8B.4

A DIFFUSIVE KALMAN FILTER AND A PARAMETERIZATION FOR NULL SPACE COVARIANCES

Xiaosong Yang * and Timothy DelSole
George Mason University, Fairfax, VA, and
Center for Ocean-Land-Atmosphere Studies, Calverton, MD

1. Introduction

A problem with the Ensemble Kalman Filter (EnKF) is filter divergence. Filter divergence occurs when the covariances specified in the filter are inconsistent with their true values (Maybeck, 1979). Since the EnKF approximates forecast covariances based on a finite sample, filter divergence is unavoidable in the EnKF. One type of filter divergence occurs when the covariances are underestimated and the filter weights the first guess too heavily. Two methods for avoiding filter divergence are covariance inflation (Anderson and Anderson, 1999) and localization (Hamill et al., 2001; Houtekamer and Mitchell, 2001). Covariance inflation attempts to avoid filter divergence by simply inflating the covariance of the ensemble. However, if only covariance inflation is applied, then it is found that filter divergence still occurs if the ensemble size is sufficiently small. This result may be understood as follows. The full model space can be split into two subspaces: the space spanned by the ensemble, which we call the ensemble space, and the complement, which we call the null space. Generally for atmospheric applications, the ensemble size is less than the model dimension, so that the ensemble does not span the full model space. Specifically, there is always a null space in atmospheric applications. The traditional EnKF updates only those vectors in the ensemble space. It follows that vectors in the null space are not updated, which is equivalent to assuming that the forecast covariance of the null space vectors vanishes. Thus, no matter how much inflation is applied, this inflation only influences the ensemble space, leaving the variances in the null space zero and hence underestimated.

Application of covariance localization might avoid the above problems because it changes the rank of the forecast covariance, and in particular it reduces or even eliminates the null space. Nevertheless, it is found empirically that covariance localization alone is not adequate to avoid filter divergence, and most applications of covariance localization also apply covariance inflation. Furthermore, both covariance inflation and covariance localization involve tunable parameters whose values are not always well known.

* *Corresponding author address:* Xiaosong Yang, Center of Ocean-Land-Atmosphere Studies, Calverton, MD, 20705;e-mail: xyang@cola.iges.org

In this work we explore an alternative method for avoiding filter divergence. The basic idea is to explicitly model the covariances in the null space. In one extreme, one could argue that the filter has no information about the covariances in the null space, since the forecast ensemble is orthogonal to it. Therefore, one could assume that the covariances in the null space should be infinite, corresponding to the limit of complete lack of knowledge. Consistent with this lack of knowledge, we assume that the vectors in the null space are uncorrelated with the vectors in the ensemble space. We call the resulting filter the Diffusive Ensemble Kalman Filter (DEnKF).

In the numerical experiments reported below, we compare the performance of the DEnKF with the traditional EnKF. We find that the DEnKF vastly improves the divergence problem, but does not eliminate it. Interestingly, the traditional EnKF with *both* covariance inflation and localization does seem to eliminate filter divergence. This fact implies that the assumption of “complete lack of knowledge” is not appropriate. But what other information is available about the null space? In the absence of instantaneous information about the null space, the only other information available is its past history. We explore an alternative approach in which the covariances in the null space equal the covariances of all previous ensembles.

2. Experimental setup

The model used here is the Lorenz-96 model (Lorenz and Emanuel, 1998). It is a nonlinear model with a state vector dimension of 40. The consecutive model states are obtained by integrating the model forward with the time interval 0.05, and a fourth-order Runge-Kutta numerical method is applied at each model time step. The “truth” is one single integration of the model. The observational data set was constructed by adding Gaussian white noise with zero mean and unit variance to the truth at each of the 40 grid points, thereby producing 40 “observations.”

The traditional EnKF used here is the mean-preserving square-root Ensemble Kalman Filter of Evensen (2004) and Sakov et al. (2007). The ensemble size used in this study is 10, corresponding to the regime in which the ensemble size is much smaller than the model dimension. The covariance inflation for all experiments in this study, when applied, is the adaptive covariance inflation algorithm proposed by Anderson (2007). The localization applied is the fifth order polynomial function of Gaspari and Cohn (1999) with half-width c, and c is a tuning parameter. In all experiments the filter errors are computed as the root mean square (RMS) of the difference between the analysis and the truth over the 40 grid points and from model time steps 3000 to 6000.

