

An Ensemble-Based Four-Dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test

CHENGSI LIU

LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, and NCAR/ESSL/MMM, Boulder, Colorado*

QINGNONG XIAO

NCAR/ESSL/MMM, Boulder, Colorado*

BIN WANG

LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

(Manuscript received 19 July 2007, in final form 27 November 2007)

ABSTRACT

Applying a flow-dependent background error covariance (**B** matrix) in variational data assimilation has been a topic of interest among researchers in recent years. In this paper, an ensemble-based four-dimensional variational (En4DVAR) algorithm, designed by the authors, is presented that uses a flow-dependent background error covariance matrix constructed by ensemble forecasts and performs 4DVAR optimization to produce a balanced analysis. A great advantage of this En4DVAR design over standard 4DVAR methods is that the tangent linear and adjoint models can be avoided in its formulation and implementation. In addition, it can be easily incorporated into variational data assimilation systems that are already in use at operational centers and among the research community.

A one-dimensional shallow water model was used for preliminary tests of the En4DVAR scheme. Compared with standard 4DVAR, the En4DVAR converges well and can produce results that are as good as those with 4DVAR but with far less computation cost in its minimization. In addition, a comparison of the results from En4DVAR with those from other data assimilation schemes [e.g., 3DVAR and ensemble Kalman filter (EnKF)] is made. The results show that the En4DVAR yields an analysis that is comparable to the widely used variational or ensemble data assimilation schemes and can be a promising approach for real-time applications.

In addition, experiments were carried out to test the sensitivities of EnKF and En4DVAR, whose background error covariance is estimated from the same ensemble forecasts. The experiments indicated that En4DVAR obtained reasonably sound analysis even with larger observation error, higher observation frequency, and more unbalanced background field.

1. Introduction

The background error covariance plays an important role in most data assimilation systems. Usually, the background error covariance model is applied, which is assumed to be homogeneous, isotropic, stationary, and

quasigeostrophic in its structure (Daley 1991) and which is determined by a few parameters that are estimated from innovation statistics (Hollingsworth and Lonnberg 1986) or the National Meteorological Center (NMC, now known as the National Centers for Environmental Prediction or NCEP) method (Parrish and Derber 1992). Although these assumptions make the error covariance statistics and their application to data assimilation much easier, it is evident that they are not always appropriate, especially under baroclinically unstable conditions.

To improve the background error covariance, two methods have recently been explored. One is to con-

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Dr. Qingnong Xiao, NCAR/ESSL/MMM, Boulder, CO 80307-3000.
E-mail: hsiao@ucar.edu

struct a more accurate background error covariance model by introducing flow-dependent features into the statistics. For example, Dee (1995) presented a scheme for the online tuning of parameters in the error covariance model based on a maximum-likelihood approach, in which a batch of simultaneous observations are used for making those parameters flow dependent. Xu et al. (2001) and Xu and Wei (2001, 2002) estimated the three-dimensional error covariance using a multilevel least squares fitting method. Fisher (2003) used a spectral method to allow spatially inhomogeneous vertical and horizontal correlation models. Derber et al. (2003) employed the recursive filter to create inhomogeneous and anisotropic background errors in the grid space. Another approach uses ensemble forecast statistics to produce flow-dependent background error covariances (Evensen 1994). Most ensemble-based data assimilation algorithms are based on Kalman filter theory (Kalman 1960), so they are generally called ensemble Kalman filtering (EnKF). There are several formulations of EnKF that have been developed in recent years, such as a double EnKF (Houtekamer and Mitchell 1998), an ensemble square root filter (EnSRF; Whitaker and Hamill 2002), an ensemble adjusted Kalman filter (EAKF; Anderson 2001), and an ensemble transform Kalman filter (ETKF; Bishop et al. 2001). In recent years, some researchers have employed ensemble-based nonsequential assimilation algorithms, which include temporal covariances within the assimilation window. Evensen and van Leeuwen (2000) presented an ensemble Kalman smoother, using only forward-in-time model integrations. Hunt et al. (2004) extended the ensemble Kalman filter to a four-dimensional assimilation form (4DEnKF) so that EnKF could handle all observations within an assimilation window. Zupanski (2005) presented the maximum-likelihood ensemble filter (MLEF), which employed the ensembles needed to calculate the Hessian preconditioning and gradient of the cost function so that the background error covariance from the ensemble statistics could be used in a variational scheme.

Comparing with the deterministic data assimilation, the ensemble-based data assimilation can easily provide the analysis probability distribution function (pdf) sampled by the ensemble as the initial condition of the ensemble forecast. EnKF has been a popular research topic in the numerical weather prediction (NWP) field during recent years. It has been proven that both 4DVAR and a Kalman filter can produce the same results at the end of an assimilation window (Hamill 2002; Li and Navon 2001) under the following conditions: (a) The observation operator and model are linear; (b) the observation and background have Gauss-

ian, unbiased random errors; (c) the model error is neglected; and (d) the same background error covariance is used. When compared with the 4DVAR technique, EnKF has its advantage; it does not need any tangent linear or adjoint model, as is necessary in 4DVAR but not easily coded.

Although many encouraging research results of EnKF have been obtained in either global data assimilation (Mitchell et al. 2002; Houtekamer and Mitchell 2005) or regional mesoscale data assimilation (Snyder and Zhang 2003), there is not much evidence to indicate that EnKF outperforms variational data assimilation systems in operational applications. Since variational data assimilation has been practically proven to be a very successful technique in operational NWP, the application of ensemble-based background error covariance statistics to variational data assimilation should be a good choice, which can help in extracting the flow-dependent error covariance from ensemble forecasts (like EnKF) to the analysis. Lorenc (2003) proposed that the control variable was preconditioned upon the perturbation of the background ensemble forecast so that the background error covariance in the variational system is flow dependent. Buehner (2005) adopted a similar hybrid ensemble 3DVAR (En3DVAR) scheme, and implemented it to a global model data assimilation system. He showed that the scheme could produce results that were similar to EnKF but better than 3DVAR. Both Lorenc (2003) and Buehner (2005) mentioned that 4DVAR with the background error covariance statistics from ensemble forecasts (En4DVAR) should be developed if the tangent-linear and adjoint models are available. There are four-dimensional ensemble-based assimilation algorithms, such as MLEF (Zupanski 2005) or 4DEnKF (Hunt et al. 2004; Fertig et al. 2007), that have shown that the background error covariance based on statistics from ensemble forecasts can be used in variational algorithms. The 4DVAR method can gain the optimal trajectory as a nonsequential assimilation algorithm, which can effectively assimilate high temporal-spatial resolution data, for example, satellite observations (Xiao et al. 2002; Simmons and Hollingsworth 2002). In addition to the advantages from the 4DVAR approach, En4DVAR can also offer benefits from the EnKF data assimilation technique, such as flow-dependent background error covariance and better ensemble initial fields. How to take advantage of the benefits from both approaches (EnKF and 4DVAR) and make the En4DVAR feasible and robust is the focus of this study. Practically, we should try to avoid the tedious tangent-linear and adjoint models in the proposed En4DVAR scheme, but without losing its nonsequential data assimilation character. Unlike

