Tellus (2003), 55A, 126–147
Printed in UK. All rights reserved

TELLUS
ISSN 0280–6495

Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model

By S. ZHANG and J. L. ANDERSON, *Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, NJ 08542, USA*

(Manuscript received 7 May 2002; in final form 10 October 2002)

ABSTRACT

The background error covariance (correlation) between model state variables is of central importance for implementing data assimilation and understanding model dynamics. Traditional approaches for estimating the background error covariance involve many heuristic approximations, and often the estimated covariance is flow-independent, i.e. only reflecting statistics of the climatological background. This study examines temporally and spatially varying estimates of error covariance in a spectral barotropic model using a Monte Carlo approach, an implementation of an ensemble square root filter called the ensemble adjustment Kalman filter (EAKF). The EAKF is designed to maintain as much information about the distribution of the prior state variables as possible, and results show that this method can produce reasonable estimates of error correlation structure with an affordable sample (ensemble) size. The impact of using temporally and spatially varying estimates of error covariance in the EAKF is examined by using the time and spatial mean error covariances derived from the EAKF in an ensemble optimal interpolation (OI) assimilation scheme. Three key results are: (1) for the same ensemble size, an ensemble filter such as the EAKF produces better assimilations since its flow-dependent error covariance estimates are able to reflect more about the synoptic-scale wave structure in the simulated flows; (2) an ensemble OI scheme can also produce reasonably good assimilation results if the time-invariate covariance matrix is chosen appropriately; (3) when using the EAKF to estimate the error covariance matrix for improving traditional assimilation algorithms such as variational analysis and OI, a relatively small ensemble size may be used to estimate correlation structure although larger ensembles produce progressively better results.

1. Introduction

The probabilistic nature of the evolution of the atmosphere has been widely recognized since the early 1960s (Gleeson, 1961; Lorenz, 1963). Representing the evolution of the atmosphere as a continuous stochastic dynamical process, the evolution of the stochastic process can be simulated using Monte Carlo methods, usually referred to as ensemble forecasting (Leith, 1974; Kalnay and Toth, 1996; Molteni et al., 1996; Houtekamer et al., 1996). In this context, data assimilation is the problem of sampling the probability of the state of a dynamical system given noisy measurements, i.e. *filtering*. In filtering theory (Jazwinski, 1970), data assimilation generates a conditional probability density function (PDF), the probability distribution of the system state given a set of observations. The data assimilation process uses model dynamics to extract the reliable information from observations.

A central issue in implementing filtering is computing the product of two distributions that represent, respectively, the information from a set of observations

*Corresponding author address: Dr. S. Zhang, GFDL/NOAA, Princeton University, P.O. Box 308, Princeton, NJ 08542, USA.
e-mail: snz@gfdl.noaa.gov

and the prior constraints from the model dynamics. The Kalman–Bucy filter is the best known approximate algorithm for solving the filtering problem (Kalman, 1960; Kalman and Bucy, 1961). Under the assumptions of linear error evolution and Gaussian error distributions, the algorithm derives a linear combination of measurements to update the state estimate. The weighting coefficient matrix (the *Kalman gain*) is determined from the prior state covariance, the observational operator, and the observational error variance. Given the assumptions, the Kalman filter update gives an optimal estimate of the system's state.

Several ensemble algorithms have been developed by modifying the Kalman filter algorithm. For example, the ensemble Kalman filter (hereafter called EnKF) (Evensen, 1994; Houtekamer and Mitchell, 1998; 2001; Burgers et al., 1998; Van Leeuwen, 1999; Keppenne, 2000; Mitchell and Houtekamer, 2000; Hamill et al., 2001; Whitaker and Hamill, 2002) accounts for the nonlinear evolution of the prior state covariance by using an ensemble sample of system states to evaluate the error covariance. In order to carry out this algorithm, the observational distribution is sampled by perturbing observations using samples from the observational error distribution. The persistent introduction of a small but significant noise into the product may destroy information about the prior relations between state variables and therefore degrade the relative performance of the algorithm (Anderson, 2001; Whitaker and Hamill, 2002). The ensemble adjustment Kalman filter (EAKF; Anderson, 2001), an implementation of an ensemble square root filter (Tippett et al., 2002) updates the prior ensemble using a linear operator derived from the product of the observational and prior distributions. The new ensemble has exactly the mean and covariance characteristics that would result if the prior sample and observational error covariance are approximated by Gaussians while maintaining information about the higher-order moment structure of the prior distribution (Anderson, 2001). Results of perfect model studies with an idealized global dry primitive equation model show that the filter is able to reconstruct the structure of the free atmosphere using only surface pressure observations at a set of randomly located points on the sphere.

The covariance between the state variables is of central importance for implementing data assimilation, since it is a measurement of the uncertainty of the system as well as the relation between state variables and observations. In principle, the covariance and other error statistics of forecasting variables can be obtained by forward integration of the Kolmogorov equation (also called the Fokker–Planck equation; Jazwinski, 1970) to obtain the probability density distribution. However, for a realistic model, direct solution of this equation is not viable due to an extreme computational cost and the lack of realistic initial conditions for probability densities. Therefore, simplifications and approximations have to be made. Traditional approaches usually analyze the climatological characteristics of error statistics using the explicit relation between physical variables under some approximate assumptions (Daley, 1991) or through the observational/forecasting time series correlation (Buell, 1960; 1971; 1972a,b; Seaman and Gauntlett, 1980; Buell and Seaman, 1983; Hollingsworth, 1986; Parrish and Deber, 1992). Therefore, the calculated correlation structures are flow-independent and, in mid-latitudes, reflect the climatological large-scale trough–ridge circulation (Seaman and Gauntlett, 1980; Buell and Seaman, 1983; Thiebaux, 1985; Hollingsworth, 1986). Cohn (1993) discussed dynamics of short-term univariate forecast error covariances using some simplified governing equations. Bouttier (1993) studied the evolution of the error covariance of geopotential height using tangent linear integration of a barotropic model starting from an idealized elliptic initial distribution of the auto-correlation of geopotential height. Ehrendorfer and Tribbia (1997) studied the characteristics of forecast error covariances through singular vectors constructed by a set of tangent linear integrations using both the Lorenz84 (Lorenz, 1984) model and a barotropic model.

Ensemble filtering algorithms approximate the error statistics using an ensemble of forecasts that represent discrete samples of the probability density distribution. This approach retains nonlinear characteristics of the error evolution because each member is advanced by separately integrating the nonlinear governing equations. Therefore, a sound ensemble filter is expected to produce useful estimates of the time evolution of covariance between the state variables. This study examines the characteristics of various estimated covariances using ensemble methods focusing on the EAKF. Due to the central role of error covariance in data assimilation, an investigation of the characteristics of the evolution of error covariance should benefit the whole data assimilation community. The knowledge of the sensitivity of the ensemble assimilation algorithms with respect to the spatially and temporally varying estimates of error covariance increases

understanding of the model dynamics and assimilation technologies.

The paper is arranged as follows: A brief description of the Monte Carlo method and the EAKF observing/assimilation simulation experiment is given in section 2. Section 3 presents details of various estimated covariance structures and section 4 examines the sensitivity of the ensemble assimilation with respect to various spatially and temporally varying estimates of error covariance. The sensitivity of the EAKF assimilation with respect to the ensemble size is also investigated in this section. Several ensemble Optimal Interpolation schemes are tested with a variety of time-invariate estimates of error covariance in section 5. Finally, a summary and discussion are presented in section 6.

2. Methodology

2.1. A Monte Carlo method for estimating error covariance

Viewing the atmosphere as a continuous stochastic dynamical process, the evolution of the atmospheric state is described by the vector stochastic differential equation (Jazwinski, 1970),

$$\mathrm{d}\mathbf{x}_t/\mathrm{d}t = \mathbf{f}(\mathbf{x}_t, t) + \mathbf{G}(\mathbf{x}_t, t)\mathbf{w}_t \tag{1}$$

Here, $\mathbf{x}_t$ is an n-dimensional vector representing the model state at time t (n is the size of the model state), $\mathbf{f}$ is an n-dimensional vector function, $\mathbf{w}_t$ is a white Gaussian process (uncorrelated in time) of dimension r with mean 0 and covariance matrix $\mathbf{S}(t)$, while $\mathbf{G}$ is an $n \times r$ matrix.

The probability density distribution of the state $\mathbf{x}_t$ completely defines the error statistics. For example, the error covariance is the second-order moment of the probability density function (PDF) of the atmospheric state, and the error covariance normalized by the standard deviations gives the error correlation. The differential equation that describes the evolution of the PDF, $p(\mathbf{x}, t)$ (the subscript t of $\mathbf{x}_t$ is dropped) is the Fokker–Planck equation (Gardiner, 1983, chapter 5; Jazwinski, 1970, chapter 4),

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{n} \frac{\partial (p f_i)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left[p(\mathbf{G}\mathbf{S}\mathbf{S}^T\mathbf{G}^T)_{ij}\right] \tag{2}$$

where $\mathbf{x}$ is a random vector which consists of n random variables, $\mathbf{x}_1, \ldots, \mathbf{x}_n$. For multivariate systems such as the atmosphere and ocean, $p(\mathbf{x}, t)$ is also called the joint probability density function. In data assimilation, the probability density of the analysed state is defined as being conditional on a set of available observations. The data assimilation problem is how to produce the analysis state using the knowledge of the conditional probability density.

For realistic numerical models, direct numerical integration of eq. (2) is unrealistic due to an extremely high computation cost. In addition, eq. (1) is a stochastic differential equation with a white Gaussian forcing function. Since the $\mathbf{w}_t$ process is only delta-correlated and not mean square Riemann integrable, eq. (1) has no mathematical meaning (chapter 4, Jazwinski, 1970). In practice, using Monte Carlo methods to simulate the continuous stochastic dynamical process is convenient. Monte Carlo methods use finite random samples as a discrete representation of the continuous stochastic process, instead of solving the stochastic differential equation. Here, the forecast model (1) is assumed to be deterministic,

$$\mathrm{d}\mathbf{x}_t/\mathrm{d}t = \mathbf{f}(\mathbf{x}_t, t) \tag{3}$$

and a set of randomly selected perturbations to a given initial state is used to form an initial ensemble of forecasts. When integrated in the model, these random samples of $\mathbf{x}$ discretely represent the distribution of the forecast probability. Approximations of the various moments of the distribution can be computed from these discrete samples. For example, the covariance matrix, a second-order central moment, is

$$\mathrm{cov}\{\mathbf{x}_i, \mathbf{x}_j\} = \frac{1}{M}\sum_{m=1}^{M}(x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j) \tag{4}$$

where $\bar{x}_i$ and $\bar{x}_j$ represent, respectively, the ensemble mean of the i-th and j-th model variables ($\mathbf{x}_i, \mathbf{x}_j$), and M is the ensemble size. The diagonal elements of the symmetric matrix cov $\{\mathbf{x}_i, \mathbf{x}_j\}$ are error variances of the model variables, while off-diagonal elements represents the error covariance between the model variables.

Miller et al. (1994, 1999) employed Monte Carlo estimates of covariance statistics to devise a generalization of the extended Kalman–Bucy filter, and compared the evolution of the PDF by the Monte Carlo method and the numerical integration of the Fokker–Planck equation using a stochastically forced double-well model. It was found that there is little difference between the solutions by the Monte Carlo method and

the numerical integration of the Fokker–Planck equation. In the following sections, an EAKF is used in a spectral barotropic model to compute error covariance estimates with various degrees of approximation by using a series of observing/assimilation simulation system experiments.

2.2. *Brief description of an ensemble adjustment Kalman filter*

Great efforts have been made to implement data assimilation to address the probabilistic nature of the dynamical/observational system of the atmosphere and ocean (Evensen, 1994; Miller et al., 1994; 1999; Houtekamer and Mitchell, 1998; 2001; Burgers et al., 1998; Van Leeuwen, 1999; Anderson and Anderson, 1999; Keppenne, 2000; Mitchell and Houtekamer, 2000; Bishop et al., 2001; Hamill et al., 2001; Whitaker and Hamill, 2002; Anderson, 2001). These studies attempted to compute a conditional probability distribution of the system state given a set of observations. The core of these filtering algorithms solves for the product of the prior distribution of the system state, which is governed by the model dynamics, and the observational error distribution (a function of the observing system which is normally given as Gaussian) (chapter 6, Jazwinski, 1970).