The log likelihood function (LLF) can be used as an online detection of filter divergence. LLF can be written as:

$$2L(N) = -\log(2\pi)^N |C| - z_N^T C^{-1} z_N , \quad (1)$$

where N is the number of observations per time step, $|\,|$ is the matrix determinant operator, z_N is the innovation vector, ***C*** is the covariance of z_N and the superscript T is the matrix transpose operator (Schweppe, 1965). As noted by Maybeck (1979), the first term of the right hand side of Equation (1) is a slowly varying negative term independent of the innovations. For the Kalman Filter, the innovation sequence is a white Gaussian sequence with mean zero and covariance matrix:

$$C = HP^f H + R , \quad (2)$$

where H is the observation operator, P^f is the covariance of the forecast ensemble, and R is the observation error covariance (see Chapter 5 of Maybeck, 1979). Thus the quadratic form $z_N^T C^{-1} z_N$ is distributed as the chi-square with N degrees of freedom χ_N^2 (see Result 4.7. of Johnson and Wichern 2002). For the purpose of detecting filter divergence, LLF can be simplified as:

$$SL(N) = -z_N^T C^{-1} z_N . \quad (3)$$

The upper 99.5% threshold value for the chi-square distribution with 40 degrees of freedom is $\chi_{40}^2(0.995) = 66.8$. Accordingly, a filter divergence is declared when LLF is less than -66.8 more than 99.5% of the time. Otherwise, we say the filter converges.

3. Comparison of Traditional and Diffusive EnKF

Figures 1a-d show a typical result for the truth, observation, forecast, and analysis by the traditional EnKF at one grid point in the Lorenz-96 model. The corresponding LLF is shown in figs 2a-d (for a longer time period). Inspection of fig. 2 shows that the filter converges only if both covariance inflation and localization are applied. In cases when the filter diverges, the analysis is weighted too heavily toward the model forecast, allowing the analysis to diverge from the observations. Interestingly, the traditional EnKF with localization only still diverges (figs. 1c and 2c) even though there is no null space.

The results for the diffusive EnKF are shown in figs. 1e and 2e. The figures show that the diffusive EnKF substantially improves the LLF, although technically divergence still occurs by our criterion. However, unlike the previous form of divergence, the divergence in this case arises because the analysis is weighted too heavily to the observations. Indeed, the analysis reveals much more high frequency noise than the truth, owing to the white noise in the observations. Perhaps this form of divergence is preferable to the other form of divergence in which the analysis rejects the observations. Nevertheless, one can always improve the LLF by inflating the covariances. Figs. 1f and 2f show the result of applying inflation to the diffusive EnKF, which is to improve the LLF further.

The diffusive EnKF can be interpreted as an extreme example of inflation for the null space. Yet, even with infinite covariances in the null space, the diffusive filter still diverged. Similarly, in the traditional EnKF with localization, there is no null space, yet the filter still diverges. Thus, an interesting conclusion from the above results is that the filter converges only when the covariance of both the ensemble space and the null space are inflated-- inflating just one subspace is not enough to avoid filter divergence.

It should be noted that the above conclusion holds for our experimental set up with relatively small ensemble size (compared to the dimension of the system). In cases with relatively larger ensemble size, the impact of the null space may not be as important. To gain insight into this issue, we show in fig. 3 the performance of the traditional and diffusive EnKF, with inflation, as a function of ensemble size. For the ensemble size 41, there is no null space, so the diffusive EnKF is identical to the traditional, and the values of RMS for the two filters are almost the same. We see that the RMS for the traditional filter decreases dramatically and eventually the filter converges after 15 ensemble members. This implies that inflation alone can allow the filter to converge if the ensemble size is sufficiently large. Equivalently, if the ensemble size is too small, then inflation alone is not enough to prevent filter divergence. In practical atmospheric applications, the dimension of the state space exceeds 100,000 whereas typical ensembles sizes are less than 100, or 0.1% of the model dimension. Therefore, ensemble size of 15 or more in the 40-variable Lorenz model is an unrealistic regime for atmospheric applications. Thus, for small ensemble sizes relative to the model dimension, the

