Some Ideas for Ensemble Kalman Filtering

Eugenia Kalnay

Department of Atmospheric and Oceanic Science
University of Maryland, College Park, MD

ABSTRACT

In this seminar we show clean comparisons between EnKF and 4D-Var made in Environment Canada, briefly describe the Local Ensemble Transform Kalman Filter (LETKF) as a representative prototype of Ensemble Kalman Filter, and give several examples of how advanced properties and applications that have been developed and explored for 4D-Var can be adapted to the LETKF without requiring an adjoint model. Although the Ensemble Kalman Filter is less mature than 4D-Var, its simplicity and its competitive performance with respect to 4D-Var suggest that it could become the method of choice.

---

1. Prelude

The WMO/THORPEX Workshop on Intercomparisons of 4D-Var and EnKF that took place in Buenos Aires, Argentina, 10-13 November 2008 was widely attended, with most major operational and research centers throughout the world sending several participants. Presentations are available at http://4dvarenkf.cima.fcen.uba.ar/. Mark Buehner et al., 2008, from Environment Canada, presented a very clean comparison of their operational 4D-Var and EnKF using the same model resolution for the inner loop as in the ensemble, and the same observations (Fig. 1). The results show that the two methods are giving comparable results, with a slight edge favorable to EnKF. In the SH, including a background error covariance based on the EnKF into the 4D-Var improved the 5-day forecast by about 10 hours (not shown).

2. Brief review of the Local Ensemble Transform Kalman Filter algorithm (Hunt et al., 2007)

This description is written as if all the observations are at the analysis time (*i.e.*, for the 3D-LETKF), but the algorithm is the same for the 4D-LETKF (Hunt *et al.*, 2007). In this case the observations are in a time interval that includes the analysis time and *H* is evaluated at the observation time.

a) LETKF **forecast step** (done globally) for each ensemble member *k*:

$$\mathbf{x}_{n,k}^b = M_{t_{n-1},t_n}\left(\mathbf{x}_{n-1,k}^a\right) \quad k = 1,...K$$

b) LETKF **analysis step** (at time *t_n*, so the subscript *n* is dropped):

$$\mathbf{X}^b = \left[\mathbf{x}_1^b - \overline{\mathbf{x}}^b,...,\mathbf{x}_K^b - \overline{\mathbf{x}}^b\right];$$

$$\mathbf{y}_k^b = H(\mathbf{x}_k^b); \mathbf{Y}^b = \left[\mathbf{y}_1^b - \overline{\mathbf{y}}^b,..,\mathbf{y}_K^b - \overline{\mathbf{y}}^b\right]$$

These computations can also be done locally or globally, which is more efficient. Here the overbar represents the ensemble average, and *M* and *H* are the nonlinear model and observation operators respectively.

Localization: choose for each grid point the observations to be used, and compute the local analysis error covariance and analysis perturbations in ensemble space:

$$\hat{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + \mathbf{Y}^{bT}\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}$$

$$\mathbf{W}^a = [(K-1)\hat{\mathbf{P}}^a]^{1/2}$$

The square root required for the matrix of analysis perturbations in ensemble space $\mathbf{W}^\alpha$ is computed using the symmetric square root (Wang *et al.* 2004). This square-root has the advantage of having a zero mean and

---

Correspondence to: Eugenia Kalnay, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD; E-mail: ekalnay@atmos.umd.edu

EnKF mean analysis vs. 4D-Var Bnmc (vs raobs)