Ensemble filters like the EAKF used here and the perturbed observation ensemble Kalman filter can be applied sequentially to individual scalar observations without loss of generality, especially when a large ensemble size is used to discretize the distribution of the state (Whitaker and Hamill, 2002). The impact of each scalar observation on each scalar component of the state vector can also be computed independently (Anderson, 2002). In what follows, a description of how the EAKF computes the impact of a scalar observation on a single state variable component is presented.

The EAKF first constructs a prior joint state/observation vector (referred to as a joint state vector), $\mathbf{z}^{\mathrm{p}} = \{x, h(\mathbf{x})\}$ with length 2, where x represents a single scalar component of the state vector at the time t and h is an operator (on the full state) that gives the expected value of the observation given the state vector $\mathbf{x}$. In the perfect model framework, the actual observation is

$$y^{\mathrm{o}} = h(\mathbf{x}) + \epsilon \tag{5}$$

where ϵ is a sample selected from an observational error distribution (assumed Gaussian here) associated with the instrument being simulated. The filtering algorithm computes the distribution of the updated joint state vector, $\mathbf{z}^{\mathrm{u}}$, which can be expressed as

$$p(\mathbf{z}^{\mathrm{u}}/y^{\mathrm{o}}) = p(y^{\mathrm{o}} \mid \mathbf{z}^{\mathrm{p}})p(\mathbf{z}^{\mathrm{p}})/\mathit{normalization} \tag{6}$$

where $p(y^{\mathrm{o}} \mid \mathbf{z}^{\mathrm{p}})$ denotes the conditional probability density distribution of the observation y^{o} given the prior joint state vector $\mathbf{z}^{\mathrm{p}}$ and $p(\mathbf{z}^{\mathrm{p}})$ represents the prior distribution of the joint state vector $\mathbf{z}^{\mathrm{p}}$. The normalization in eq. (6) guarantees that the total probability of all possible states is 1 and does not need to be computed for the EAKF algorithm. Equation (6) expresses how a new observation, y^{o}, modifies the probability distribution of the prior joint state, $p(\mathbf{z}^{\mathrm{p}})$.

The computation of the EAKF for assimilating a single observation is schematically shown in Fig. 1. First, the EAKF approximates the numerator in eq. (6) as a product of two Gaussians: the observational distribution ("obs PDF" curve in Fig. 1) and a Gaussian approximation to the prior with the ensemble sample mean and covariance (the solid curve labeled "prior

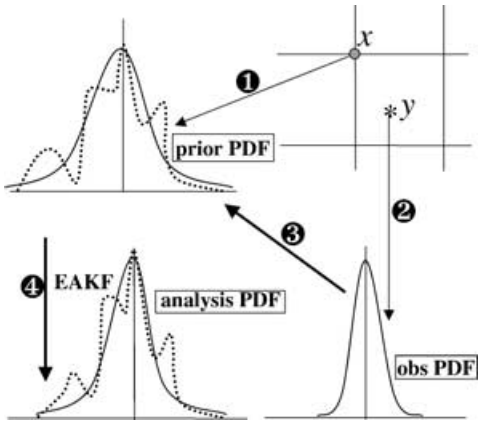

Fig. 1. Cartoon of how the ensemble adjustment Kalman filter (EAKF) updates the estimate of the probability distribution of a single state variable, x, when given a single observation, y. The dotted curves in the prior and analysis PDF are schematic representations of the ensemble estimate of the PDF (generated, for instance, by doing a kernel approximation) and are meant to suggest that the ensemble has meaningful non-Gaussian structure. A Gaussian fit to the prior PDF is represented by the solid curve. The update procedure in eq. (8), making use of the forward observation operator in eq. (5), gives the analysis Gaussian shown by the solid curve; the observation error is assumed to be Gaussian. Finally, application of a linear operator as in eq. (9) leads to a new ensemble with sample mean and covariance identical to the solid analysis curve, but retaining non-Gaussian structure as reflected by the dotted curve.

PDF" in Fig. 1). The mean of the prior joint state vector is $\overline{\mathbf{z}^{\mathrm{p}}}$ and the covariance matrix is

$$\Sigma^{\mathrm{p}} = \begin{pmatrix} \sigma_x^2 & \mathrm{cov}(x, y) \\ \mathrm{cov}(x, y), & \sigma_y^2 \end{pmatrix} \tag{7}$$

in which σ_x^2 and σ_y^2 are the prior ensemble estimates of the variances of the state variable and of the observation variable respectively and $\mathrm{cov}(x, y)$ represents the background covariance between the model grid and observation location.

In this Gaussian approximation, the updated covariance and mean of the joint state vector are

$$\begin{aligned} \Sigma^{\mathrm{u}} &= \left[(\Sigma^{\mathrm{p}})^{-1} + \mathbf{H}^T \left(\sigma_{\mathrm{y}}^{\mathrm{o}}\right)^{-2} \mathbf{H}\right]^{-1} \\ \overline{\mathbf{z}^{\mathrm{u}}} &= \Sigma^{\mathrm{u}} \left[(\Sigma^{\mathrm{p}})^{-1} \overline{\mathbf{z}^{\mathrm{p}}} + \mathbf{H}^T \left(\sigma_{\mathrm{y}}^{\mathrm{o}}\right)^{-2} \mathbf{y}^{\mathrm{o}}\right] \end{aligned} \tag{8}$$

where an overbar denotes the ensemble mean, a superscript "p" or "u" represents the prior/updated distribution, and $\sigma_{\mathrm{y}}^{\mathrm{o}}$ is the observational error standard deviation of the instrument. $\mathbf{H}$ is a 1×2 matrix in which the second element is 1 and the other is 0, so that the estimated observation value calculated from the joint state vector are $y = \mathbf{Hz}$.

Given the updated mean ($\overline{\mathbf{z}^{\mathrm{u}}}$) and covariance ($\Sigma^{\mathrm{u}}$), the EAKF uses a linear operator $\mathbf{A}$ to update each prior ensemble member as

$$\mathbf{z}_i^{\mathrm{u}} = \mathbf{A}^T \left(\mathbf{z}_i^{\mathrm{p}} - \overline{\mathbf{z}^{\mathrm{p}}}\right) + \overline{\mathbf{z}^{\mathrm{u}}}, i = 1, \ldots, M \tag{9}$$

where M represents the ensemble size, so that the updated ensemble has exactly the mean $\overline{\mathbf{z}^{\mathrm{u}}}$ and covariance Σ^{u} while maintaining much of the non-Gaussian information from the prior distribution, as shown in the dotted curve of "analysis PDF" in Fig. 1. As stated above only 2×2 matrices are required in order to evaluate eqs. (8) and (9). Once eq. (8) is evaluated, $\mathbf{A}$ is solved using the computed Σ^{u} and $\overline{\mathbf{z}^{\mathrm{u}}}$ (Anderson, 2001).

Like other filtering techniques, the EAKF can experience filter divergence (Jazwinski, 1970) in which the distribution produced by the filter drifts away from the truth. In order to avoid filter divergence, the EAKF uses a covariance inflation parameter (denoted by γ hereafter) to "broaden" the prior distribution and enhance the impact of the observations in the product (Anderson, 2001). Also, extensive testing of the effects of γ in ensemble-based data assimilation was made in Hamill et al. (2001) and Whitaker and Hamill (2002).

The EAKF algorithm can be summarized by the following steps:

Step 1. Advance the numerical model (3) to the time of the next observation for each ensemble member to form an ensemble that samples the prior distribution of the state variables. Then each observation is processed for each variable at each grid point for assimilation.

Step 2. Compute Σ^{p} in eq. (8) using the prior ensemble estimates of the observation at the observational location and the model gridpoint. A covariance inflation factor γ may be applied to broaden the spread of the prior ensemble [each member's departure from the mean is increased by the factor γ as $\gamma(\mathbf{x}_i^{\mathrm{p}} - \overline{\mathbf{x}^{\mathrm{p}}})$].

Step 3. Compute $\mathbf{A}$ (Appendix A in Anderson, 2001) and update each ensemble member by eq. (9).

Step 4. Repeat steps 2 and 3 for each state variable.

In the ensemble adjustment filtering algorithm, the adjusted ensemble maintains much of the information about higher-order moments of the prior distribution while having exactly the mean and covariance of the product of two Gaussians. The non-Gaussian information is still useful to represent the prior distribution, although the prior ensemble is scaled by the covariance inflater. Unlike the Kalman filter update equation, the update formula [eq. (9)] does not require perturbing the observations. In addition, only 2×2 matrix operations are required, so computation and storage requirements are small. Therefore the algorithm has good potential to be applied to realistic atmospheric and oceanic models.

2.3. Observing/assimilation system simulation experiments using a global barotropic spectral model

Here, observing/assimilation system simulation experiments are conducted using a perfect model assumption. For simplicity, the primary investigation is carried out on a univariate system. The model chosen is a non-divergent barotropic model in which the vorticity equation is represented on the sphere by spherical harmonic functions with a triangle truncation of 42 wavenumbers. A pseudo-spectral method with a physical space grid consisting of 128 longitudes and 64 Gaussian latitudes for a total of 8192 grid points is used to compute products and is the representation of the state used for data assimilation. A time step of 1800 s is used with a third-order Adams–Bashforth time differencing scheme which is initialized with a single forward step followed by a single leapfrog step. A ∇^8 diffusion on the streamfunction is applied with

a constant factor, so that the smallest resolved wave is damped with an e-folding time of 2 d. A forcing must be added to the model to induce interesting long-term variability and to produce a quasi-realistic simulation of the northern hemisphere zonal flow. In this case, the zonal flow spherical harmonic components are relaxed towards the observed time mean zonal flow for the period November through March 1991–92, with an e-folding time of approximately 20 d.

In order to simulate a worldwide observational network, 600 randomly chosen locations on the surface of the sphere are used to produce observations of streamfunction every 12 h. Observational error is simulated by adding a sample of white noise with a standard deviation of 10^6 m^2 s^{-1} to the "truth". The "truth" is a long control run of the model starting from an initial streamfunction generated from the NCEP reanalysis on 1 November 1991. The ensemble initial conditions are produced by adding random samples of a Gaussian with 10^6 m^2 s^{-1} standard deviation to the unperturbed streamfunction initial field at each grid point. This study only adds the spatially uncorrelated random perturbations to produce the ensemble initial conditions. The impact of the spatially correlated perturbations on ensemble-based filtering algorithms would be another research topic in this field. The system is spun up by conducting an assimilation out to 500 d with 20 ensemble members and a $\Delta\phi \times \Delta\lambda \sec\phi$ observational window, where $\Delta\phi$ and $\Delta\lambda$ are the latitudinal and longitudinal width (20° in this study) and $\sec\phi$ is the latitudinal adjustment factor of the longitudinal width. An observation is only allowed to impact the state variables within the window. The window means that the analysis of each model gridpoint at the middle latitude uses about 40 nearby observations. More sophisticated methods for limiting the impact of remote observations on state variables are a topic of ongoing research in ensemble filtering (Hamill et al., 2001). Applying a smoothly varying distance-dependent weight to reduce the prior sample correlation between state and observation variables (Hamill et al., 2001) would be expected to improve further the performance of the filtering algorithms applied here (Anderson, 2001).

3. Error correlation estimates

3.1. Variations of error correlation in time and space

Estimates of the error correlation in this study are based on the computation of eq. (4) in the EAKF assimilation run over the 100-d period from day 500 to day 600. The error correlation is a 8192×8192 matrix in which each column represents the distribution of the streamfunction auto-correlation for a particular reference point to all the model grid points. Three reference points (177°E, 3°S), (177°E, 42°N), and (177°E, 82°N) were chosen to represent the situation for the low-, middle- and high-latitudes.

