Data Assimilation Using an Ensemble Kalman Filter Technique

P. L. HOUTEKAMER AND HERSCHEL L. MITCHELL

Direction de la Recherche en Météorologie, Atmospheric Environment Service, Dorval, Quebec, Canada

(Manuscript received 18 December 1996, in final form 10 June 1997)

ABSTRACT

The possibility of performing data assimilation using the flow-dependent statistics calculated from an ensemble of short-range forecasts (a technique referred to as ensemble Kalman filtering) is examined in an idealized environment. Using a three-level, quasigeostrophic, T21 model and simulated observations, experiments are performed in a perfect-model context. By using forward interpolation operators from the model state to the observations, the ensemble Kalman filter is able to utilize nonconventional observations.

In order to maintain a representative spread between the ensemble members and avoid a problem of inbreeding, a pair of ensemble Kalman filters is configured so that the assimilation of data using one ensemble of short-range forecasts as background fields employs the weights calculated from the other ensemble of short-range forecasts. This configuration is found to work well: the spread between the ensemble members resembles the difference between the ensemble mean and the true state, except in the case of the smallest ensembles.

A series of 30-day data assimilation cycles is performed using ensembles of different sizes. The results indicate that (i) as the size of the ensembles increases, correlations are estimated more accurately and the root-mean-square analysis error decreases, as expected, and (ii) ensembles having on the order of 100 members are sufficient to accurately describe local anisotropic, baroclinic correlation structures. Due to the difficulty of accurately estimating the small correlations associated with remote observations, a cutoff radius beyond which observations are not used, is implemented. It is found that (a) for a given ensemble size there is an optimal value of this cutoff radius, and (b) the optimal cutoff radius increases as the ensemble size increases.

1. Introduction

For two decades, statistical interpolation (e.g., Lorenc 1981), and more recently the closely related three-dimensional variational (3DVAR) algorithm (e.g., Parrish and Derber 1992), have been the foremost data assimilation methods for operational numerical weather prediction (NWP). Applied multivariately, with different observational errors for different types of observations and using a short-range forecast (typically 6 h) as background, these methods have proved capable of combining information from model forecasts and the heterogeneous set of available observations. They have supported a remarkable increase in forecast accuracy over the period, as shown by, for example, Kalnay et al. (1990, Figs. 4 and 5), Hollingsworth and Lönnberg (1990), and Mitchell et al. (1993, Fig. 14).

Of the various aspects of statistical interpolation, the one that has received the most attention in the literature has been the determination and specification of the forecast and observation error statistics (see, e.g., Hollingsworth and Lönnberg 1986; Lönnberg and Hollingsworth 1986; Bartello and Mitchell 1992; Polavarapu 1995; and references therein). Consideration of statistical interpolation in the context of Kalman filtering validates this preoccupation, since, as discussed, for example, by Cohn and Parrish (1991), the analysis of data is the same in the two methods. Rather, it is in the specification of the forecast-error statistics, or more precisely covariances, that the two methods differ. On the one hand, the Kalman filter gives a systematic way to calculate the time evolution of the forecast-error statistics according to the dynamics of the forecast model. In contrast, the statistics used in statistical interpolation and the 3DVAR algorithm are generally taken to be isotropic and largely homogeneous with little variation in time (see e.g., Rabier and McNally 1993 and the references cited above). Despite the demonstrated deficiencies of these restrictions [due to, e.g., data discontinuities (Cohn and Parrish 1991; Daley 1992b), baroclinic zones (Jørgensen 1987), and fronts (Desroziers and Lafore 1993)], it has not yet been possible to eliminate them in operational data assimilation systems.

Following the approach of stochastic–dynamic prediction, proposed by Epstein (1969), Evensen (1994) recently considered the use of an ensemble to estimate the forecast-error statistics. Using Monte Carlo methods to generate the ensemble and assuming four available measurements, he considered data assimilation in the context of a two-layer quasigeostrophic ocean model on

Corresponding author address: Recherche en Prévision Numérique, 2121 Route Trans-Canadienne, Dorval, PQ H9P 1J3, Canada. E-mail: Peter.Houtekamer@ec.gc.ca

© 1998 American Meteorological Society

a 17 × 17 grid. In comparisons with the extended Kalman filter, the new method (termed an ensemble Kalman filter) was shown to work well, requiring ensembles having of the order of 100 members. In a further study (Evensen and van Leeuwen 1996), a 500-member ensemble Kalman filter was used to assimilate gridded Geosat altimeter data for the Agulhas current into a two-layer quasigeostrophic model on a 51 × 65 grid.