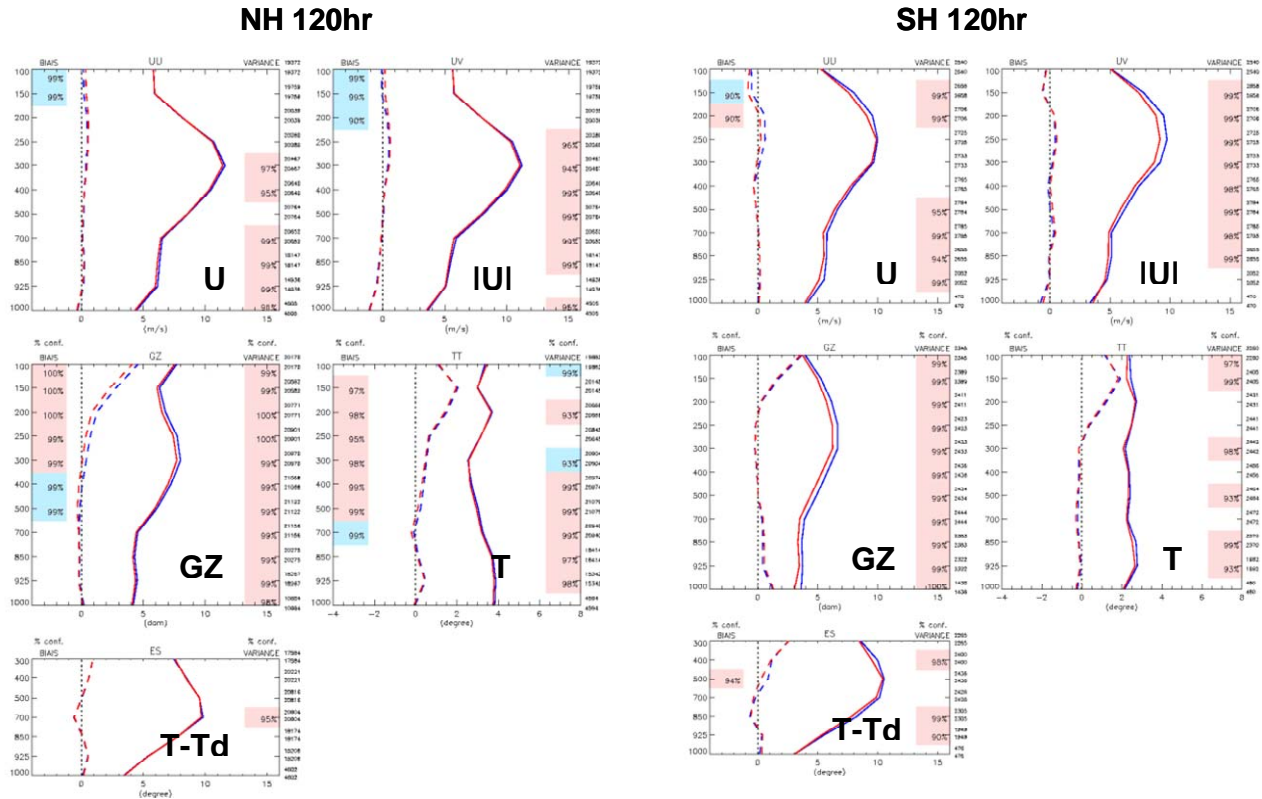### NH 120hr   SH 120hr

Fig. 1. Comparison of bias and standard deviation of 5-day forecasts for February 2007 in the NH and SH verified against rawinsondes for zonal and total wind speed, geopotential height, temperature and dew point depression. Blue and pink colors on the left and right side of each panel indicate the results are better for 4D-Var and for EnKF, respectively, with a level of significance of at least 95%.   From Buehner et al., 2008, http://4dvarenkf.cima.fcen.uba.ar/Download/Session_7/Intercomparison_4D-Var_EnKF_Buehner.pdf

being closer to the identity than the square-root matrix obtained by Cholesky decomposition. As a result the analysis perturbations (chosen in different ways in different EnKF schemes) are also closest to the background perturbations (Ott *et al.* 2002). Note that W^α can also be considered a *matrix of weights* since multiplying the forecast ensemble perturbations at each grid point by W^α gives the grid point analysis ensemble perturbations.

Local analysis in ensemble space:

$$\overline{w}^a = \hat{P}^a Y^{bT} R^{-1}(y^o - \overline{y}^b)$$

Note that $\overline{\mathbf{w}}^a$, in the analysis ensemble space, is a vector of weights, which when multiplied by the matrix X^b of forecast perturbations gives the grid point analysis increment.

$$\mathbf{W}^a \leftarrow \mathbf{W}^a \oplus \overline{\mathbf{w}}^a$$

Here the analysis $\overline{\mathbf{w}}^a$ is added to each column of W^α to get the analysis ensemble in ensemble space. The new ensemble analyses are the *K* columns of

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a + \overline{\mathbf{x}}^b$$

Global analysis ensemble:

The analysis ensemble columns for each grid point are gathered together to form the new global analysis ensemble $\mathbf{x}^a_{n,k}$, and the analysis cycle can proceed.

3. Adaptation of 4D-Var techniques into EnKF

4DVar and EnKF are essentially solving the same problem since they minimize the same cost function using different computational methods. These differences lead to several advantages and disadvantages for each of the two methods (see, for example, Lorenc 2003; Table 7 of Kalnay et al. 2007a; discussion of Gustafsson 2007; response of Kalnay et al. 2007b).

A major difference between 4D-Var and the EnKF is the dimension of the subspace of the analysis increments (analysis minus background). 4D-Var corrects the background forecast in a subspace that has the dimension of the linear tangent and the adjoint models used in the minimization algorithm, and this subspace is generally much larger than the local subspace of corrections in the EnKF determined by the ensemble size K-1. It would be impractical to try to overcome this apparent EnKF disadvantage by using a very large ensemble size. Fortunately, the localization of the error covariances carried out in the EnKF in order to reduce long distance covariance sampling errors, substantially addresses this problem by greatly increasing the number of degrees of freedom available to fit the data. As a result, experience so far has been that the quality of the EnKF analyses with localization increases with the number of ensemble members, but that there is little further improvement when the size of the ensemble is increased beyond about 100. The observation that 50-100 ensemble members are sufficient for the EnKF seems to hold for atmospheric problems ranging from the storm- and meso-scales to the global-scales.

There are a number of additional attractive advantages of 4D-Var. They include the ability to assimilate observations at their right time (Talagrand and Courtier 1987), the fact that within the data assimilation window 4D-Var acts as a smoother (Thépaut and Courtier 1991), ability to use the adjoint model to estimate the impact of observations on the analysis (Cardinali et al. 2004) and on the forecasts (Langland and Baker 2004), the ability to use long assimilation windows (Pires et al. 1996), the computation of outer loops correcting the background state when computing nonlinear observation operators, the ability to use a lower resolution simplified model in the inner loop (see discussion of Fig. 3 later), and the possibility of accounting for model errors by using the model as a weak constraint (Trémolet 2007). In the rest of this section we discuss how these advantages that have been developed for 4D-Var systems can also be adapted and used in the LETKF, a prototype of EnKF.