Figure 2 displays the instantaneous spatial structure of the streamfunction auto-correlation for the middle-latitude reference point (marked by a star) for day 30, day 50, day 70 and day 90. The streamfunction spatial correlation has a temporally varying character which can be denoted by $C(\mathbf{r}_0, \mathbf{r}, t)$, where $\mathbf{r}_0$ and $\mathbf{r}$, respectively, represent the location vectors of the reference/correlated points and t represents the time. The spatial correlation at middle latitudes has a wave-train character. This wave-train character is more clearly shown in Fig. 3, where a daily evolution of $C(\mathbf{r}_0, \mathbf{r}, t)$ during the period from day 51 to day 56 is presented. The temporally varying character of $C(\mathbf{r}_0, \mathbf{r}, t)$ reflects the oscillation of the synoptic and/or planetary scales, i.e. the estimated correlation structure is flow-dependent. The temporally and spatially varying character of $C(\mathbf{r}_0, \mathbf{r}, t)$ is also shown for the low (Fig. 4) and high (Fig. 5) latitudes. However, instead of wave-train structures, the streamfunction spatial correlation at low and high latitudes shows an approximately symmetric structure, i.e., the positive correlation centered at the reference point and negative correlation areas surrounding the positive center. In addition, the positive correlation center around the reference point at low latitudes appears smaller than at middle and high latitudes, and the distribution at low latitudes shows more localized noise. These phenomena demonstrate the different dynamics of flows at different latitudes.

3.2. Estimate of anisotropic error correlation

The previous subsection presented a temporally and spatially varying estimate of the streamfunction spatial correlation. In many traditional data assimilation schemes (Optimal Interpolation and some implementations of Variational Analysis, for instance), a time-invariant error correlation is used. This subsection examines the time-invariant spatial correlation character of the streamfunction, estimated by the EAKF observing/assimilation system simulation experiment.

The time mean spatial structure of the streamfunction correlation, $\overline{C}(\mathbf{r}_0, \mathbf{r})$ is estimated by conducting a time average of $C(\mathbf{r}_0, \mathbf{r}, t)$ over a long period.

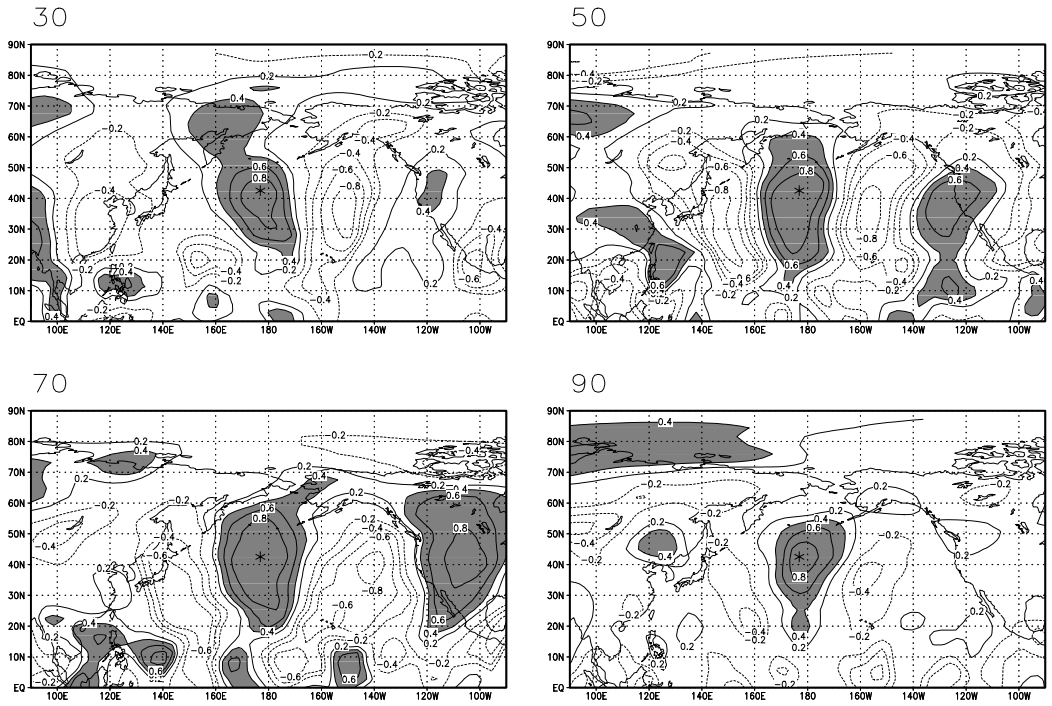

Fig. 2. The time-evolution of the correlation distribution over the domain from 0 to 90°N and 90°E to 270°E for the reference point (177°E, 42°N) (marked by an asterisk) at 20-d intervals starting from day 30 during the 100-d assimilation period from day 500 to day 600, using 20 ensemble members. The contour interval is 0.2 and values greater than 0.4 are shaded.

Figure 6 presents the structures of $\overline{C}(\mathbf{r}_0, \mathbf{r})$ over the 100-d period from day 500 to day 600 for the high (the reference point at 177°E, 82°N) (panel a), middle (the reference point at 177°E, 42°N) (panel b) and low (the reference point at 177°E, 3°S) (panel c) latitudes. Time averaging filters much of the small-scale localized "noise" and keeps the large-scale characters seen in the time-varying correlation patterns. The resulting streamfunction correlation distributions mainly reflect the flow-independent character at different latitudes. For example, the trough–ridge dynamics in the middle latitudes is represented by a set of positive/negative correlation centers alternately arrayed along the great circle of the sphere (panel b), while the correlation patterns in low and high latitudes (panels a and c) are approximately axisymmetric. The correlation patterns at the middle latitudes are consistent with the propagating path of the planetary waves derived from a barotropic vorticity equation by Hoskins and Karoly (1981). These characteristics are shown more clearly by Fig. 7, which shows the zonal mean of the distributions of the streamfunction auto-correlation over all reference points located at the same latitude (a total of 128 gridpoints), denoted by $[\overline{C}][r_0(\phi), \mathbf{r}]$ where $r_0(\phi)$ means that the location of the zonal mean reference point is only dependent on the latitude of gridpoints. Again, the auto-correlation distribution at the high (panel a) and low (panel c) latitudes is approximately isotropic, while the one at the middle latitudes has a wave-train (trough–ridge) character.

3.3. Estimate of the isotropic error correlation

In order to derive a globally isotropic streamfunction spatial correlation structure, $C_0(r)$, where r is the distance between a correlated point and the reference point, the isotropic correlation structure is first calculated at each latitude, denoted by $C_0[r_0(\phi), r]$ using $[\overline{C}][r_0(\phi), \mathbf{r}]$, which is available from the previous subsection. The distance from every correlated point to the reference point is computed and the correlation coefficients are re-arrayed by the distance order (from small to large), and then a linear interpolation is used to produce the curve of the isotropic correlation coefficient

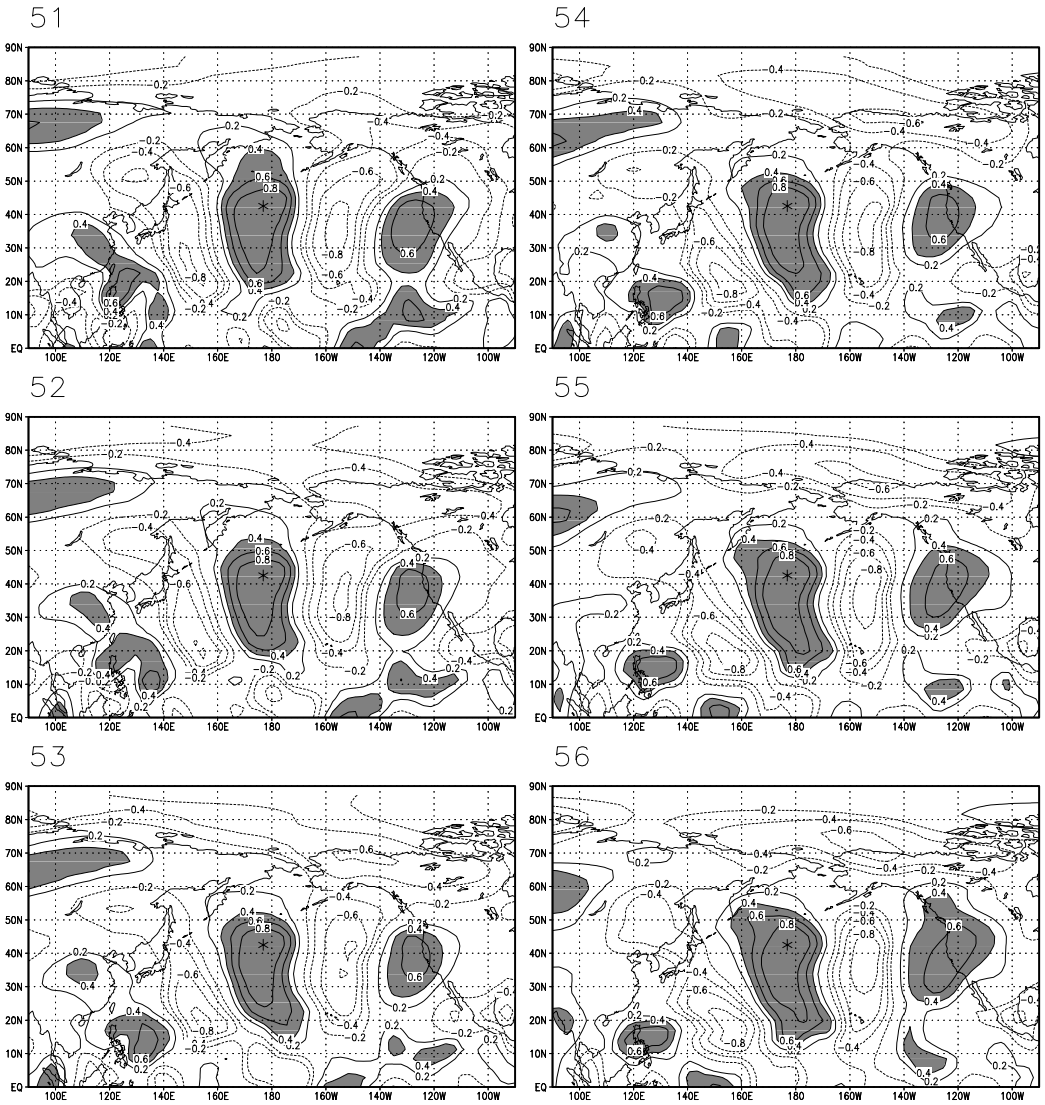

Fig. 3. Same as Fig. 2 except for a daily interval from day 51 to day 56.

$C_0[r_0(\phi), r]$ for a specific latitude ϕ in equal-distance space. Figure 8 shows the correlation coefficient as a function of distance (r) for different latitudes: panel a shows the situation beyond 70°N(S), panel b between 30°N(S) and 60°N(S) and panel c between 20°S and 20°N (each curve in the panels represents a different latitude in the corresponding region). Panels a and c show a nearly axisymmetric correlation structure at high and low latitudes, while panel b exhibits a more oscillatory correlation coefficient with respect to r. This oscillation of the correlation at the middle latitudes is related to the strong anisotropic character over there (quasi-longitudinal wave-train, panel b in Fig. 7, for instance). For example, the ratio of the major axis (in the zonal direction) and the minor axis (in the meridional direction) of the 'elliptic' correlation centers in panel b of Fig. 7 governs the amplitude of the oscillation while the model resolution (gridpoint distance, approximately 300 km in this case) is responsible for the oscillation scale (when the gridpoint space

