

## 4D-Var or Ensemble Kalman Filter?

Eugenia Kalnay<sup>1\*</sup>, Hong Li<sup>1</sup>, Takemasa Miyoshi<sup>2</sup>, Shu-Chih Yang<sup>1</sup>, and Joaquim Ballabrera-Poy<sup>3</sup>

<sup>1</sup>University of Maryland, College Park, MD, 20742-2425

<sup>2</sup> Numerical Prediction Division, Japan Meteorological Agency, Tokyo, Japan

<sup>3</sup> Institut de Ciències del Mar (CSIC), Barcelona, Spain

TELLUS A, in press

### ABSTRACT

We consider the relative advantages of two advanced data assimilation systems, 4D-Var and ensemble Kalman filter (EnKF), currently in use or under consideration for operational implementation. With the Lorenz (1963) model we explore the impact of tuning assimilation parameters such as the assimilation window length and background error covariance in 4D-Var, variance inflation in EnKF, and the effect of model errors and reduced observation coverage. For short assimilation windows EnKF gives more accurate analyses. Both systems reach similar levels of accuracy if long windows are used for 4D-Var. For infrequent observations, when ensemble perturbations grow nonlinearly and become non-Gaussian, 4D-Var attains lower errors than EnKF. If the model is imperfect, the 4D-Var with long windows requires weak constraint. Similar results are obtained with a quasi-geostrophic channel model. EnKF experiments made with the primitive equations SPEEDY model provide comparisons with 3D-Var and guidance on model error and “observation localization”. Results obtained using operational models and both simulated and real observations indicate that currently EnKF is becoming competitive with 4D-Var, and that the experience acquired with each of these methods can be used to improve the other. A table summarizes the pros and cons of the two methods.

\* Corresponding author

ekalnay@atmos.umd.edu

Tel: +1-301-405-5370; Fax: 301-314-9482

3431 CSS building

University of Maryland

College Park, MD 20742-2425

## 1. Introduction

The data assimilation community is at a transition point, with a choice between variational methods, for which there is considerable experience, and ensemble methods, which are relatively new. Many centers use 3D-Var (e.g., Parrish and Derber, 1992), which is an economical and accurate statistical interpolation scheme that does not include the effects of “errors of the day”, although there are several proposed 3D-Var schemes that incorporate some aspects of flow-dependent forecast errors (e.g., Riishojgaard, 1998; Corazza et al., 2002, <http://ams.confex.com/ams/pdfpapers/28755.pdf>; De Pondeca et al, 2006, [www.emc.ncep.noaa.gov/officenotes/newernotes/on452.pdf](http://www.emc.ncep.noaa.gov/officenotes/newernotes/on452.pdf)). Several centers (ECMWF, France, UK, Japan, Canada) have switched to 4D-Var (e.g., Rabier et al., 2000), which requires the development and maintenance of an adjoint model and is computationally much more expensive, but which has proven to be significantly more accurate than 3D-Var in pre-operational tests leading to their implementation. In addition to its demonstrated higher accuracy, 4D-Var was developed and implemented because it allows the assimilation of synoptic data such as satellite radiances at their correct observation time, and because further improvements, such as weak constraint formulations, could be incorporated later. One of the potentially promising extensions to 4D-Var is the use of reduced rank Kalman filters to estimate the analysis error covariance, but tests in a high resolution NWP system showed no significant benefit (Fisher and Hollingsworth, 2004, [ams.confex.com/ams/84Annual/techprogram/paper\\_74522.htm](http://ams.confex.com/ams/84Annual/techprogram/paper_74522.htm)).

Research on Ensemble Kalman Filtering (EnKF) started with Evensen (1994), Evensen and van Leeuwen (1996), Burgers et al. (1998), and Houtekamer and Mitchell (1998). Their methods can be classified as *perturbed observations* (or stochastic) EnKF, and are essentially ensembles of data assimilation systems. A second type of EnKF is a class of *square root* (or deterministic) filters (Anderson, 2001; Whitaker and Hamill, 2002; Bishop et al, 2001; see review of Tippett et al, 2003), which consist of a single analysis based on the ensemble mean, and where the analysis perturbations are obtained from the square root of the Kalman Filter analysis error covariance. Whitaker and Hamill (2002) concluded that square root filters are more accurate than perturbed observation filters because they avoid the additional sampling error introduced by perturbing the observations with random errors, but Lawson and Hansen (2004) suggested that perturbed observations filters could handle nonlinearities better than the square root filters. The three square root filters discussed by Tippett et al. (2003) assimilate observations *serially* (as suggested by Houtekamer and Mitchell, 2001), which increases their efficiency by avoiding the inversion of large matrices. Zupanski (2005) proposed the Maximum Likelihood Ensemble Filter where a 4D-Var cost function with possibly nonlinear observation operators is minimized within the subspace of the ensemble forecasts. A review of EnKF methods i

neighboring grid points. The Kalman Filter equations are solved for each grid point using as basis the singular vectors of the ensemble within the local volume. This method, known as Local Ensemble Kalman Filter (LEKF) allows processing all the observations within the local volume simultaneously, and since the analysis at each grid point is done independently from other grid points, it allows for parallel implementation. Hunt (2005) and Hunt et al. (2007) developed the Local Ensemble Transform Kalman Filter (LETKF) using an approach similar to Bishop et al. (2001) but performed locally as in Ott et al. (2004). Since the LETKF does not require an orthogonal basis, its computational cost is reduced when compared to the original LEKF. In the LETKF localization is based on the selection of observations that are assimilated at each grid point rather than on a local volume, allowing for more flexibility than the LEKF (Fertig et al., 2007b). Keppenne and Rienecker (2002) developed a similar local EnKF for ocean data assimilation.

Hunt et al. (2004) extended EnKF to four dimensions, allowing the assimilation of asynchronous observations, a procedure also suggested by Evensen (2003, section 4.6), and by Lorenc (2003), that becomes particularly efficient in the LETKF formulation. This method (4D-EnKF) expresses an observation as a linear combination of the ensemble perturbations at the time of the observation. The same linear combination of ensemble members can then be used to move the observation forward (or backward) in time to the analysis time. This simple method gives the Ensemble Kalman Filter the ability of 4D-Var to assimilate observations at their right time, but without iterations and allowing the use of future observations when available (e.g., within reanalysis). Although 4D-Var transports the observations in the subspace of the tangent linear model rather than the ensemble subspace, Fertig et al. (2007a) found that 4D-Var and 4D-LETKF yield similar results when the 4D-Var assimilation window is sufficiently long, and when the 4D-LETKF is performed frequently enough.

In the Meteorological Service of Canada, where perturbed observations EnKF was pioneered for the atmosphere, preoperational tests indicated that 4D-Var yielded forecasts clearly superior to those of 3D-Var, whereas EnKF forecasts were only comparable to 3D-Var (Houtekamer et al., 2005). Until recently there was no clear evidence that EnKF could outperform an operational 3D-Var analysis, let alone 4D-Var. However, in the last year there have been a number of encouraging new results. In an intercomparison organized by NCEP, Whitaker et al. (2007) and Szunyogh et al. (2007) showed that the application of the Ensemble Square Root Filter (EnSRF) and the LETKF to the NCEP global forecasting system (GFS) at resolution T62/L28, using all operationally available atmospheric observations (except for satellite radiances), yields better forecasts than the operational 3D-Var using the same data. Houtekamer and Mitchell (2006) tested a number of changes to the configuration that became operational in the current cycle.

The purpose of this paper is to compare some of the advantages and disadvantages of these two methods based on recent experience. In section 2 we discuss experimental results with the very nonlinear Lorenz (1963) model, which although simple, bring up several important aspects of practical optimization. Section 3 contains a discussion of the characteristics of different approaches of EnKF and the experience acquired implementing the Local Ensemble Kalman Filter developed at the University of Maryland on a quasi-geostrophic channel model and a low-resolution primitive equations model using both perfect model and reanalysis “observations”. It also contains a summary of recent results obtained using the LETKF at the University of Maryland and with the Earth Simulator of Japan. In section 4 we discuss questions of efficiency, model error and nonlinearity, and summarize arguments in favor and against the two methods. We conclude by adapting and modifying a table presented by Lorenc (2003) in view of these results.

## 2. Experiments with the Lorenz (1963) model

In this section we compare the performance of 4D-Var, EnKF, and the Extended Kalman Filter (EKF). The EKF allows the Kalman filter to be applied to nonlinear models via the linear tangent and adjoint models. We follow the classic papers of Miller et al. (1994) and Pires et al (1996) and use the Lorenz (1963) model for a simulation of data assimilation. This model has only 3 degrees of freedom, so (unlike experiments with realistic models) it is possible to implement the Ensemble Kalman Filter at full rank, or even with more ensemble members than the size of the model<sup>1</sup>. Its numerical cost allows tuning the 4D-Var background error covariance and window length. First, perfect model experiments are performed observing all variables over a short interval (8 time steps) during which perturbations grow essentially linearly, and over a longer interval (25 time steps) allowing perturbations to develop nonlinearly. Then we compare the impacts of observing only subsets of variables and of model errors. The results illustrate several characteristics of the schemes that in more realistic settings require special attention or tuning.

### 2.1 D-Var: Tuning the window length and background error covariance

In the standard formulation of 4D-Var (e.g., Rabier and Liu, 2003) the analysis is obtained by minimizing a cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{i=0}^N [H_i(\mathbf{x}_i) - \mathbf{y}_i]^T \mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i] = J^b + J^o \quad (1)$$

computed over an assimilation window of length  $t_N - t_0$ , where  $\mathbf{x}_0^b$  is the background or first guess at  $t_0$ ,  $\mathbf{y}_i$  and  $\mathbf{R}_i$  are vector of observations made at time  $t_i$  and its corresponding observation error covariance,  $\mathbf{B}$  is the background error covariance,  $\mathbf{x}_i = M_i(\mathbf{x}_0)$  is the model state at the observation time  $t_i$  obtained by integrating the nonlinear model  $M_i$ , and  $H_i$  is the (nonlinear) observation operator at time  $t_i$  that maps model variables to observation variables. The control variable is the model state vector  $\mathbf{x}_0$  at the beginning of the window  $t_0$ . This is a *strong constraint* minimization in which the analysis valid at  $t_N$  is given by the model forecast  $\mathbf{x}_N = M_N(\mathbf{x}_0)$ .