Fig. 2. Schematic showing that the 4D-LETKF finds the linear combination of the ensemble forecasts at t_n that best fits the observations *throughout* the assimilation window $t_{n-1} - t_n$. The white circles represent the ensemble of analyses (whose mean is the analysis $\bar{\mathbf{x}}^a$), the full lines represent the ensemble forecasts, the dashed line represents the linear combination of the forecasts whose final state is the analysis, and the grey stars represent the asynchronous observations. The cross at the initial time of the assimilation window t_{n-1} is a *no-cost Kalman smoother*, *i.e.*, an analysis at t_{n-1} improved using the information of "future" observations within the assimilation window by weighting the ensembles at t_{n-1} with the weights obtained at t_n. The *smoothed analysis ensemble* at t_{n-1} (not shown in the schematic) can also be obtained at no cost using the same linear combination of the ensemble forecasts valid at t_n given by $\mathbf{W}^a$. (Adapted from Kalnay *et al.* 2007b).

a) 4D-LETKF and no-cost smoother

As indicated by Figure 2, the same weighted combination of the forecasts with weights given by the vector $\bar{\mathbf{w}}^a$ is valid at any time of the assimilation interval. This provides a smoothed analysis mean that (like in 4D-Var) is more accurate than the original analysis because it uses all the future data available throughout the assimilation window (Kalnay *et al.* 2007b; Yang *et al.* 2008a). It should be noted that, as in 4D-Var, the smoothed analysis at the beginning of the assimilation window is an improvement over the filtered analysis computed using only past data. At the end of the assimilation interval only past data is used so that (like in 4D-Var) the smoother coincides with the analysis obtained with the filter. Similarly we can use the matrix
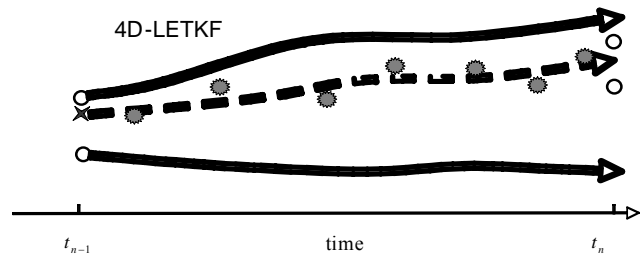
W^α and apply it to the forecast perturbations $X^b W^\alpha$ to provide an associated uncertainty evolving with time (Ross Hoffman, pers. comm., 2008). The updating of the uncertainty is critical for the "Running in Place" method described next, but the uncertainty is not updated in the "outer loop" approach.

b) Application of the no-cost smoother to the acceleration of the spin-up

   4D-Var has been observed to spin up faster than EnKF (e.g. Caya *et al.* 2005), presumably because of its smoothing properties that allow finding the initial conditions at the beginning of the assimilation window that will best fit all the observations. The fact that we can compute a no-cost smoother allows the development of an efficient algorithm, called *Running in place* by Kalnay and Yang (2008), that should be useful in rapidly evolving situations. For example, when radar measurements first detect the development of a severe storm, then the current EnKF estimate of the atmospheric state and its uncertainty are no longer useful. In other words, *while formally the EnKF members and their average are still the most likely state and best estimate of the uncertainty given all the past data, these EnKF estimates are no longer likely at all*. At the start of severe storm convection, the dynamics of the system change substantially, and the statistics of the processes become non-stationary. In this case, as in the spin-up case in which there are no previous observations available, the running in place algorithm ignores the rule "use the data and then discard it" and recycles a few times the new observations.

Running in place algorithm:

   This algorithm is applied to each assimilation window during the spin-up phase. The LETKF is "cold-started" with any initial ensemble mean and perturbations at t_0. The "running in place" loop at time t_n (initially t_0) is as follows: 1. Integrate the ensemble from t_n to t_{n+1}, perform a standard LETKF analysis and obtain the analysis weights for the interval $[t_n, t_{n+1}]$, saving the mean square observations minus forecast (OMF) computed by the LETKF; 2. Apply the no-cost smoother to obtain the smoothed analysis ensemble at t_n by using these weights; 3. Perturb the smoothed analysis ensemble with small zero-mean random Gaussian perturbations, a method similar to additive inflation. Typically the perturbations have amplitudes equal to a small percentage of the climate variance; 4. Integrate the perturbed smoothed ensemble to t_{n+1}. While the forecast fit to the observations continues to improve according to a criterion such as

$$\frac{OMF^2(iter) - OMF^2(iter+1)}{OMF^2(iter)} > \varepsilon,$$

go to step 2 and perform another iteration. If not, replace t_n with t_{n+1} and go to step 1.

Fig. 3.   Comparison of the spin-up of a quasi-geostrophic model simulated data assimilation when starting from random initial conditions. Observations (simulated radiosondes) are available every 12 hours, and the analysis RMS errors are computed comparing with a nature run. Black line: original LETKF with 40 ensemble members, and no prior statistical information. Blue line: 4D-Var with optimal background error covariance. Red line: LETKF "running in place" with $\varepsilon = 5\%$ and 40 ensemble members. Green line: as the red line but with 20 ensemble members.