Related work with a global atmospheric data assimilation system utilizing a 23-level T63 spectral primitive equations model has been performed by Houtekamer et al. (1996a). Working with an eight-member ensemble, the forecast-error statistics (variances and correlation lengths) obtained from the ensemble and averaged over different regions were compared to the corresponding statistics utilized in the statistical interpolation algorithm. These latter statistics had been obtained, assuming regional and seasonal homogeneity, using the traditional method (e.g., Hollingsworth and Lönnberg 1986; Daley 1991, section 4.3) from radiosonde observation innovations. Although the spread in the ensemble tended to be too small, the possibility of estimating forecast-error statistics from an ensemble showed promise.

The purpose of this study is to further examine the possibility of using ensembles, generated using Monte Carlo methods, to calculate spatially and temporally varying forecast-error covariances for the purpose of performing data assimilation. These flow-dependent statistics will be calculated at each point directly from the ensemble. As in Evensen (1994), they will not be parameterized in terms of simple correlation models, as is normally done, and, of course, they need not be either homogeneous or isotropic. Unlike the previous implementations of the ensemble Kalman filter, a cutoff radius will be used for purposes of data selection and it will be shown that this approach deals effectively with the practical computational problems that arise due to the limited size of the ensemble. In addition, we will examine the correlation structures generated by the ensemble Kalman filter and perform a systematic study of the effect of ensemble size on filter performance.

Our study will be performed in the context of a simplified atmospheric model using simulated radiosonde and satellite observations. These components of the data assimilation system and various aspects of the ensemble Kalman filter are described in the next section.

2. The experimental environment

In a pilot study for a new data assimilation method, intended to be used eventually to obtain analyses of the atmospheric state, one needs an environment that has similar characteristics to the system formed by the atmosphere, a global forecast model, and the observational network. We have decided to use a T21 global spectral model together with a subset of the current observational network. We simulate observations by applying random perturbations to the (known) true state. Here the true state is obtained from a long integration with the model, which is considered perfect. The actual assimilations are performed using a pair of ensemble Kalman filters. With a view toward the real-time constraints associated with operational atmospheric data assimilation, we focus on small- and moderate-sized ensembles.

a. The model

The nonlinear model used in this study is the three-level, quasigeostrophic global spectral model of Marshall and Molteni (1993). It has a resolution of T21, includes orography, and is driven by empirical forcing functions. The global model has been tuned to describe a perpetual winter situation in the Northern Hemisphere. As a measure of the error variance, we use the domain-averaged streamfunction error squared north of 20°N at 50 kPa. Thanks to its ease of use, transparent coding, and realism, the model has by now been used for a large number of predictability studies (e.g., Barkmeijer et al. 1993; Molteni and Palmer 1993; Houtekamer and Derome 1995; Lin and Derome 1996).

b. The observational network

Observations are performed by radiosondes and satellites. Each day the observations are taken at 0000 and 1200 UTC at which times the analysis is performed. We have one network for 0000 UTC and another for 1200 UTC.

1) RADIOSONDES

To obtain a reduced radiosonde network that is more or less realistic, we performed the following procedure.

We started with a list of 671 and 568 radiosonde locations for 0000 and 1200 UTC, respectively. These radiosonde positions have been used by Houtekamer and Derome (1995) and are presented in their Fig. 2. These networks are rather dense in comparison with the number (483) of spectral coefficients at each level of the T21 model. We decided to randomly select only 9% of these sites. For convenience of calculation, we shifted the position of each radiosonde to the nearest point of a reduced Gaussian grid (the original Gaussian grid being the 64 × 32 quadratic grid used by the T21 model). At most, one radiosonde is allowed at each point of the reduced grid. In this way, we obtained 57 (49) radiosondes for the analyses at 0000 (1200) UTC. The positions are presented in Fig. 1.