For the first set of experiments all three model variables are observed, so that  $H$  and its tangent linear operator  $\mathbf{H}$  are equal to the identity matrix  $\mathbf{I}$ . In order to obtain the minimum of  $J$  by iterative methods the most efficient computation of the gradient requires the integration of the adjoint model  $\mathbf{M}_i^T$  (transpose of the linear tangent model

---

<sup>1</sup> Evensen (1997) compared for this model a nonlinear solution of the 4D-Var cost function minimized over a very long assimilation period (many cycles) using a method of gradient descent with a weak constraint, with 1000-member ensembles of Kalman filter and Kalman smoother.

$\mathbf{M}_i$ ). Although  $\mathbf{M}_i$  and  $\mathbf{M}_i^T$  are linear with respect to perturbations, they do depend on the evolving model state (e.g., Kalnay, 2003, p. 213). The simulated observations contain errors in accordance to  $\mathbf{R} = 2\mathbf{I}$  as in Miller et al. (1994). The corresponding observational error standard deviation ( $\sqrt{2}$ ), is an order of magnitude smaller than the natural variability, and all successful data assimilation experiments have RMS errors smaller than this value even with sparse observations. The experiment design consists of a perfect model scenario where the initial conditions are chosen as a random state of the truth run. The performance of each assimilation experiment is measured using the Root Mean Square (RMS) of the difference between the analysis and the true solution. We performed experiments assimilating observations of the three variables every 8 time steps of length 0.01, and every 25 steps as in Miller et al (1994). The shortest decorrelation time scale for the model is about 16 steps, so that during 8 steps perturbations generally evolve linearly and there is little difficulty in obtaining an accurate analysis. 25 steps is an interval long enough to introduce some nonlinear evolution of perturbations and the problems associated with the presence of multiple minima. The corresponding optimal 3D-Var background error covariances obtained iteratively (Yang et al., 2006) have eigenvalues 0.10, 1.11, 1.76 for 8 time steps (a size comparable to that of the observational errors), and are an order of magnitude larger for 25 steps (0.67, 9.59, 14.10).

In order to be competitive with a full rank EnKF, 4D-Var requires longer assimilation windows, but this risks introducing the problem of multiple minima (Pires et al, 1996). As shown in Table 1, with observations every 8 steps, lengthening the window of assimilation reduces the RMS analysis error (at the expense of increased computational cost) up to 32 steps. Beyond that, errors become larger because of the problem of multiple minima (Miller et al. 1994). This problem can be overcome with the Quasi-static Variational Assimilation (QVA) approach proposed by Pires et al. (1996), where short windows are used initially and progressively increased to the maximum, while performing quasi-static adjustments of the minimizing solution.

Table 1 compares the performance of the assimilation as a function of the assimilation window length (either fixing the window, or with the QVA approach). With 8 steps the improvement of the analysis RMS errors with window length extends up to a window of about 48 steps. For observations every 25 steps, the optimal window length is between 50 and 125 steps, but the results are strongly sensitive to the choice of background error covariance. In operations, if a new analysis is needed every 6 hours, and if the optimal window length is a few days, the computation of overlapping windows is required (Fisher et al., 2005).

The minimization of (1) requires an estimation of the background error covariance  $\mathbf{B}$ . If 4D-Var is started from a 3D-Var analysis, the use of  $\mathbf{B}=\mathbf{B}_{3DV}$  (the optimal 3D-Var background error covariance) is a reasonable and widely used estimation. However, if 4D-Var is cycled, the background  $\mathbf{x}_0^b$  is provided by the 4D-Var analysis  $\mathbf{x}_N$  of the previous cycle, and after a transient of a few cycles, this background should be more accurate than the 3D-Var forecast. Therefore,  $\mathbf{B}$  should be significantly smaller than  $\mathbf{B}_{3DV}$ . In our experiments we obtained  $\mathbf{B}_{3DV}$  by iteratively optimizing the 3D-Var data

assimilation as in Yang et al. (2006). We also computed the background error covariance from the 4D-Var analysis errors (which cannot be done in practice) and found that its structure was very similar to that of  $\mathbf{B}_{3DV}$  with a magnitude about 20 times smaller for the case of 25 steps, suggesting that the use in 4D-Var of a covariance proportional to  $\mathbf{B}_{3DV}$  and tuning its amplitude is a good strategy to estimate  $\mathbf{B}$ . Table 2 shows that tuning the amplitude of the background error covariance has a large impact on the results. Similar results were obtained with the more complex quasi-geostrophic model of Rotunno and Bao (1996) (section 3.1). The ratio of the size of  $\mathbf{B}_{3DV}$  to the optimal  $\mathbf{B}_{4DV}$  is larger than would be expected in a more realistic model because in a perfect model scenario the improvement over 3D-Var obtained using 4D-Var with optimal parameters is larger than in the presence of model errors (see sections 2.3 and 3.2).

## 2.2 Kalman Filter: Formulation and tuning the covariance inflation

4D-Var is next compared with both Extended and Ensemble Kalman Filter, which are briefly described here. The Extended Kalman Filter (EKF, e.g., Ide et al, 1997) consists of a forecast step,

$$\mathbf{x}_n^b = M_n(\mathbf{x}_{n-1}^a) \quad (2a)$$

$$\mathbf{B}_n = \mathbf{M}_n \mathbf{A}_{n-1} \mathbf{M}_n^T + \mathbf{Q}_n \quad (2b)$$

and an analysis step,

$$\mathbf{x}_n^a = \mathbf{x}_n^b + \mathbf{K}_n (\mathbf{y}_n - H \mathbf{x}_n^b) \quad (3a)$$

$$\mathbf{A}_n = (\mathbf{I} - \mathbf{K}_n \mathbf{H}) \mathbf{B}_n \quad (3b)$$

where  $\mathbf{K}_n$  is the Kalman gain matrix given by two equivalent formulations,

$$\mathbf{K} = \mathbf{B} \mathbf{H} (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H})^{-1} = (\mathbf{B}^{-1} + \mathbf{H} \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H} \mathbf{R}^{-1}, \quad (4)$$

and  $\mathbf{A}_n$  is the new analysis error covariance. Here  $n$  denotes the analysis step (in our case the analysis is done at the observation time, either 8 or 25 model time steps),  $M_n$  is the nonlinear model that provides the forecast or background  $\mathbf{x}_n^b$  at step  $n$  starting from the previous analysis  $\mathbf{x}_{n-1}^a$ ,  $\mathbf{M}_n$  and  $\mathbf{M}_n^T$  are the linear tangent and adjoint models,  $\mathbf{B}_n$  is the background error covariance at the time of the analysis, and  $\mathbf{Q}_n$  is the covariance of the model errors (assumed to be zero here). As in the 4D-Var experiments, we used  $\mathbf{H}_n = \mathbf{I}$  and  $\mathbf{R}_n = 2\mathbf{I}$ .

The Ensemble Kalman Filter (EnKF) is similar to EKF, the main difference being that an ensemble of  $K$  forecasts

$$\mathbf{x}_{n,k}^b = M_n(\mathbf{x}_{n-1,k}^a), k = 1 \dots K \quad (5a)$$

is carried out in order to estimate the background error covariance  $\mathbf{B}_n$ . Defining the

forecast ensemble mean as  $\bar{\mathbf{x}}_n^b = \frac{1}{K} \sum_{k=1}^K \mathbf{x}_{n,k}^b$ , and  $\mathbf{X}_n^b$  as the  $M \times K$  matrix whose columns are

the  $K$  ensemble perturbations  $\mathbf{x}_{n,k}^b - \bar{\mathbf{x}}_n^b$ , and  $M$  is the dimension of the state vector, then

$$\mathbf{B}_n = \frac{1}{K-1} \mathbf{X}_n^b \mathbf{X}_n^{bT}. \quad (5b)$$

In the analysis step we used the second formulation in (4) and solved

$$[\mathbf{I} + \mathbf{B}_n \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}] (\bar{\mathbf{x}}_n^a - \bar{\mathbf{x}}_n^b) = \mathbf{B}_n \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}_n - H \bar{\mathbf{x}}_n^b) \quad (6a)$$

iteratively to obtain  $\bar{\mathbf{x}}_n^a$ , the new analysis ensemble mean.

Alternatively, the analysis mean can be obtained from (3a) with

$$\mathbf{K}_n = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}_n^{-1})^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{X}_n^b [(\mathbf{H} \mathbf{X}_n^b)^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{X}_n^b) + (K-1) \mathbf{I}]^{-1} (\mathbf{H} \mathbf{X}_n^b)^T \mathbf{R}^{-1} \quad (6b)$$

where the matrix inversion is performed in the  $K \times K$  space of the ensemble perturbations.

We used the Ensemble Transform Kalman Filter (ETKF) approach to obtain the analysis perturbations (Bishop et al., 2001; Hunt, 2005; Hunt et al., 2007):

$$\mathbf{A}_n = \frac{1}{K-1} \mathbf{X}_n^a \mathbf{X}_n^{aT} = \mathbf{X}_n^b \hat{\mathbf{A}}_n \mathbf{X}_n^{bT}$$

where  $\hat{\mathbf{A}}_n = [(\mathbf{H} \mathbf{X}_n^b)^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{X}_n^b)]^{-1}$ , a  $K \times K$  matrix, represents the new analysis covariance in the  $K$  space of the ensemble forecasts. The new analysis perturbations are then obtained from the columns of

$$\mathbf{X}_n^a = \mathbf{X}_n^b [(K-1) \hat{\mathbf{A}}_n]^{1/2}. \quad (7)$$

Since  $\mathbf{H} \mathbf{x}_{n,k}^b \simeq H \mathbf{x}_{n,k}^b - H \bar{\mathbf{x}}_n^b \simeq H \mathbf{x}_{n,k}^b - \overline{H \mathbf{x}_n^b}$  and  $\mathbf{H}$  always appears multiplying a perturbation vector or matrix, in EnKF it is possible to use the full nonlinear observation operator, without the need for its Jacobian or adjoint (e.g., Houtekamer and Mitchell, 2001; Evensen, 2003; Lorenc, 2003).