MLEF (Zupanski 2005) or 4DEnKF (Hunt et al. 2004; Fertig et al. 2007), our approach (En4DVAR) adopts the incremental and preconditional scheme in the variational algorithm so that it can be more easily incorporated into many of the operational centers that using variational assimilation as their operational data assimilation system. It may also be more easily encourage the variational data assimilation research community to include the advantages in ensemble-based data assimilation algorithms.

This paper is arranged as follows: In the next section, we briefly review EnKF and 4DVAR, and propose an En4DVAR technique. The new En4DVAR formulation uses background perturbations in observation space during the minimization iteration process. As derived in the following section, the En4DVAR avoids tangent linear and adjoint models while keeping the characteristics of 4DVAR. In section 3, some proof-of-concept tests with simple designs are presented. We evaluate the performance of the En4DVAR scheme and compare it with other data assimilation techniques using a one-dimensional shallow water model. The last section presents a summary and conclusions from this study.

2. En4DVAR formulation

a. EnKF analysis algorithm

The proposed En4DVAR scheme combines the necessary components from EnKF and 4DVAR. EnKF is a Kalman filter but using a Monte Carlo ensemble of short-range forecasts to estimate the covariance of the forecast error (Evensen 1994). Assuming there are N ensemble members, the background error is estimated by

$$\mathbf{X}'_b = \frac{1}{\sqrt{N-1}} (\mathbf{x}_{b1} - \bar{\mathbf{x}}_b, \mathbf{x}_{b2} - \bar{\mathbf{x}}_b, \dots, \mathbf{x}_{bN} - \bar{\mathbf{x}}_b), \quad (1)$$

where \mathbf{x} is a state vector (to ensemble-based assimilation, a matrix that each column represents one ensemble member state vector). The background error covariance is therefore approximately calculated by

$$\mathbf{B} \approx \mathbf{X}'_b \mathbf{X}'_b{}^T. \quad (2)$$

The final analysis of the EnKF is

$$\mathbf{X}_a = \mathbf{X}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{X}_b). \quad (3)$$

In the formulation of (1)–(3), subscript a denotes analysis, subscript b means background, N is the ensemble number, \mathbf{X}'_b is the matrix whose column is the normal-

ized deviations from the ensemble mean, \mathbf{B} is the background error covariance, \mathbf{O} is the observation error covariance, \mathbf{y} is the observation vector, \mathbf{T} represents the matrix transpose, H is the observation operator, and \mathbf{H} is the tangent linear observation operator.

In the EnKF implementation, Eqs. (1) and (2) are transformed to observation space as follows:

$$\mathbf{H}\mathbf{X}'_b \approx \frac{1}{\sqrt{N-1}} (\mathbf{H}\mathbf{x}_{b1} - \overline{\mathbf{H}\mathbf{x}_b}, \mathbf{H}\mathbf{x}_{b2} - \overline{\mathbf{H}\mathbf{x}_b}, \dots, \mathbf{H}\mathbf{x}_{bN} - \overline{\mathbf{H}\mathbf{x}_b}), \quad (4)$$

$$\mathbf{B}\mathbf{H}^T \approx \mathbf{X}'_b(\mathbf{H}\mathbf{X}'_b)^T, \quad \text{and} \quad (5)$$

$$\mathbf{H}\mathbf{B}\mathbf{H}^T \approx \mathbf{H}\mathbf{X}'_b(\mathbf{H}\mathbf{X}'_b)^T. \quad (6)$$

b. Incremental 4DVAR algorithm

Most variational data assimilation systems adopt an incremental approach (Courtier et al. 1994) and the preconditioning technique (Gilbert and Lemarechal 1989). Introducing the preconditioning matrix \mathbf{U} , and defining

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T, \quad (7)$$

we obtain

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{U}\mathbf{w}. \quad (8)$$

In 4DVAR, we introduce the forecast model, M , and observation vectors at different times (\mathbf{y}_i). The innovations at different times (with subscript i) are calculated by

$$\mathbf{d}_i = \mathbf{H}M(\mathbf{x}_b) - \mathbf{y}_i. \quad (9)$$

The cost function in control variable space becomes

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} \sum_{i=0}^I (\mathbf{H}\mathbf{M}\mathbf{U}\mathbf{w} + \mathbf{d}_i)^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{M}\mathbf{U}\mathbf{w} + \mathbf{d}_i), \quad (10)$$

where I is the total number of time levels on which observations are available. The gradient of the cost function with respect to the control variables is

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^I \mathbf{U}^T \mathbf{M}^T \mathbf{H}^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{M}\mathbf{U}\mathbf{w} + \mathbf{d}_i). \quad (11)$$

In Eqs. (10) and (11), \mathbf{M} is the tangent linear model and \mathbf{M}^T is the adjoint model. Calculation of the cost function gradient using Eq. (11) involves the integrations of the linear and adjoint models. Therefore, development of the tangent linear and adjoint models is usually necessary for the realization of the 4DVAR approach in data assimilation (Talagrand and Courtier 1987; Zupanski 1993; Zou and Kuo 1996; Xiao et al. 2000).

c. En4DVAR algorithm

In the variational algorithm, Lorenc (2003) employed the perturbation in an ensemble forecast that was calculated by Eq. (1) to precondition the control variable of the variation:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{X}'_b \mathbf{w}. \quad (12)$$

Defining the innovation as

$$\mathbf{d} = H(\mathbf{x}_b) - \mathbf{y}, \quad (13)$$

the cost function in control variable space is written as

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} (\mathbf{H}\mathbf{X}'_b \mathbf{w} + \mathbf{d})^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{X}'_b \mathbf{w} + \mathbf{d}), \quad (14)$$

where \mathbf{w} is the control variable and \mathbf{d} is an innovation that is the difference between the background and the observation in observation space. As Lorenc (2003) mentioned, this scheme is easily extended to 4DVAR form and the forward model is included if the background forecast is not at the same time as the observation and the sum notation will be added to the observation cost function. If the scheme is applied to 4DVAR, it will use a flow-dependent background error covariance at the beginning of the assimilation window and implicitly evolve within the window as 4DVAR does if the tangent linear model and adjoint model are available (Buehner 2005). In this study, we try to extend Eq. (14) to formulate an En4DVAR scheme.

