# 6.11 COMPARISON OF ENSEMBLE-BASED AND VARIATIONAL-BASED DATA ASSIMILATION SCHEMES IN A QUASI-GEOSTROPHIC MODEL

Shu-Chih Yang[1*], Matteo Corazza[2], Alberto Carrassi[3], Eugenia Kalnay[1], and Takemasa Miyoshi[4]
[1] University of Maryland, College Park, Maryland
[2] ARPAL CFMI-PC, Italy
[3] University of Ferrara, Italy
[4] Japanese Meteorological Agency

## 1 INTRODUCTION

Data assimilation estimates the best analysis state to initialize a numerical model by combining the information of the model forecast state and observations. A good data assimilation system can improve the forecast skill and provide better use of observation system. The three dimensional variational analysis (3D-Var), widely used in many operational centers, has been considered as an economically and statistically reliable method. However, the flow-independent characteristic of 3D-Var (it does not include "errors of the day") limits the quality of the corrections to the first guess.

Four-dimensional variational analysis (4D-Var) is an optimal technique to seek the initial condition leading to a forecast that best fits the observations within an assimilation interval or window (Talagrand and Courtier 1987; Courtier et al, 1994; Rabier et al 2000). The minimization of the 4D-Var cost function (defined over a time interval) requires the gradient of the cost function, which in turn involves the use of the tangent linear and adjoint models. During the past decade, experience with ensemble forecasting has suggested that ensemble forecast and data assimilation could be combined in a natural way (i.e., in Ensemble Kalman Filter, Evensen 1994). Experiments based on the ensemble-based data assimilations have indicated the possibility that they could become an alternative method to 4D-Var (Whitaker and Hamill 2002, Anderson 2001, Ott et al. 2004; Kalnay et al. 2005). The Kalman Filter can optimally update (predict) the model state and forecast error covariance. The error statistics of the forecast is flow-dependent and estimated nonlinearly by ensemble members, which may be beneficial for analyses in situations where nonlinearity is strong and statistics exhibit some non-normality (Hamill 2003; Yang et al. 2005).

Compared with 4D-Var, an ensemble-based scheme is easy to implement and maintain, since it does not require the development and maintenance of the tangent linear and adjoint models. Moreover, it creates an ensemble of possible analysis states, providing information on both the mean analysis and the uncertainty. It automatically generates the initial ensemble perturbations, while 3D-Var/4D-Var schemes can only provide the mean analysis and an additional procedure is needed to start an ensemble prediction system. The information about the flow state, such as the uncertainties associated with flow instabilities, will be propagated through the ensemble-based data assimilation cycle, while each data assimilation cycle is an independent cycle in the variational-based methods. Reviews of the ensemble-based data assimilations are available in Hamill (2003) and Evensen (2003). Lorenc (2003) and Kalnay et al. (2005) discuss the pros and cons of ensemble Kalman filter and 4D-Var.

In this study, we would like to consider how the characteristics of data assimilation methods impact results. Two variational and two ensemble-based schemes have been implemented in a channel quasigeostrophic (QG) model (Rotunno and Bao, 1996) that has been widely used for testing and comparing problems related to data assimilation and adaptive observations (e.g., Morss et al, 2001, Hamill et al 2000, Hamill and Snyder 2000, Corazza et al, 2003). The model variables are potential temperature at the bottom and top levels, and potential vorticity at the interior 5 levels. "Rawinsonde observations" provide velocity components and temperatures at all levels. In this study, we would like to compare how the data assimilation schemes perform when offered the same noisy observations. The two variational schemes are 3DVAR (developed by Morss, 1998) and 4DVAR (newly developed for this study). The two ensemble schemes are the Local Ensemble Kalman Filter (LEKF, based on Ott et al, 2004), and a hybrid system using an ensemble of bred vectors added to the regular 3DVAR system (Corazza et al, 2002) to augment the background error covariance with information on the "errors of the day". The object of this study is to explore the differences between the variational-based and ensemble-based data assimilation methods and discuss considerations that would be applicable to a more realistic scenario (if implemented operationally).

## 2 DATA ASSIMILATION SCHEMES

* *Corresponding author address*: Shu-Chih Yang, 4353 CSS building, Univ. of Maryland, College Park, 20742; email: scyang@atmos.umd.ed.

### *2.1 3D-Var*

The 3D-Var system, developed by Morss (1999) is based on the approach of Parrish and Derber (1992). The analysis increment from 3D-Var is obtained solving (1) with the conjugate gradient method.

$$[\mathbf{I}+\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}](x_a-x_b)=\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}(y-\mathbf{H}x_b) \qquad (1)$$

The 3D-Var background error covariance **B** is derived using "the NMC method" of Parrish and Derber (1992). In this study, results from obtained with 3D-Var where the amplitude of **B** has been optimized, will be taken as the baseline for comparison with more advanced methods.