   Running in place was tested with the LETKF in a quasi-geostrophic, QG, model (Fig. 3, adapted from Kalnay and Yang 2008). When starting from a 3D-Var analysis mean, the LETKF converges quickly (not shown), but from random initial states it takes 120 cycles (60 days) to reach a point in which the ensemble perturbations represent the "errors of the day" (black line in Fig. 3). From then on the ensemble converges quickly, in about 60 more cycles (180 cycles total).
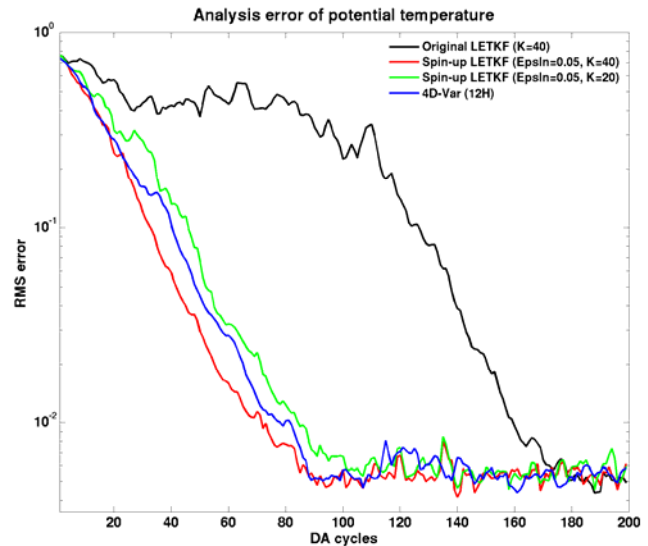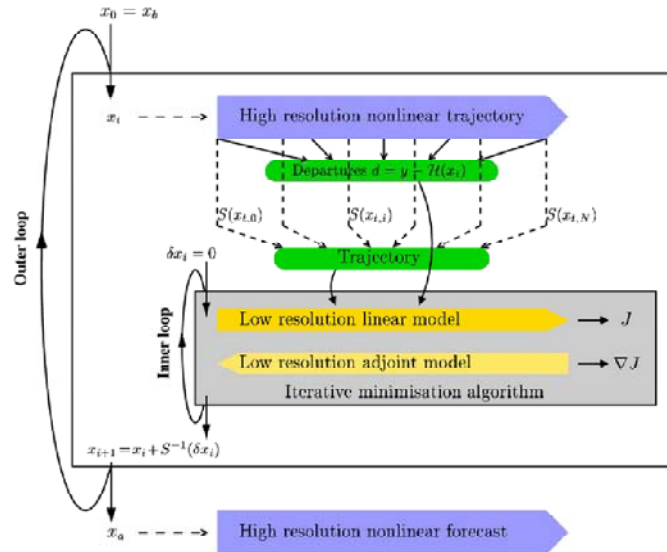
By contrast, the 4D-Var started from the same initial mean state, but using as background error covariance the 3D-Var B scaled down with an optimal factor, converges twice as fast, in about 90 cycles (blue line in Fig. 3). The running in place algorithm with $\varepsilon = 5\%$ (red line) converges about as fast as 4D-Var, and it only takes 2-4 iterations per window even though it does not have the benefit of any prior statistical information.

c) "Outer loop" and dealing with nonlinear ensemble perturbations

A disadvantage of the EnKF is that the Kalman Filter equations used in the analysis assume that the ensemble perturbations are Gaussian, so that when windows are relatively long and perturbations become nonlinear, this assumption breaks down and the EnKF is not optimal. By contrast, 4D-Var is recomputed within an assimilation window until the initial conditions that minimize the cost function for the nonlinear model integration in that window are found. In many operational centres (*e.g.* the National Centers for Environmental Prediction, NCEP, and the European Centre for Medium-range Weather Forecasts, ECMWF) the minimization of the 3D-Var or 4D-Var cost function is done with a linear "inner loop" that improves the initial

Fig. 4. Schematic of how the 4D-Var cost function is minimized in the ECMWF system. (From Yannick Trémolet, August 2007 class on Incremental 4D-Var at University of Maryland summer Workshop on Applications of Remotely sensed data to Data Assimilation).

conditions minimizing a cost function that is quadratic in the perturbations. In the 4D-Var "outer loop" the nonlinear model is integrated from the initial state improved by the inner loop, and the linearized observational increments are recomputed for the next inner loop (Fig. 4).

The ability of including an outer loop increases significantly the accuracy of both 3D-Var and 4D-Var analyses (Arlindo da Silva, pers. comm., 2006), so that it would be important to develop the ability to carry out an equivalent "outer loop" in the LETKF. This can be done by considering the LETKF analysis for a window as an "inner loop", and, using the no-cost smoother, adapting the 4D-Var outer loop algorithm to the EnKF. The method was tested with the Lorenz (1963) model with short and long windows as in Kalnay et al. 2007a. The results (Table 1) suggest that it should be possible to deal with nonlinearities and obtain results comparable or better than 4D-Var by methods such as an outer loop and running in place.

d) Adjoint forecast sensitivity to observations without adjoint model

Langland and Baker (2004) proposed an adjoint-based procedure to assess the observation impact on short-range forecasts without carrying out data-denial experiments. This adjoint-based procedure can evaluate the impact of any or all observations assimilated in the data assimilation and forecast system on a selected measure of short-range forecast error. In addition, it can be used as a diagnostic tool to monitor the quality of observations, showing which observations make the forecast worse, and can also give an estimate of the relative importance of observations from different sources. Following a similar procedure, Zhu and Gelaro (2008) showed that this adjoint-based method provides accurate assessments of the forecast sensitivity with respect to most of the observations assimilated, and detected that the way certain AIRS

RMSE analysis error	4D-Var	LETKF (inflation factor)	LETKF with less than 3 "outer loop" iterations
Window=8 steps (perturbations are approximately linear)	0.31	0.30 (1.05)	0.27 (1.04)
Window=25 steps (perturbations are nonlinear)	0.53	0.66 (1.28)	0.48 (1.08)

Table 1. Comparison of 4D-Var and LETKF for the Lorenz (1963) 3-variable model. 4D-Var has been simultaneously optimized for the window length (Kalnay et al., 2007a, Pires et al. 1996) and the background error covariance, and the full nonlinear model is used in the minimization. LETKF is performed with 3 ensemble members (no localization is needed for this problem), and inflation is optimized. For the 25 steps case, running in place further reduces the error to a remarkably low value of about 0.39.

humidity channels were used actually made forecasts worse. Unfortunately, this powerful method to estimate observation impact requires the adjoint of the forecast model which is complicated to develop and not always available.