Similar to Lorenc's (2003) definition of the cost function as shown in Eq. (14), we can use the background perturbation to precondition the 4DVAR control variables. Equation (10) is then rewritten as

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} \sum_{i=0}^I (\mathbf{H}\mathbf{M}\mathbf{X}'_b \mathbf{w} + \mathbf{d}_i)^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{M}\mathbf{X}'_b \mathbf{w} + \mathbf{d}_i). \quad (15)$$

This is the cost function of the proposed En4DVAR scheme. The gradient of the cost function with respect to the control variables becomes

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^I \mathbf{X}'_b{}^T \mathbf{M}^T \mathbf{H}^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{M}\mathbf{X}'_b \mathbf{w} + \mathbf{d}_i). \quad (16)$$

Although En4DVAR can be implemented by calculating the gradient using Eq. (16) and the flow-dependent background error covariance structure from the ensemble forecast can be obtained at the beginning of the 4DVAR time window, the tangent linear model \mathbf{M} and its adjoint \mathbf{M}^T have to be employed in the 4DVAR minimization.

To further formulate the En4DVAR scheme, we adopt the idea of EnKF and introduce the perturbation

in observation space. Similar to Eq. (4), the background error (estimated from the ensemble perturbations) can be transformed to observation space via

$$\mathbf{H}\mathbf{M}\mathbf{X}'_b \approx \frac{1}{\sqrt{N-1}} (H\mathbf{M}\mathbf{x}_{b1} - H\mathbf{M}\bar{\mathbf{x}}_b, H\mathbf{M}\mathbf{x}_{b2} - H\mathbf{M}\bar{\mathbf{x}}_b, \dots, H\mathbf{M}\mathbf{x}_{bN} - H\mathbf{M}\bar{\mathbf{x}}_b). \quad (17)$$

The gradient of the cost function (16) is reformulated as

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^I (\mathbf{H}\mathbf{M}\mathbf{X}'_b)^T \mathbf{O}^{-1} (\mathbf{H}\mathbf{M}\mathbf{X}'_b \mathbf{w} + \mathbf{d}_i). \quad (18)$$

In Eq. (18), the background error in observation space is calculated just once using ensemble forecasts outside the minimization iteration, so that the computational and coding costs are greatly reduced. We can see that the adjoint model \mathbf{M}^T is elegantly avoided in (18) by the transformation of the background error to observation space in (17). Moreover, Eq. (17) indicates that En4DVAR does not need a linear approximation in the forward model and observation operators.

According to (12), the analysis increment is actually a linear combination of the predicted ensemble perturbations. The coefficient of the linear combination, \mathbf{w} , can be estimated by minimizing the cost function in (14). After preconditioning, multiplying \mathbf{U} by its adjoint can be implemented by use of a recursive filter (Lorenc 1992; Hayden and Purser 1995). However, the recursive filter cannot be applied to Eqs. (16) or (18), because some error correlation models, for example, a Gaussian model, cannot be used. In Eq. (18), $\mathbf{H}\mathbf{M}\mathbf{X}'_b$ is a $m \times N$ dimensional matrix (m is the observation dimension). Since m is usually small (especially in regional models), the calculation of $\mathbf{H}\mathbf{M}\mathbf{X}'_b$ or $(\mathbf{H}\mathbf{M}\mathbf{X}'_b)^T$ is not too expensive. Moreover, a reduction of the dimensions in the En4DVAR control state vector makes the minimization cost even less.

Comparing 4DVAR and En4DVAR, some properties of En4DVAR can be exposed. In the 4DVAR scheme, the \mathbf{B} matrix is in full rank. The 4DVAR control variable \mathbf{w} is an n -dimensional state vector that is the same as the \mathbf{x}_b dimension (n is the model's degree of freedom). In En4DVAR, on the other hand, \mathbf{B} is estimated from the ensemble forecasts as shown in Eq. (2); it will be a degeneration matrix and its rank is not larger than the ensemble number N and its smallest eigenvalue is 0. If the estimated \mathbf{B} is applied to the no-precondition 4DVAR scheme, the condition number of the cost function becomes larger and the minimization is difficult to converge. A discussion of the condition number of the cost function and the minimization convergence can be found in Bouttier and Courtier (2007).

If using the perturbation precondition as shown in Eq. (12), the En4DVAR control vector dimension is N and the minimization of the cost function is in N -dimensional space. Therefore, the degenerated \mathbf{B} matrix can make the En4DVAR minimization converge efficiently.

Although the minimization can be realized in subspace spanned by an ensemble, the analysis is hardly extracting the full information from the control variable in Eq. (12) because the model space is usually larger than the ensemble space. This is similar to the so-called sample error problem in ensemble-based data assimilation. Many studies have shown that the sample error is a challenge to ensemble-based data assimilation. Several approaches, such as the Schur product (Houtekamer and Mitchell 2001; Lorenc 2003; Buehner 2005), local truncation (Houtekamer and Mitchell 1998), an inflation factor (Anderson and Anderson 1999), and a hybrid scheme (Hamill and Snyder 2001; Lorenc 2003) have been proposed to relax this problem. In an En4DVAR implementation with high-dimensional space, it is necessary to use localization to reduce the sample error. If localization is not used and the model is linear, the covariance at any later time in the assimilation window will be equal to those obtained from evolving the initial ensemble members. That is, the covariance will be restricted to the low-dimensional subspace spanned by the evolved ensemble members. However, if localization is applied to the covariance at the beginning of the window, the standard 4DVAR or En4DVAR schemes using the tangent linear and adjoint models [Eq. (16)] will implicitly evolve these covariances, which, due to the localization, span a much higher dimensional space. But this approach will come with a larger computation cost due to tangent linear and adjoint model integration. Lorenc (2003) and Buehner (2005) have indicated that the Schur operator can be used in ensemble-based variational algorithms. Similar to the truncated spectral expansion idea in appendix B of Buehner (2005), the Schur operator after EOF decomposition can be adopted for En4DVAR localization, and the computer cost can be reduced. We will employ this technique for localization in our real-model experiment with the Weather Research and Forecasting model (WRF).