### *2.2 4D-Var*

The cost function in the 4D-Var system is the generalization of the formula used in 3D-Var, considering observations distributed in a time interval. The operators include the time-dependent model states and observations, which allow comparing their differences at the appropriate time. Through the minimization of the cost function, 4D-Var seeks the initial condition such that its forecast best fits the observations within the assimilation interval.

For practical applications, the incremental form (Courtier et al, 1994) of 4D-Var is used in the cost function (2) and its gradient (3). In (2) and (3), $\delta x(t_0)$ is the difference between initial and first guess (forecast) states in model grid coordinate and $d(t_i)$ is the difference between forecast state and observations at observation coordinate at observation time, $t_i$. **L** is the tangent linear model that maps a set of small perturbations from $t_0$ to $t_i$ through forward integration and $\mathbf{L}^{\mathbf{T}}$ is the adjoint operator of the linear tangent operator that maps the adjoint perturbations backward in time.

$$J(\delta x(t_0))=\frac{1}{2}(\delta x(t_0))^T\mathbf{B}^{-1}(\delta x(t_0)) + \frac{1}{2}\sum_{i=0}^{n}[\mathbf{H}_i\mathbf{L}(t_0,t_i)\delta x(t_0)-d(t_i)]^T\mathbf{R}^{-1}[\mathbf{H}_i\mathbf{L}(t_0,t_i)\delta x(t_0)-d(t_i)] \qquad (2)$$

$$\nabla J(\delta x(t_0))=\mathbf{B}^{-1}(\delta x(t_0))+\sum_{i=0}^{K}\mathbf{L}^T(t_i,t_0)\mathbf{H}^T\mathbf{R}^{-1}[\mathbf{H}\mathbf{L}(t_0,t_i)\delta x(t_0)-d(t_i)] \qquad (3)$$

A variable transformation is applied in (2) and (3) in order to avoid the difficulty of the inverting **B**. The transform operator, **U**, is chosen to be the square root of the inverse of **B** ($\mathbf{B}^{-1}=\mathbf{U}^T\mathbf{U}$). The variable transformation not only enables us to compute the cost function directly, but also contributes to efficient minimization by preconditioning. The cost function and its gradient are reformulated with the preconditioned variable, $\delta v$, and $\delta x=\mathbf{U}^{-1}\delta v$.

$$J(\delta v(t_0))=\frac{1}{2}\delta v(t_0)^T\delta v(t_0) + \frac{1}{2}\sum_{i=0}^{n}[\mathbf{H}\mathbf{L}(t_0,t_i)\mathbf{U}^{-1}\delta v(t_0)-d(t_i)]^T\mathbf{R}^{-1}[\mathbf{H}\mathbf{L}(t_0,t_i)\mathbf{U}^{-1}\delta v(t_0)-d(t_i)] \qquad (4)$$

$$\nabla J(\delta v(t_0))=\delta v(t_0)+\sum_{i=0}^{K}(\mathbf{U}^{-1})^T\mathbf{L}^T(t_i,t_o)\mathbf{H}^T\mathbf{R}^{-1}[\mathbf{H}\mathbf{L}(t_0,t_i)\mathbf{U}^{-1}\delta v(t_0)-d(t_i)] \qquad (5)$$

The minimization is made with respect to the variable $\delta v$, which is initially set to zero, and only $\mathbf{U}^{-1}$ is needed. As in Morss et al (2001) and Parrish and Derber (1992), **B** is built by constructing the horizontal error covariance at each level and linked with a vertical correlation matrix with the assumption that there is no correlation between different wavenumbers. Therefore, we can define $\mathbf{U}^{-1}$ as:

$$\mathbf{U}^{-1}=\mathbf{S}\hat{\mathbf{C}}^{1/2}\mathbf{E} \qquad (6)$$

In (6), **S** is the operator that maps from spectral to model coordinates, $\hat{\mathbf{C}}^{1/2}$ is the square root of the horizontal background error variance of each wavenumber, and **E** is the square root of the vertical correlation matrix. In our case, $\hat{\mathbf{C}}^{1/2}$ will be a diagonal matrix, and **E** is simply a 7x7 matrix. Note that the control variable is $\delta v(t_0)$ expressed in spectral coordinates. Given a guess for $\delta v(t_0)$, the cost function and its gradient are computed using (4) and (5). Then, the quasi-Newton minimizer finds a better guess for $\delta v(t_0)$, which is used to compute the next cost function and its gradient value. This process is repeated till the minimization criterion is satisfied. In addition, we approximate the innovation vector with the full nonlinear model as

$$\mathbf{H}\mathbf{L}(t_0,t_i)\mathbf{U}^{-1}\delta v-d(t_i)\approx\mathbf{H}M(x(t_0))-y(t_i) \qquad (7)$$

This is based on the assumption that the analysis state is close to the background state.