Radiosondes observe streamfunction values at the three model levels (20, 50, and 80 kPa). At the same levels, the horizontal derivatives of the streamfunction (u and v) are also observed. This yields a total of nine reported values per radiosonde and a total of 513 and

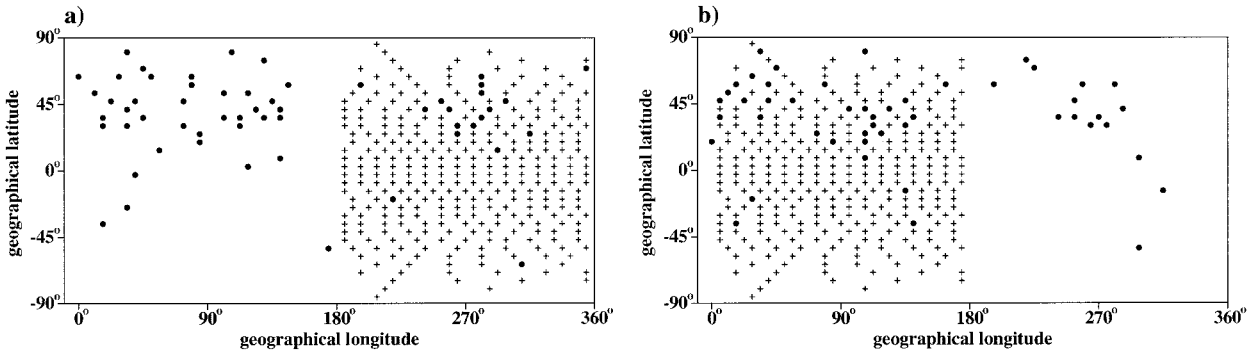

FIG. 1. (a) The observational network at 0000 UTC. Dots indicate the positions of radiosonde observations. The plus signs indicate the positions of satellite observations. (b) As in (a) but at 1200 UTC.

441 observed values at 0000 and 1200 UTC, respectively.

For the observational errors we have for each component of the wind ($m^2\ s^{-2}$)

$$\mathbf{R}_{\text{wind}} = \overline{(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})^{\mathrm{T}}(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})} = \begin{pmatrix} 7.0 & 0.0 & 0.0 \\ 0.0 & 5.4 & 0.4 \\ 0.0 & 0.4 & 4.1 \end{pmatrix}. \tag{1}$$

These values have been modeled after the values used in the operational analysis at the Canadian Meteorological Centre.

For the streamfunction errors we start from the covariances for the heights (m^2) used by the operational analysis:

$$\mathbf{R}_{\text{height}} = \overline{(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})^{\mathrm{T}}(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})} = \begin{pmatrix} 260 & 80 & 20 \\ 80 & 80 & 20 \\ 20 & 20 & 30 \end{pmatrix}. \tag{2}$$

We then divide by the value of f_0^2/g^2, with $f_0 = 2\Omega \sin 45°$, in order to obtain covariances for streamfunction values. (Note $\Omega = 7.292 \times 10^{-5}$ rad s^{-1} and $g = 9.81$ m s^{-2}.)

2) SATELLITES

The satellites report thicknesses at the points of the reduced Gaussian grid. Observations are simulated for the Western Hemisphere at 0000 UTC and for the Eastern Hemisphere at 1200 UTC, as shown in Figs. 1a,b. The satellite observations that were coincident with radiosondes (i.e., over land) have been removed. Altogether we simulate 306 satellite soundings at 0000 UTC and 291 soundings at 1200 UTC. Each sounding consists of two thicknesses: one for the difference between the streamfunction at 20 and 50 kPa, and another for the difference between the values at 50 and 80 kPa.

For the error covariance matrix, $\mathbf{R}_{\text{sat}}$, we were guided once more by the values used by the operational system. We specified (m^2),

$$\mathbf{R}_{\text{sat}} = \overline{\begin{pmatrix} \epsilon_{20} - \epsilon_{50} \\ \epsilon_{50} - \epsilon_{80} \end{pmatrix}(\epsilon_{20} - \epsilon_{50},\ \epsilon_{50} - \epsilon_{80})} = \begin{pmatrix} 1300 & -202 \\ -202 & 450 \end{pmatrix}. \tag{3}$$

Using the same factor f_0^2/g^2 as before, these covariances are then converted to streamfunction.

c. Initial forecast-error covariances

To start the experiments at the initial time t_0, we use a first guess (background) that is not equal to the true state. In fact, the difference is obtained as a realization of a multivariate probability distribution. The same distribution will be used for the generation of an ensemble of first-guess fields valid at t_0. Our choice for the multivariate distribution is based on the forecast error statistics (in particular a vertical covariance matrix and a horizontal correlation function) used by the operational system.