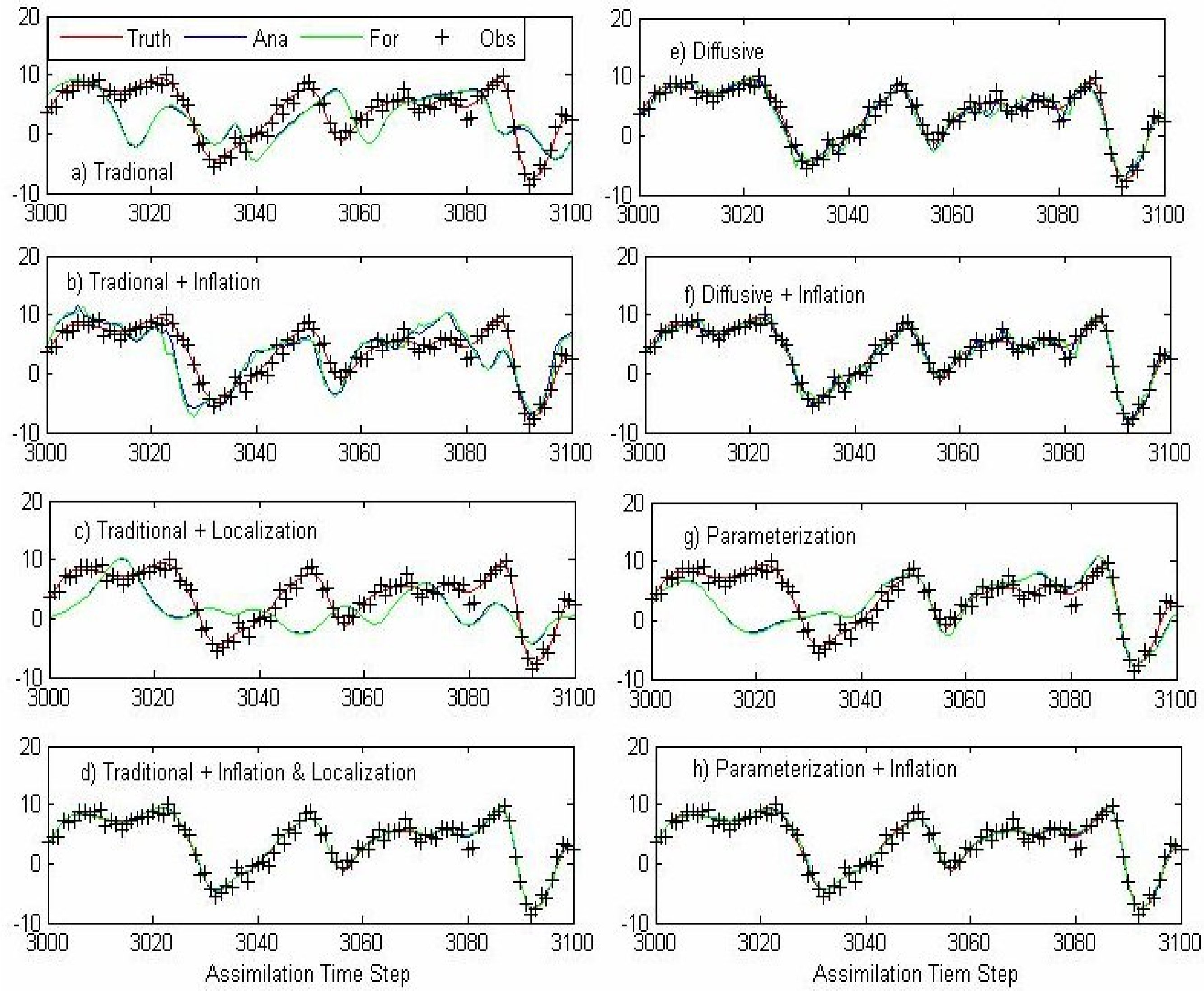

Fig.1 Time series of the true solution, the model forecast, the analysis and the observation at one grid point for a) Traditional EnKF without inflation and localization, b) Traditional EnKF with inflation only, c) Traditional EnKF with localization only, d) Traditional EnKF with localization and inflation, e) Diffusive EnKF without inflation, f) Diffusive EnKF with inflation, g) Null space parameterization without inflation, and h) Null space parameterization with inflation.

diffusive EnKF may be an attractive alternative to the traditional EnKF.