Figure 1 is a schematic representing the characteristics of an EnKF with  $K$  ensemble forecasts (black circles,  $K=2$  in the schematic). Although the forecasts states have the model dimension  $M$ , the ensemble lies on an attractor of much lower dimension (dotted curve). The ensemble of  $K \ll M$  forecasts defines the subspace spanned by the background error covariance (solid line) of dimension  $K-1$  (denoted “error subspace” by Nerger et al., 2005), which is an approximation of the model attractor. Only the projection of the observations (crosses) onto the error subspace is assimilated in EnKF. The analysis ensemble members (white circles) are also defined within the error subspace, i.e., they are linear combinations of the ensemble forecasts obtained using the Kalman filter equations in the error subspace. The analysis ensemble mean is the best estimate of the analysis, and its spread is the best estimate of the analysis error, so that for a given ensemble size, the analysis ensemble provides the optimal initial conditions for the next ensemble of forecasts (Ehrendorfer and Tribbia, 1997).



Because of nonlinearities, even with a perfect model ( $\mathbf{Q}_n = 0$ ) both EKF and EnKF analyses can drift away from the real solution due to an underestimation of the true forecast error covariance. In addition, the EnKF is also affected by the lack of representation of the background error covariance outside the subspace defined by the ensemble forecasts (Figure 1). Miller *et al.* (1994) suggested a Monte Carlo approach of adding perturbations to avoid the underestimation of the forecast errors in the EKF. Similarly, Corazza *et al.* (2002) found that “refreshing” bred vectors by adding to them random perturbations after the analysis solved the related problem that bred vectors tend to collapse into a too small subspace (Wang and Bishop, 2003), and improved the performance of bred vectors in estimating the “errors of the day”. Yang *et al.* (2006) tested two approaches to avoid “filter divergence” in the EKF. The first one is the multiplicative variance inflation suggested by Anderson (2001), in which the background error covariance is multiplied by  $(1 + \Delta)$ , and the second method is to enhance the analysis error covariance matrix by adding to the diagonal elements random perturbations uniformly distributed between 0 and 1, multiplied by a coefficient  $\mu$  before performing the time integration. With the Lorenz model it was found for EKF multiplicative inflation alone did not converge. Here, for the EnKF, we multiplied the background ensemble perturbation by  $(1+\delta)$ , equivalent to multiplying the background error covariance by  $(1+\delta)^2$ .

In the 4D-Var and EKF formulations for the Lorenz (1963) model, a seemingly minor approximation in the adjoint model of keeping the nonlinear model trajectory constant within a single time step (without updating it at every substep of the Runge-Kutta time scheme) resulted in a substantial deterioration of about 50% in the analysis errors (Yang *et al.*, 2006). The Runge-Kutta time scheme, which requires 4 estimations of the time derivative per time step, is too expensive to be used in operational applications, so other time schemes such as leap-frog are used, and this particular problem does not appear. However, it is common to make even stronger approximations of the adjoint by either keeping the trajectory constant or interpolating it in time within the adjoint integration. The results with the Lorenz model suggest that any approximation to the exact adjoint can significantly increase the 4D-Var analysis errors. Since the EnKF method uses nonlinear model integrations, it is not affected by this problem, although it still requires the use of variance inflation (Anderson, 2001).

Table 3 shows that if the number of ensemble members  $K$  is larger than  $M$ , the size of the model, EnKF becomes more accurate than EKF, but for realistically large models we always have  $K \ll M$ . With a long interval between observations (25 steps) there were short episodes of large analysis errors, so that we found useful to perform a “sanity check” (Miller *et al.*, 1994). Whenever  $\|\mathbf{y} - H\bar{\mathbf{x}}_b\|$  was found to be greater than 5, indicating that forecast errors were growing faster than suggested by EnKF, the background error covariance was taken as the average of the EnKF estimate and  $\mathbf{B}_{3D-Var}$ . This simple “hybrid” approach, which increases the size of the background error covariance beyond its EnKF Gaussian estimation when the observational increments are unexpectedly large, had a significant positive impact.

For short observation intervals (8 steps) EnKF and 4D-Var with long windows and optimal  $\mathbf{B}$  give similar optimal results. However, it is striking that with longer intervals (25 steps), the optimal 4D-Var yields significantly lower errors than those that could be obtained with a full rank EnKF, even using a hybrid system. This is because the KF assumption that forecast perturbations are Gaussian becomes inaccurate when they grow nonlinearly (see discussion in section 4.2). Although for linear perfect models KF and 4D-Var solve the same problem, for nonlinear models the EnKF analysis is still constrained to the error subspace (Figure 1), whereas 4D-Var finds iteratively the initial condition for a nonlinear forecast that best fits the observations. This advantage of strong-constraint 4D-Var, which is present even without nonlinear observation operators, disappears in the presence of model errors (section 2.3).

So far we presented tests observing all variables. Table 4 shows results obtained when the observation coverage is reduced by observing only one or two variables. The impact of reduced observations is similar in the 4DVar with optimal window length and in the EnKF, but generally 4D-Var is worse than EnKF even with optimal window length.

### 2.3 Imperfect model experiments

Handling model errors in data assimilation is a subject of considerable current research (e.g., Dee and DaSilva, 1998; Dee and Todling, 2000; Tremolet, 2005; Andersson et al., 2005; Keppenne et al., 2005; Baek et al., 2006; Danforth et al., 2007; Whitaker et al., 2007a). In the final set of experiments with the Lorenz (1963) model, we allowed for an imperfect forecast model, by reducing the parameter  $r$  from 28 in the nature run used to create the observations, to  $r=26$  in the forecast model. This increases the forecast errors by an order of magnitude, and the optimal 3D-Var background error covariance is two orders of magnitude larger.

In order to account for model errors, in the 4D-Var experiments we used a weak constraint approach as in Tremolet (2005), modifying the cost function (1) by including a model bias  $\beta$  assumed to be constant within the assimilation window:

$$J(\mathbf{x}_0, \beta) = \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{i=0}^N [H_i(\mathbf{x}_i + \beta) - \mathbf{y}_i]^T \mathbf{R}_i^{-1} [H_i(\mathbf{x}_i + \beta) - \mathbf{y}_i] \quad (8)$$

$$+ \frac{1}{2} \beta^T \mathbf{Q}^{-1} \beta = J^b + J^o + J^\beta$$

The bias  $\beta$  shifts the trajectory so that  $H(\mathbf{x}_i + \beta)$  best fits the observations within each assimilation window. We tuned the amplitude of  $\mathbf{Q}$ , the bias error covariance, by making it proportional to the background error covariance, although this may not be optimal (Tremolet, 2005).

Several different approaches have been suggested to deal with model errors within EnKF. The simplest is to increase the multiplicative inflation (Anderson, 2001), which reduces the weight given to the imperfect model compared to the observations. Additive inflation

was found to be more effective than multiplicative inflation by Whitaker et al. (2007a). Baek et al. (2006) showed how to correct constant model error by augmenting the model state with the model bias, and Danforth et al. (2007) proposed a low order approach to estimate bias, diurnal and seasonal errors, and state dependent model errors. Li et al. (2007) compared these methods on the SPEEDY model. For simplicity, in the experiments presented here we only tested increasing the EnKF multiplicative inflation discussed in section 1.2 beyond the value needed in a perfect model.

Table 5 (with observations every 8 steps) shows that the RMS analysis error of the EnKF in the presence of model errors becomes quite large (1.19), but that increasing inflation, although not optimal, reduces the errors substantially (to 0.81). With model errors, strong constraint 4D-Var becomes less sensitive to the background error covariance and increasing the window only reduces the error very slightly, from 0.84 to 0.83, confirming that model errors strongly limit the length of the assimilation window of the 4D-Var (Schröter et al., 1993). The introduction of a model bias in the cost function as in (8) has little effect for short windows but improves substantially the results for long assimilation windows, reducing the RMS error to 0.71 when  $\mathbf{Q}$  is optimally tuned. Note that with the long assimilation window, more observations are available to better define the unbiased increment  $y - h(x_i + \beta)$ , and the 4D-Var results show the benefits after the model trajectory is corrected with  $\beta$ . In EnKF there is only one set of observation to estimate the best analysis state, and multiplicative inflation is not optimal to deal with a state-dependent bias like this. Despite its simplicity, the inflation method results are an improvement, but a more sophisticated strategy that accounts for model error based on recent training would be expected to further improve the results.

Table 6 (with observations every 25 steps) indicates that for infrequent observations, EnKF with inflation alone gives RMS errors similar to those of strong constraint 4D-Var, about 1.04. Increasing the number of ensemble members in EnKF or using weak constraint with 4D-Var improves their results to similar levels, and there is no advantage to windows longer than 50.

In summary, the three methods are able to reach similar levels of accuracy for the very nonlinear Lorenz (1963) system, but each of them requires considerable tuning and empirical techniques, such as “sanity checks” and inflation for the Kalman Filter, and optimization of  $\mathbf{B}$  and long windows with the QVA approach for 4D-Var, without which the results are significantly worse. In the presence of model errors, weak constraint for 4D-Var is somewhat more effective than simple multiplicative inflation in EnKF. We found that EnKF was the easiest method to implement even for the Lorenz (1963) model, because EKF and 4D-Var required the computation of the linear tangent and adjoint models using the nonlinear model corresponding to each substep of the Runge-Kutta time integration, and because 4D-Var required the estimation of the background error covariance.

### 3 Comparisons of EnKF, 3D-Var and 4D-Var in QG and PE models

### 3.1 Quasi-geostrophic channel model

We now present comparisons of 3D-Var (developed by Morss et al. 2000), 4D-Var and LETKF, using the Rotunno and Bao (1996) quasi-geostrophic channel model in a perfect model set-up. We found that for this model multiplicative inflation alone is not enough to prevent filter divergence. Instead, as in Corazza et al. (2002; 2007), random perturbations are added after the analysis with a standard deviation of 5% of the natural variability of the model. For optimal 4D-Var results, the background error covariance  $\mathbf{B}$  had to be tuned as discussed in the previous section, with an optimal value of about  $0.02\mathbf{B}_{3DV}$ .