Liu and Kalnay (2008) proposed an ensemble-base sensitivity method able to assess the same forecast sensitivity to observations as in Langland and Baker (2004), but without using the adjoint model.
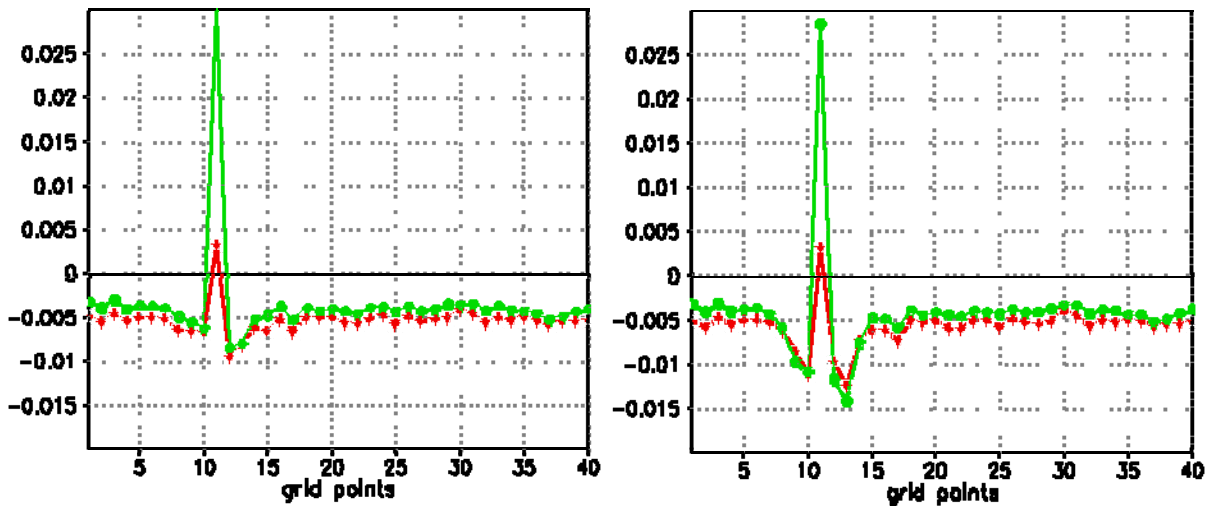

Fig. 5. Time average (over the last 7000 analysis cycles) of the contribution to the reduction of forecast errors from each observation location. Left: the observation at the 11th grid point has 0.8 random errors rather than the specified value of 0.2. Right: the observation at the 11th grid point has random errors as specified but it has a bias of 0.5 rather than 0.0, as specified. Ensemble sensitivity method: green line with closed circles; adjoint method: red line with plus signs. Adapted from Liu and Kalnay (2008).

Figure 5 shows the result of applying this method to the Lorenz (1996) 40-variables model. In this case there were observations at every point every 6 hours created from a "nature" run by adding Gaussian observational errors of mean zero and standard deviation 0.2. At the location 11, however, the standard deviation of the errors was increased to 0.8, (Fig. 5, left-hand panel) without "telling" the data assimilation system about the observation problem in this location. In Fig. 5 (right-hand panel), the standard deviation was kept at its correct value, but a bias of 0.5 was added to the observation at the 11th grid point, still assuming that the bias was zero in the data assimilation. As shown in the figure, both the adjoint and the ensemble-based sensitivity were able to identify that the observations at grid point 11 had a deleterious impact on the forecast. They both show that the neighboring points improved the forecasts more than average by partially correcting the effects of the 11th point observations.

Although the cost function in this example was based on the Eulerian norm, appropriate for a univariate problem, the method can be easily extended to an energy norm, allowing the comparison of the impact of winds and temperature observations (or any other type of observation such as radiances) on the forecasts.

e) Use of a lower resolution analysis

The inner/outer loop used in 4D-Var was introduced in subsection c), where we showed that a similar outer loop can be carried out in EnKF. We now point out that it is common practice to compute the inner loop minimization, shown schematically in Figure 4, using a simplified model (Lorenc, 2003), which usually has lower resolution and simpler physics than the full resolution model used for the nonlinear outer loop integration. The low-resolution analysis correction computed in the inner loop is interpolated back to the full resolution model (Figure 4). The use of lower resolution in the minimization algorithm of the inner loop results in substantial savings in computational cost compared with a full resolution minimization, but it also degrades the analysis.