For the vertical covariances of the streamfunction errors, we start from the error covariance matrix (m^2) for height,

$$\mathbf{P}^f_{\text{height}}(t_0) = \overline{(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})^{\mathrm{T}}(\epsilon_{20},\ \epsilon_{50},\ \epsilon_{80})} = \begin{pmatrix} 390 & 80 & 20 \\ 80 & 120 & 20 \\ 20 & 20 & 50 \end{pmatrix}, \tag{4}$$

and convert to streamfunction as after (2).

For the horizontal correlation function, we specify a sum of two third-order autoregressive functions, as in the operational system (Mitchell et al. 1990)

$$\rho(r) = \frac{1}{1+\alpha}\left[\left(1 + cr + \frac{c^2r^2}{3}\right)\exp(-cr) + \alpha\left(1 + \frac{cr}{N} + \frac{c^2r^2}{3N^2}\right)\exp\left(\frac{-cr}{N}\right)\right], \tag{5}$$

where $\alpha = 0.2$ and $N = 3$. The scale parameter c is set to 11.5 rad^{-1}.

As discussed in section 2d, these values will be used at some initial time, t_0, to generate an isotropic and homogeneous error (with respect to a ''true'' state) for an ensemble of first-guess (background) fields.

The sensitivity of our results to these initial conditions is of interest. The asymptotic stability of the Kalman filter is guaranteed only if the dynamics are linear and the system is observable and controllable (Cohn and Dee 1988; Daley 1991, 382; Ménard 1994, Chapter 2; Ghil and Todling 1996). In this study, the dynamics are nonlinear and we have no model error and therefore these theoretical results do not apply. The sensitivity of our results to the initial conditions will be examined experimentally in section 3.

d. Simulation concepts

The simulation of the true atmospheric state is denoted by the symbol $\boldsymbol{\Psi}_t(t)$, where the subscript t stands for ''true.'' It is obtained from a long model integration that is interrupted every 12 h. The true atmospheric state at the initial time t_0, $\boldsymbol{\Psi}_t(t_0)$, is itself the product of a long integration.

Every 12 h from time t_0 onward, we simulate the (imperfect) observations that are available in a real situation. These simulated observations, $\mathbf{O}_s(t)$, where the subscript s stands for ''simulated,'' are generated from the true state using

$$\mathbf{O}_s(t) = \mathbf{H}\boldsymbol{\Psi}_t(t) + \text{a random vector.} \qquad (6)$$

Here $\mathbf{H}$ interpolates the true state to the observed quantities. The random vector (actually a specific realization) is obtained from a multivariate normal distribution having the error statistics of the observations.

To start up our experiments, we require a field $\boldsymbol{\Psi}_c^f(t_0)$, which represents the best available estimate of the true field $\boldsymbol{\Psi}_t(t_0)$. In a real situation the latter is, of course, unknown. The field $\boldsymbol{\Psi}_c^f(t_0)$ is generated by adding a random perturbation field to the true state $\boldsymbol{\Psi}_t(t_0)$:

$$\boldsymbol{\Psi}_c^f(t_0) = \boldsymbol{\Psi}_t(t_0) + \text{a random field,} \qquad (7)$$

where the superscript f indicates that this field is to be used as a forecast (first-guess) field. The random field (again actually a specific realization) is generated as discussed in the appendix using the 3D covariance structure given in section 2c. The field $\boldsymbol{\Psi}_c^f(t_0)$ will serve as a central field for the initial ensemble of first-guess fields, hence the subscript c.

e. Generation and use of the ensembles

We are now in a position to generate the initial ensemble members. As discussed at the end of this section, it will prove useful to have a *pair* of N-member ensembles. To obtain the pair of N-member ensembles of first-guess fields at time t_0, we use the following equation:

$$\boldsymbol{\Psi}_{i,j}^f(t_0) = \boldsymbol{\Psi}_c^f(t_0) + \text{random field}_{i+(j-1)N}, \qquad (8)$$

where the random fields are generated using the same method and prescribed covariance structure as for (7). Here the index i is over the N ensemble members and $j = 1$ or 2, so we require a total of $2N$ different realizations of the random field.