With simulated rawinsonde observations every 12 hrs, Figure 2 shows that the LETKF is more accurate than 4D-Var with 12 hr windows, and comparable with 4D-Var with a 24 hr window. The advantage that 4D-Var has for longer windows (48 hr) is analogous to that observed with the 3-variable perfect Lorenz model. As with the Lorenz model, seemingly minor approximations of the adjoint model were found to result in a significant deterioration of the 4D-Var results. As noted before with the Lorenz model, and in section 3.2, the difference in performance between 3D-Var and the more advanced methods, is much larger in perfect model simulations than in real operational applications, where the presence of model error is more important than effects such as parameter optimization and the accuracy of the adjoint.

These computations were performed on a single processor Alpha EV6/7 617MHz computer, with the following wall clock timings for 200 days of simulated data assimilation: 3D-Var – 0.5 hours, LETKF - 3 hours, 4D-Var (12 hr window) – 8 hours, 4D-Var (48hr window) – 11.2 hours. These estimates of the computational requirements are not representative of the relative computational costs that would be attainable in an operational set up with optimized parallelization, since the LETKF is designed to be particularly efficient in massively parallel computers (Hunt et al., 2007).

### 3.2 Global Primitive Equations model (SPEEDY)

In this subsection we discuss several results obtained by Miyoshi (2005) who developed and tested 3D-Var, the EnSRF approach of Whitaker and Hamill (2002), and the LEKF of Ott et al. 2004 using the SPEEDY global primitive equations model of Molteni (2003). The SPEEDY model has a horizontal spectral resolution of T30 with 7 vertical levels and a simple but fairly complete parameterization of physical processes, which result in a realistic simulation of the atmospheric climatology. Miyoshi, 2005 first performed “perfect model” simulations, and then created realistic atmospheric “soundings” by sampling the NCEP-NCAR Reanalysis (Kistler et al., 2001).

#### 3.2.1 Observation localization

It is well known (e.g., Lorenc, 2003), that the main disadvantage of EnKF is that the use of a limited ( $K \sim 100$ ) number of ensemble members inevitably introduces sampling errors in the background error covariance  $\mathbf{B}$ , especially at long distances. Several Ensemble Kalman Filters now incorporate a *localization* of the background error covariance

(Houtekamer and Mitchell, 2001) to handle this problem by multiplying each term in  $\mathbf{B}$  by a Gaussian shaped correlation that depends on the distance between points (a Schur product suggested by Gaspari and Cohn, 1999). This approach, easy to implement on systems that assimilate observations *serially* (Tippett et al., 2003) has been found to improve the performance of the assimilation. In the LEKF/LETKF spurious long distance correlation of errors due to sampling are avoided by the use of a region of influence (a local volume), beyond which the correlations are assumed to be zero, so that the localization has a top-hat and not a Gaussian shape. Miyoshi (2005) found that the LEKF performed slightly worse than EnSRF using a Gaussian localization of  $\mathbf{B}$ . Since in the LEKF it is not possible to efficiently localize  $\mathbf{B}$  with a Schur product, Miyoshi, following a suggestion of Hunt (2005), multiplied instead *the observation error covariance  $\mathbf{R}$  by the inverse of the same Gaussian*, thus increasing the observational error of observations far away from the grid point. This has an effect similar to the localization of  $\mathbf{B}$  and it was found that with this “observation localization” the performance of the LEKF became as good as that of the EnSRF. Whitaker et al. (2007) obtained similar results when comparing EnSRF and LETKF with observation localization.

### 3.2.2 Model errors

With identical twin experiments (observations derived from a “nature” run made with the same model as the forecasts), Miyoshi (2005) obtained EnKF RMS analysis errors much lower than those obtained with an optimized 3D-Var. However, when using atmospheric “soundings” derived from the NCEP Reanalysis, the advantage of EnKF with respect to 3D-Var became considerably smaller (Figure 3, full lines). An attempt to apply the method of Dee and da Silva (1998) at full resolution to correct for model bias led to filter divergence because of sampling problems. Following Danforth et al. (2007), 6 hr model errors including the time average and leading EOFs representing the errors associated with the diurnal cycle were first estimated. The Dee and da Silva (1998) method was then used to estimate the time evolving *amplitude* of these error fields, thus reducing by many orders of magnitude the sampling problem.

Figure 3 compares the geopotential height analysis errors for both 3D-Var and EnKF, with (dashed lines) and without (full lines) bias correction. Without bias correction, the EnKF (open circles) is only marginally better than 3D-Var (closed circles). With bias correction the EnKF (open squares) has a substantial advantage over 3D-Var (closed squares). Estimating the amplitude of the bias correction within EnKF (full line with triangles) does not improve the results further. Miyoshi (2005) also found that assimilating Reanalysis moisture soundings improved not only the moisture, but the winds and temperature analysis as well, whereas humidity information is generally known to have a low impact on the analysis of other large-scale variables in 3D-Var (Bengtsson and Hodges, 2005). This indicates that EnKF provided “tracer” information to the analysis, since perfect model simulations confirmed that the improvement observed in winds and temperature analysis when assimilating humidity was not due to a reduction of the model bias.

### *3.3 Recent LETKF experiments with global operational models*

Szunyogh et al. (2005) implemented the LEKF on the NCEP Global Forecasting System (GFS) model at T62/L28 resolution in a perfect model set-up, with very accurate analyses. It was found that, even with a very small number of observations, the LEKF was able to accurately analyze a gravity wave present in the nature run. This suggests that the localization of the LEKF/LETKF is apparently able to maintain well the atmospheric balance. Szunyogh et al. (2007) have tested the assimilation of real observations with encouraging results. Liu et al. (2007) coupled the LETKF with 40 ensemble members on a version of the NASA finite volume GCM (fvGCM) and compared the results of assimilating simulated rawinsonde observations with those obtained using the NASA operational PSAS, a form of 3D-Var computed in observational space (Cohn et al., 1998). Figure 4, adapted from their study, shows that the globally averaged LETKF analysis errors were about 30-50% smaller than those of PSAS.

Miyoshi and Yamane (2007) implemented the LETKF on the Atmospheric GCM for the Earth Simulator (AFES) in Japan with a resolution of T159/L48. Figure 5, adapted from their study for a perfect model scenario, shows that after a single LETKF analysis step, the analysis error decreases with increasing ensemble size, (and saturates at about 320 members, not shown). After 10 days of data assimilation with LETKF, the saturation with respect to the number of ensemble members is faster, taking place at about 80 members, and there is less dependence on the number of ensemble members than on whether variance inflation is used or not. Figure 6 presents one month of global RMS analysis errors for the surface pressure, and confirms that with 80 ensemble members the error is close to saturation. Ratios of the analysis error and the analysis spread (not shown) were very close to 1, as were the corresponding ratios for the forecasts, suggesting that the system is behaving as expected. Tests comparing the current JMA 4D-Var with a 100 members LETKF ensemble assimilating the operational observations for August 2004 indicated their results were indistinguishable in the NH and 4D-Var was slightly better in the SH.

At the time of this writing, comparisons have been carried out with the NCEP GFS model at T62/L28, using NCEP operational 3D-Var (Spectral Statistical Interpolation, SSI, Parrish and Derber, 1992), EnSRF and LETKF, assimilating all operational non-radiance observations. Results (Whitaker et al, 2007, Szunyogh et al., 2007) indicate that the EnKF are similar to each other and superior to the SSI. With the assimilation of radiances LETKF was still superior to SSI but the gap between the two methods was reduced (Whitaker, pers. comm., 2007). Tests comparing the operational 4D-Var and an experimental LETKF have shown the two methods giving similar results in the NH and tropics and 4D-Var having an advantage in the SH.

## **4 Summary and Discussion**

As indicated in the introduction, EnKF is a relatively young area of research in data assimilation, and until recently there was no clear evidence that it could outperform an operational 3D-Var analysis, let alone 4D-Var. Whitaker et al. (2007) reporting on a

comparison between 3D-Var and EnKF organized by NCEP have recently shown for the first time that at the resolution of T62/L28 and using the observations operationally available at NCEP (except

Because of this, recently there were no examples of EnKF outperforming 3D-Var when using observations. It has been generally assumed that EnKF is much more sensitive than 4D-Var to the problem of model errors. However, recent results from Miyoshi and Whitaker et al. (2007a) suggest that when the model errors are addressed with suitable approaches, the advantages of EnKF with respect to 3D-Var become apparent. The results agree with those in section 2, where we found that in the presence of model errors, EnKF with strong inflation gave analysis errors similar to those obtained with strong constraint 4D-Var with a short window, but that weak constraint 4D-Var reduced these errors. Whitaker et al. (2007) found that additive inflation performed better than multiplicative inflation in their system, so that their EnSRF with the Thompson et al. (2004) EP GFS yielded better scores in the northern and southern extratropics as well as the tropics. Houtekamer (2006, pers. comm.) found that adding random perturbations to the model errors to the analysis, so that these perturbations evolve naturally during the 6-hour forecast, improved their results. With a QG model, they also found that additive inflation after the analysis is better than multiplicative inflation (section 3.1). Inflation through random perturbations added before the model integration is apparently advantageous because it allows the ensemble to explore regions of phase space that lie outside the analysis subspace and thus to overcome the tendency of the unforced ensemble to collapse towards the dominant unstable direction. Randomly increasing the ensemble. Multiplicative inflation, by contrast, does not change the ensemble spread. The low-order approach of Danforth et al. (2007) may also be advantageous to correct not only biases but also state-dependent model errors.

Current research is also showing considerable interest in the development of 4D-Var with *weak model errors*, i.e., allowing for model errors. Tremolet (2005) obtained very encouraging preliminary results, and further comparisons between 4D-Var and EnKF will be required when these constraint systems are implemented.

Although EnKF does not require linearization of the model, it (e) Tj 50 0 0 50e



forecasts. This avoids the linearization of the observation operators while making the minimization problem better conditioned than in 4D-Var so that typically only 2-3 iterations are needed.

### *4.3 Relative advantages and disadvantages*

In summary, the main advantages of EnKF are that: i) it is simple to implement and model independent; ii) it automatically filters out, through nonlinear saturation, fast processes such as convection that, if exactly linearized, would dominate error growth; and iii) it provides optimal initial ensemble perturbations, for a given ensemble size, because they represent the analysis error covariance (Ehrendorfer and Tribbia, 1997). Since the number of ensemble members required is similar to that used for ensemble prediction, the cost of the 6-hour ensemble integration is covered by its use in longer ensemble predictions. The main disadvantage of EnKF is the low dimensionality of the ensemble, which introduces sampling errors in the estimation of the background error covariance. The background covariance localization solves this problem in EnKF when observations are assimilated serially. Although this is not computationally feasible in local EnKF methods able to handle many observations simultaneously like the LETKF, an equivalent approach is a simple “observation localization” where the observations error covariance is increased for observations far from the grid point being analyzed (Miyoshi, 2005; Hunt et al., 2007).