3. Experiments with a one-dimensional shallow water model

In this section, we present preliminary results of the En4DVAR scheme and compare them with the results from other data assimilation schemes based on a discrete version of the one-dimensional shallow water model:

$$\frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} - fv + \frac{\partial \Phi}{\partial x} = 0, \quad (19)$$

$$\frac{\partial v}{\partial t} + U \frac{\partial v}{\partial x} + fu = 0, \quad \text{and} \quad (20)$$

$$\frac{\partial \Phi}{\partial t} + U \frac{\partial \Phi}{\partial x} + \Phi \frac{\partial u}{\partial x} = 0. \quad (21)$$

The comparisons among variational and ensemble data assimilation schemes are made under the condition that makes the Kalman filter and 4DVAR equivalent, which is described in section 1.

a. Experiment design

The advection terms in (19)–(21) are linearized with constant U (17 m s^{-1}). The Coriolis parameter f is $1.03 \times 10^{-4} \text{ s}^{-1}$. The model is discretized with a uniformly spaced grid, using a horizontal resolution of 300 km and 20 grid points ($D = 20$). The time step for the integration is 600 s. We solve (19)–(21) numerically with an Euler backward-in-time integration scheme.

The initial height field is defined as

$$\Phi_j^0 = \Phi_0 - f\bar{U}d + Af \sin\left[\frac{4\pi(j-1)}{D-1}\right] \sin\left(\frac{8\pi}{17}\right), \quad (22)$$

where $\Phi_0 = 5.5 \times 10^4 \text{ m}^2 \text{ s}^{-2}$, $\bar{U} = 20 \text{ m s}^{-1}$, and $A = (1.5/\pi) \times 10^{-7} \text{ m}^2 \text{ s}^{-2}$. We use j as the grid number. The initial v component of the wind is derived from the height with the geostrophic relation, and the u wind is the horizontal derivative of the v wind. For a quasigeostrophic solution to (19)–(21), the velocity component u is at least an order of magnitude smaller than v , and we will therefore completely ignore u in the presentation of our results.

Figure 1 shows a schematic diagram of our experimental design. The 1D shallow water model is integrated for 132 h with the specified initial conditions, and the results are taken as the “true” atmospheric-state evolution. With the simulated true atmospheric state, we set up an observing network that provides height and v wind observations at each grid point for every 6 h. The simulated observation errors are specified as $5 \text{ m}^2 \text{ s}^{-2}$ and 0.03 m s^{-1} , respectively. The control evolution is obtained from the integration of the 1D shallow water model by adding normal random perturbations to the specified initial field. The standard deviations (spreads) of the normal random perturbations for the height and the v component of the wind are $80 \text{ m}^2 \text{ s}^{-2}$ and 0.1 m s^{-1} , respectively. With the same standard deviations of the normal random perturbations, we construct 50 members of the ensemble initial conditions, and the 50-member ensemble forecast is obtained with the same 1D shallow water model integrations.

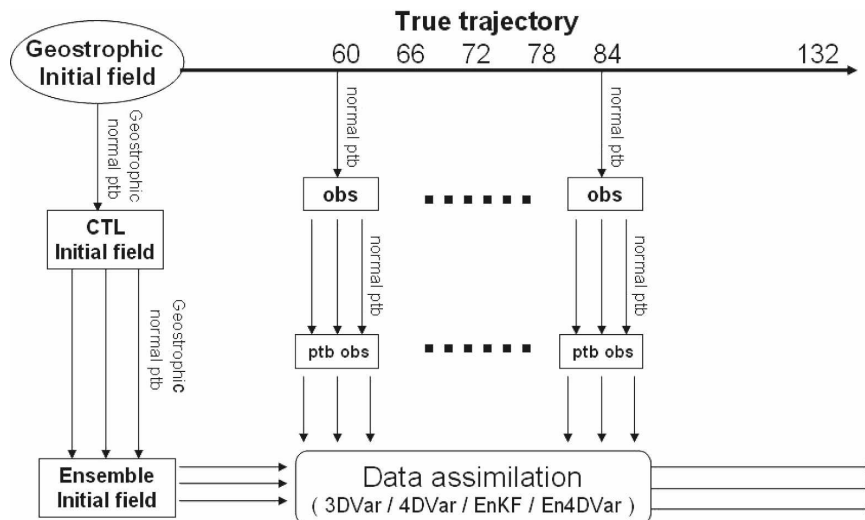


FIG. 1. The experimental design flow chart. The true trajectory is obtained from a geostrophic initial field. The initial perturbation (noted as ptb) is added to the model in order to obtain the control (noted CTL) initial field. The data assimilation window is set between 60 and 84 h. The observations (noted as obs) are available at every 6 h by adding normal perturbations to the true state. A set of similar perturbations is added to the CTL initial field and the observations. The ensemble initial field and the perturbation observations are provided to the ensemble-based data assimilation, e.g., EnKF and En4DVAR. For deterministic data assimilations (3DVAR and 4DVAR), the ensemble background mean and ensemble observation mean are adopted.

After 60 h of integration, the spreads of the ensemble forecast for height and the v component of the wind are reduced to $20 \text{ m}^2 \text{ s}^{-2}$ and 0.1 m s^{-1} , respectively. Since the ensemble forecast at 60 h is more balanced, we use the result as a background field for all assimilation experiments (i.e., data assimilation starts at 60 h). The assimilation window is set at 24 h. A 48-h forecast is run after the data assimilation. Different data assimilation schemes (3DVAR, 4DVAR, EnKF, and En4DVAR, proposed in this paper) are applied. All variational schemes use the steepest descent method for minimization of the cost function (Snyman 2005). The localization is not applied in the following ensemble-based assimilation experiments because they are performed in a low-dimensional space. Applying localization gives few positive impacts in these experiments. However, when we implement En4DVAR in real-dimensional space, the Schur operator localization scheme can reduce noise effectively. The results using En4DVAR localization in real-dimensional space will be shown in a future paper.

b. Preliminary evaluations of the En4DVAR scheme

With the 50 members of ensemble perturbations, the background error statistics are computed and transformed to observation space using Eq. (17). The gradi-

ent of the cost function in En4DVAR is calculated by Eq. (18) in the minimization procedure. Figure 2 shows the variations of the cost function and gradient norm with iterations (solid line). With 20 iterations, the minimization converges well; the cost function value is reduced by over one order after 20 iterations. The analysis fields after 20 iterations of minimization are very close to the “true” fields (Fig. 3). In this experiment, the first-guess fields before minimization are from the normal random perturbations to the true fields. Figures 3a and 3b indicate that there are obvious deviations around the true fields in the first guess of the height and the v component of the wind. However, the deviations of the En4DVAR analysis are close to zero in either positive or negative perturbed locations, indicating that the En4DVAR scheme works robustly.