Every 12 h from time t_0 onward, we generate perturbed sets of observations to be assimilated into the different ensemble first-guess fields. These are generated using

$$\mathbf{O}_{i,j}(t) = \mathbf{O}_s(t) + \text{random vector}_{i+(j-1)N}, \qquad (9)$$

where, as for (6), the random vector realizations are obtained from a multivariate normal distribution having the error statistics of the observations. Note that in a real situation, the exact observational error statistics are unknown. Our assumption that these statistics *are* known allows us to generate perturbed sets of observations, $\mathbf{O}_{i,j}$, having the same multivariate distribution as the simulated observations, $\mathbf{O}_s$.

The forecast-error covariance statistics are to be determined from an ensemble of first-guess fields. To do this, we require that the spread among the ensemble members be representative of the difference between the ensemble mean and the true state. Such representative ensembles can, at least in theory, be used to determine forecast-error statistics for the purposes of data assimilation. However, special care must be taken to maintain the representativeness of the ensembles as the assimilation cycles proceed.

To show one potential problem, suppose we have an N-member ensemble of first-guess fields. Computing covariances from the ensemble, one might determine weights for data assimilation. Now one would like to estimate the analysis error distribution that results from using these weights. To test the quality of the weights, we need an independent ensemble of first-guess fields and an independent ensemble of perturbed observations with statistics representative of the actual first-guess and observation errors. If we assimilate these observations into this ensemble of first-guess fields, we obtain an ensemble of analyses. The differences between these analyses will be representative of the analysis error distribution. Thus, we need two ensembles of first-guess fields, one to compute the weights and another to obtain a representative ensemble of analysis errors. If the same ensemble was to be used for both purposes, then the estimation of the gain (i.e., weights) and the test of its quality would be based on exactly the same information. Such a dependent test would likely give an underestimate of the uncertainty in the analysis; that is, the spread in the ensemble of analyses would be too small and, in particular, it would be smaller than the difference be-

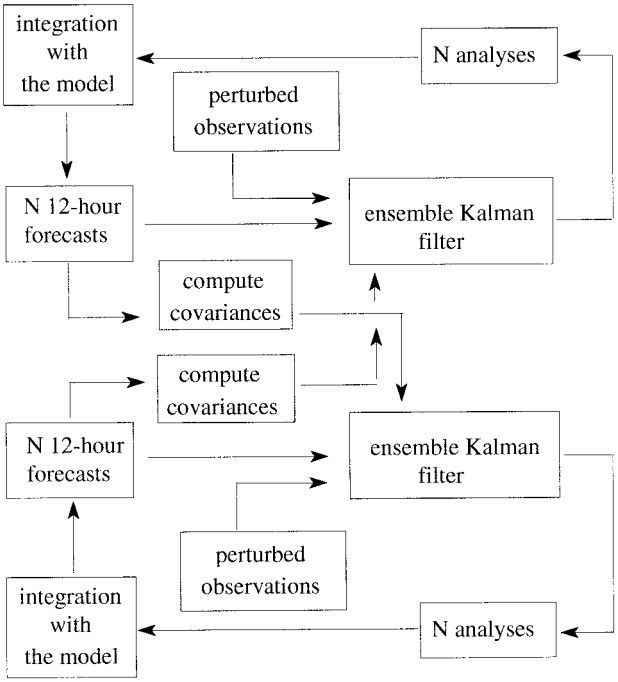

FIG. 2. A setup with a pair of ensemble Kalman filters. Data are assimilated into each ensemble using the covariances computed from the other ensemble.

tween the ensemble mean and the truth. This will be demonstrated experimentally in section 3.

If the weights could be obtained from an ensemble of infinite size, they would of course be optimal. In that case, an independent ensemble would not be needed to confirm this and a single infinite ensemble could ''simply'' be used. This corresponds to the situation with the conventional Kalman filter where the gain matrix is not at all degraded by sampling error.