The main advantages of 4D-Var are its ability to assimilate asynchronous observations, (like 4DEnKF) and the fact that, when using a long enough window, its performance converges to that of full rank EKF. Fisher et al. (2005) has shown that long-window 4D-Var with model error is equivalent to a Kalman smoother where the initial guesses for the state and forecast error covariance are “forgotten”. He argues that the approach of long overlapping windows and weak constraint should be advantageous compared to EnKF, because the latter is severely rank-reduced. An additional advantage of 4D-Var is that it allows for the assimilation of observations with serially-correlated errors by including such time correlations in  $\mathbf{R}$ . Järvinen et al. (1999) showed that this correlation results in less weight given to the mean of the observations (if biased) and more weight to their time evolution. However, this time correlation can also be included within the 4DEnKF formulation. The main disadvantage of 4D-Var is that it requires the development and maintenance of the adjoint model, including special consideration of how to represent fast processes (such as convection) that become nonlinear and quickly saturate in the nonlinear model. Both methods require additional development of advanced approaches to the treatment of model errors, including weak constraint for 4D-Var and efficient estimates of state-dependent bias in EnKF. Like 4D-Var, EnKF has a few tuning “handles” that need to be explored, including the number of ensemble members, the strength and characteristics of the covariance localization, the handling of model errors, the use of multiplicative or additive inflation, and its adaptive estimation using observational increments. In the next few years more experiments with real observations will build up the EnKF experience needed for operational implementation. Fortunately, because the problems solved in EnKF and 4D-Var are very closely related, researchers can take advantage and share advances made in either method.

We conclude the paper by adapting and updating in Table 7 a very useful table of advantages and disadvantages from Lorenc (2003), with additions [in brackets] and comments (*in italics*).

## 5 Acknowledgments

We are very grateful to our colleagues from the Chaos/Weather group at the University of Maryland, including Profs. B. Hunt, E. Kostelich, E. Ott, I. Szunyogh, J. Yorke, and current and former students, including M. Corazza, C. Danforth, E. Fertig, J. Harlim, J. Liu, D.J. Patil, who had a strong influence on this research. B. Hunt suggested the observation localization method, and F. Molteni provided the SPEEDY model. The review and suggestions of Olivier Talagrand significantly improved the original draft of the paper, as well as comments from Herschel Mitchell. We have incorporated detailed corrections suggested by Nils Gustafsson, although we respond separately to his very stimulating and valid comments. The careful and detailed reviews of Erik Andersson and an anonymous reviewer significantly improved the paper.

This research was partially supported by NSF ATM9328402, NASA NNG04GK29G and NNG04GK78A and NOAA NA04OR4310103.

## 6 References

- Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.
- Andersson, E., M. Fisher, E. Holm, L. Isaksen, G. Radnoti and Y. Tremolet, 2005: Will the 4D-Var approach be defeated by nonlinearity? ECMWF Tech Memo 479. Available from [www.ecmwf.int/publications](http://www.ecmwf.int/publications).
- Bengtsson, L. and Hodges, K. I. 2005. On the impact of humidity observations in numerical weather prediction. *Tellus* 57A, 701-708.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme in the Ensemble Kalman Filter. *Mon. Wea. Rev.*, **126**, 1719-1724.
- Cohn, S., Da Silva, A., Guo, J., Sienkiewicz, M., and Lamich, D., 1998: Assessing the Effects of Data Selection with the DAO Physical-Space Statistical Analysis System. *Mon. Wea. Rev.*, **126**, 2913-2926.
- Corazza, M., E. Kalnay, D.J.Patil, E. Ott, J.A.Yorke, B.R. Hunt, I Szunyogh and M. Cai, 2002: Use of the breeding technique in the estimation of the background error covariance matrix for a quasi-geostrophic model. Paper 6.4 in the AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction, Orlando, FA, January 14-17 2002. Available from [ams.confex.com/ams/pdfpapers/28755.pdf](http://ams.confex.com/ams/pdfpapers/28755.pdf).
- Corazza, M., E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt, and J. A. Yorke, 2003: Use of the breeding technique to estimate the structure of the analysis "error of the day". *Nonlinear Processes in Geophysics*, **10**, 233-243.
- Corazza, M., E. Kalnay and S.-C. Yang, 2007: An implementation of the Local Ensemble Kalman Filter in a quasigeostrophic model and comparison with 3D-Var. *Nonlinear Processes in Physics*, **14**, 89-101.
- Danforth, C. M., E. Kalnay, and T. Miyoshi: Estimating and Correcting Global Weather Model Error," *Monthly Weather Review*, **134**, 281-299.
- Davolio, S. and A. Buzzi. 2004: A Nudging Scheme for the Assimilation of Precipitation Data into a Mesoscale Model. *Weather and Forecasting*: Vol. 19, No. 5, pp. 855–871.
- Dee, D. P. and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Raoy. Meteor. Soc.*, **126**, 269-295.
- Dee, D.P. and R. Todling, 2000: Data assimilation in the presence of forecast bias: the GEOS moisture analysis.*Mon. Wea. Rev.*, **128**, 3268-3282.
- De Pondeca, M. S. F. V., Purser, R. J., Parrish, D. F., and Derber, J. C., 2006: Comparison of strategies for the specification of anisotropies in the covariances of a three-dimensional atmospheric data assimilation system, NOAA/NCEP Office Note 452, 13 pp, 2006. Available from [www.emc.ncep.noaa.gov/officenotes/newernotes/on452.pdf](http://www.emc.ncep.noaa.gov/officenotes/newernotes/on452.pdf).

- Derber, J., R. Puser, W-S. Wu, R. Treadon, M. Pondeva, D. Parrish, and D. Kleist, 2003: Flow-dependent Jb in grid-point 3D-Var. Available from [http://www.ecmwf.int/publications/library/ecpublications/\\_pdf/seminar/2003/sem2003\\_derber.pdf](http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2003/sem2003_derber.pdf)
- Ehrendorfer, M. and J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, **54**, 286-313.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10143-10162.
- Evensen, G., and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Aghulas current using the ensemble Kalman Filter with a quasi-geostrophic model. *Mon. Wea. Rev.*, **124**, 85-96.
- Evensen, G., 1997: Advanced data assimilation for strongly nonlinear dynamics. *Mon. Wea. Review*, **125**, 1342-1354.
- Evensen, G., 2003: The ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocen Dyn.*, **53**, 343-367.
- Fertig, E., J. Harlim and B. Hunt, 2007a: A comparative study of 4D-Var and 4D Ensemble Kalman Filter: Perfect Model Simulations with Lorenz-96. *Tellus*, **59**, 96-101.
- Fertig, E., B. Hunt, E. Ott and I. Szunyogh, 2007b: Assimilating nonlocal observations with a Local Ensemble Kalman Filter. *Tellus*, accepted.
- Fisher, M., and A. Hollingsworth, 2004: Evaluation of reduced rank Kalman filters. Paper J1.10 of AMS 16<sup>th</sup> Conference on Numerical Weather Prediction, 84<sup>th</sup> AMS Annual Meeting, Seattle, Wash, January 12-15 2004. Available from [ams.confex.com/ams/84Annual/techprogram/paper\\_74522.htm](http://ams.confex.com/ams/84Annual/techprogram/paper_74522.htm)
- Fisher, M., M. Leutbecher and G. Kelly, 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Q. J.R. Meteorol. Soc.*, 3235-3246.
- Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723-757.
- Houtekamer, P. L. and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Wea. Rev.*, **126**, 796-811.
- Houtekamer, P. L. and Herschel L. Mitchell. 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*: Vol. 129, pp. 123-137.
- Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Mon. Wea. Rev.*, **133**, 604-620.
- Houtekamer, P. L. and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Q. J. Roy. Met. Soc.*, **131**, 3269-3290.
- Hunt, B. R., 2005: An efficient implementation of the Local Ensemble Kalman Filter. <http://arxiv.org/abs/physics/0511236>.
- Hunt, B. R., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J.A. Yorke and A. V. Zimin, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273-277.
- Hunt, B. R., E. J. Kostelich and I. Szunyogh, 2007: Efficient data assimilation for

- spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. *Physica D*, in press.
- Ide, K., P. Courtier, M. Ghil and A. Lorenc, 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181-189.
- Järvinen, H., E. Andersson and F. Bouttier, 1999: Variational assimilation of time sequences of surface observations with serially correlated errors, *Tellus*, **51A**, 469-488.
- Kalnay, E., 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge, 341 pp.
- Keppenne, C. and Rienecker, H. 2002. Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.* **130**, 2951– 2965.
- Keppenne, C., M. Rienecker, N. Kurkowski and D. Adamec, 2005: Ensemble Kalman Filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlinear Processes in Geophysics*, **12**, 491-503.
- Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bull. Amer. Meteor. Soc.*, **82**(2), 247-267.
- Lawson, W. G. and J. A. Hansen, 2004: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.*, **132**, 1966-1981.
- Li, Hong, E. Kalnay, T. Miyoshi and C. Danforth, 2007: Ensemble Kalman Filter in the presence of model errors. Available from <http://ams.confex.com/ams/pdfpapers/120175.pdf>
- Liu, J., E. Fertig, H. Li, E. Kalnay, B. Hunt, E. Kostelich, I. Szunyogh, and R. Todling, 2007: Comparison between Local Ensemble Transform Kalman Filter and PSAS in the NASA finite volume GCM: perfect model experiments. Available from <http://arxiv.org/ftp/physics/papers/0703/0703066.pdf>.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183-3203.
- Lorenz, E., 1963: Deterministic Non-periodic Flow, *J. Atmos. Sci.*, **20**, 130 - 141.
- Miller, R., M. Ghil, and F. Gauthiez, 1994: Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, **51**, 1037-1056.
- Miyoshi, T., 2005: Ensemble Kalman filter experiments with a Primitive-Equation global model. Doctoral dissertation, University of Maryland, College Park, 197pp. Available at <https://drum.umd.edu/dspace/handle/1903/3046>.
- Miyoshi, T. and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.*, in press.
- Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parameterizations. I: model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, **20**, 175-191.
- Morss, R. E., K. A. Emanuel and C. Snyder, 2001: Idealized adaptive observations strategies for improving numerical weather prediction. *J. Atmos. Sci.*, **58**, 210-234.