### *2.3 Hybrid schemes*

In the 3D-Var scheme, the background error covariance is time-independent and therefore does not include "errors of the day". The breeding method proposed by Toth and Kalnay (1993 and 1997) can generate the fast growing perturbations in a nonlinear dynamic system for ensemble prediction. Bred vectors, defined as the differences between perturbed and non-perturbed nonlinear runs, carry information on "errors of the day", since both are related to flow instabilities. This was confirmed by Corazza et al (2003) who found that the background error strongly projects on the subspace spanned by 20-bred vectors (BV). Therefore, it is reasonable to use BVs to provide the structure of evolving "errors of the day" and augment the time-independent background error covariance in the 3D-Var scheme. The new background error covariance (7) is a linear combination of the 3D-Var background error covariance with an error covariance constructed from a BV ensemble weighted by a coefficient $\alpha$. The additional covariance **B'** is given by the inner product of BV ensemble that represent flow dependent growing perturbations.

$$\mathbf{B} = (1-\alpha)\mathbf{B}_{3DVAR} + \alpha\mathbf{B}' \quad (7)$$

$$\mathbf{B}' = \mathbf{b}\mathbf{b}^T \quad (8)$$

Corazza et al. (2002) proposed two approaches to augment the background error covariance in the 3D-Var scheme: by directly applying the full rank (global) **B**' or by localizing the covariance structure.

i.*Global approach*

In the global approach, the augmented part is directly inserted in (1). In (7), $\alpha$ is a tuning parameter. The analysis increment is considered to be as coming from two parts and is solved separately to obtain the correction from regular 3D-Var (**B**) and from BV error covariance (**B**') system. Then, corrections from both are linearly combined with the weighting coefficient $\alpha$.

$$(\mathrm{I} + ((1-\alpha)\mathrm{B}_{3DVAR} + \alpha c\mathrm{bb}^{\mathrm{T}})\mathrm{H}^T\mathrm{R}^{-1}\mathrm{H})(x_a - x_b) = ((1-\alpha)\mathrm{B}_{3DVAR} + \alpha c\mathrm{bb}^{\mathrm{T}})\mathrm{H}^T\mathrm{R}^{-1}(y - \mathrm{H}x_b) \quad (9)$$

In (9) $c$ is a normalization factor that ensures the error covariance has a reasonable magnitude. Results from Corazza et al. (2002) suggested that the optimal value for $\alpha$ is 0.4, but in our experiments with more careful tuning, it increases to 0.7 (section 3). In addition, a small amount of random perturbations is added to the bred vectors for every breeding cycle in order to "refresh" BVs and slightly perturb the subspace spanned by BVs. This reduces the tendency of BVs to converge to a too small dimension (Wang et al, 2003). The perturbations are created at the observation locations and converted back to the grid points with the transpose of the observation operator. We found that this not only accelerates the analysis error convergence but also leads to a lower rms error level. By default, the size of the random perturbation has the amplitude of the observation error. In order to avoid excessive noise added when the flow is less unstable, the size of the random perturbation is made adaptive, i.e., proportional to the amplitude of the observation increment $(y - \mathbf{H}x_b)$.

A similar hybrid system by Hamill and Snyder (2000) showed that the ensemble-based background error covariance could completely replace the background error covariance of 3D-Var if the ensemble is large enough. However, since the errors of the day lie on a low-dimension attractor due to the flow instabilities, the need for a large ensemble to sustain the rank of the error covariance can be lessened if the ensemble perturbations are chosen to represent the shape of background error. Corazza et al. (2003) found that the *local* dimensionality computed based on 10 BVs is always lower than six (out of a maximum of 10) and has an average value equal to about four. These values do not change when the number of bred vectors is doubled, indicating that ten of them are able to provide most of the necessary (local, not global) information.

*ii. Local approach*

The fact that the local dimension of the bred vectors is low suggests that a local approach may lead to a good description of the background error with a relatively small ensemble (Kalnay and Toth, 1994, Ott et al, 2004). In the local approach, (1) is reformulated based on a local patch. The local background error covariance, $\hat{\mathbf{B}}$, has a form like (8), but is constructed by means of the local BV ensemble, $\hat{\mathbf{b}}$. The local patch we choose has a size of 7x7 by 7 levels and then each local BV is reshaped to a long vector with a dimension of L (equal to 7x7x7 in this case). The local domain we choose is small enough to allow for a singular value decomposition so that

$$\hat{\mathbf{b}} = \mathbf{V}\mathbf{S}\mathbf{W}^T \quad (10)$$

With K bred vectors, the matrix **V** has a dimension of L by K, storing the K left singular vectors columnwise. The K corresponding singular values are stored on the diagonal elements of the matrix **S.** An Ensemble-Dimension analysis (Patil et al. 2001) is used to define the number Kc of leading singular vectors that will be retained. In the experiement with 10 bred vectors, typical Kc ranges from 3 to 7. We then assume that the local analysis increment can be represented by the linear combination of these Kc singular vectors as indicated by (11) with the Kc coefficients stored in the column vector $\mathbf{\Gamma}$.