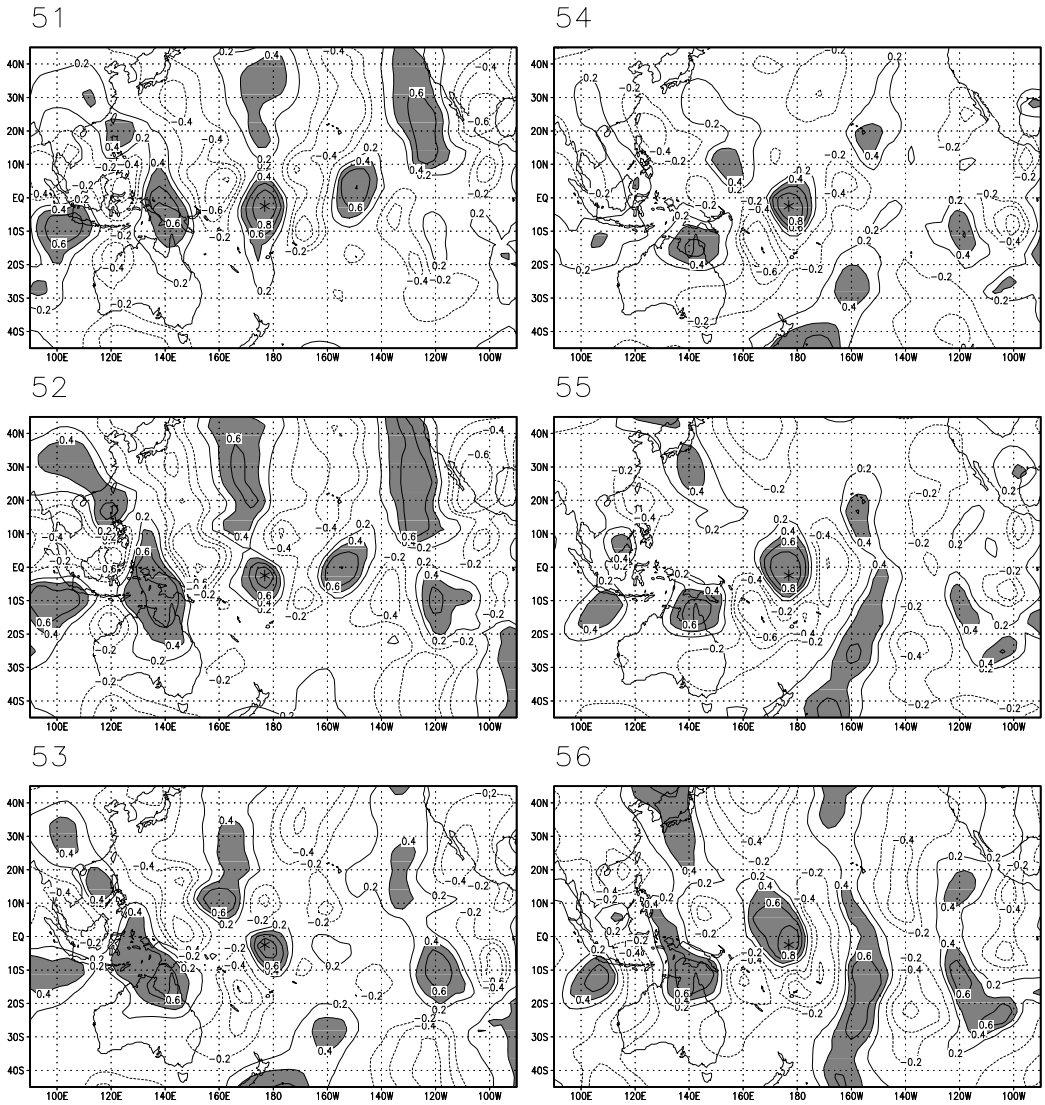

Fig. 4. The daily evolution of the correlation distribution over the domain from 45°S to 45°N and 90°E to 270°E for the reference point (177°E, 3°S) from day 51 to day 56. Details are as in Fig. 2.

is converted to the distance space, the correlation coefficients appear jumping around through the resolved interval) (panel b in Fig. 8)

A globally isotropic streamfunction correlation structure is obtained by averaging results from all 64 latitudes, as shown in Fig. 9 (solid line). The globally isotropic correlation function can be approximately fitted by a theoretical correlation model (Thiebaux, 1976; 1985) as

$$\rho(r) = \left(\cos(cr) + \frac{\sin(cr)}{Lc}\right) \mathrm{e}^{-\frac{r}{L}} \tag{10}$$

where c and L are two free parameters that control the shape of the curve. The correlation coefficients at every 100 km over the first 4000 km (41 values) are used to fit the curve (determining the parameters c and L) by the least-square estimate. Since eq. (10) is a highly nonlinear function, a Newton numerical method

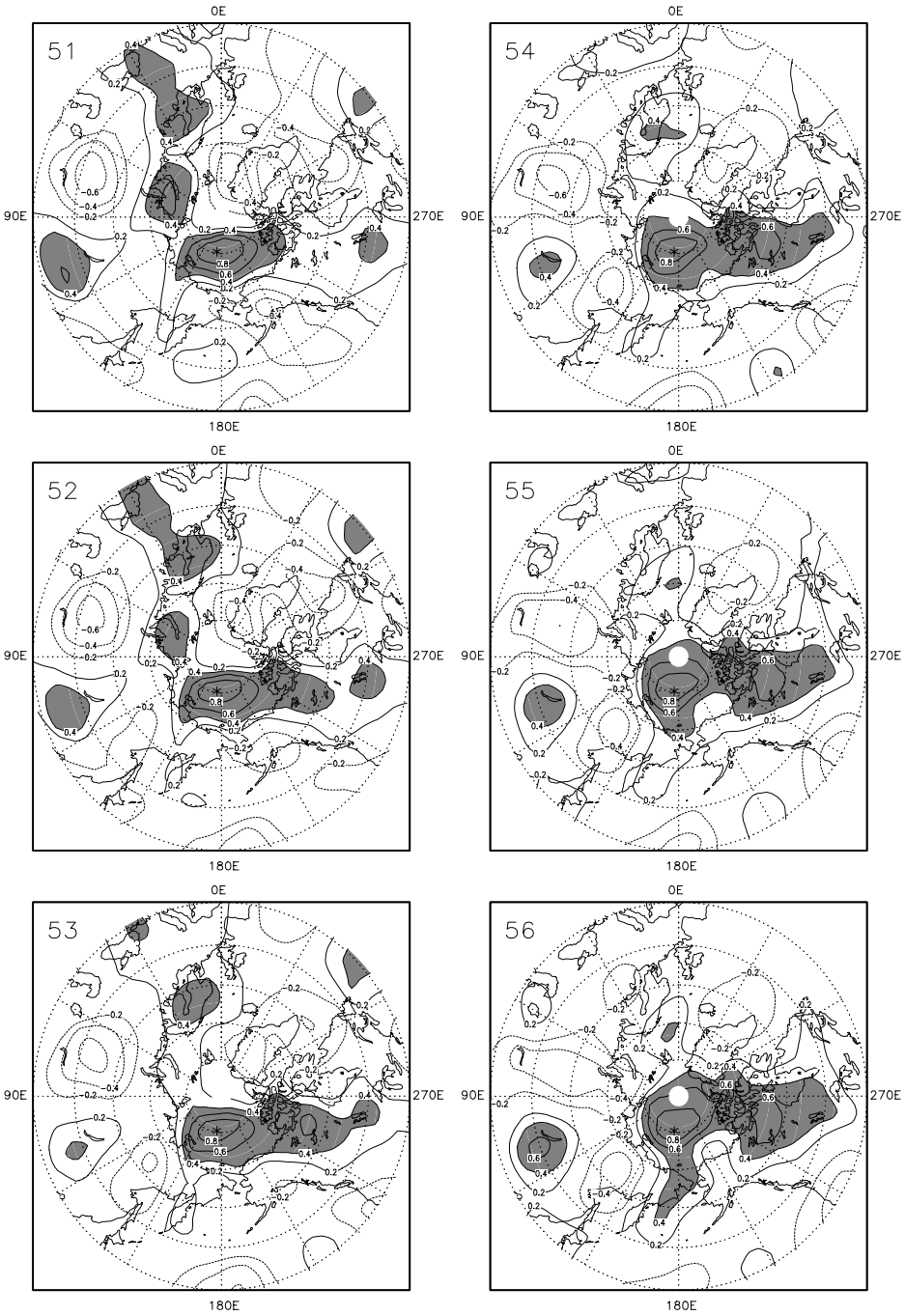

Fig. 5. Same as Fig. 4 except for the reference point (177°E, 82°N) and over the domain of 40°N to 90°N.

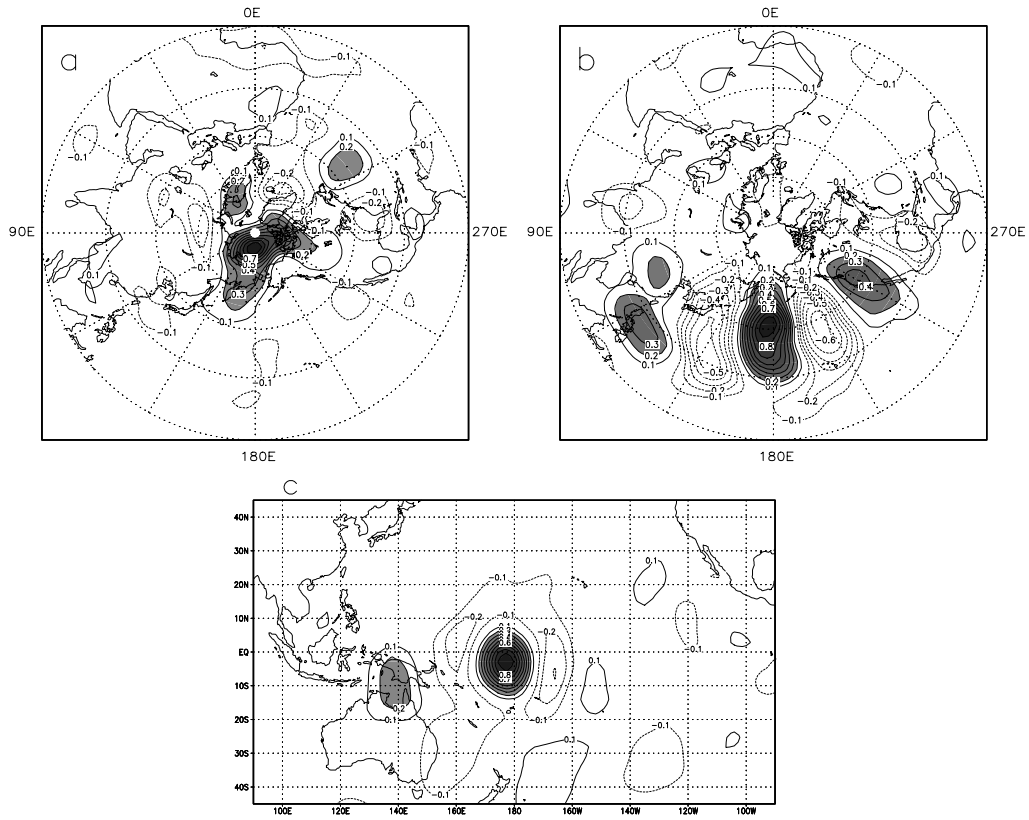

Fig. 6. The distributions of the time-averaged correlation over days 500 to 600) for the reference points (a) (177°E, 82°N), (b) (177°E, 42°N) and (c) (177°E, 3°S). The contour interval is 0.1 and with values greater than 0.2 shaded. The domain for (a) and (b) is the northern hemisphere and for (c) is 45°S to 45°N and 90°E to 270°E. Details are as in Fig. 2.

(Kincaid and cheney, 1996) is employed to approach the approximate stationary point of the Euclidean distance between the 41 derived correlation values and the theoretical values, $\rho(r)$, with respect to c and L. The estimated c and L are respectively 1.25×10^{-6} m^{-1} and 1.25×10^{6} m, as shown by the dashed line in Fig. 9. The solid and dashed lines in Fig. 9 show that the numerically derived globally isotropic correlation function from the EAKF observing/assimilation system simulation experiment is close to the theoretical correlation model. However, section 4.2 will show that the small discrepancy, especially over the short distance, between $C_0(r)$ and $\rho(r)$, will produce the different assimilation results.