- Nerger, L. W. Hiller and J. Scroeter, 2005: A comparison of error subspace Kalman filters. *Tellus*, **57A**, 715-735.
- Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, 2002: Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering. ArXiv:archive/paper 020358, <http://arxiv.org/abs/physics/020358>.
- Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415-428.
- Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, **120**, 1747-1763.
- Pires, C., R. Vautard and O. Talagrand, 1996: On extending the limits of variational assimilation in chaotic systems. *Tellus*, **48A**, 96-121.
- Rabier, F., Jarvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A., 2000: The ECMWF operational implementation of four-dimensional variational physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170.
- Rabier, F. and Z. Liu, 2003: Variational assimilation: theory and overview. Available at [http://www.ecmwf.int/publications/library/ecpublications/\\_pdf/seminar/2003/sem2003\\_rabier.pdf](http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2003/sem2003_rabier.pdf)
- Riishojgaard, L.-P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus* **50A**, 42-57.
- Rotunno, R. and Bao, J. W., 1996: A case study of cyclogenesis using a model hierarchy. *Mon. Wea. Rev.*, **124**, 1051-1066.
- Schröter, J., U. Seiler and M. Wenzel, 1993: Variational assimilation of Geosat into an eddy-resolving model of the Gulf Stream extension area, *J. Phys. Ocean.*, **23**, 925-953.
- Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. A. Yorke, 2005: Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus*, **57A**, 528-545.
- Szunyogh, I., E. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, and J. A. Yorke, 2007: Assessing a local Ensemble Kalman Filter: Assimilating Real Observations with the NCEP Global Model. Under revision in *Tellus*.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters. *Mon. Wea. Rev.*, **131**, 1485-1490.
- Tremolet, Yannick, 2005: Accounting for an imperfect model in 4D-Var. ECMWF Tech Memo #477. Available from [www.ecmwf.int/publications](http://www.ecmwf.int/publications).
- Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140-1158.
- Whitaker, J. S. and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Wea. Rev.*, **130**, 1913-1924.
- Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song and Z. Toth, 2007: Ensemble Data Assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, under revision, available from .
- Yang, S-C., M. Corazza, A. Carrassi, and E. Kalnay, 2007: Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic

- model. AMS 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface. Available from <http://ams.confex.com/ams/pdfpapers/101581.pdf> .
- Yang, S-C., D. Baker, H. Li, M. Huff, G. Nagpal, E. Okereke, J. Villafañe, E. Kalnay and G. Duane, 2006: Data assimilation as synchronization of truth and model: experiments with the 3-variable Lorenz system. *J. Atmos. Sci.* 63, 2340-2354.
- Zupanski, M., 2005: Maximum likelihood ensemble filter: theoretical aspects. *Mon. Wea. Rev.* 133, 1710-1726.

## Tables

Table 1: Impact of the window length on the RMS Analysis error for 4D-Var when x, y, z are observed every 8 or 25 steps (perfect model experiments).

Obs./ 8 steps	Win=8	16	24	32	40	48	56	64	72
Fixed window	0.59	0.51	0.47	0.43	0.62	0.95	0.96	0.91	0.98
QVA (starting with short window)	0.59	0.51	0.47	0.43	0.42	0.39	0.44	0.38	0.43



**Table 3:** Comparison of the RMS error in perfect model experiments obtained observing x, y and z every 8 steps and every 25 steps, using EnKF with 3 or 6 members and optimal inflation factors, and EKF with optimal random and inflation factors (from Yang et al., 2005). The best results obtained with 4D-Var optimizing *simultaneously* the window length and the background error covariance, are also included. The best results obtained with 3D-Var are 0.64 and 1.02 for observations every 8 and 25 steps respectively (Yang et al. 2006).

a) Observations and analysis every 8 time steps			
EnKF, 3 Members	EnKF, 6 members	EKF from Yang et al (2005)	Optimal 4D-Var (W=Window)
0.30 ( $\delta = 0.04$ )	0.28 ( $\delta = 0.02$ )	0.32 ( $\mu = 0.02, \delta = 0$ )	0.31 (W=48)
b) Observations and analysis every 25 time steps			
EnKF, 3 members	EnKF, 6 members	EKF from Yang et al (2005)	Optimal 4D-Var (W=Window)
0.71 ( $\delta = 0.39$ ) 0.61 (hybrid + $\delta = 0.12$ )	0.59 ( $\delta = 0.13$ ) 0.55 (hybrid, + $\delta = 0.04$ )	0.63 ( $\mu = 0.1, \delta = 0.05$ )	0.53 (W=75)

**Table 4:** Comparison of EnKF and 4D-Var with different subsets of variables observed every 8 steps (perfect model experiments).

RMS Obs. type)	EnKF (inflation)	4D-Var (Window=8)	4D-Var (Optimal Window~48)
x	0.82 (.05)	1.10 (0.3 B <sub>3DVAR</sub> )	0.84 (0.15 B <sub>3DVAR</sub> )
y	0.49 (.02)	0.75 (0.3 B <sub>3DVAR</sub> )	0.49 (0.07 B <sub>3DVAR</sub> )
z	>5	>5	>5
x, y	0.39 (.03)	0.61 (0.3 B <sub>3DVAR</sub> )	0.42 (0.08 B <sub>3DVAR</sub> )
x, z	0.41 (.05)	0.77 (0.2 B <sub>3DVAR</sub> )	0.44 (0.03 B <sub>3DVAR</sub> )
y, z	0.31 (.03)	0.58 (0.2 B <sub>3DVAR</sub> )	0.35 (0.04 B <sub>3DVAR</sub> )
x, y, z	0.30 (.04)	0.50 (0.3 B <sub>3DVAR</sub> )	0.31 (0.04 B <sub>3DVAR</sub> )

Table 5: Comparison of EnKF and 4D-Var with complete observations every 8 steps but using an imperfect model (the parameter  $r$  is 28 in the nature run and 26 in the forecast model).  $B$  is the 3D-Var background error covariance.

EnKF, 3 ensemble members, no hybrid, 8 steps observations							
Inflation coeff.	0.1	0.2	0.3	0.4	0.5	0.6	0.7
RMS	1.19	0.97	0.89	0.83	0.81	0.81	0.81

Strong constraint

4D-Var	100B	10B	B	0.1B	0.01B
Win=8	0.84	0.84	0.84	0.84	0.87
Win=16	0.83	0.83	0.84	0.87	0.93
Win=24	0.93	0.93	0.93	0.96	0.99
Win=32	1.01	1.01	1.01	1.02	1.03

Weak constraint

4D-Var	Q=0.001B	Q=0.005B	Q=0.01B	Q=0.05B	Q=0.1B	Q=0.5B
Win=8	0.84	0.83	0.83	0.87	0.92	1.05
Win=16	0.81	0.77	0.75	0.81	0.87	1.00
Win=24	0.87	0.76	0.73	0.79	0.84	0.93
Win=32	0.91	0.75	0.71	0.77	0.80	0.90

Table 6: As Table 5, but with observations every 25 steps.

EnKF, 3 ensemble members, no hybrid, 25 steps observations							
Inflation coeff.	0.1	0.2	0.5	0.7	0.9	0.97	1.0
RMS	2.84	2.02	1.26	1.13	1.07	1.04	1.06

EnKF, 6 ensemble members, no hybrid, 25 steps observations							
Inflation coeff	0.1	0.2	0.5	0.7	0.9	0.98	1.0
RMS	1.61	1.16	0.99	0.96	0.95	0.95	0.95

Strong constraint

4D-Var	No Jb	100B <sub>3DV</sub>	10B <sub>3DV</sub>	5B <sub>3DV</sub>	B <sub>3DV</sub>	0.3B <sub>3DV</sub>	0.1B <sub>3DV</sub>
Win=25	1.04	1.04	1.05	1.06	1.08	1.08	1.07
Win=50	1.17	1.17	1.16	1.16	1.16	1.17	1.20
Win=75	1.26	1.22	1.27	1.36	1.38	1.29	1.70
Win=100	1.47	1.48	1.56	1.50	1.59	1.54	1.64

Weak constraint

4D-Var	Q=0.001B	Q=0.005B	Q=0.01B	Q=0.05B	Q=0.1B	Q=0.5B
Win=25	1.01	0.93	0.91	0.96	1.02	1.15
Win=50	1.11	0.97	0.91	0.98	1.03	1.09
Win=75	1.17	1.05	1.00	1.04	1.13	1.19
Win=100	1.31	1.06	1.03	1.15	1.20	1.35

Table 7: Adaptation of the table of advantages and disadvantages of EnKF and 4D-Var (Lorenc, 2003). Parentheses indicate disadvantages, square brackets indicate clarifications and italics indicate added comments.