$$(\hat{x}_a - \hat{x}_b) = \mathbf{V}\Gamma \quad (11)$$

With the orthogonality property of the local background error covariance, we can project (1) in a KcxKc space, indicated as (12) and solve for $\mathbf{\Gamma}$.

$$[\hat{\mathbf{B}}'^{-1} + (\mathbf{HV})^{\mathbf{T}}\hat{\mathbf{R}}^{-1}(\mathbf{HV})]\Gamma = (\mathbf{HV})^{\mathbf{T}}\hat{\mathbf{R}}^{-1}(\hat{y} - H(\hat{x}_b)) \quad (12)$$

In (12), $\hat{\mathbf{B}}'^{-1}$ is now simply a diagonal matrix whose elements are the inverse of the square of **S** and is trivial to compute. This formula is similar to the singular evolutive extended Kalman (SEEK) filter introduced by Pham et al (1998) that computes Kalman Filter in the rank-reduced space spanned by several leading singular vectors. The analysis increment obtained according to (12) will also be linearly combined with the analysis increment from 3D-Var with a weighting coefficient. As in the global approach, bred vectors will be "refreshed" with a small amount of random perturbations. Unlike the global approach, the random perturbations are now introduced at the model grid points.

### *2.4 Local ensemble Kalman Filter*

The LEKF system implemented for this work is based on the scheme described in Ott et al. (2004) and we refer to this paper for a detailed

description. Here we only describe the specific implementation related to the use of the QG-model adopted in all the experiments of this study. The LEKF system is based on the use of an ensemble of models and on the decomposition of the representation of the flow on local domains, where the data assimilation step and the construction of the new vectors of the ensemble is performed simultaneously. The process can be schematically represented as follows:

1. Localize the *K* ensemble members after removing the (global) ensemble mean from each ensemble member;
2. Define the local background error covariance matrix $\mathbf{P}_l^b$ by means of the local vectors (here *b* represents background and *l* represents the *l*-th local patch);
3. Decompose the local vectors in order to find the most unstable directions of the local space. Now, $\mathbf{P}_l^b$ is transformed onto a K-dimensional space spanned by those directions and becomes $\hat{\mathbf{P}}_l^b$, a diagonal matrix in this new K-dimension coordinates.
4. Solve the local analysis error covariance in the K-dimension coordinates. The local analysis ensemble is then derived for two parts: the mean analysis state and the ensemble perturbations. Note that the local analysis perturbations are a linear combination of the local background perturbations (also in the K-dimension).
5. Transform the local analysis ensemble back to the global space but using only the analysis at the center of this local patch to update the global analysis ensemble.

The main issues related to the technical implementation of the LEKF for the QG-model are due to the structure of the "Observation Operator" $\mathbf{H}$, which maps from the model space, where the model variables are potential vorticity *q* and potential temperature $\theta$ at the top and bottom boundaries, to the observational space, where the observation variables are the horizontal velocity components *u* and *v*, and temperature *T*. This mapping requires the computation of the streamfunction and the solution of a Poisson equation. In a global approach, as in the case of the original 3D-Var developed by Morss (1998), this problem can be solved in the entire domain and, in particular, the Poisson solver can be simplified working in the spectral space. This is not true in the framework of a local domain, since the fast Fourier transform cannot be applied. Moreover, even adopting a finite difference approach for the Poisson solver, it is not possible to avoid large errors due to the dominance of the lateral boundary conditions.

For these reasons we have chosen to solve $\mathbf{H}$ using a different approach, following the scheme described by Corazza et al. (2005). This solution is based on the fact that $\mathbf{H}$ can be split in two parts: $\mathbf{H}=\overline{\mathbf{H}}\tilde{\mathbf{H}}$,where $\tilde{\mathbf{H}}$ is the operator that maps from *q* and $\theta$ to *u*, *v* and *T* in the global domain, and $\overline{\mathbf{H}}$ is the operator mapping from the model domain (the model grid) to the observational grid. The localization of $\overline{\mathbf{H}}$ is then trivial, since only an interpolation is required, and this leads to implementing the LEKF by first applying the global operator $\tilde{\mathbf{H}}$ to all the ensemble vectors, and then performing the rest of the data assimilation as if the model variables were *u*, *v* and *T*. Once the new global vectors are computed a global $\tilde{\mathbf{H}}^T$ can be applied in order to go back to the QG-model variables.

| | |
|---|---|
| Horizontal dimension of the local domain | *l*=3 |
| Number of members in the ensemble | *K*=30 |
| Method to update the global field | average |
| Amplitude of the random perturbations | ~5% vectors amplitude |

Table 1**:** Settings adopted for the LEKF system for the simulations described in this work.