3.4. Sensitivity of error correlation estimate to ensemble size

This subsection evaluates the sensitivity of estimated error correlation from the EAKF with respect to ensemble size, using ensembles with 100 and 200 members. Figure 10 presents a daily evolution of the estimated streamfunction spatial correlation using 100 (left) and 200 (right) members during the period from day 51 to day 53 for the middle-latitude reference point (177°E, 42°N). Generally, comparing with the results using 20 members (left panel in Fig. 3), the estimated correlation using 100/200 members has a similar structure but less noise. The difference between the 100-member correlation estimate (left panels in Fig. 10) and the 200-member corrlation estimate (right panels in Fig. 10) is much less than the difference between the 100-member correlation estimate and the 20-member correlation estimate (left panels in Fig. 3). The same results are found for both high and low latitudes (not shown). This means a small ensemble size may introduce some noise in the correlation estimate. Considering that increasing the ensemble size will greatly increase the computational cost, in practice

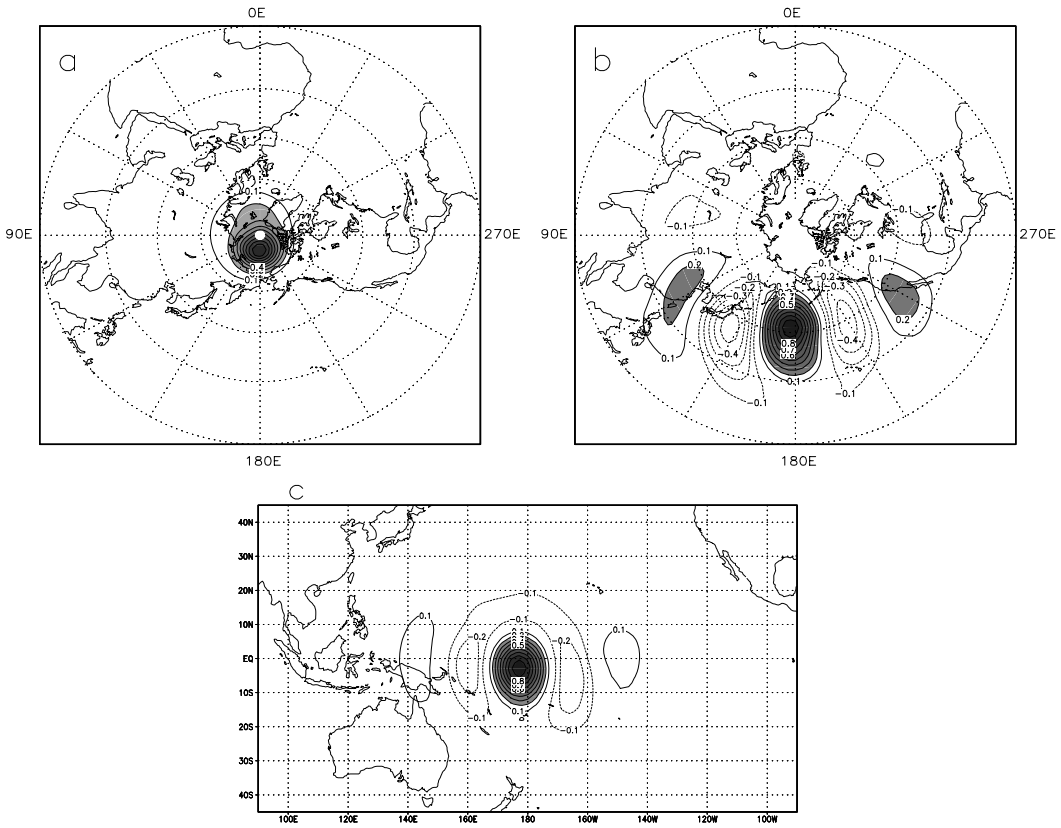

Fig. 7. Same as Fig. 6 except for the zonal mean of correlation distributions of all reference points located at the same latitude.

as a trade-off ones have to choose a practical ensemble size (20 for instance in this study) to estimate the correlation. Figure 11 shows the time averaged estimated spatial correlations using 100 members, which are similar to the results using 20 members (Fig. 6). The globally isotropic correlation functions derived with 20-member (solid line in Fig. 9) and 100-member (dotted line in Fig. 9) ensembles also are quite similar except for small differences before 300 km and between 1000 and 2000 km. This implies that for this barotropic model framework, ensemble size may have a little larger impact on the correlation estimates at these scales. These results reveal that the ensemble-based filtering algorithms, such as EAKF, may be able to produce a reasonable correlation distribution of the model state using a relatively small sample ensemble (20 in this case). The reasons include that the spatial correlation of the state variables is governed by the dynamics of various scale flows such as propagating waves. More detailed discussions will be given in section 5, where estimates of correlation and covariance using the ensemble technique are compared.

4. Sensitivity of the ensemble adjustment filtering assimilation to different correlation structures

4.1. Motivation and experiment design

All modern data assimilation algorithms are closely related. For example, Miller (1998) derived the formulation of the Kalman–Bucy filter starting from the Euler–Lagrange equations of a weak constraint four-dimensional variational problem. In addition, *Optimal interpolation* (OI) (Gandin, 1963) is a simplification of the Kalman–Bucy filter with an error covariance matrix that does not vary in time, while the

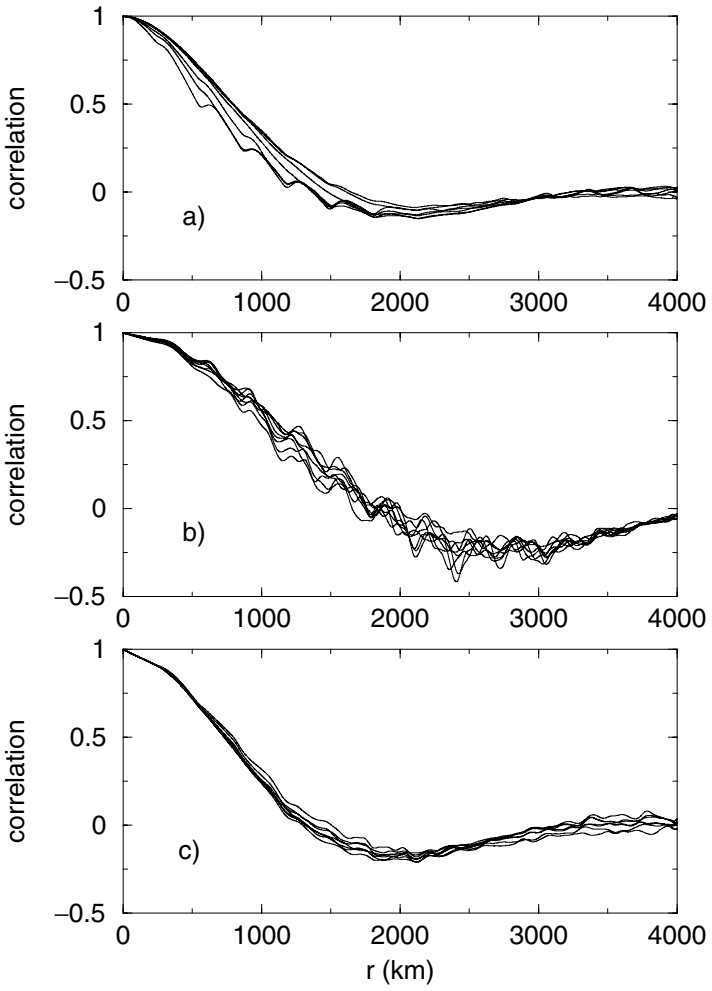

Fig. 8. Isotropic correlation functions in distance space for (a) poleward of 70°N(S), (b) between 30°N(S) and 60°N(S) and (c) between 20°S and 20°N. Each curve in (a), (b), and (c) represents a model-defined latitude in that region. Details are as in Fig. 2.

Kalman–Bucy filter is a simplification of the nonlinear (ensemble) filter to the case of linear dynamics and linear observational operators. Examining the sensitivity of a nonlinear filtering assimilation algorithm such as the EAKF to the temporal and spatial variation of error correlation can increase understanding of the related data-assimilation algorithms.

In this section, the EAKF assimilation algorithm is applied using different error correlation structures. To do this, every analysis step (precisely, step 2 in section 2.2) only computes the error variances (diagonal elements of Σ^p), while the error covariance (off-diagonal elements of Σ^p) between a model gridpoint and an observational location is obtained using the computed variances and a previously computed correlation structure $\{\rho(r), C_0(r), [\overline{C}][r_0(\phi), \mathbf{r}]$ or $\overline{C}(\mathbf{r}_0, \mathbf{r})$, derived in last section$\}$. In addition, the sensitivity of the EAKF data-assimilation algorithm to the accuracy of error correlation estimates is examined by comparing the assimilation results using different accuracy error correlation estimates from different ensemble sizes. Finally, the relative importance of increasing ensemble size and enhancing the accuracy of correlation estimates is investigated by conducting

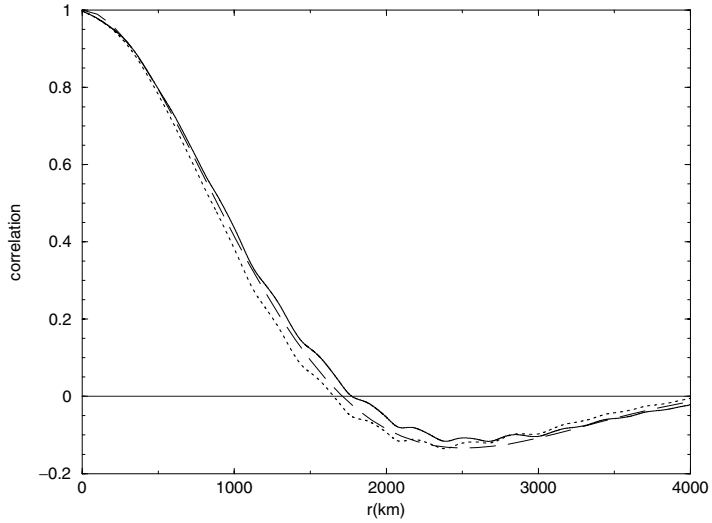

Fig. 9. Globally isotropic correlation functions in distance space, estimated by 20 ensemble members (solid) and 100 ensemble members (dotted), and the theoretical model $\rho(r) = [\cos(cr) + \sin(cr)/Lc]e^{-\frac{r}{L}}$ (dashed) (Thiebaux, 1976; Thiebaux, 1985) with the free parameters c and L set to 1.25×10^{-6} m^{-1} and $L = 1.25 \times 10^6$ m as suggested by least squares using the coefficients at every 100 km over the first 4000 km, derived by a 20-member ensemble. A Newton numerical method (Kincaid and cheney, 1996) is employed to approach the approximate stationary point of the Euclidean distance between the 41 derived correlation values and the theoretical values, $\rho(r)$, with respect to c and L.

assimilation experiments using these error correlation estimates with different accuracy combined with different assimilation ensemble sizes.

4.2. Comparisons of assimilation results

Starting from the end of the 500-d spin-up run described in section 2.3, data assimilation is conducted for another 500 d for five cases using $\rho(r)$, $C_0(r)$, $[\overline{C}][r_0(\phi), \mathbf{r}]$, $\overline{C}(\mathbf{r}_0, \mathbf{r})$ and $C(\mathbf{r}_0, \mathbf{r}, t)$ as the estimates of the background error correlation between the model state variables. The error statistics of assimilation results using 20 ensemble members are listed in Table 1; the statistical results of the 500 d of ensemble forecasts (without assimilation) starting from the ensemble initial conditions described in section 2.3 are also included as case 0. Column 3 of Table 1 gives the values of the covariance inflation factor (γ) that gave a reasonable ratio of the time-averaged root mean square (RMS) error of the ensemble mean (RmsEm) to the mean of the RMS (MRms) error from the individual ensemble members. A good covariance inflation is chosen empirically by testing a number of values. The ratio in column 6 is normalized by the factor $\sqrt{[(M+1)/2M]}$ (Anderson, 1996), the expected value of the ratio for ensemble size M. Values of the normalized ratio close to unity imply that the ensemble has a spread (standard deviation) that is approximately consistent with the truth. For example, in order to determine the value of γ for case 5, several trials were conducted, in which the resulting normalized ratios are 1.128, 1.104, 1.091, 1.080 and 1.113 for $\gamma = 1.00$, 1.02, 1.04, 1.06 and 1.10. The case with $\gamma = 1.06$ was selected for case 5 for comparison in Table 1.