As a comparison, we carried out an experiment using the standard 4DVAR approach. The gradient of the cost function is calculated by the adjoint model. The background error covariance matrix \mathbf{B} is constructed assuming that the error correlation is Gaussian and the error variance is homogenous. Using the same minimization procedure, the variations of the cost function and gradient norm with iterations are shown in Fig. 2 (dashed line). Both En4DVAR and standard 4DVAR can achieve the minimization of the defined cost function. The analyses of the height and the v component of

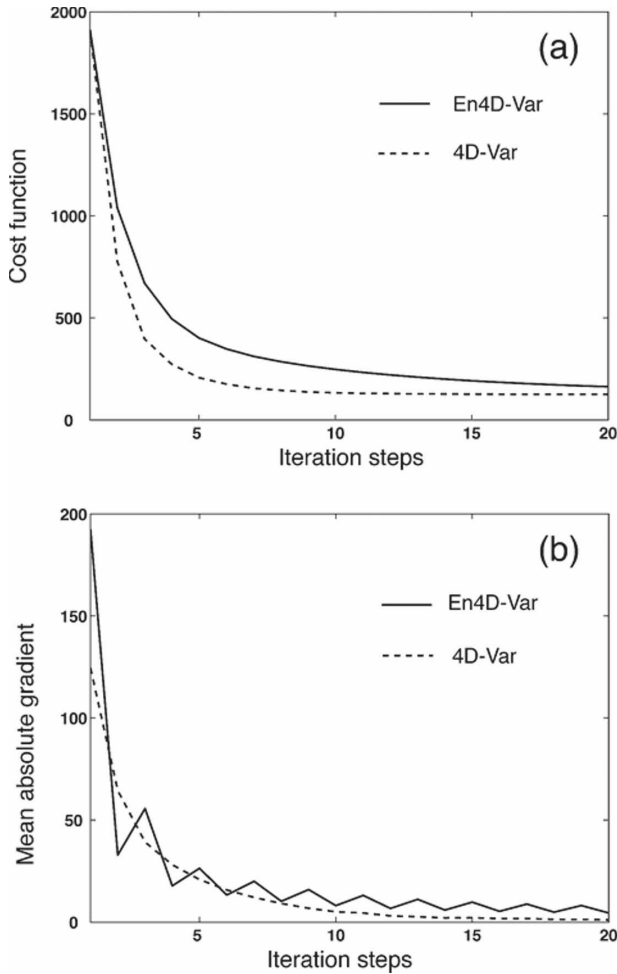


FIG. 2. Variations of (a) the cost function and (b) its gradient with respect to iterations for En4DVAR (solid line) and 4DVAR (dashed line) experiments.

the wind in both schemes are very close to each other. (We therefore omitted the similar figure for the standard 4DVAR method.) However, the costs of the standard 4DVAR and En4DVAR are different. Because the background error covariance in En4DVAR is calculated by ensemble forecasts and its condition number is larger than that of the standard 4DVAR in this experiment, more iteration steps are required in En4DVAR to achieve the same minimization. However, the computing time of every iteration step in En4DVAR is far less than in 4DVAR because the adjoint model is not needed to calculate the gradient.

To further evaluate the proposed En4DVAR scheme, we also conducted an experiment using Eq. (16) to calculate the gradient of the cost function. This still uses the adjoint model in the formulation, but the ensemble background error covariance and preconditioning are adopted. It is indicated that the gradients of

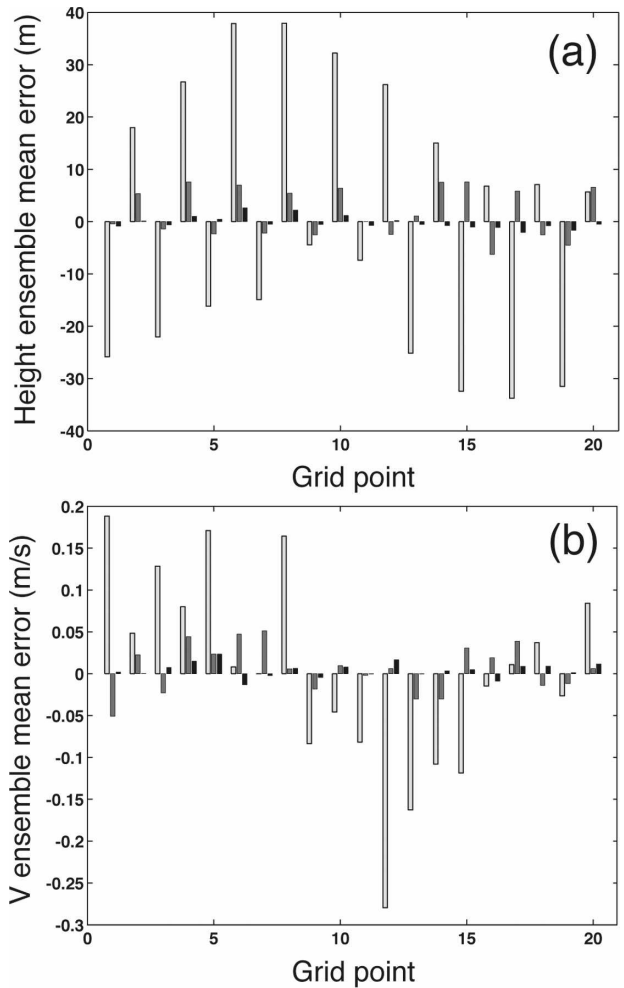


FIG. 3. En4DVAR related (a) ensemble mean height error and (b) ensemble mean v -component wind error at the first analysis time. The light-gray bar is the background ensemble mean error (mean - true). The gray bar is the observation ensemble mean error (mean - true). The black bar is the analysis ensemble mean error (mean - true).

the cost function calculated by using Eqs. (16) and (18) are almost identical. The variations of the cost function and the gradient with iterations are similar to the solid line in Fig. 2. The height difference between the formulations of En4DVAR in Eqs. (16) and (18) is less than $10^{-5} \text{ m}^2 \text{ s}^{-2}$. The v component of the wind difference is less than 10^{-7} m s^{-1} . If the models and observation operators in Eqs. (16) and (18) were linear, the gradients calculated by both equations should be the same. However, there is still a slight difference in the gradient of the cost function as calculated by Eqs. (16) and (18) because a very weak nonlinear term in the shallow water model exists.