The choice of the main parameters adopted for the LEKF implementation follows the study performed by Corazza et al. (2005), where different versions of this system using the same QG-model discussed in this work are presented. In Table 1 the most important choices are summarized. In particular, the local domain are chosen to be columns of 7×7×7 grid points (*l*=3) and no vertical localization is performed. In the final part of the data assimilation process the global fields are built as the average of the contributions provided by the local domains containing the considered grid point.

In order to maintain the space spanned by the ensemble sufficiently large, we have chosen to add random perturbations to the global vectors after the data assimilation step. In particular, using an approach similar to the one described in section 2.3.i, in order to simulate the contribution introduced by observational errors, random noise is generated in the observational space and then transferred to the model space by applying the global operator $\mathbf{H}^T$. Even though the relative amplitude of the perturbations respect to the original vectors is small (of the order of 5% of the field variability), this method provides much better results than simple inflation, at least for this implementation.

## 3 RESULTS

In the following results, we have 64 rawinsonde observations randomly distributed in the domain

and the analysis cycle is 12-hour. The observation covariance is constructed in the same fashion as Dey and Morone (1985), where the observation errors are assumed to be uncorrelated between different observations and between wind and temperature. Error correlations are included in the vertical and only for the same variable. In our experiments, the magnitude of the observation errors is doubled respect to that used in Morss (1998) and Corazza et al (2002).

### *3.1 3D-Var vs. the hybrid schemes*

Corazza (2002) obtained that augmenting the background error covariance with an ensemble of 20 global bred vectors reduced by 22% the 3D-Var mean squared error of the potential vorticity at level 3. As a benchmark, we also tested the algorithm replacing the bred vectors with random vectors that do not provide any information about the errors of the day, since no forward integration is involved in random vectors. Figure 1 shows the results in terms of the rms error of the enstrophy norm using 20 bred (or random) vectors with different weighting ranging from 0.1 to 0.8. The hybrid system with bred vectors blew up if the weighting coefficient was 0.9 or larger.

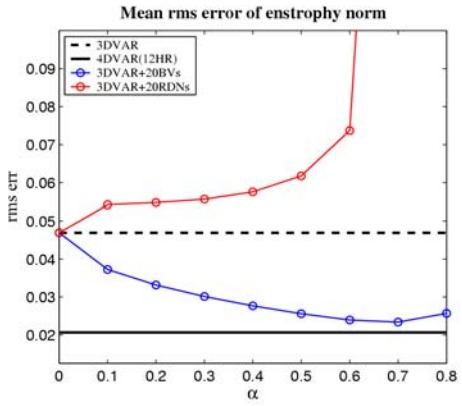


**Figure 1 Mean rms error of enstrophy norm for global hybrid 3DVAR system with 20 bred vectors (blue line) and 20 random noise vectors (red line). The dashed line denotes the rms error level of standard 3D-Var and the solid line denotes the 4D-Var error level using 12-hour assimilation window. Results are averaged from day 45 to day 200. The hybrid scheme returns to regular 3D-Var as $\alpha$ equals to 0.**

It is clear that the reason for the improvement obtained by the hybrid scheme is that it captures the shapes of the background error covariance dominated by flow dependent properties. In Corazza et al (2003), it was demonstrated that the forecast error strongly projects on the subspace spanned by bred vectors. This allows us to correct the background state by moving it closer to the observations along the unstable subspace. As indicated in (9), the analysis corrections associated with flow-dependent property are computed in the bred vectors subspace ($\mathbf{bb}^T$). On the other hand, the random vectors, which do not provide useful information, worsen the correction and degrade the rms error, as could be expected. We also notice that the hybrid system with random vectors tends to worsen the large spikes that occur in the results of the 3D-Var system. Large error spikes occur when the flow undergoes a rapid growth of the dynamic instability, which is not expected in a time-independent background error covariance matrix, such the one used in 3D-Var. In contrast, such large analysis error growth can be avoided in the ensemble-based data assimilation schemes, since they are able to detect the "errors of the day".

We note that the optimal weighting coefficient is higher ($\alpha$=0.7) than the one ($\alpha$=0.4) found in Corazza et al. (2002). This is due to the increase of the observation error in our experiments, which encourages the system to trust the background error covariance more and give more weights to the BV's structures. In addition, adding a random perturbation of constant size is not optimal. Instead, it is preferable to make the amplitude adaptive depending on the observational increment that indicates if larger random perturbations are required because the background flow is more unstable and uncertain. It is notable in Figure 1 that the results of hybrid scheme with weighting coefficients larger than 0.5 are competitive with those obtained by 4D-Var with a 12-hour window time. However, there is huge difference in the computational cost between these two schemes. The 3D-Var used about 1 hour and the hybrid scheme about 1.4 hour, but the 4D-Var needed 36 hours to compute 400 analysis cycles under the same computational configuration.