Table 1 shows that all five assimilation cases, including the theoretical isotropic correlation model (case 1) and estimated correlation structures (cases 2–5), reduce both RmsEm and MRms greatly from the pure ensemble forecasts without assimilation (which essentially represent the 'climatological' variability of the model). The use of temporally and spatially varying correlation estimates (the basic EAKF algorithm) produces the smallest assimilation errors (case 5). However, all estimated correlation structures in cases 2–5 improve the assimilation results from the use of theoretical correlation model $\rho(r)$ (case 1). The percentage of assimilation error reduction of cases 2–5 from case 1 is presented in Fig. 12. Although Fig. 9 showed a small difference between $C_0(r)$ (the estimated isotropic correlation structure from EAKF) and $\rho(r)$ (the theoretical correlation model, Thiebaux, 1976; 1985), the use of $C_0(r)$ reduced both RmsEm and MRms by around 6% (case 2) compared

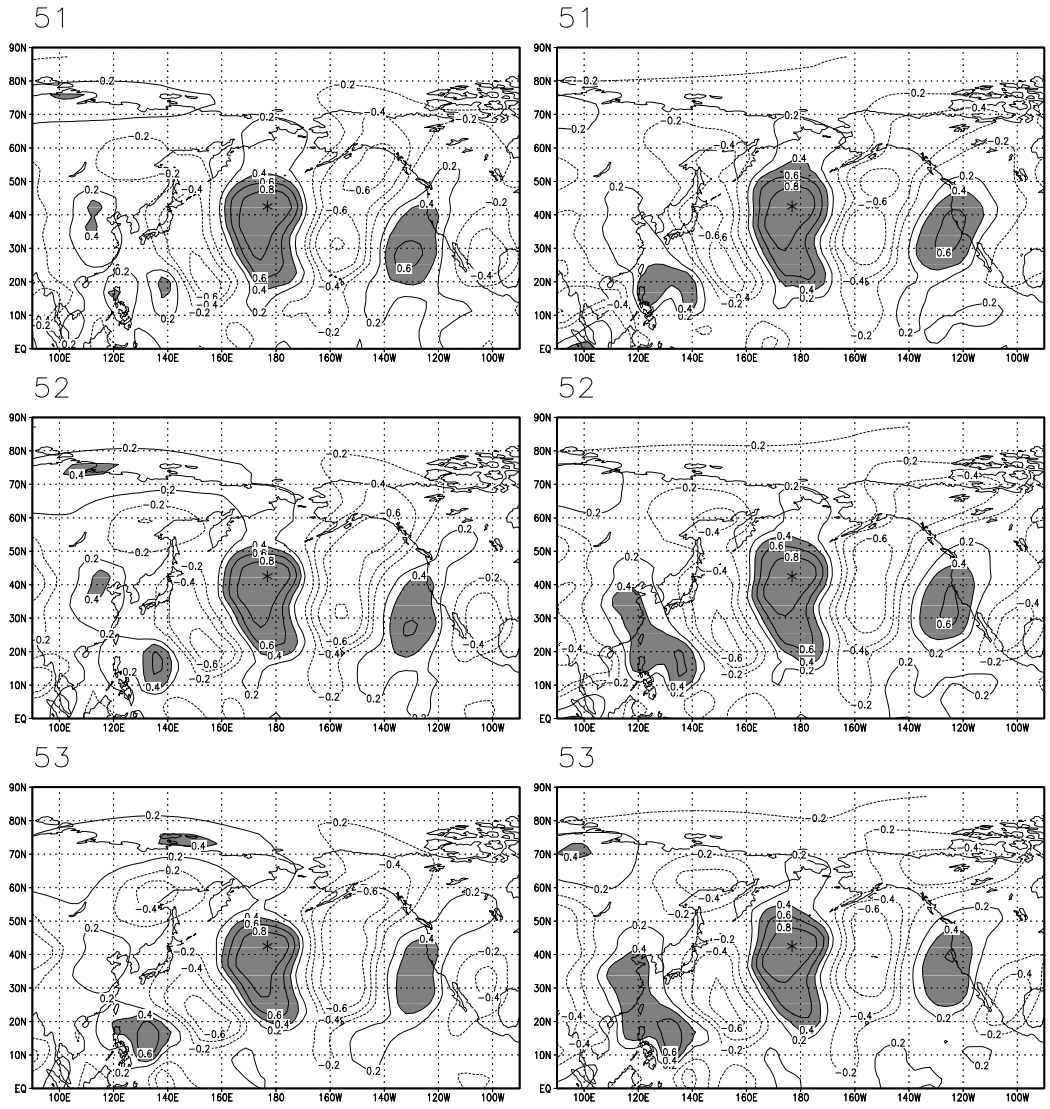

Fig. 10. Same as Fig. 3 except for using 100 ensemble members (left) and 200 ensemble members (right) from day 51 to day 53.

to the use of $\rho(r)$ (case *1*). This means that using the ensemble filter to estimate the isotropic correlation structure is able to provide more consistency between the assimilation model and the correlation model, so that the assimilation process can extract more signal from observations. Particularly, the small difference between $C_0(r)$ and $\rho(r)$ at small distances ($<$300 km in this case) may play a more significant role to reduce the assimilation errors. Using $[\overline{C}][r_0(\phi), \mathbf{r}]$, in which an anisotropic correlation structure is only latitudinally dependent, reduced assimilation errors more (by 10%) (case 3). As more spatial variation was considered [case 4, where each location has a time-invariant anisotropic correlation structure $\overline{C}(\mathbf{r}_0, \mathbf{r})$] the assimilation errors were further reduced by 12% and 15% for RmsEm and MRms, respectively. Finally fully considering the temporal and spatial variation of error correlation, the EAKF had a best improvement of the assimilation results (reduced RmsEm and MRms by 17% and 22%, respectively)

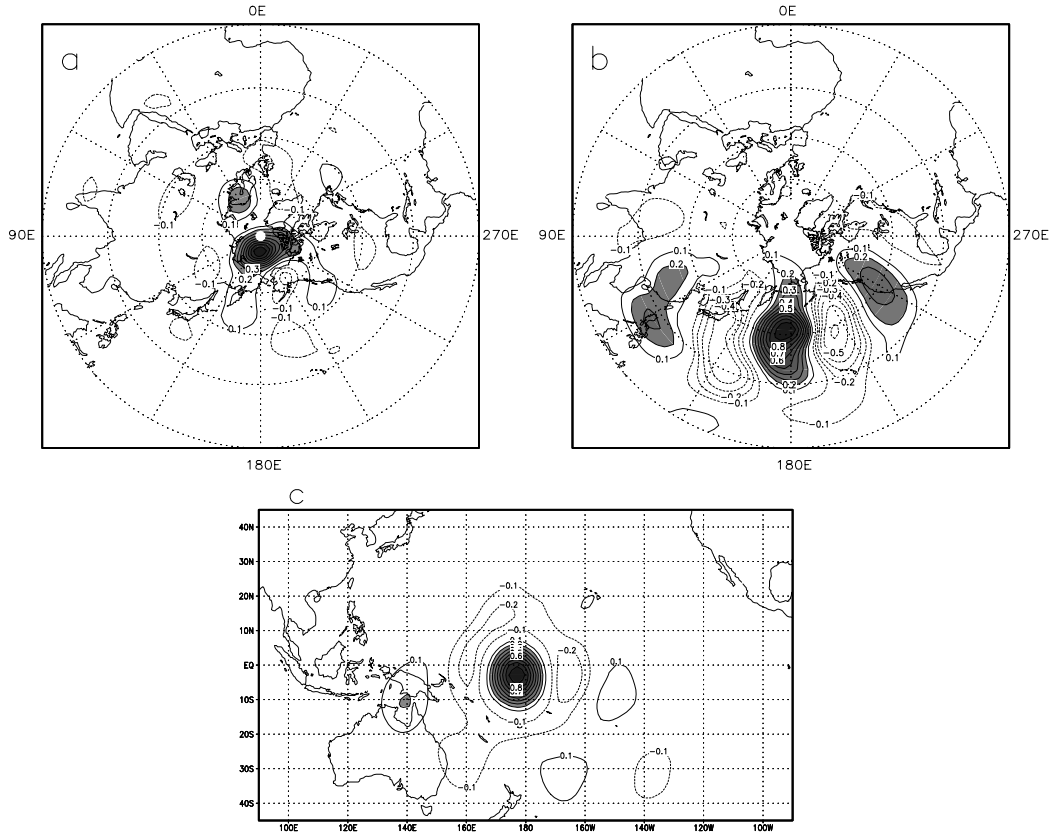

Fig. 11. Same as Fig. 6 except for using 100 ensemble members.

Table 1. *Error statistics of 500-d EAKF assimilation from day 500 to day 1000 using different correlation structures*

Case no.	Correlation model	Cov-inflate γ	RmsEm ($m^2\ s^{-1}$)	MRms ($m^2\ s^{-1}$)	Normalized ratio
0	-	-	16.32×10^5	22.0×10^5	1.024
1	$\rho(r)$	1.020	1.94×10^5	2.64×10^5	1.015
2	$C_0(r)$	1.015	1.83×10^5	2.48×10^5	1.020
3	$[\overline{C}][r_0(\phi), \mathbf{r}]$	1.015	1.74×10^5	2.37×10^5	1.015
4	$\overline{C}(\mathbf{r}_0, \mathbf{r})$	1.015	1.71×10^5	2.25×10^5	1.050
5	$C(\mathbf{r}_0, \mathbf{r}, t)$	1.060	1.60×10^5	2.05×10^5	1.080

RmsEm, RMS of ensemble mean.
MRms, mean of RMS.
Case 0, ensemble forecasts (without assimilation).
Case 1, using a theoretical correlation model described in section 3.3 (dashed line in Fig. 9).
Case 2, using a globally isotropic correlation structure derived in section 3.3 (solid line in Fig. 9).
Case 3, using the anisotropic correlation structure for each latitude derived in section 3.2 (Fig. 7).
Case 4, using the anisotropic correlation structure for each observational location derived in section 3.2 (Fig. 6).
Case 5, using the temporally and spatially varying correlation estimate in the EAKF algorithm.

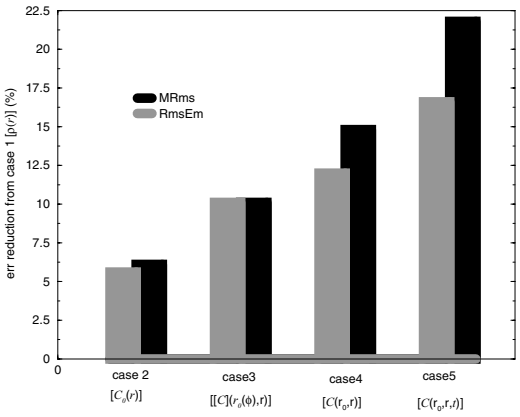

Fig. 12. Percentage of assimilation error reduction by using the EAKF estimated background error correlation structures (cases 2–5 in Table 1) from the theoretical correlation model $\rho(r)$ (case 1 in Table 1).

from the use of a theoretical isotropic correlation model.

These results showed that considering the temporal and spatial variations of error correlation can significantly improve the assimilation results.

The relative importance of improved correlation estimates from larger ensemble sizes as opposed to other improvements from increased ensemble size can be examined. 100-d assimilations using different ensemble sizes (20 or 100) and estimated correlation structures previously generated by various ensemble sizes are conducted, starting from the end of the 500-d spin-up run. As described in the beginning of this section, the values of γ are chosen by experimentation as 1.01 for all cases with 100 members, 1.015 for cases 2–4 with 20 members and 1.06 for case 5 with 20 members. Error statistics are exhibited in Table 2, which shows that the assimilation results are improved by either increasing ensemble size or using refined correlation estimates. Table 3 gives the percentage of error reduction by increasing assimilation ensemble size or using the refined correlation estimate. For example, columns 2 and 3 tell that using the 20-member correlation estimate but increasing assimilation ensemble size from 20 to 100 reduces RmsEm and MRms by 10% and 6%, while columns 6 and 7 tell that using 20 members as an assimilation ensemble size but changing the correlation estimate from cor_{20} to cor_{100} reduce both RmsEm and MRms only by 2%. Table 3 showed that increasing assimilation size improves the assimilation results much more than using refined correlation estimates.

