Nonlin. Processes Geophys. Discuss.,
doi:10.5194/npg-2016-44-RC2, 2016

Nonlinear Processes in Geophysics
Discussions

Interactive comment on "An Estimate of Inflation Factor and Analysis Sensitivity in Ensemble Kalman Filter" *by* G. Wu

Anonymous Referee #2

Received and published: 22 November 2016

[12pt]article graphicx graphics amsmath amssymb hyperref latexsym [utf8]inputenc authblk

1 General Comments

This paper presents a new technique in estimating model error covariance inflation factor which is widely used in ensemble-based filters. An inflated model error covariance is necessary to arrest the divergence of the filter. This paper relies on estimating the inflation factor from an objective function inspired from the domain of generalized cross validation (GCV) techniques widely used in the field of machine-learning. The author also shows that this method, in comparison to a basic Ensemble Kalman Filter,

considerably improves the root-mean-squared error and enhances the influence of observations on the analysis when applied to the Lorenz 96 model.

However, it is well known that a basic Ensemble Kalman Filter (EnKF) falls short on many accounts and a mere improvement with respect to it does not give much credence to this new technique. Even introducing a simple constant multiplicative inflation factor to the basic EnKF considerably improves the analysis. The author should address the following questions :

1) How does this method fare when compared to simple multiplicative inflation techniques like setting a constant inflation factor in the basic ensemble kalman filter ? It will be more interesting to see this method pitted against other sophisticated inflation schemes.
2) How does the time-series of the inflation factor look like ?
3) What are the improvements in other statistical measures like spatial correlation, correlation coefficients, measure of ensemble spread, etc ?
4) What are the computational challenges in estimating the inflation factor ?
5) How does this method fare in presence of sparse observations ?

It is clear that English is not the native language of the author. One of the suggestion is to seek help and revise the language of this paper. This paper, in its present version, is more of a curious exercise and needs to be considerably revised to make it publishable. If these revisions are made, this paper may be accepted for publication.

2 Specific Comment

1) “The objective function needs to be minimized to estimate the inflation parameter” is not explicitly mentioned when it is introduced in the main text. It is however casually mentioned in the Discussion section.
2) What is the motivation of generating observations at every 4 time-steps ?
3) The simplest observation error covariance matrix is a diagonal R. There are many in-situ observation systems in which R is diagonal. What happens when R is diagonal ? Also, what is the motivation behind choosing that particular expression of R ? What is the harm in introducing a parmeter in the expression of R which may be conveniently tuned to set R diagonal ?
4) P4 L1 : What does the author mean by **"gradually important"** ?
5) P5 L9-10: Kindly elaborate on what the “favorable properties” of GCV are ?
6) P6 L20-21 : What does the author mean by “ The EnKF assimilation result is . . . **sufficiently close** to the corresponding true state . . .” ?
7) Please elaborate the flowchart in the main text as well.
8) What is N in Fig 1 ?
9) P13, L 8-9 : “The model forecast changes very much along with the change in F . . ..”. This is a very general statement. Please be a little more specific. Also cite references that show the model to be chaotic for $F > 3$.

3 Technical Comment

1) There are many notable absence of articles, usage of wrong prepositions and a few grammatical corrections to point out. One such instance is in P13 L8-9. “The model forecast . . . and is chaos with integer values of F larger than 3”. This should be “ The model forecast . . . and is **chaotic for** integer values of $F > 3$”.
2) P13 L22: “ The variety of the analysis RMSE . . .”. It is not clear what the author

wants to convey.