The number of bred vectors needed to represent a field of growing errors in a local domain is much smaller than the number needed in a global domain scheme (Kalnay et al, 2002, Ott et al, 2004). We can therefore expect the local approach to be able to capture the growing instabilities more efficiently than the global approach with the same number of bred vectors. Although one might expect the local hybrid approach to have errors similar to the global approach (section 2.3), the fact that the local scheme is more flexible and requires fewer ensemble members than the global scheme makes it more efficient. Figure 2 compares the results obtained using the global and local approaches. Results from 3D-Var and 4D-Var are also included for reference. The experiment for the local approach uses a local domain made of 7x7 grid points by 7 levels. The obtained local analysis is added to the regular 3D-Var analysis using a weight equal to 0.4. With 10 BVs, we found that the result of the local approach in fact outperforms those obtained by the global approach by successfully avoiding the large error spikes that characterize the 3D-Var system. Our results show that the local approach with 10 BVs has a better performance than the global approach with 20 BVs and even the

4D-Var with a 12-hour assimilation window time. The computational time for the local hybrid system is about 12 hours. In operational, this scheme is easy to implement and also computed parallelly, so that we can reduce the computational time. Our results show that this local hybrid approach has improved the regular 3D-Var by about 70%, which is much more accrurate than the hybrid system with 10 global bred vectors (table 1).

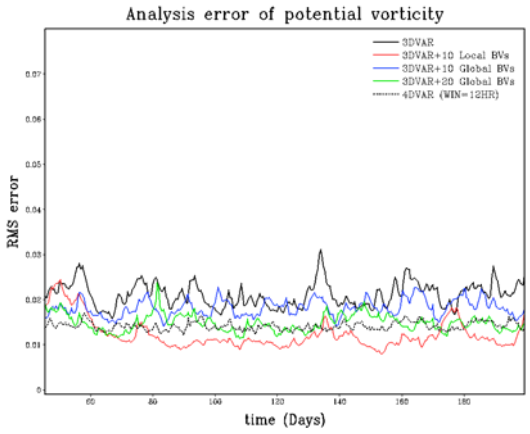


**Figure 2 Time series of rms error of potential vorticity from 3D-Var, global (10 and 20 BVs) and local (10 BVs) hybrid schemes and 4D-Var (12 hour window) systems.**

### *3.2 4D-Var and LEKF*

In this subsection, we focus on the comparison of the data assimilation schemes considered to be the most advanced.

In 4D-Var, **B** in (2) is usually set equal to $\mathbf{B}_{3D\text{-}Var}$. However, the performance of 4D-Var is sensitive to the size of **B**. When 4D-Var is cycled, it provides a better first guess than 3D-Var, and therefore $\mathbf{B}_{3D\text{-}Var}$ is too large. Therefore, we retained the structure of $\mathbf{B}_{3D\text{-}Var}$ but tuned its amplitude, and found that for a 12 hr assimilation window, $\mathbf{B=0.4\ B}_{3D\text{-}Var}$ was optimal. Reducing the amplitude of **B** should also benefit the hybrid scheme using a large number of ensemble members (20BVs), but would degrade the performance of the regular 3D-Var and of the hybrid system using a small ensemble. Longer windows are more beneficial (but more costly) for 4D-Var (Pires et al, 1996), and for longer windows the sensitivity to **B** decreases. Our results show that a 36-hour assimilation window is the optimal configuration for our 4D-Var system. For assimilation windows longer than 48 hours, the results become worse because the minimizer has difficulties to distinguish between multiple minima of the cost function.

Figure 3 shows the potential vorticity rms errors using LEKF, 36-hour 4D-Var, 3D-Var and the local hybrid scheme with 10 BVs. Both 4D-Var and LEKF, after an initial spindown, are characterized by stable error and they both outperform 3D-Var. The local hybrid scheme, though more accurate than the 4D-Var with the 12-hour assimilation window, is somewhat worse than the 36-hour 4D-Var. Although the initial spin-down of the error is faster for the variational schemes, the LEKF converges to the lowest rms error level. The long spin-down of LEKF is due to the choice of the initial condition (zonal mean state in our experiments), which is too far away from the truth, so that the shapes of instability captured by ensemble members are not representative of the real errors. However, in practice, the initial condition is much closer to the truth and the transition period should not be a serious concern. For example, one can use the analysis from 3D-Var as the initial guess to start LEKF. Also, our results suggest that adding random perturbations to the ensemble members reduces the rank deficiency problem of ensemble Kalman filter.