Yang *et al.* (2008b) took advantage that in the LETKF the analysis ensemble members are a weighted combination of the forecasts, and that the analysis weights W^α are much smoother (they vary on a much larger scale) than the analysis increments or the analysis fields themselves. They tested the idea of interpolating the weights but using the full-resolution forecast model on the same quasi-geostrophic model discussed before. They performed full resolution analyses and compared the results with a computation of the LETKF analysis (*i.e.*, the weight matrix W^α) on coarser grids, every 3 by 3, 5 by 5 and 7 by 7 grid points, corresponding to an analysis grid coverage of 11%, 4% and 2% respectively, and with the interpolation of the analysis increments. They found that interpolating the weights did not degrade the analysis compared with the full resolution, whereas

Fig. 6. The time series of the RMS analysis error in terms of the potential vorticity from different DA experiments. The LETKF analysis from the full-resolution is denoted as the black line and the 3D-Var derived at the same resolution is denoted as the grey line. The LETKF analyses derived from weight-interpolation with different analysis coverage are indicated with blue lines. The LETKF analyses derived after the first 20 days from increment-interpolation with different analysis coverage are indicated with the red lines. Adapted from Yang *et al.* (2008b).

interpolating the analysis increments resulted in a serious degradation (Fig. 6). The use of a symmetric square root in the LETKF ensures that the interpolated analysis has the same linear conservation properties as the full resolution analysis. The results suggest that interpolating the analysis weights computed on a coarse grid without degrading the analysis can substantially reduce the computational cost of the LETKF. Although the full resolution ensemble forecasts are still required, they are also needed for ensemble forecasting in operational centers. We note that the fact that the weights vary on large scales, and that the use of a coarser analyses with weight interpolation actually improves slightly the analysis in data sparse regions, suggesting that smoothing the weights is a good approach to filling data gaps such as those that appear in between satellite orbits. (Yang *et al.* 2008b, Lars Nerger, *pers. comm.* 2008). Smoothing the weights, both in the horizontal and in the vertical may also reduce sampling errors and increase the accuracy of the analyses.
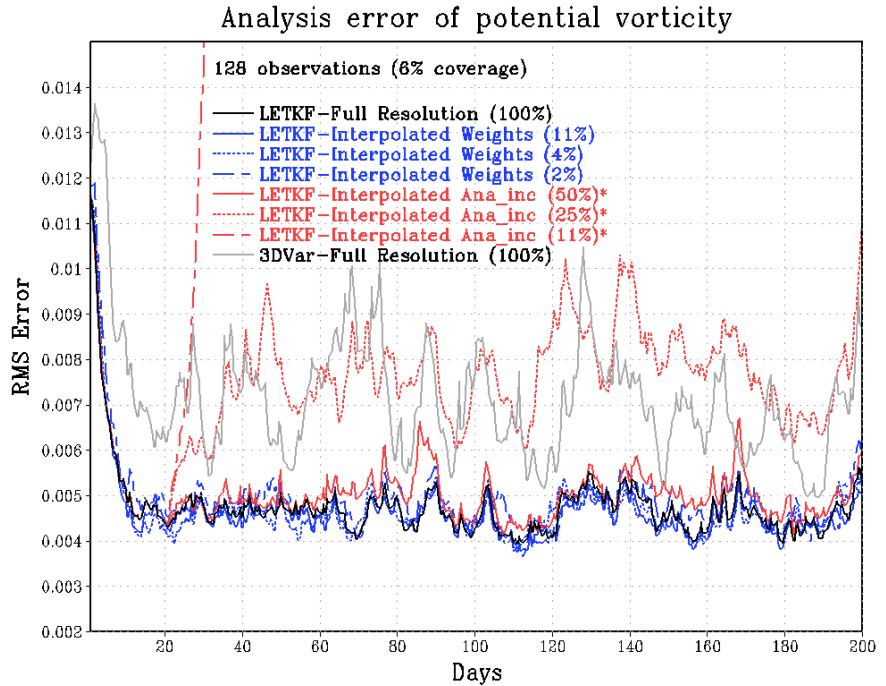
f) Model error

Model error can seriously affect the EnKF because, among other reasons, the presence of model biases cannot be detected by the EnKF original formulation, and the ensemble spread is the same with or without model bias (Li 2007). For this reason, the most widely used method for imperfect models is to increase the multiplicative or additive inflation (*e.g.* Whitaker *et al.* 2007). Model biases can also be taken into account by estimating the bias as in Dee and da Silva (1998) or its simplified approximation (Radakovich *et al.* 2001). More recently, Baek et al. (2007) pointed out that model bias could be estimated accurately augmenting the model state with the bias, and allowing the error covariance to eventually correct the bias. Because the bias was assumed to be a full resolution field, this required doubling the number of ensemble members in order to reach convergence.

In the standard 4D-Var, the impact of model bias cannot be neglected within longer windows because the model (assumed to be perfect) is used as a strong constraint in the minimization (*e.g.* Andersson *et al.* 2005). Trémolet (2007) has developed several techniques allowing for the model to be a weak constraint in order to estimate and correct model errors. Although the results are promising, the methodology for the weak constraint is complex, and still under development.

Fig. 7. Comparison of the analysis error averaged over two months for the zonal velocity in the SPEEDY model for several simulations with the radiosonde observations available every other point. The yellow line corresponds to a perfect model simulation with the observations extracted from a SPEEDY model "nature run". The red is the control run, in which the observations were extracted from the NCEP-NCAR Reanalysis, but the same multiplicative inflation was used as in the perfect model case. The blue line and the black solid lines correspond to the application of optimized multiplicative and additive inflation respectively. The long-dashed line was obtained correcting the bias with the Dee and DaSilva (1998) method, and combining it additive inflation. The short-dashed is as the long-dashed but using the Danforth *et al.* (2007) low dimensional method to correct the bias, and the green line is as the long-dashed line but using the simplified Dee and DaSilva method. Adapted from Li (2007).

Li (2007) compared several methods to deal with model bias (Fig. 7), including a "Low-dimensional" method based on an independent estimation of the bias from averages of 6 hour estimated forecast errors started from a reanalysis (or any other available good quality analysis). This method was applied to the SPEEDY (Simplified Parameterizations primitivE-Equation Dynamics) model assimilating simulated observations from the NCEP-NCAR (National Centers for Environmental Prediction-National Center for Atmospheric Research) Reanalysis, and it was found to be able not only to estimate the bias, but also the errors in the diurnal cycle and the model forecast errors linearly dependent on the state of the model (Danforth *et al.* 2007; Danforth and Kalnay 2008).

