Tellus (2007), 59A, 778–780
Printed in Singapore. All rights reserved

TELLUS

Response to the discussion on "4-D-Var or EnKF?" by Nils Gustafsson

By EUGENIA KALNAY[1*], HONG LI[1], TAKEMASA MIYOSHI[2], SHU-CHIH YANG[1] and JOAQUIM BALLABRERA-POY[3], [1]*University of Maryland, College Park, 3431 CSS building, MD 20742-2425, USA;* [2]*Numerical Prediction Division, Japan Meteorological Agency, Tokyo, Japan;* [3]*Institut de Ciències del Mar (CSIC), Barcelona, Spain*

Gustafsson (2007) discussion of our paper '4-D-Var or EnKF?' is very stimulating and brings up several interesting and valid points. We completely agree that the most useful question should be 'How can ideas from EnKF and 3–4-D-Var be best combined?' In the response to his comments, we point out that many of the ideas originally developed for 4-D-Var can be adapted within the EnKF framework.

General comments

1. We strongly agree that operational centres should carry comprehensive comparisons not possible within the infrastructure of a university setup.

2. This is a very good point. The EnKF analysis is a linear combination of balanced trajectories, but localization affects the balance between the geopotential and wind increments. The 'observation localization' approach in which the observation error is increased for observations according to their distance to the grid point being analyzed (section 3.2.1, Hunt et al., 2007), should maintain the geostrophic balance of the increments better than the standard "background localization" approach, especially for long waves. LETKF localization based on physical distance has been found to be better than box localization (Fertig et al., 2007b). Even using box localization, Liu et al. (2007), found that the LETKF analyses were considerably more balanced than those obtained with PSAS (Cohn et al., 1998), using the same simulated observations.

3. We agree that 4-D-Var, in principle, has a higher dimensionality than EnKF. 4-D-Var final analysis increments project strongly on the leading (growing) final singular vectors, which are similar to the leading Lyapunov or bred vectors, whereas decaying modes are more important at the beginning of the assimilation window (e.g. Pires et al., 1998). Being lower rank, it is essential for the EnKF to capture growing errors, which are the ones that dominate forecast errors. In QG model experiments, the 'stochastic seeding' (in lieu of variance inflation) was found to be particularly effective when added at the observation locations, and can be interpreted as a simple representation of the stochastic forcing introduced by nonlinearities and observation errors (Corazza et al., 2007).

4. We agree that accounting for model errors is an important area of research for both 4-D-Var and EnKF. The large improvement obtained with long windows in 4-D-Var in a perfect model simulation does require an exact adjoint. For short windows and real forecasts, given the presence of significant model errors in every nonlinear model, it is not surprising that the impact of more subtle changes such as the discretization of the linear model dynamics was found to be small for the HIRLAM model.

Discussion on table 7

1. We agree that 3-D-Var avoids data selection complications and that balances were vastly improved within 3-D-Var global data selection compared with OI. However, the main advantage of 3-D-Var with respect to OI is a more accurate estimation of the multivariate background error covariance based on, for example, the NMC method. The use of ensembles maintains balance because the analysis is essentially a linear combination of the ensemble forecasts (see also response to general comment 2).

2. We agree that the existence of automatic compilers (e.g. TAMC) to obtain the linear and adjoint models is extremely helpful. However, our experience with a simple model such as the QG channel model was that subtle errors not captured by the standard tests were present in the linear and adjoint models obtained with this compiler, and almost a year of effort was needed to find and correct these errors. By contrast, students frequently code their own version of the LETKF with a simple model, and then adapt it from one model to another.

*Corresponding author.
e-mail: ekalnay@atmos.umd.edu
DOI: 10.1111/j.1600-0870.2007.00263.x

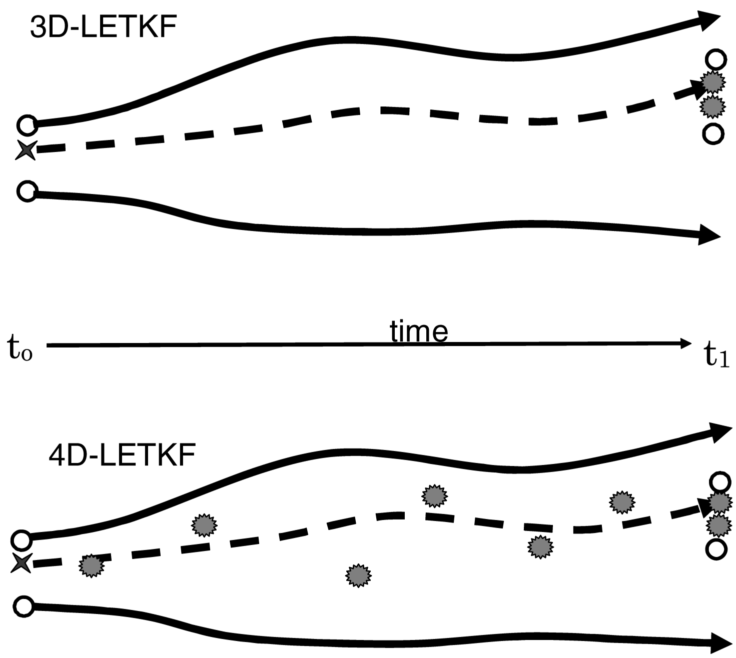

Fig. 1. Schematic showing that the 3D-LETKF finds the linear combination of the ensemble forecasts that best fits the observations at the analysis time t_1, whereas 4D-LETKF finds the linear combination of the forecasts that best fits the observations throughout the assimilation window $t_0 - t_1$. The white circles represent the ensemble of analyses, the full lines represent the ensemble forecasts, the dashed lines represent the linear combination of the forecasts whose final state is the analysis, and the grey stars represent the observations. The cross at the initial time of the assimilation window t_0 is a no-cost Kalman smoother, i.e., an improved analysis at t_0 that uses the information of the "future" observations within the assimilation window by weighting the ensembles at t_0 with the same weights as the analysis at t_1.