Figure 4 shows vertical profiles of the rms errors of potential vorticity, indicating that the advantage of LEKF over 4D-Var is larger at lower levels. The hybrid scheme improves substantially over 3D-Var, but the vertical distribution of rms error is evidently influenced by the 3D-Var errors. For LEKF, the improvement is uniform in the vertical. The largest differences between LEKF and 4D-Var appear in the analysis fields of the bottom potential temperature. In our 4D-Var system, the adjoint operator includes all physical processes. The adjoint of the Ekman pumping forcing, which is the dominant physical process at the bottom level, may be the origin of the larger bottom temperature errors.

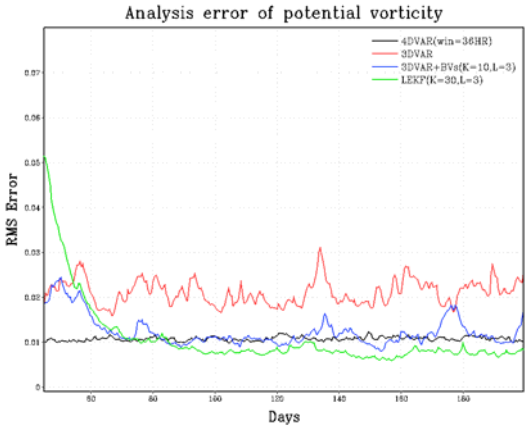


**Figure 3 Time series of rms errors of potential vorticity from 3D-Var, the local hybrid schemes, 4D-Var systems and LEKF.**

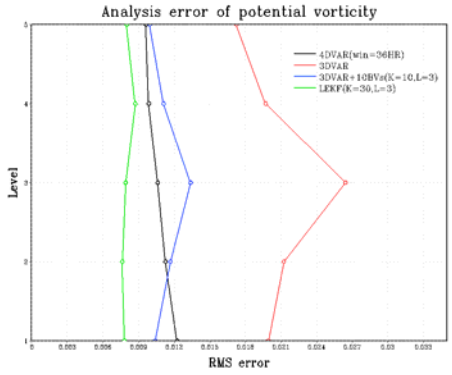


**Figure 4 The Vertical profiles of rms errors of potential vorticity from 3D-Var, the local hybrid schemes, 36-hour 4D-Var systems and LEKF. Results are averaged from day 75 to day 200.**

We summarize our experiments in Table 1 by the mean rms error for the last 250 analysis cycles. In this table, it clearly shows that the schemes involved with localizations give the best corrections in either the hybrid systems or the advanced data assimilation schemes.

## 4 SUMMARY AND DISCUSSION

In this study, data assimilation schemes related to variational and ensemble methods are implemented for a quasi-geostrophic model. Four different schemes are discussed: 3D-Var, 4D-Var, 3D-Var hybridized with bred vectors, and LEKF. The goal of this study is not only to compare individual performances but also to try to understand advantages and disadvantages for practical implementation in operational systems, as we know that 4D-Var has already been implemented in several operational centers, and Ensemble Kalman Filters are regarded as the possible next phase data assimilation systems.

Our results are discussed in terms of the rms error in potential vorticity (which is a model variable) for all the data assimilation experiments. Given the same rawinsonde observations, the LEKF scheme outperformed 4D-Var in both accuracy and computational cost. This scheme converged to the lowest rms error level, and the 4D-Var with a long assimilation window of 36 hour reached the second lowest error level. However, the slow convergence rate of the LEKF suggests that that an initial guess sufficiently close to the true state of the atmosphere (such as that obtained from a 3D-Var assimilation cycle) is a key factor for fast convergence. Results from 4D-Var experiments suggest that the amplitude of the background error covariance and the length of the assimilation window greatly influence its performance. However, for a high-resolution model with full physics, a long assimilation window is too costly. The minimization process requires special care with long windows by gradually lengthening the window length in order to deal with multiple minima (Pires et al., 1996). Also, the minimization problem is more difficult if the cost function is ill-conditioned and the chosen preconditioner cannot effectively reduce the condition number, so that the minimizer cannot converge, resulting into too many iterations.

We also found that the hybrid system with BVs can effectively improve the regular 3D-Var system by suppressing the spurious error spikes. With 20 bred vectors (global approach), the hybrid system is competitive with 12-hour window 4D-Var, and has a very low computational cost. A local hybrid system with 10 bred vectors competes even with the 4D-Var with an optimal 36-hour window. The success of the hybrid scheme using a rather small ensemble suggests that including the background error covariance from 3D-Var provides an important base for the representation of the average behavior of background error. This may be also helpful for methods using ensemble Kalman filters that suffer from rank-deficiency or sampling problems (Hamill and Snyder, 2002). Our results also indicate that a small amount of random perturbations can be helpful for stabilizing the scheme and reducing the error level for all ensemble-related schemes. Variational schemes, including the hybrid one, have larger vertical dependence of the error than the LEKF.