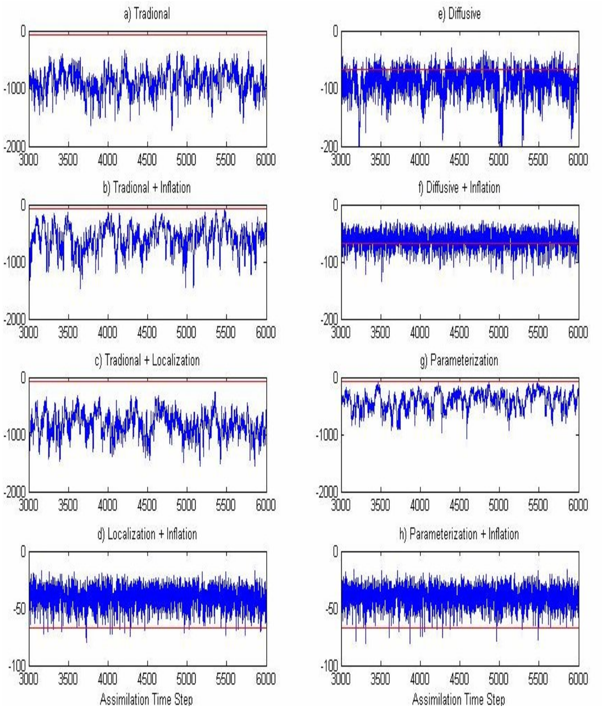

Fig.2 Time series of the log likelihood function (LLF) for a) Traditional EnKF without inflation and localization, b) Traditional EnKF with inflation only, c) Traditional EnKF with localization only, and d) Traditional EnKF with localization and inflation, e) Diffusive EnKF without inflation, f) Diffusive EnKF with inflation, g) Null space parameterization without inflation, and h) Null space parameterization with inflation. Ensemble size is 10 for all experiments. Red line indicating the threshold value of LLF.

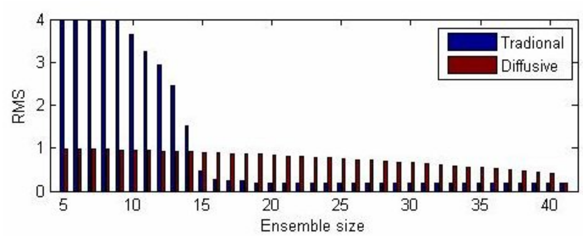

Fig.3 The root square mean (RMS) of the difference between the true solution and the analysis as a function of ensemble size for the traditional EnKF with inflation (Blue bars) and the diffusive EnKF with inflation (Red bars). Results are averaged over the 3000 to 6000 assimilation time step.

4. A parameterization for the null space covariance

The traditional and diffusive EnKFs define two extreme cases for representing the null space covariance: the traditional essentially treats the null space covariance as vanishing while the diffusive treats it as infinite. The fact that the diffusive EnKF diverges implies that the assumption of "complete lack of knowledge" is not appropriate. Furthermore, the fact that the traditional EnKF with localization and inflation converges implies that there must exist some information about the null space covariances from the available ensemble. In the absence of instantaneous information about the null space, the only other information available is its past history. We now consider an alternative approach in which the covariances in the null space equal the covariances of all previous ensembles. Let P^f_{clm} be the average of all previous forecast covariances P^f. Then, we consider a new forecast covariance matrix defined as

$$\tilde{P}^f = P^f + U_{null} P^f_{clm} U_{null}\,, \qquad (4)$$

where U_{null} is the projection matrix to the null space. P^f_{clm} is the climatology of the model forecast error covariance.

The rank of the above matrix increases with time; it will usually become full rank when the number of time steps exceeds the dimension of the model divided by the ensemble size. One might also apply localization or the hybrid method of Wang et al. (2007) to increase the rank of the matrix, but here we show results only after 3000 time steps, well after the above covariance becomes full rank. The traditional EnKF is applied using the covariance $\tilde{P}^f$ to update the mean analysis, and using the covariance P^f from the forecast ensemble to update the analysis perturbations. The results of applying this new filter with parameterized covariances in the null space are shown in figs. 1g and figs. 2g. We see that the filter still diverges, and the sense of the divergence is that the analysis is weighted too heavily toward the forecast. This suggests that the forecast covariances are underestimated and that the filter could benefit by inflating the forecast covariances. Figs. 1h and 2h show the results of applying inflation to the new filter, which reveals that the filter with inflation converges.