It should be noted that the adjoint model calculation is usually time consuming. The advantage of En4DVAR

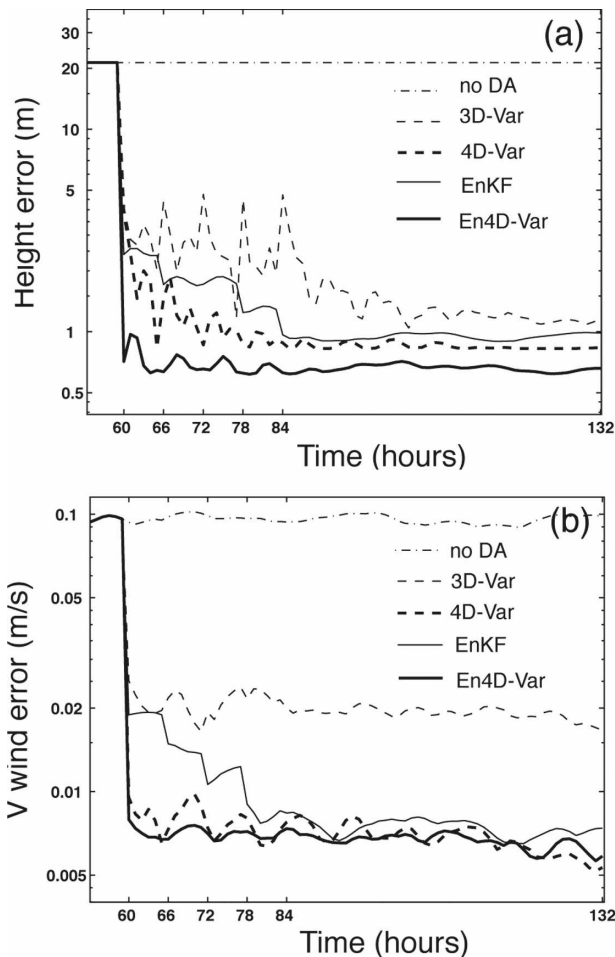


FIG. 4. Mean analysis and forecast errors of different data assimilation schemes during the assimilation (from 60 to 84 h) and a 48-h forecast (from 84 to 132 h): (a) height field and (b) v -component wind field. (Dashed-dotted line is for the CTL experiment when no observation is assimilated, the thin-dashed line is for 3DVAR, the thick-dashed line is for 4DVAR, the thin solid line for EnKF, and the thick solid line for En4DVAR.)

using background perturbation in observation space is that the adjoint model integration is avoided in its minimization procedure. We found that the computation cost of each iteration with the adjoint model using Eq. (16) is over 300 times more expensive than the scheme without the adjoint model using Eq. (18). As En4DVAR without the adjoint model can achieve almost the same minimization as when using the adjoint model, the reduction in computational costs shows great potential for 4DVAR in real-time applications.

To evaluate the performance of the developed En4DVAR scheme, we calculated the variations of the absolute errors (compared with the truth) in assimilations and subsequent forecasts from experiments using 3DVAR, 4DVAR, and EnKF, and compared them

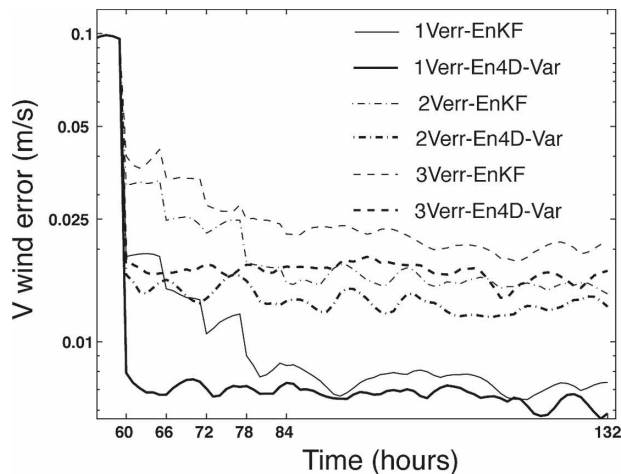


FIG. 5. The v -component wind mean analysis and forecast error trajectories for different v -component wind observation errors in EnKF and En4DVAR. Thin lines are for EnKF and thick lines for En4DVAR. Solid line is for the original v -component wind observation error, dashed-dotted line is for doubling the v -component wind observation error, and dashed line is for tripling the v -component wind observation error.

with the results from En4DVAR (Fig. 4). Since the background error covariance in 3DVAR is static, its absolute error between the analysis and truth ($A - T$) is not reduced except at the first analysis time (at 60 h). In fact, the errors for the height are slightly increased in the 3DVAR cycling analyses at 66, 72, 78, and 84 h. Although the same background error covariance was used in 4DVAR as in 3DVAR, the absolute error of the forecast in the 4DVAR experiment is reduced because 4DVAR has the capability to implicitly develop background error covariance information in the assimilation window. Comparing the results of EnKF and En4DVAR, it is seen that the analysis errors of EnKF (compared with the truth) are sequentially reduced at every analysis time. En4DVAR, on the other hand, maximally reduced its analysis error at 60 h (the beginning of the assimilation window). We believe this is a result of the impact of the future observations at 66, 72, 78, and 84 h.

c. Some sensitivity studies of En4DVAR and its comparison with EnKF

EnKF and En4DVAR obtained similar results at the end of the assimilation window (at 84 h), as shown in Fig. 4. We choose the results of EnKF and En4DVAR in Fig. 4 as control experiments in the studies of this section. Two more sets of experiments for both EnKF and En4DVAR are conducted to test the sensitivity of the v observation errors (Fig. 5). If the error of the v