#### Advantages (disadvantages) of EnKF

- Simple to design and code.
- Does not need a smooth forecast model [i.e., model parameterizations can be discontinuous].
- Does not need perturbation [linear tangent] forecast and adjoint models.
- Generates [optimal] ensemble [initial perturbations that represent the analysis error covariance].
- Complex observation operators, for example rain, coped with automatically, but sample is then fitted with a Gaussian.
  - *Nonlinear observation operators are possible within EnKF, e.g., MLEF.*
- Covariances evolved indefinitely (only if represented in ensemble)
  - *Under-representation should be helped by “refreshing” the ensemble.*
- (Sampled covariance is noisy) and (can only fit N data)
  - *Localization reduces the problem of long-distance sampling of the “covariance of the day” and increases the ability to fit many observations.*
  - *Observation localization can be used with local filters.*

#### Advantages (disadvantages) of 4D-Var

- [Can assimilate asynchronous observations]
  - *4D-EnKF can also do it without the need for iterations. It can also assimilate time integrated observations such as accumulated rain.*
- Can extract information from tracers
  - *4D-EnKF should do it just as well*
- Nonlinear observation operators and non-Gaussian errors [can be] modeled
  - *Maximum Likelihood Ensemble Filter allows for the use of nonlinear operators and non-Gaussian errors can also be modeled.*
- Incremental 4D-Var balance easy.
  - *In EnKF balance is achieved without initialization for perfect models. For real observations, digital filtering may be needed.*
- Accurate modeling of time-covariances (but only within the 4D-Var window)
  - *Only if the background error covariance (not provided by 4D-Var) includes the errors of the day, or if the assimilation window is long.*

## List of Figures

**Figure 1:** Schematic of EnKF with an ensemble of  $K=2$  forecasts  $\mathbf{x}_1^f$  and  $\mathbf{x}_2^f$  lying on a local attractor (dotted line) which indicates the direction of the “errors of the day”. They are assumed to be Gaussian, with mean  $\bar{\mathbf{x}}^f$  and background error covariance  $\mathbf{B}$ . The analysis is performed within the error subspace defined by the ensemble forecasts, an approximation of the local attractor (solid line). Only the projection  $\hat{\mathbf{y}}$  of the observations  $\mathbf{y}$  (with error covariance  $\hat{\mathbf{R}}$ ) on the ensemble subspace is assimilated. The analysis ensemble  $\mathbf{x}_1^a, \mathbf{x}_2^a$  used as initial conditions for the following forecast, and the analysis mean  $\bar{\mathbf{x}}^a$  are also linear combinations of the ensemble forecasts. A 3D-Var analysis, by contrast, does not include information on the errors of the day.

**Figure 2:** Analysis error in potential vorticity for 100 days of data assimilation using rawinsondes with a 3% observational density randomly distributed in the model domain. All the data assimilation systems, 3D-Var, LETKF (with 30 ensemble members, local volumes of 9x9 horizontal grid points and the full vertical column, and random perturbations of 5% size compared to the natural variability added to the model variables after the analysis), and 4D-Var (12 hour, 24 hour and 48 hour windows) have been optimized. All the experiments are based on a perfect model simulation.

**Figure 3:** Analysis geopotential height RMS errors versus pressure (hPa) in the SPEEDY model using realistic (Reanalysis) observations, either neglecting the presence of model errors (full lines) or correcting them using a constant bias estimation obtained from the time average of 3D-Var analysis increments (dotted lines). The closed circles and squares correspond to 3D-Var, and the open circles and squares to EnKF. The line with triangles corresponds to a bias correction in which the amplitude of the bias is estimated at each EnKF analysis step.

**Figure 4:** Comparison of the globally averaged RMS analysis errors for the zonal wind (left, m/sec) and temperature (right, K) using PSAS (a 3D-Var scheme, dashed line) and the LETKF (solid line) on a finite volume GCM. The same observations (geopotential heights and winds from simulated rawinsondes) are used by both systems. Adapted from Liu et al. (2007).

**Figure 5:** Dependence on the number of LETKF ensemble members of the surface pressure RMS analysis error with the AFES model in a perfect model simulation on the Earth Simulator. Solid line: after a single analysis step; short dashes: analysis errors after 10 days, without using forecast error covariance inflation; long dashes: analysis errors after 10 days, with forecast error covariance inflation (adapted from experiments by Miyoshi and Yamane, 2007).

**Figure 6:** One month time evolution of the LETKF/AFES analysis error for the surface pressure with 10, 20, 40 and 80 ensemble members, in a perfect model simulation using a T159/L48 model (adapted from Miyoshi and Yamane, 2007).

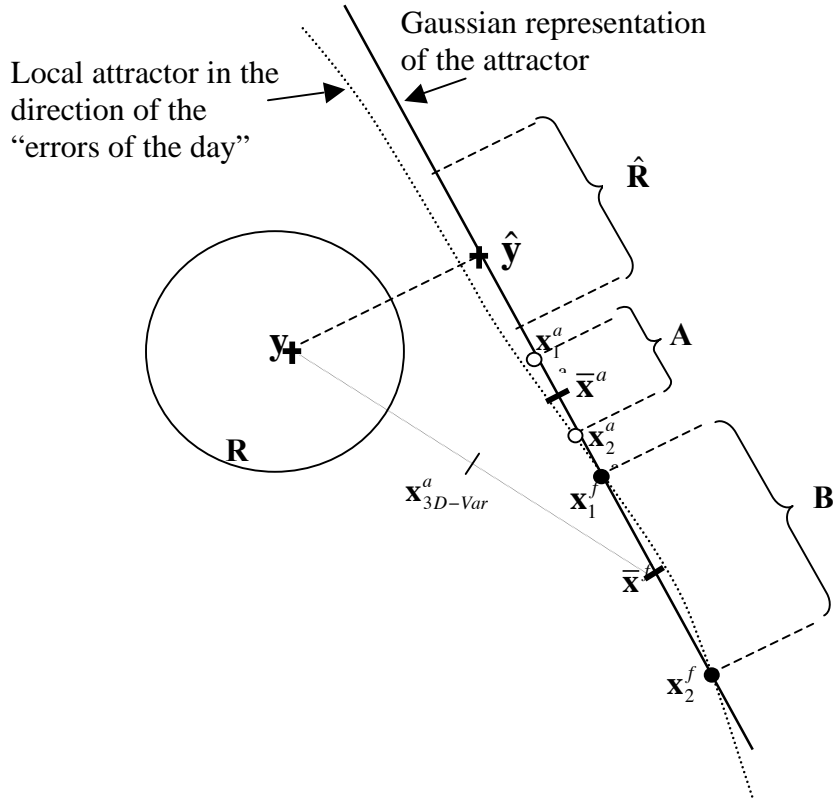


Figure 1: Schematic of EnKF with an ensemble of  $K=2$  forecasts  $\mathbf{x}_1^f$  and  $\mathbf{x}_2^f$  lying on a local attractor (dotted line) which indicates the direction of the “errors of the day”. They are assumed to be Gaussian, with mean  $\bar{\mathbf{x}}^f$  and background error covariance  $\mathbf{B}$ . The analysis is performed within the error subspace defined by the ensemble forecasts, an approximation of the local attractor (solid line). Only the projection  $\hat{\mathbf{y}}$  of the observations  $\mathbf{y}$  (with error covariance  $\hat{\mathbf{R}}$ ) on the ensemble subspace is assimilated. The analysis ensemble  $\mathbf{x}_1^a, \mathbf{x}_2^a$  used as initial conditions for the following forecast, and the analysis mean  $\bar{\mathbf{x}}^a$  are also linear combinations of the ensemble forecasts. A 3D-Var analysis, by contrast, does not include information on the errors of the day.

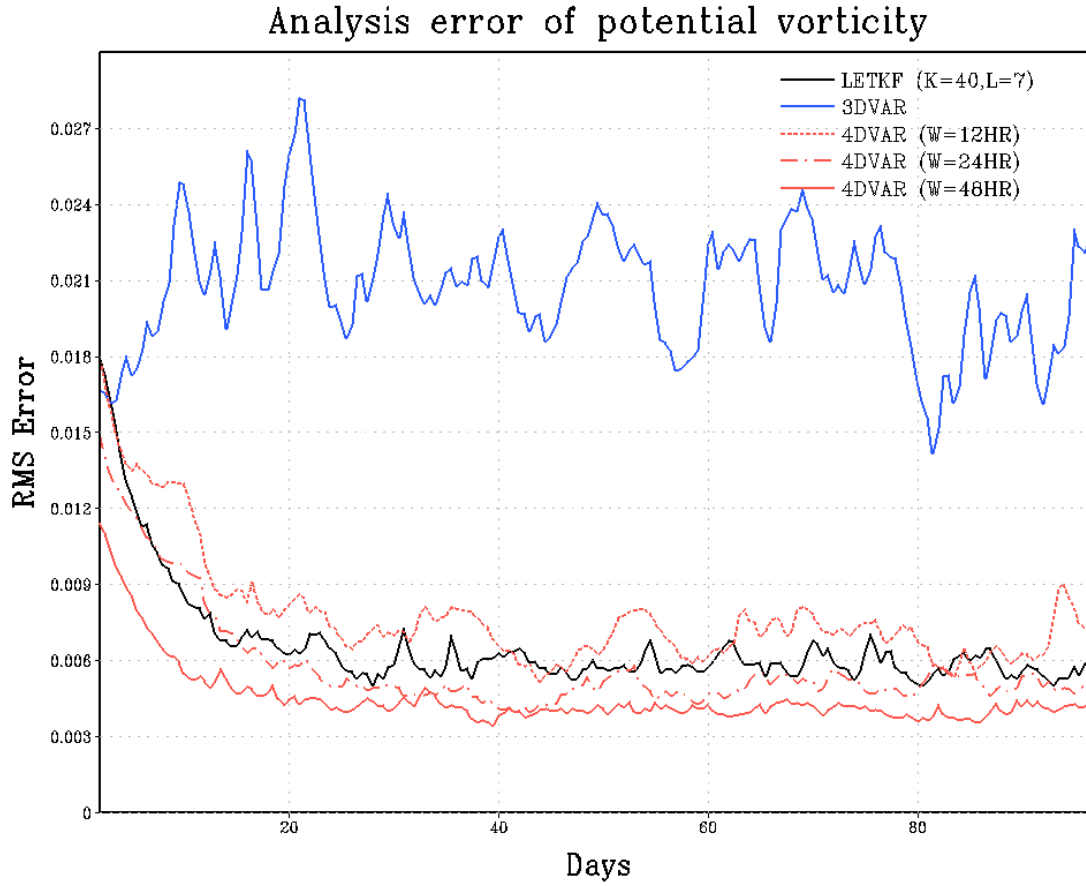


Figure 2: Analysis error in potential vorticity for 100 days of data assimilation using rawinsondes with a 3% observational density randomly distributed in the model domain. All the data assimilation systems, 3D-Var, LETKF (with 30 ensemble members, local volumes of 9x9 horizontal grid points and the full vertical column, and random perturbations of 5% size compared to the natural variability added to the model variables after the analysis), and 4D-Var (12 hour, 24 hour and 48 hour windows) have been optimized. All the experiments are based on a perfect model simulation.