More experiments have to be performed, in particular in order to take into account more sophisticated features in the schemes and, most importantly, in order to abandon a perfect model approach. As far as the LEKF is concerned, it is important to understand the potential improvement that can arise from the introduction of observation localization (Miyoshi, 2005) and the impact due to a vertical localization of the model domain. In fact, for the QG-model this latter feature was of minor importance; though we expect that for a real model, including boundary layer description and more sophisticated physics, vertical localization can have a major impact on the performance of the LEKF.

**Table 1 Mean rms error of the potential vorticity for different data assimilation scheme, averaged from the last 250 analysis cycles. (GBV: the hybrid system with global bred vector, LBV: the hybrid system with local bred vector and W: the assimilation window time used in 4D-Var)**

| | 3DVar | 3DVar +20GBV | 3DVar +10GBV | 3DVar +10LBV | 4DVar (w=12hr) | 4DVar (w=36hr) | LEKF |
|---|---|---|---|---|---|---|---|
| Rms error | 0.0468 | 0.0206 | 0.0347 | 0.0137 | 0.0189 | 0.0120 | 0.0068 |

***Acknowledgement***
We are deeply grateful to Rebecca Morss for providing the QG-model and the 3D-Var system. S.-C. Yang was supported by a NASA grant NNG004GK78A.

**REFERENCES**

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.

Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4DVAR, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**,1367-1387.

Corazza, M., E. Kalnay, D. J. Patil, E. Ott, J. Yorke, I. Szunyogh, M. Cai, 2002: Use of the breeding technique in the estimation of the background error covariance matrix for a quasigeostrophic model. AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction, Orland, Florida, 154-157.

------., 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.

----------, E. Kalnay, and S.-C. Yang, 2005: An implementation of the Local Ensemble Kalman Filter for a simple quasi-geostrophic model. Results and differences with a 3D-Var data assimilation system, to be submitted.

Dey, C. and L. L. Morone, 1985: Evolution of the national meteorological center global data assimilation system: January 1982-December 1983. *Mon. Wea. Rev.*, **113**, 304-318.

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.

------, 2003: The ensemble Kalman filter: theoretical formulation and pratical implementation. *Ocean Dynamics.* **53**, 343-367.

Hamill, T. M., 2003: Ensemble-based data assimilation: a review. University of Colorado and NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA.

------, and C. Snyder, 2000: A hybrid Ensemble Kalman Filter 3D Variational analysis scheme. *Mon. Rev. Rev.*, **128**, 2905-2919.

------, C. Snyder and R. E. Morss. 2000: A Comparison of probabilistic forecasts from bred, singular-Vector, and perturbed observation ensembles. *Mon. Wea. Rev.*, **128**, 1835–1851.

Houtekamer, P. L. and H.L. Mitchell, 1998: Data assimilation using ensemble Kalman filter technique. *Mon. Rev. Rev.*, **126**, 796-811.

Kalnay, E., Corazza, M., and Cai, M.: Are bred vectors the same as Lyapunov vectors?, in AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction, pp. 173–177, Orlando, Florida, 2002.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang and J. Ballabrer, 2005: 4D-Var or Ensemble Kalman Filter? Submitted to Physica D.

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

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a Primitive-Equation global model. Doctoral dissertation, University of Maryland, College Park, 197pp. Available at http://www.atmos.umd.edu/ ~ekalnay.

Morss, R. E. 1998: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. PhD thesis, Massachusetts Institute of Technology. 225 pp.

------, K. A. Emanuel and C. Snyder, 2001: Idealized adaptive observation strategies for improving numerical weather prediction. *J. Atmos. Sci.*, **58**, 210–232.

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.

Pam, D. T., Verron, J. and Roubaud, M.-C., 1998: Singular Evolutive Extended Kalman filter with EOF initialization for data assimilation in oceanography. *J. Marine Systems*, **16**, 323-340.

Parrish, D. and J. Derber, 1992: The National Meteorology 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, H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational impletation of four-dimensional variational assimilation. I: experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**,1143-1170.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meterological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**,1311-1328.

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.

Yang, S.-C., D. Baker, H. Li, M. Huff, G. Nagpal, E. Okereke, J. Villafane, E. Kalnay and G. Duane, 2005: Data assimilation as synchronization of truth and model: experiments with the 3-variable Lorenz system. Submitted to J. Atmos. Sci.

Yang, S.-C., 2005: Bred Vectors In The NASA NSIPP Global Coupled Model and Their Application To Coupled Ensemble Predictions And Data Assimilation. PhD thesis. Appendix B: Errors of the day, bred vectors and singular vectors in a QG atmospheric model: implications for ensemble forecasting and data assimilation. University of Maryland, 174 pages. Available at http:// www.atmos.umd.edu/~ekalnay