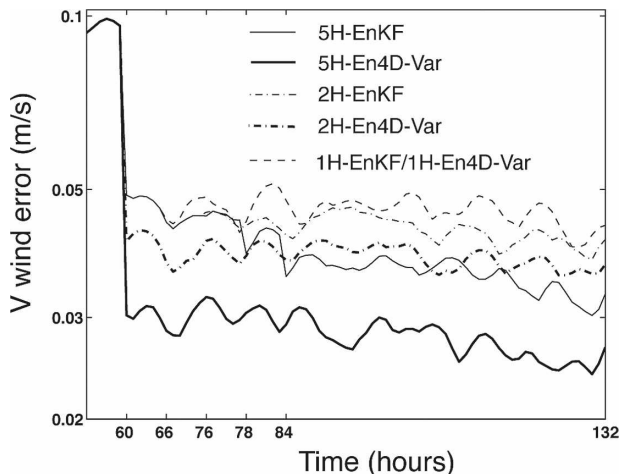


FIG. 6. The v -component wind mean analysis and forecast error trajectories in EnKF and En4DVAR with height observations but no v -component wind observations. The frequencies of the designed height observations are, respectively, five observation times at 60, 66, 72, 78, and 84 h (solid lines); two observation times at 60 and 66 h (dashed-dotted lines); and one observation time at 60 h (dash lines). The thin lines are for EnKF and thick lines are for En4DVAR.

wind observation is amplified, for example, the contribution of the v wind observation to the analysis should be reduced accordingly. Due to their multivariate features in both the EnKF and En4DVAR schemes, the impact of the height observation on the v wind analysis increment increases with the amplified v observation error. When the v observation error is tripled, which is around the magnitude of the v background error, the v analysis error for either EnKF or En4DVAR is still far less than the background error due to the impact of the observation height. The larger the v observation error, the smaller the analysis error of En4DVAR than EnKF is at the end of the window. It seems that En4DVAR can extract more information from the height observations than EnKF can in the v wind analysis.

We also conducted experiments to compare the sensitivities of EnKF and En4DVAR to the observation frequencies. In this set of experiments, the control experiments are 5H-EnKF and 5H-En4DVAR, which assimilate only height observations at 60, 66, 72, 78, and 84 h (five times) without any v observations. In this case, the v wind analysis increment is completely extracted from the height observations. This is equivalent to assigning an infinite v observation error in the 5H-EnKF and 5H-En4DVAR experiments. From Fig. 6, it can be seen that the analysis error in 5H-En4DVAR is far less than that in 5H-EnKF. Comparing 5H-EnKF and 5H-En4DVAR in Fig. 6 with the results of different v errors in Fig. 5, the improvement of the v analysis

in En4DVAR compared with EnKF is increased with the increase of the v observation error.

As sensitivity studies to different observation frequencies, two more sets of experiments are carried out. In 2H-EnKF–2H-En4DVAR, the height observations are set at 60 and 66 h (two times), and in 1H-EnKF–1H-En4DVAR the height observation is set at 60 h (one time only). As Fig. 6 shows, the lower the height observation frequency is, the closer the results of EnKF and En4DVAR are. When only one time–height observation is assimilated, En4DVAR, which is degenerated to En3DVAR, obtained almost the same error variation as EnKF (so only one line is shown in Fig. 6).

To test the assimilation ability of EnKF and En4DVAR under the unbalanced conditions, we designed experiments with univariate analysis, in which each variable analysis's increments will be produced by its own observations and background. The control experiments still adopt the EnKF and En4DVAR design (with results shown in Fig. 4) but with a univariate algorithm. Comparing Fig. 7 and Fig. 4, it is shown that the error of the univariate analysis is slightly larger than that of the multivariate analysis because the multivariate analysis can use more information from other variable observations. In Fig. 7, the univariate analyses of EnKF and En4DVAR are also similar at the end of the assimilation window.

When the v wind observation error is magnified 3 times, the analysis is more unbalanced because the geotropic relation in height and wind observations is further reduced (their errors are uncorrelated). As shown in Fig. 7, height error variations appear to oscillate after the analysis. These oscillations do not appear in the experiments with the multivariate algorithm (Fig. 4). In the assimilation window, the oscillation of the error variations in the En4DVAR experiment is less than that in EnKF. The improvement of the analysis and forecast in En4DVAR is larger than that in EnKF in unbalanced situations compared to the control experiments.

4. Summary and conclusions

Hybrid—ensemble-based and variational data assimilation—techniques have been a popular research topic in recent years. Its performance in three-dimensional global model data assimilation has been evaluated by Buehner (2005). As 3DVAR is gradually updated to 4DVAR in many research and operational centers, ensemble-based four-dimensional variational algorithms should be developed. After reviewing the classical EnKF algorithm and Lorenc's (2003) ensemble-based variational algorithm, we have proposed an ensemble-

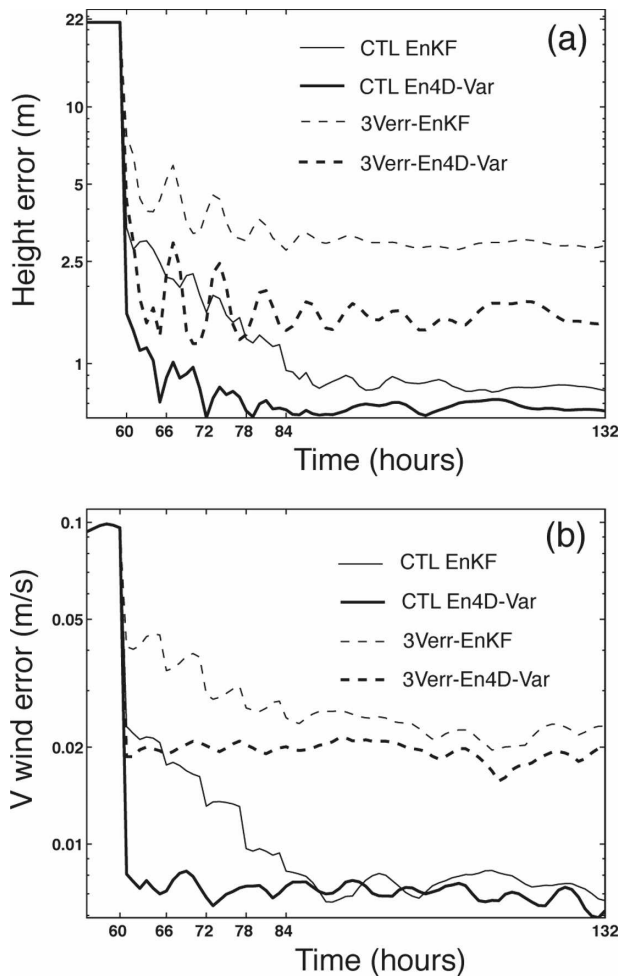


FIG. 7. The analysis and forecast errors in EnKF (thin lines) and En4DVAR (thick lines) for the (a) height field and (b) v -component wind field. The figure is similar to Fig. 4, but using a univariate algorithm (solid line). The dashed lines show the results with tripling the v -component wind observation error.

based four-dimensional variational algorithm. This scheme uses background perturbation in observation space to calculate the gradient during the minimizing procedure so that it does not need a tangent linear model or an adjoint model in its formulation.