5. Ensemble OI assimilation scheme with different covariance matrices

Using time-invariate correlation structures $\rho(r)$, $C_0(r)$, $[\overline{C}][r_0(\phi), \mathbf{r}]$ and $\overline{C}(\mathbf{r}_0, \mathbf{r})$ and time-invariate variance estimates, the EAKF assimilation algorithm can be degraded to an ensemble OI scheme. In this case, each ensemble member is independently adjusted by a time-invariate covariance matrix (Σ^{p}). In this section, assimilation results from various

Table 2. *Error statistics of 100-d EAKF assimilation from day 500 to day 600 using different ensemble sizes and estimated correlations*

Case no. (correlation model)	20 members				100 members			
	RmsEm ($\text{m}^2\ \text{s}^{-1}$)		MRms ($\text{m}^2\ \text{s}^{-1}$)		RmsEm ($\text{m}^2\ \text{s}^{-1}$)		MRms ($\text{m}^2\ \text{s}^{-1}$)	
	cor_{20}	cor_{100}	cor_{20}	cor_{100}	cor_{20}	cor_{100}	cor_{20}	cor_{100}
2 $[C_0(r)]$	1.77×10^5	1.73×10^5	2.36×10^5	2.30×10^5	1.60×10^5	1.55×10^5	2.22×10^5	2.16×10^5
3 $[[\overline{C}](r_0(\phi), \mathbf{r})]$	1.81×10^5	1.72×10^5	2.31×10^5	2.21×10^5	1.61×10^5	1.53×10^5	2.15×10^5	2.06×10^5
4 $[\overline{C}(\mathbf{r}_0, \mathbf{r})]$	1.85×10^5	1.77×10^5	2.30×10^5	2.20×10^5	1.73×10^5	1.62×10^5	2.19×10^5	2.07×10^5
5 $[C(\mathbf{r}_0, \mathbf{r}, t)]$	1.67×10^5	-	2.07×10^5	-	-	1.22×10^5	-	1.59×10^5

cor_{20}, estimated correlation distribution using 20 ensemble members.
cor_{100}, estimated correlation distribution using 100 ensemble members.

Table 3. *Reduction of errors by increasing ensemble size from 20 to 100 and/or using estimated correlation structures from different ensemble sizes for 100-d assimilation from day 500 to day 600*

Case no. (correlation model)	Reduction of errors									
	Increasing ensemble size from 20 to 100				Refining cor-estimate from 20 to 100 members					
	cor_{20}		cor_{100}		20-member		100-member		Both	
	RmsEm	MRms	RmsEm	MRms	RmsEm	MRms	RmsEm	MRms	RmsEm	MRms
2 $[C_0(r)]$	10%	6%	11%	6%	2%	2%	3%	3%	13%	8%
3 $\{[\overline{C}][r_0(\phi), \mathbf{r}]\}$	11%	7%	11%	7%	5%	4%	5%	4%	16%	11%
4 $[\overline{C}(\mathbf{r}_0, \mathbf{r})]$	7%	5%	9%	6%	4%	4%	7%	6%	13%	10%
5 $[C(\mathbf{r}_0, \mathbf{r}, t)]$	-	-	-	-	-	-	-	-	27%	23%

ensemble OI schemes using different covariance matrices that are computed using different ensemble sizes (20 and 100) and different correlation structures (cases 1–4 in Table 1) are examined ($\gamma = 1.01$). Figure 13 exhibits the estimated time-invariate variances from the EAKF using 20-member (top) and 100-member (bottom) ensembles over a global domain. Larger variances of the streamfunction in the barotropic model are found over the subtropics and south polar area, and the estimated variance using the 100-member ensemble is larger than that using the 20-member ensemble. The covariance matrices Σ^{p}_{20m} and Σ^{p}_{100m} for the four different correlation structure cases are used in place of the sample prior covariance in step 2 in the EAKF algorithm described in section 2.2 to carry out different ensemble OI schemes. The error statistics of assimilation results during day 500 to day 1000 using 20 ensemble members are listed in Table 4, which shows that for the Σ^{p}_{20m} case, both the RmsEm (column 3) and the MRms (column 6) of all four ensemble OI cases are worse than for the EAKF (see case 5 in Table 1). Among these four cases, the assimilation errors from those two cases that have anisotropic correlation structure (cases 3–4) are smaller than those with isotropic correlation structure (cases 1 and 2). Case 4, which fully considers the spatial variation of covariance structure, gives the best ensemble OI assimilation result.

Table 4 also shows that the use of the refined covariance matrix generated by 100 ensemble members

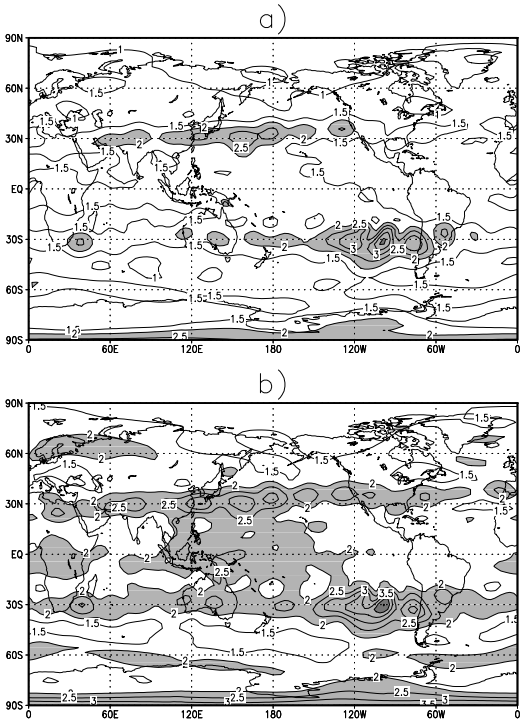

Fig. 13. The distributions of the time-averaged standard deviation over days 500–600 estimated by the EAKF using (a) 20-member and (b) 100-member ensembles. The contour interval is 0.5 ($\times 10^5$ m^2 s^{-1}). Values greater than 2 ($\times 10^5$ m^2 s^{-1}) are lightly shaded and values greater than 4 ($\times 10^5$ m^2 s^{-1}) are heavily shaded.

Table 4. *Error statistics of 20-member ensemble OI assimilation from day 500 to day 1000 using different* Σ^{pa}[a]

Case no.	Correlation model	RmsEm ($m^2\ s^{-1}$)			MRms ($m^2\ s^{-1}$)		
		Σ^{p}_{20m}	Σ^{p}_{100m}	Reduction	Σ^{p}_{20m}	Σ^{p}_{100m}	Reduction
1	$\rho(r)$	3.07×10^5	1.89×10^5	38%	5.27×10^5	2.53×10^5	52%
2	$C_0(r)$	3.14×10^5	1.71×10^5	46%	5.41×10^5	2.38×10^5	56%
3	$[\overline{C}][r_0(\phi), \mathbf{r}]$	2.46×10^5	1.66×10^5	33%	3.93×10^5	1.92×10^5	51%
4	$\overline{C}(\mathbf{r}_0, \mathbf{r})$	2.28×10^5	1.9×10^5	17%	3.58×10^5	2.01×10^5	44%

[a]The percent reduction in error resulting from the use of covariances estimated from 100 member ensembles, rather than 20 member ensembles, is also shown.
Σ^{p}_{20m}, estimated covariance matrix using 20 ensemble members.
Σ^{p}_{100m}, estimated covariance matrix using 100 ensemble members.

greatly reduces the assimilation errors for all cases; RmsEm and MRms are reduced on average by 35% and 52% respectively. Generally, the magnitude of the assimilation error reduction due to the use of Σ^{p}_{100m} with isotropic correlation structure $[\rho(r), C_0(r)]$ is greater than that with anisotropic correlation structure $\{[\overline{C}][r_0(\phi), \mathbf{r}], \overline{C}(\mathbf{r}_0, \mathbf{r})\}$, in which the use of Σ^{p}_{100m} with $C_0(r)$ causes the greatest error reduction (46% for RmsEm and 56% for MRms) and the use of Σ^{p}_{100m} with $\overline{C}(\mathbf{r}_0, \mathbf{r})$ produces the smallest error reduction (17% for RmsEm and 44% for MRms). This is consistent with the distributions of the ensemble RMS errors shown in Fig. 14 for case 2 (top), case 3 (middle) and case 4 (bottom) at the end of the assimilation (day 1000) using Σ^{p}_{20m} (left) and Σ^{p}_{100m} (right) are presented. In addition, using Σ^{p}_{100m}, $\overline{C}(\mathbf{r}_0, \mathbf{r})$ does not give the best OI results. This phenomenon may come from the use of a uniform value of γ and therefore can be improved by refining the γ value. Note that the difference between Σ^{p}_{20m} and Σ^{p}_{100m} in case 1 is only due to the difference in the accuracy of the error variance estimate using different ensemble sizes (shown in Fig. 13). This implies that although the EAKF is able to produce a reasonable correlation distribution with a relatively small ensemble size (20 in this case), the enhanced accuracy of variance estimates by a larger ensemble size may substantially improve the performance of related assimilation algorithms. In other words, in ensemble-based filters, it may be easier to approximate the correlation distribution than the variance magnitude of the model state variables using a relatively small ensemble size, since the former reflects the relationship of the motion status at different spatial locations while the latter reflects a local variability. The spatial relationship of the motion status is governed by the dynamics of various scale flows (usually grid-scale, propagating waves, for instance). Therefore the spatial correlation distribution can be reasonably estimated once the ensembles members used reasonably samples these scale motions. A local variability is, however, related to many other complicated factors (including sub-grid scales) such as turbulence, instability, the wave-flow interactions etc. Then more samples are required to represent these charracteristics for gaining the reasonable magnitude of the variance that includes these factors.

The results and analyses suggest that when using an ensemble-based filter such as the EAKF to estimate the covariance matrix for use in traditional assimilation algorithms, a relatively small ensemble size may be used to estimate correlation structure, while estimating the variances of the model variables with a larger ensemble size may upgrade the assimilation algorithms.

Consistent with the previous study of Hamill and Snyder (2000), as the observational network becomes sparser, the use of a flow-dependent covariance improves the assimilation results more significantly. Reducing the observational network to 300 observations over the global domain, gives RmsEm for ensemble OI (case 4) and EAKF (case 5) as 4.2×10^5 and 2.5×10^5 $m^2\ s^{-1}$, and the MRms for ensemble OI and EAKF are 7.0×10^5 and 3.2×10^5 $m^2\ s^{-1}$. For this sparser observational network, the magnitude of the assimilation error reduction (40% for RmsEm and 54% for MRms) is much greater than that for the network with 600 observations (29% for RmsEm and 42%).

The results above also confirm that under some circumstances an ensemble OI scheme may produce an acceptable assimilation result if used with an appropriately estimated background error covariance matrix.