The results obtained by Li (2007) accounting for model errors within the LETKF, presented in Figure 7 indicate that: a) additive inflation is slightly better than multiplicative inflation, b) methods to estimate and correct model bias (e.g. Dee and da Silva 1998; Danforth *et al.* 2007) should be combined with inflation, which is more appropriate in correcting random model errors. The combination of the Low-Dimensional method with additive inflation gave the best results, and was substantially better than the results obtained assuming a perfect model (Fig. 7). We note that the approach of Baek *et al.* (2007) of correcting model bias by augmenting the state vector with the bias can be used to determine other parameters, such as surface fluxes, observational bias, nudging coefficients, etc. It is similar to increasing the control vector in the
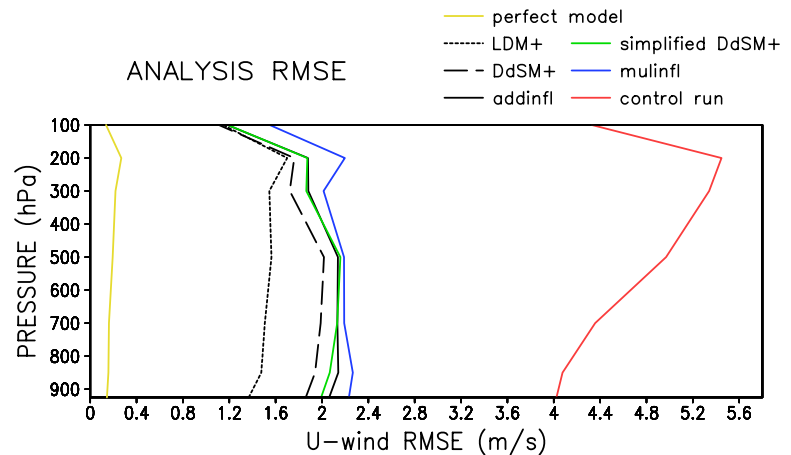
variational approach, and is only limited by the number of degrees of freedom that are added to the control vector and sampling errors in the augmented background error covariance.

4. Summary and discussion

4D-Var and the EnKF are the most advanced methods for data assimilation. 4D-Var has been widely adopted in operational centres, with excellent results and much accumulated experience. EnKF is less mature, and has the disadvantage that the corrections introduced by observations are done in a space of lower resolution, since they depend on the ensemble size, although this problem is ameliorated by the use of localization. The main advantages of the EnKF are that it provides an estimate of the forecast and analysis error covariances, and that it is much simpler to implement than 4D-Var. Recent "clean" comparisons between the operational 4D-Var and EnKF systems in Environment Canada, using the same model resolution and observations, indicated that the forecasts had essentially identical scores, whereas the 4D-Var using a background error covariance based on the EnKF gave a 10-hour improvement in the 5-day forecasts in the Southern Hemisphere (Buehner *et al.* 2008).

It is frequently stated that the best approach should be a hybrid that combines "the best characteristics" of both EnKF and 4D-Var (e.g. Lorenc 2003). Unfortunately this would also bring the main disadvantage of 4D-Var to the hybrid system, *i.e.*, the need to develop and maintain an adjoint model. This makes the hybrid approach attractive to operational centres that already have appropriate linear tangent and adjoint models, but not otherwise.

In this review we have instead focused on the idea that the advantages and new techniques developed over the years for 4D-Var, can be adapted and implemented within the EnKF framework, without requiring an adjoint model. The LETKF (Hunt *et al.* 2007) was used as a prototype of the EnKF. It belongs to the square root or deterministic class of the EnKF (*e.g.* Whitaker and Hamill 2002) but simultaneously assimilates observations locally in space, and uses the ensemble transform approach of Bishop et al. (2001) to obtain the analysis ensemble as a linear combination of the background forecasts. We showed how the LETKF could be modified to include some of the most important 4D-Var advantages: a no-cost smoothing algorithm, useful not only to use "future" (as in reanalysis) but also to accelerate spin-up and handle nonlinear, non-Gaussian ensemble perturbations, and how to implement an "outer loop" within the LETKF. Taking advantage that the LETKF calculates the analysis weights valid throughout the data assimilation window that linearly combine the forecast perturbations to compute the analysis ensemble, we computed the LETKF on coarse grids and interpolated the weights to the full resolution grid. Yang *et al.* (2008) found that the weight interpolation from a coarse resolution grid did not degrade the analysis, suggesting that the weights vary on large scales and smoothing them can increase the accuracy of the analysis, and that weight interpolation is ideal to fill up analysis data voids.

One of the most powerful applications of the adjoint model is the ability to estimate the impact of classes of observations on the short range forecast (Langland and Baker 2004), and we showed how this "adjoint sensitivity" can be computed within the LETKF without an adjoint model (Liu and Kalnay, 2008). Finally, Li (2007) compared several methods used to correct model errors and showed that it is advantageous to combine methods that correct the bias, such as that of Dee and da Silva (1998) and the low-dimensional method of Danforth *et al.* (2007), with methods like inflation that are more appropriate to account for random model errors. This is an alternative to the weak constraint method (Trémolet 2007) to deal with model errors in 4D-Var, and involves the addition of a relatively small number of degrees of freedom to the control vector. Li *et al.* (2008) also developed a method to optimally estimate both the inflation coefficient for the background error covariance and the actual observation error variances (not shown here).