3. Miyoshi and Aranami (2006) have tested the potential of assimilation of rain in a regional model, with somewhat encouraging results. We agree this is an important emerging area of research.

4. This is another good point. As indicated above, a digital filter has been used in EnKF to initialize the background trajectories (Whitaker et al., 2006; Szunyogh et al., 2007), and it is not clear whether it is possible to include it as weak constraint. It would be simple to include a weak balance constraint in the LETKF analysis by assuming that the time derivative of, say, surface pressure or horizontal divergence, should be small, and penalizing the ensemble members whose time derivative is large, but this idea has not yet been tested.

5. We agree.

Example of flow-dependence in 4-D-Var and 4-D-LETKF no-cost smoothing

The example presented by Gustafsson (2007) in section 4 is very revealing. It shows that in a rapidly developing situation, the introduction of observations at the end of a 6-hr assimilation window results in (a) an analysis increment which is tight and anisotropic, in contrast to that obtained with 3-D-Var and (b) an improvement in the estimate of the state at *the start of the assimilation window*, reflecting an increase of the baroclinicity that leads to the rapid storm development. This is an excellent example of the *smoothing* properties of 4-D-Var: although constrained by the state-independent initial background error covariance, 4-D-Var is able to improve the initial state using 'future' observations within the assimilation window. Such properties would be particularly useful in a reanalysis mode, in which observations obtained for a time later than the analysis time are indeed available but have not been yet been used in the 3-D-Var reanalyses (NCEP-NCAR, NCEP II, ERA-40 and JRA-25).

This ability to use 'future' observations within the current assimilation window can be easily implemented within LETKF/4-D-LETKF. Figure 1 shows a schematic of how this is done. The 4-D-LETKF analysis at the end of the assimilation window is given by the linear combination of ensemble trajectories that best fits the observations throughout the assimilation window. The same linear combination can be used at the beginning of the assimilation window to obtain a smoothed analysis that takes advantage of the future data, and is more accurate than the original analysis. Tests with the Lorenz 40-variable and the QG models show that the analysis obtained by this cost-free smoother is indeed more accurate than the original LETKF analysis. Furthermore, the final time analysis ensemble (which provides the uncertainty for each variable at each grid point) is also given by a linear combination of the ensemble forecasts. The same weights can be used at the initial time of the assimilation window to obtain estimates of the uncertainty of the initial smoother analysis, also useful for reanalysis. As in 4-D-Var, the improved initial analysis does not improve the final state: the forecast starting from the smoother's analysis is the same as the analysis at the end of the interval (assuming linearity of the perturbations).

Concluding remarks and acknowledgments

We agree completely with Nils Gustafsson that the optimal approach is to combine the best ideas of 4-D-Var and EnKF. We believe that in most cases, ideas developed in the context of 4-D-Var, for which there is so much experience, can be easily adapted and included within EnKF. An additional example of the application to EnKF of variational developments is the adaptive estimation of observational errors and inflation parameters within EnKF, using the methodology developed by Desroziers et al. (2005) and applied by Navascues et al. (2006). Kalnay et al. (2007) show that it is possible to adaptively estimate within LETKF the optimal inflation parameter and the observation error variance using the statistics of observation minus forecast, analysis minus background, and observations minus analysis. Similarly, the diagnostic estimate of the 'influence of observations' introduced by Cardinali et al. (2004) can be easily computed 'on the fly' within EnKF. A hybrid approach extending the 3-D-Var to allow for flow-dependent background error covariance (e.g. Corazza et al., 2002; De Pondeca et al., 2006) is certainly worth exploring because of its efficiency.

We are most grateful to Nils Gustafsson for his comments and to Harald Lejenäs for making this interaction possible. Discussions with colleagues and several students from the chaos/weather group at the University of Maryland (Brian Hunt, Istvan Szunyogh, Eric Kostelich, Ed Ott, Junjie Liu, Elana Fertig and Jim Yorke) as well as with Jeff Whitaker and Milija Zupanski have strongly influenced this response.

References

Cardinali, C., Pezzuli, S. and Andersson, E. 2004. Influence matrix diagnostic of a data assimilation system. *Q. J. R. Meteo. Soc.* **130**, 2767–2786.

Desroziers G., Berre, L., Chapnik, B. and Poli, P. 2005. Diagnosis of observation, background and analysis error statistics in observation space. *Q. J. R. Meteor. Soc.* **131**, 3385–3396.

Gustafsson, Nils 2007. Discussion. *Tellus* **59A**, this issue.

Kalnay, E., Li, H. and Miyoshi, T. 2007. Adaptive estimation of background and observation errors within the Local Ensemble Transform Kalman Filter. Preprints of the 11th Symposium on Integrated Observing and Assimilation Systems, Paper 7.2A, 87th AMS Annual Meeting, San Antonio, TX, January 2007. Available at: http://ams.confex.com/ams/pdfpapers/120864.pdf.

Miyoshi, T. and Aranami, K. 2006. Applying a Four-dimensional local ensemble transform Kalman filter (4-D-LETKF) to the JMA nonhydrostatic model (NHM). *SOLA* **2**, 128–131.

Navascues, B., Lindskog, M., Yang, X. and Amstrup, B. 2006. Diagnosis of error statistics in the HIRLAM 3D-Var. HIRLAM Technical Report # 66. Available at: HIRLAM.org.