The performance of En4DVAR was evaluated by use of shallow water model experiments. The convergence of the En4DVAR cost function needs a few more iterations than that of traditional 4DVAR. However, each iteration of the En4DVAR minimization needs far less computation time than that of traditional 4DVAR. A great advantage of En4DVAR is that it can be implemented without a tangent linear model or an adjoint model. Our experiments have indicated that En4DVAR without using tangent linear and adjoint models produced an analysis result that is similar to

the approach that used tangent linear and adjoint models, but with much less computation cost. Among the experiments with different data assimilation schemes (3DVAR, 4DVAR, EnKF, and En4DVAR), En4DVAR produced comparable and reasonably sound analysis results. The sensitivity experiments in our comparison of En4DVAR and EnKF showed that En4DVAR is a sophisticated data assimilation algorithm due to its nonsequential analysis character, especially under the unbalance conditions.

Since the experiments in this paper were carried out in low-dimensional space with a nearly linear model, some basic issues of the ensemble-based assimilation scheme and 4DVAR can be clearly examined. When En4DVAR is applied in a real NWP environment, the dimension of the real atmospheric model is far larger than the ensemble dimension so that the analyses are prone to incremental noise. Techniques that are more sophisticated will have to be considered, such as localization, dealing with model error, and analysis balance, etc. To further test the capability of En4DVAR, we are implementing it using a real NWP data assimilation system (WRF model and WRF data assimilation system) and conducting experiments with real observational data. We are testing the Schur operator in En4DVAR to perform the localization. The localization technique can reduce the sample noise. In addition, it is important to have an appropriate assimilation window length and analysis time when implementing En4DVAR in real atmospheric model. Some preliminary results are encouraging, among them that the En4DVAR can be a choice for real atmosphere data assimilation. The implementation details of the En4DVAR scheme with a real atmospheric model and its experimental results will be reported upon in the future.

Acknowledgments. We are grateful to Chris Snyder, Jeffery Anderson and Hui Liu (NCAR) for their critical comments on the paper. This research is supported by NOAA grant under Contract 05111076.

REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- , and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.
- Bouttier, F., and P. Courtier, cited 2007: Data assimilation concepts and methods. Meteorological Training Course Lecture Series, ECMWF, Reading, United Kingdom. [Available online at <http://www.ecmwf.int/newsevents/training/>]

- rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts.html.]
- Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operation NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043.
- Courtier, P., J. N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4DVAR, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1128–1145.
- Derber, J. C., R. J. Purser, W.-S. Wu, R. Treadon, M. Ponca, D. Parrish, and D. Kleist, 2003: Flow dependent Jb in a global grid-point 3D-var. *Proc. Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, United Kingdom, ECMWF, 125–134. [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2003/sem2003_derber.pdf.]
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 143–162.
- , and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852–1867.
- Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D ensemble filter: Perfect model simulations with Lorenz-96. *Tellus*, **59A**, 96–100.
- Fisher, M., 2003: Background error covariance modeling. *Proc. Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, United Kingdom, ECMWF, 45–63. [Available online http://www.ecmwf.int/newsevents/meetings/annual_seminar/seminar2003_presentations/Fisher.pdf.]
- Gilbert, J. C., and C. Lemarechal, 1989: Some numerical experiments with variable storage quasi-Newton algorithms. *Math. Programming*, **45B**, 407–435.
- Hamill, T. M., 2002: Ensemble-based data assimilation. *Proc. Workshop on Predictability*, Reading, United Kingdom, ECMWF, 83–112. [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2002/sem02_hamill.pdf.]
- , and C. Snyder, 2001: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919.
- Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields application to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.
- Hollingsworth, A., and P. Lonnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- , and —, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289.
- Hunt, B. R., and Coauthors, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277.
- Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, **82D**, 35–45.
- Li, Z., and I. M. Navon, 2001: Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. *Quart. J. Roy. Meteor. Soc.*, **127**, 661–683.
- Lorenc, A. C., 1992: Iterative analysis using covariance functions and filters. *Quart. J. Roy. Meteor. Soc.*, **118**, 569–591.
- , 2003: The potential of the ensemble Kalman filter for NWP: A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.
- Parrish, D., and J. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Simmons, A. J., and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **128**, 647–677.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.
- Snyman, J. A., 2005: *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer Publishing, 257 pp.
- Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Xiao, Q., X. Zou, and B. Wang, 2000: Initialization and simulation of a landfalling hurricane using a variational bogus data assimilation scheme. *Mon. Wea. Rev.*, **128**, 2252–2269.
- , —, M. Ponca, M. A. Shapiro, and C. S. Velden, 2002: Impact of GMS-5 and GOES-9 satellite-derived winds on the prediction of a NORPEX extratropical cyclone. *Mon. Wea. Rev.*, **130**, 507–528.
- Xu, Q., and L. Wei, 2001: Estimation of three-dimensional error covariances. Part II: Analysis of wind innovation vectors. *Mon. Wea. Rev.*, **129**, 2939–2954.
- , and —, 2002: Estimation of three-dimensional error covariances. Part III: Height–wind forecast error correlation and related geostrophy. *Mon. Wea. Rev.*, **130**, 1052–1062.
- , —, A. V. Tuiyi, and E. H. Baker, 2001: Estimation of three-dimensional error covariances. Part I: Analysis of height innovation vectors. *Mon. Wea. Rev.*, **129**, 2126–2135.
- Zou, X., and Y. H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*, **124**, 2859–2882.
- Zupanski, M., 1993: Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, **121**, 2396–2408.
- , 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726.