Recall that the best filter results were obtained from the traditional EnKF with both covariance localization and inflation. It is interesting to compare this best filter with the above filter with parameterized null space covariances. However, covariance localization involves a tunable parameter, namely the half-width of the localization. Fig. 4 shows the RMS as a function of the localization half-width for the two filters. Note that the RMS of the filter with parameterized null space covariances does not depend on a half-width, but for comparison purposes we plot its constant value at each half-width. The figure shows that the two filters have only marginally different RMS values, although the extra tuning parameter in the traditional EnKF allows it to have slightly better performance, provided that the truth is known. We conclude that use of parameterized covariances in the null space can give filter performance comparable to the best tuned traditional EnKF with inflation and localization.

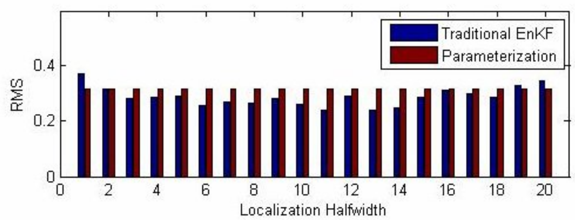

Fig.4 The root square mean (RMS) of the difference between the true solution and the analysis as a function of Gaspari-Cohn localization half-width for the traditional EnKF with inflation and localization (Blue bars) and the null space covariance parameterized (NSCP) EnKF with inflation (Red bars). Ensemble size is 10 for both filters. Results are averaged over the 3000 to 6000 assimilation time step.

5. Conclusions

In this study, the diffusive EnKF is proposed. This filter assumes that the covariances in the space complementary to the ensemble are infinite, corresponding to "complete lack of knowledge." Numerical experiments using small ensemble sizes demonstrate that the diffusive EnKF dramatically improves, but does not eliminate, filter divergence. In fact, the diffusive EnKF exhibits divergence due to weighting too heavily toward the observations, which in practice may be preferable to divergence due to weighting too heavily to the forecast. We believe that the diffusive EnKF also has educational value in elucidating the role of the null space in filter divergence. In particular, the fact that the diffusive filter still diverges, and the traditional filter with inflation and localization converges, implies that the assumption of "complete lack of knowledge" is inappropriate. These considerations lead us to explore an alternative approach in which the covariances in the null space equal the covariances of all previous ensembles. This latter filter still diverged, but it converged when covariance inflation was applied. Furthermore, the performance of this filter was comparable to the best performance from the traditional EnKF with tuned inflation and localization, even though the new filter contains no tunable parameters.

Reference

Anderson, J. L. and Anderson, S. L., 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741-2758.

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210-224.

Evensen, G., 2004: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 54, 539-560.

Gaspari, G., and Cohn, S. E., 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc*., **125**, 723-757.

Hamill, T. M., Whitaker, J. S. and Snyder C., 2001: Distance-dependent filtering of background-error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev*., **129**, 2776-2790.

Houtekamer, P. L., and Mitchell, H. L., 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev*., **129**, 123-137.

Lorenz, E. N., and Emanuel, K. A., 1998: Optimal sites for supplementary weather observations: simulation with a small model. *J. Atmos. Sci.*, **55**, 399-414.

Maybeck, P. S., 1979: Stochastic models, estimation and control, Vol. 1, Academic Press, 423p.

Johnson, R. A., and Wichern, D. W., 2002: Applied Multivariate Statistical Analysis, Pearson Education Asia, 767p.

Sakov, P., Oke, P. R., and Corney, S. P., 2007: Only mean-preserving ensemble transformations should be used in ensemble square root filters. *Mon. Wea. Rev*., in press.

Schweppe, F. C., 1965: Evaluation of likelihood functions for Gaussian signals. IEEE Transactions on Information Theory, 61-70.

Wang, X., Hamill, T. M., Whitaker, J. S., and Bishop, C. H., 2007: A comparison of Hybrid Ensemble Transform Kalman Filter—optimum interpolation and Ensemble Square Root Filter analysis schemes. *Mon. Wea. Rev*., **135**, 1055-1076.