6. Summary and discussion

Estimating the background error covariance between model state variables is a key issue for

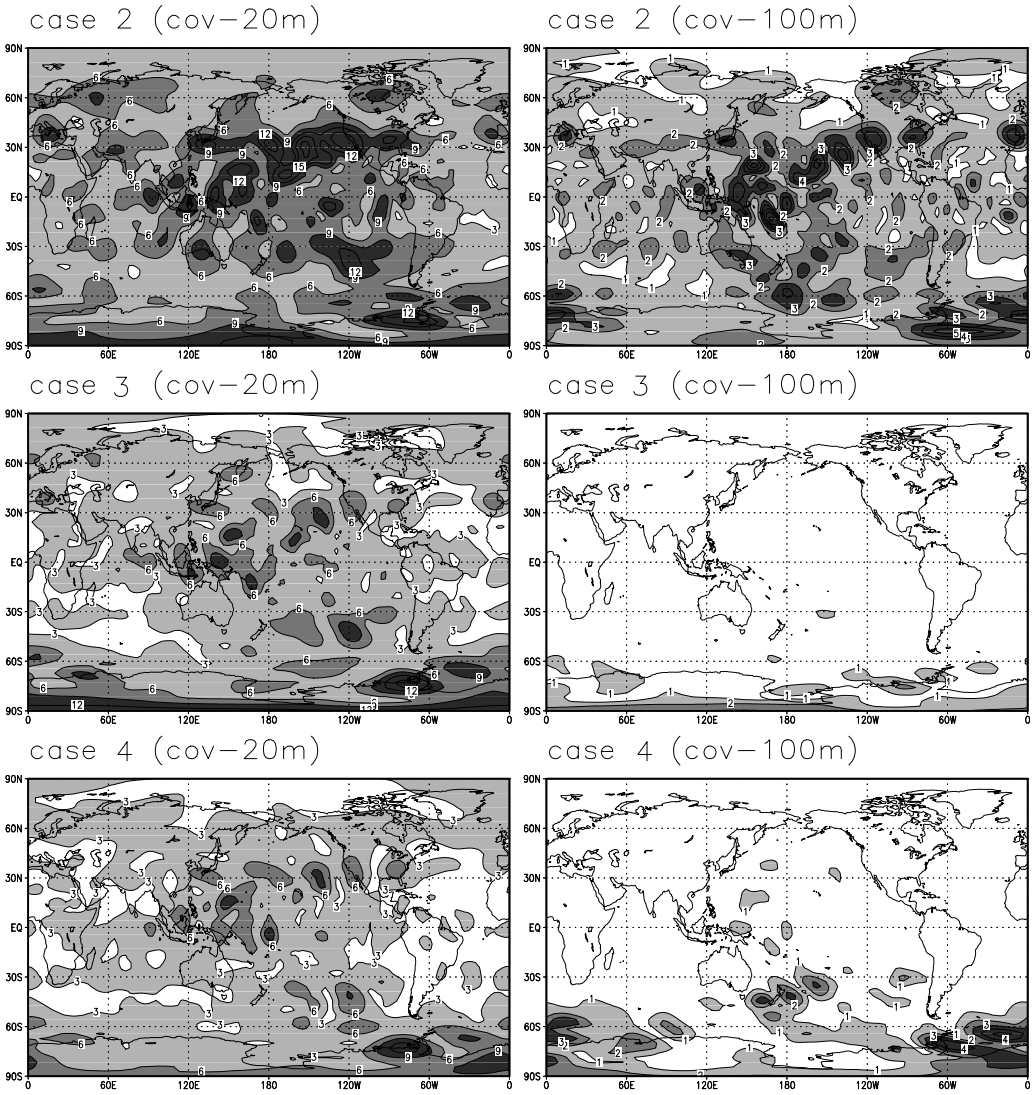

Fig. 14. Distributions of ensemble RMS errors at day 1000 over the global domain for the 20-member ensemble OI assimilation of case 2 (top), case 4 (middle) and case 5 (bottom) using the estimated covariance matrix from 20-member (Σ^{p}_{20m}, left) and 100-member (Σ^{p}_{100m}, right) ensembles. The shading boundaries are 3, 6 and 9 ($\times 10^5$ m^2 s^{-1}) for the $\mathbf{\Sigma}^{p}_{20m}$ situation and 1, 2 and 3 ($\times 10^5$ m^2 s^{-1}) for the Σ^{p}_{100m} situation. The contour interval is 3 ($\times 10^5$ m^2 s^{-1}) for the former and 1 ($\times 10^5$ m^2 s^{-1}) for the latter.

implementing data assimilation and understanding model dynamics. Using the Monte Carlo approach, without solving the stochastic differential equations, this study has estimated the background error covariances generated by the application of an ensemble adjustment Kalman filter (EAKF) to a barotropic spectral model. Results showed that the EAKF can produce reasonably accurate estimates of the error correlation structure with a practical ensemble size (20 in this case). The use of flow-dependent correlation structures generated by the EAKF was shown to be superior to the use of a variety of time-averaged flow-independent correlation structures in this context.

An ensemble optimal interpolation (OI) assimilation scheme was designed using a time-invariate covariance matrix to adjust a prior ensemble at every observation time. The covariance matrix for the ensemble OI is constructed from an estimated isotropic or anisotropic error correlation structure (obtained from the EAKF assimilation) combined with a time-averaged error variance distribution. Five different flow-independent covariance matrices were generated using five different error correlation estimates. The impact of the use of temporally and spatially varying error covariance estimates was evaluated by comparing the ensemble OI assimilations to the EAKF results. The examination produced three key findings: (1) For a given ensemble size, an ensemble filter such as EAKF produces the best assimilation since its flow-dependent error covariance estimates are able to reflect more about the synoptic-scale wave structure in the assimilated flows; (2) an ensemble OI scheme may also produce reasonably good assimilation results if the flow-independent covariance matrix is appropriately chosen; (3) when using the EAKF to estimate the error covariance matrix for improving traditional assimilation algorithms such as variational analysis and OI, a relatively small ensemble size may be used to estimate the error correlation structure while the variances of the model variables estimated by a larger ensemble size may upgrade the related assimilation algorithms.

A stationary (flow-independent) error covariance matrix is the first-order approximation of error covariance matrix. Due to the rapid advances on computer resource and assimilation methodology, some authors have begun to pay attention to the trial that introduces the temporally varying information of error covariance matrix into data assimilation (Ghil et al., 1981; Dee, 1991; Bouttier, 1993; Ehrendorfer and Tribbia, 1997; Hamill and Snyder, 2000). The results of this study further showed that the use of the temporally varying information of error covariance matrix may significantly improve the results of data assimilation. Therefore more attention on this issue may bring more benefits for this community.

This study only evaluated the impacts of error covariances in a univariate system. Considering more complicated and realistic models and observational networks (Houtekamer and Mitchell, 2001; Hamill et al., 2001) will be required to extend further understanding of the EAKF and other ensemble-based assimilation methods. Initial results suggest that the EAKF will continue to perform well in more realistic models. Results have been promising in applications in the dry dynamical core of a global primitive equation model and in a fully parameterized global numerical weather prediction (NWP) model. Of particular interest in these multivariate models is the structure of the cross-correlation structure between state variables of different kinds. For instance, the cross-correlation structure of moisture variables with other variables can shed light on the potential for assimilating measurements of precipitation. A follow-on study will extend the results presented here to global NWP models and realistic observing systems.

7. Acknowledgements

The authors thank Tony Rosati, Matt Harrison, Brian Soden, and Shree Khare for their comments on earlier versions of this manuscript. Thanks go to Dr. Ngar-Cheung Lau for his suggestions that were useful for improving the original manuscript. Thanks also go to two anonymous reviewers for thorough and helpful comments and suggestions.

REFERENCES

Anderson, J. L. 2002. A local least squares framework for ensemble filtering. *Mon. Wea. Rev.* in pres.

Anderson, J. L. 2001. An ensemble adjustment kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.

Anderson, J. L. and Anderson, S. L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127**, 2741–2758.

Anderson, J. L. 1996. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate* **9**, 1518–1530.

Bishop, C. H., Etherton, B. J. and Majumdar, S. 2001. Adaptive sampling with the ensemble transform Kalman filter, part I. *Mon. Wea. Rev.* **129**, 420–436.

Bouttier, F. 1993. The dynamics of error covariances in a barotropic model. *Tellus* **45A**, 408–423.

Buell, C. 1960. The structure of two-point wind correlations in the atmosphere. *J. Geophys. Res.* **65**, 3353–3366.

Buell, C. 1971. Two-point wind correlations on an isobaric surface in a non-homogeneous non-isotropic atmosphere. *J. Appl. Meteorol.* **10**, 1266–1274.

Buell, C. 1972a. Correlation functions for wind and geopotential on isobaric surface. *J. Appl. Meteorol.* **11**, 51–59.

Buell, C. 1972b. Variability of wind with distance and time on an isobaric surface. *J. Appl. Meteorol.* **11**, 1085–1091.

Buell, C. and Seaman, R. 1983. The 'scissors effect': anisotropic and ageostrophic influences on wind correlation coefficients. *Aust. Meteorol. Mag.* **31**, 77–83.

Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.* **126**, 1719–1724.

Cohn, S. E. 1993. Dynamics of short-term univariate forecast error covariances. *Mon. Wea. Rev.* **121**, 3123–3149.

Daley, R. 1991. *Atmospheric data analysis*. Cambridge University Press, New York, 457 pp.

Dee, D. P. 1991. Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. R. Meteorol. Soc.* **117**, 365–384.

Ehrendorfer, M. and Tribbia, J. 1997. Optimal prediction of forecast error covariances through singular vectors. *J. Atoms. Sci.* **54**, 286–313.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**, 10143–10162.

Gandin, L. S. 1963. Objective analysis of meteorological fields. *Gidrometeorolicheskoe Izdatelstvo,* Leningrad. English translation by: Israel Program for Scientific Translations, 242 pp. [NTIS N6618047, Library of Congress QC9 6, G3313].

Gardiner, C. W. 1983. *Handbook of stochastic methods for physics, chemistry, and the natural sciences*. Springer-Verlag, Berlin, 442 pp.

Ghil, M., Cohn, S., Tavantzis, J., Bube K. and Isaacson, E. 1981. Applications of estimation theory to numerical weather prediction. In *Dynamical meteorology: Data assimilation methods,*. Bengtsson et al., eds. Springer-Verlag, New York, 139–224.

Gleeson, T. A. 1961. A statistical theory of meteorological measurements and predictions. *J. Meteorol.* **18**, 192–198.

Hamill, T. M. and Snyder, C. 2000. A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.* **128**, 2905–2919.

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

Hollingsworth, A. and Lonnberg, P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **38A**, 111–136.

Hoskins, B. J. and Karoly, D. J. 1981. The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atoms. Sci.* **38**, 1179–1196.

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

Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.

Houtekamer, P. L., Lefaivre, L. and Derome, J. 1996. The RPN ensemble prediction system. In: Proc. ECMWF Seminar on Predictability, Vol. II, Reading, UK, 121–146.

Jazwinski, A. H. 1970. *Stochastic processes and filtering theory*. Academic Press, New York, 376 pp.

Kalman, R. 1960. A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D* **82**, 35–45.

Kalman, R. and Bucy, R. 1961. New results in linear filtering and prediction theory. *Trans. ASME, Ser. D* **83**, 95–109.

Kalnay, E. and Toth, Z. 1996. Ensemble prediction at NCEP. Preprints *11th Conf. on Numerical Weather Prediction*, Am. Meteorol. Soc., Norfolk, VA, 191–120.

Keppenne, C. L. 2000. Data assimilation into a primitive equation model with a parallel ensemble Kalman filter. *Mon. Wea. Rev.* **128**, 1971–1981.

Kincaid, D. and Cheney, W. 1996. *Numerical Analysis*, 2nd Ed., Brooks/Cole Publishing Co., CA, USA, 804 pp.

Leith, C. E. 1974. Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.* **102**, 409–418.

Lorenz, E. N. 1963. Deterministic non-periodic flow. *J. Atoms. Sci.* **20**, 130–141.

Lorenz, E. N. 1984. Irregularity: A fundemental property of the atmosphere. *Tellus* **21**, 739–759.

Miller, R. N. 1998. Introduction to the Kalman filter. In: *Proceedings of the ECMWF Seminar on Data Assimilation,*. European Center for Medium Range Weather Forecasting, Shinfield Park, Reading, UK, 9–11 September 1996, 47–59.

Miller, R. N., Carter, E. F. and Blue, S. T. 1999. Data assimilation into nonlinear stochastic models. *Tellus* **51A**, 167–194.

Miller, R. N., Ghil, M. and Gauthiez, P. 1994. Advanced data assimilation in strongly nonlinear dynamical system. *J. Atoms. Sci.* **51**, 1037–1056.

Mitchell, H. L. and Houtekamer, P. L. 2000. An adaptive ensemble Kalman filter. *Mon. Wea. Rev.* **128**, 416–433.

Molteni, F. R. B., Palmer, T. N. and Petroliagis, T. 1996. The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. R. Meteorol. Soc.* **122**, 73–119.

Parrish, D. F. and Deber, J. C. 1992. The national meteorological center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.

Seaman, R. and Gauntlett, F. 1980. Directional dependence of zonal and meridional wind correlation coefficients. *Aust. Meteorol. Mag.* **28**, 217–321.

Thiebaux, H. J. 1976. Anisotropic correlation functions for objective analysis. *Mon. Wea. Rev.* **104**, 994–1002.

Thiebaux, H. J. 1985. On approximations to geopotential and wind-field correlation structures. *Tellus* **37A**, 126–131.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2002. Ensemble square-root filters. *Mon. Wea. Rev.* in press.

Van Leeuwen, P. J. 1999. Comment on "Data assimilation using an ensemble Kalman filter technique." *Mon. Wea. Rev.* **127**, 1374–1377.

Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* in press.