In summary, we have emphasized that the EnKF can profit from the methods and improvements that have been developed in the wide research and operational experience acquired with 4D-Var. Given that operational tests comparing 4D-Var and the LETKF indicate that the performance of these two methods is already very close (*e.g.* Miyoshi and Yamane 2007, Buehner *et al.* 2008), and that the LETKF and other EnKF methods are much simpler to implement, their future looks bright.

References

Bishop, C.H., B.J. Etherton and S.J. Majumdar, 2001. Adaptive Sampling with Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Weather Rev.*, **129**, 420-436.

Buehner, M., C. Charette, Bin He, Peter Houtekamer, Herschel Mitchell, 2008. Intercomparison of 4-D Var and EnKF systems for operational deterministic NWP. Available from

 http://4dvarenkf.cima.fcen.uba.ar/Download/Session_7/Intercomparison_4D-Var_EnKF_Buehner.pdf

Burgers, G., P.J. van Leeuwen and G. Evensen, 1998. On the analysis scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.*, **126**, 1719-1724.

Caya, A., J. Sun and C. Snyder, 2005. A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.*, **133**, 3081-3094

Danforth, C.M., E. Kalnay and T. Miyoshi, 2006 Estimating and Correcting Global Weather Model Error. *Mon. Weather Rev.*, **134**, 281-299.

Danforth, C.M. and E. Kalnay, 2008. Using Singular Value Decomposition to parameterize state dependent model errors. *J. Atmos. Sci.*, **65**, 1467-1478

Dee, D.P. and A.M. da Silva, 1998. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124**, 269-295.

Evensen, G., 2003. The ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean 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.

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.*, **131**, 3235-3246.

Gustafsson, N., 2007. Response to the discussion on "4-D-Var or EnKF?". Tellus A, 59, 778-780.

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

Houtekamer, P.L. and H.L. Mitchell, 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Weather Rev.*, **126**, 796-811.

Houtekamer, P.L. and H.L. Mitchell. 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Weather Rev.*, **129**, 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. Weather Rev.*, **133**, 604-620.

Houtekamer, P.L. and H.L. Mitchell, 2005. Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.*, **131**, 3269-3290.

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*, **230**, 112-126.

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.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang and J. Ballabrera-Poy, 2007a. 4D-Var or Ensemble Kalman Filter? *Tellus A*, **59**, 758-773.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang and J. Ballabrera-Poy, 2007b. Response to the Discussion on "4D-Var or EnKF?" by Nils Gustaffson. *Tellus A*, **59**, 778-780.

Kalnay, E. and S-C. Yang, 2008. Accelerating the spin-up in EnKF. *Arxiv: physics:Nonlinear/0.806.0180v1*.

Langland, R.H. and N.L. Baker, 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201.

Li, H., 2007. Local ensemble transform Kalman filter with realistic observations. Ph. D. thesis. Available at http://hdl.handle.net/1903/7317

Li, Hong, E Kalnay and T Miyoshi, 2008. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. . *Q. J. R. Meteorol. Soc.,* **134**, in press.

Liu, J., and E Kalnay, 2008. Estimating observation impact without adjoint model in an ensemble Kalman Filter. *Q. J. R. Meteorol. Soc.,* **134**, 1327-1335.

Lorenc, A.C., 2003. The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Q. J. R. Meteorol. Soc.*, **129**, 3183-3203.

Lorenz, E., 1963. Deterministic Non-periodic Flow, *J. Atmos. Sci.*, **20**, 130-141.

Mitchell, H.L., P.L. Houtekamer and G. Pellerin, 2002. Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter *Mon. Weather Rev.*, **130**, 2791–2808.

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. Weather Rev.*, **135**, 3841-3861.

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.

Pires, C., R. Vautard and O. Talagrand, 1996. On extending the limits of variational assimilation in chaotic systems. *Tellus*, **48A**, 96-121.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf and A. Simmons, 2000. The ECMWF operational implementation of four-dimensional variational physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170.

Radakovich, J.D., P.R. Houser, A.M. da Silva and M.G. Bosilovich, 2001. Results from global land-surface data assimilation methods. *Proceeding of the fifth symposium on integrated observing systems,* 14-19 January 2001, Albuquerque, NM. 132-134.

Rotunno, R. and J.W. Bao, 1996. A case study of cyclogenesis using a model hierarchy. *Mon. Weather Rev.*, **124**, 1051-1066.

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, 2008. Assessing a local Ensemble Kalman Filter: Assimilating Real Observations with the NCEP Global Model. *Tellus*, in press.

Talagrand O. and P. Courtier, 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill and J.S. Whitaker, 2003. Ensemble Square Root Filters. *Mon. Weather Rev.*, **131**, 1485-1490.

Trémolet, Y., 2007. Model-error estimation in 4D-Var. *Q. J. R. Meteorol. Soc.*, **133**, 1267–1280.

Whitaker, J.S. and T.M. Hamill, 2002. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.*, **130**, 1913–1924.

Yang, S-C, M. Corazza, A. Carrassi, E. Kalnay and T. Miyoshi, 2008. Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic model. *Mon. Weather Rev.*, under revision.

Yang, S-C, E. Kalnay, B. Hunt and N. Bowler, 2008. Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *Q. J. R. Meteorol. Soc.*, in press.

Zupanski, M., 2005. Maximum likelihood ensemble filter: theoretical aspects. *Mon. Weather Rev.*, **133**, 1710-1726.