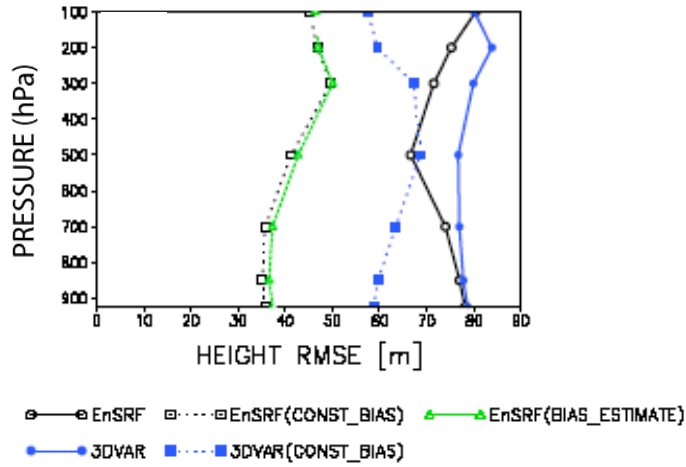


Figure 3: Analysis geopotential height RMS errors versus pressure (hPa) in the SPEEDY model using realistic (Reanalysis) observations, either neglecting the presence of model errors (full lines) or correcting them using a constant bias estimation obtained from the time average of 3D-Var analysis increments (dotted lines). The closed circles and squares correspond to 3D-Var, and the open circles and squares to EnKF. The line with triangles corresponds to a bias correction in which the amplitude of the bias is estimated at each EnKF analysis step.



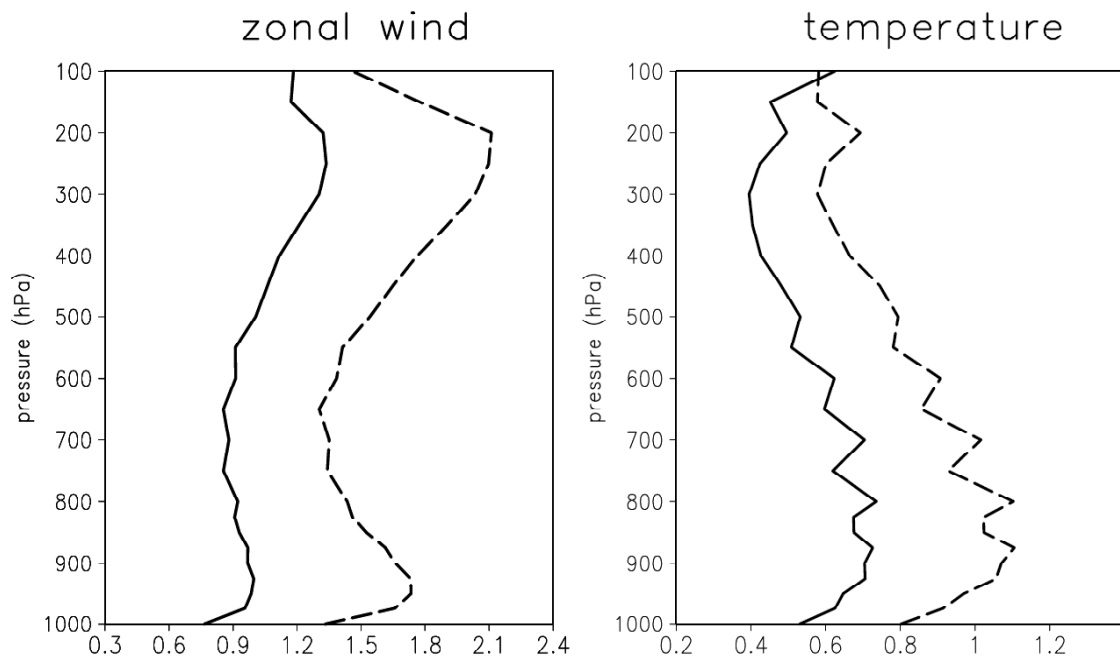


Figure 4: Comparison of the globally averaged RMS analysis ~~and~~ for the z

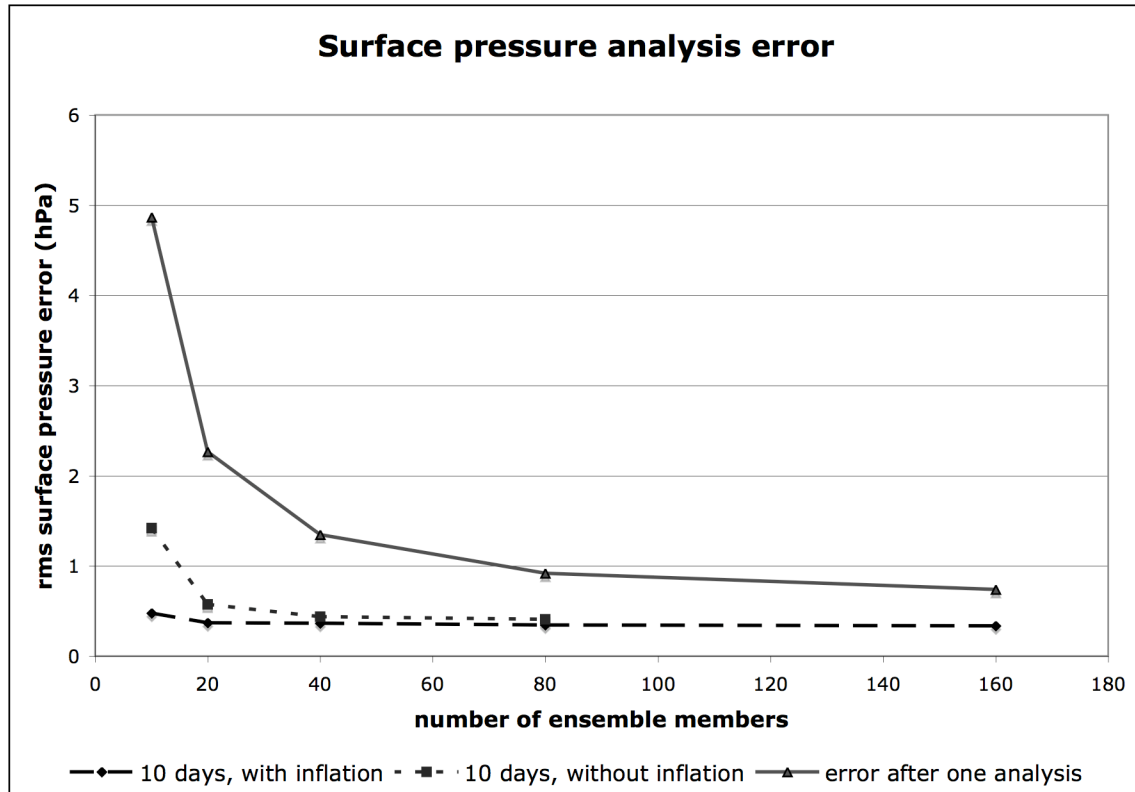


Figure 5: Dependence on the number of LETKF ensemble members of the surface pressure RMS analysis error with the AFES model in a perfect model simulation on the Earth Simulator. Solid line: after a single analysis step; short dashes: analysis errors after 10 days, without using forecast error covariance inflation; long dashes: analysis errors after 10 days, with forecast error covariance inflation (adapted from experiments by Miyoshi and Yamane, 2007).

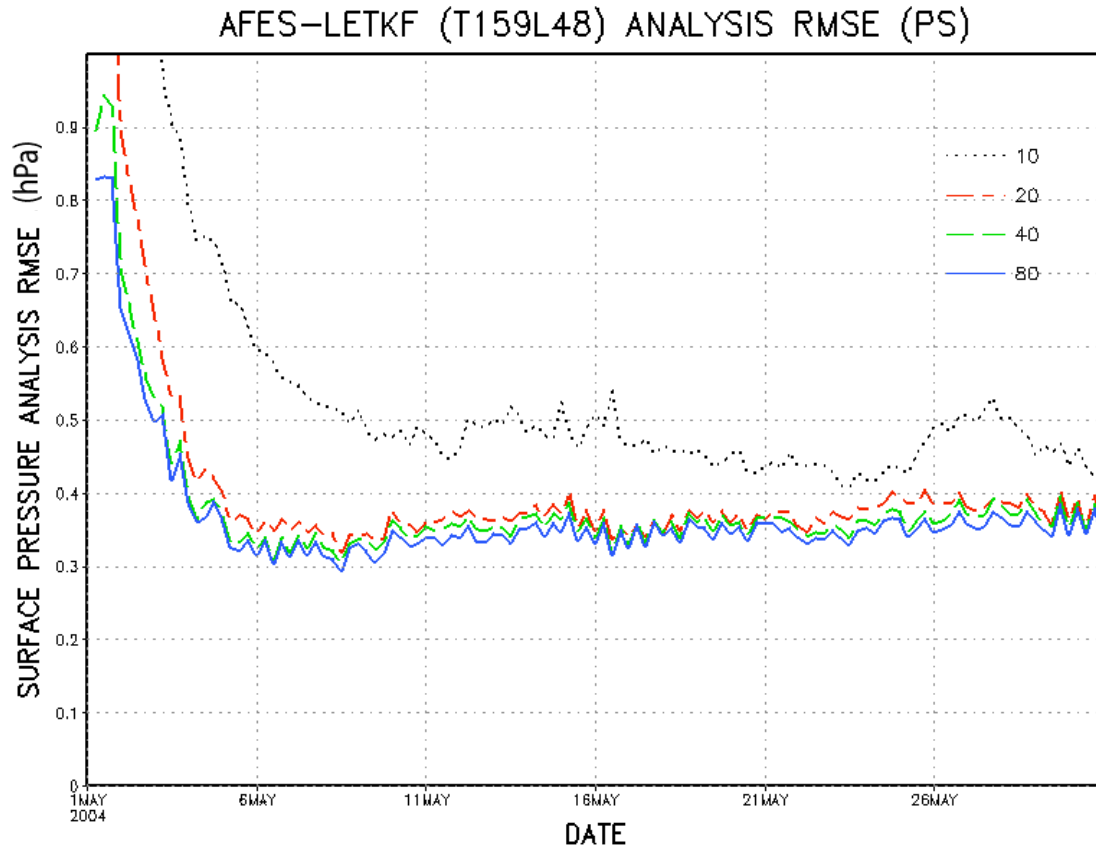


Figure 6: One month time evolution of the LETKF/AFES analysis error for the surface pressure with 10, 20, 40 and 80 ensemble members, in a perfect model simulation using a T159/L48 model (adapted from Miyoshi and Yamane, 2007).