest you to be more caution and state that these experiments offer a partial assessment of the scheme.

3 Minor points or comments related to the major points

1. page 867, l.5-7: "However, Gauss-Newton iterations may not converge, not even locally." Yes, it is important that you mention it. However, in practice (which is also important for this journal), for a well designed system failures to converge are rare.
2. page 868, l.7 "work is relatively cheap": The EnKS is wonderful as it is computationally cheap. But in high-dimensional systems, it has a huge storage requirement which has been warned against (Cosme et al., 2010, and earlier references).
3. page 867, l.17-18: "It is well known that weak constraint 4DVAR is equivalent to the Kalman smoother in the linear case." This is only true for the analysis within the data assimilation window.
4. page 880-886: I believe the discussion on the impact of the hyper-parameters should also depend on the outcome of a long cycling of the experiment. You may

not have to achieve a high precision minimization to address properly the nonlinear effects within the data assimilation window and propagate later the ensemble (hence the errors) through the window.

5. page 887, l.17: “and have shown that it is capable of handling strongly nonlinear problems”: in the absence of cycling, it is difficult to really conclude. Cycling is important for the L63 and the QG model. That said, the numerical experiments are convincing enough for the case of a single nonlinear minimization. Please mitigate your statements.

References

Bocquet, M. and Sakov, P.: Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems, Nonlin. Processes Geophys., 19, 383–399, doi:10.5194/npg-19-383-2012, 2012.

Bocquet, M. and Sakov, P.: Joint state and parameter estimation with an iterative ensemble Kalman smoother, Nonlin. Processes Geophys., 20, 803–818, doi:10.5194/npg-20-803-2013, 2013.

Bocquet, M. and Sakov, P.: An iterative ensemble Kalman smoother, Q. J. R. Meteorol. Soc., 140, 1521–1535, doi:10.1002/qj.2236, 2014.

Chen, Y. and Oliver, D. S.: Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother, Math. Geosci., 44, 1–26, 2012.

Chen, Y. and Oliver, D. S.: Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification, Comput. Geosci., 17, 689–703, 2013.

Cosme, E., Brankart, J.-M., Verron, J., Brasseur, P., and Krysta, M.: Implementation of a Reduced-rank, square-root smoother for ocean data assimilation, Ocean Modelling, 33, 87–100, 2010.

Cosme, E., Verron, J., Brasseur, P., Blum, J., and Auroux, D.: Smoothing problems in a Bayesian framework and their linear Gaussian solutions, Mon. Wea. Rev., 140, 683–695, 2012.

Desroziers, G., Camino, J.-T., and Berre, L.: 4DEnVar: link with 4D state formulation of variational assimilation and different possible implementations, Q. J. R. Meteorol. Soc., 140, 2097–2110, doi:10.1002/qj.2325, in press, 2014.

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer-Verlag Berlin Heildelberg, second edn., 2009.

Lorenc, A. C., Bowler, N. E., Clayton, A., M., Fairbarn, D., and PRing, S. R.: Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for gloabl NWP., Mon. Wea. Rev., 132, 2495–2502, doi:10.1175/MWR-D-14-00195.1, 2014.

Nerger, L., Janjić, T., Schröter, J., and Hiller, W.: A regulated localization scheme for ensemble-based Kalman filters, Q. J. R. Meteorol. Soc., 138, 802–812, doi:10.1002/qj.945, 2012.

Nerger, L., Schulte, S., and Bunse-Gerstner, A.: On the influence of model nonlinearity and localization on ensemble Kalman smoothing, Q. J. R. Meteorol. Soc., 140, 2249–2259, doi:10.1002/qj.2293, 2014.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, A.: A local ensemble Kalman filter for atmospheric data assimilation, Tellus A, 56, 415–428, 2004.

Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, Comput. Geosci., 15, 225–237, 2011.

Sakov, P., Oliver, D., and Bertino, L.: An iterative EnKF for strongly nonlinear systems, Mon. Wea. Rev., 140, 1988–2004, 2012.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 865, 2015.