Our actual configuration is displayed in Fig. 2. Here we have a pair of N-member ensembles. The covariances computed from each ensemble are used to assimilate data into the other ensemble. In this way, each of the two ensemble Kalman filters uses different ensembles of first-guess fields for the estimation of the weights and the estimation of the analysis error.

f. Data assimilation algorithm

For assimilating data we use an algorithm that is very similar to the one used by Evensen (1994). We just list the main equations and draw attention to characteristic properties of the algorithm. The basic equation for the analysis at time t is

$$\boldsymbol{\Psi}^a_{i,j} = \boldsymbol{\Psi}^f_{i,j} + \mathbf{K}_{j'}(\mathbf{O}_{i,j} - \mathbf{H}\boldsymbol{\Psi}^f_{i,j}), \qquad i = 1, \ldots, N, \tag{10}$$

where the superscript a indicates analyses and all quantities apply at time t. Here, j' represents the ensemble that is complementary to ensemble j, that is, $j' = 2$ for $j = 1$ and $j' = 1$ for $j = 2$. The statistics for the assimilation of ensemble j are thus computed from the complementary ensemble j'. The Kalman gain $\mathbf{K}_j$ calculated from ensemble j at time t is given by

$$\mathbf{K}_j = \mathbf{P}^f_j\mathbf{H}^{\mathrm{T}}[\mathbf{H}\mathbf{P}^f_j\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1}, \tag{11}$$

where $\mathbf{R}$ is the observation error covariance matrix, $\mathbf{P}^f_j$ is the 12-h forecast-error covariance matrix of ensemble j, and again all quantities apply at time t.

First of all, we note that the Kalman gain $\mathbf{K}_j$ at time t is the same for all N members of ensemble j'. As pointed out by Evensen (1994) and Houtekamer et al. (1996a), this permits an important computational economy to be realized as compared to the cost of doing N independent analyses. Furthermore, for each point of the horizontal (Gaussian) analysis grid, the analysis is performed independently for each vertical column of three analysis points. This allows for an analysis algorithm that is completely parallel. Exploiting these two points, the analysis part of the experimental Canadian ensemble forecast system (Houtekamer and Lefaivre 1997) is now running conveniently on a powerful workstation.

Unlike what should theoretically be done for a Kalman filter but similar to what is done in traditional statistical interpolation, we perform a data selection. For each vertical column of analysis points, all the data within a given horizontal distance, $r_{\max}$, and only that data, is used. We shall see that this is, in fact, a convenient way to eliminate observations that are only weakly correlated with the analysis point. A standard result from statistical theory states that if two normal distributions have correlation ρ, an estimate $\hat{\rho}$ of ρ based on N pairs has variance

$$\overline{(\rho - \hat{\rho})^2} = \frac{1}{N}(1 - \rho^2)^2 \approx \frac{1}{N}(1 - \hat{\rho}^2)^2 \tag{12}$$

(Gandin 1965, 34–35; Stuart and Ord 1987, 330). It follows that the accurate estimation of these weak correlations would, in any case, require thousands of ensemble members. Furthermore, the use of a cutoff radius, $r_{\max}$, greatly reduces the order of the matrices that we have to deal with, rendering the matrix inversion in (11) feasible. In fact, rather than computing an explicit inverse, (11) is solved using a Cholesky decomposition where only a single decomposition needs to be done for each column of analysis points.

As in Evensen (1994), we do not need to compute or store the full matrix $\mathbf{P}^f_j$. In fact, for the term $\mathbf{P}^f_j\mathbf{H}^{\mathrm{T}}$ of (11), the relevant matrix elements are computed using

$$\mathbf{P}^f_j\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{\Psi}^f_{i,j} - \overline{\boldsymbol{\Psi}^f_j})[\mathbf{H}(\boldsymbol{\Psi}^f_{i,j} - \overline{\boldsymbol{\Psi}^f_j})]^{\mathrm{T}}, \tag{13}$$

where

$$\overline{\boldsymbol{\Psi}^f_j} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Psi}^f_{i,j}. \tag{14}$$

For the term $\mathbf{HP}_j^f\mathbf{H}^T$ of (11), the required elements of the *global* matrix $\mathbf{HP}_j^f\mathbf{H}^T$ were computed as

$$\mathbf{HP}_j^f\mathbf{H}^T = \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{H}(\boldsymbol{\Psi}_{i,j}^f - \overline{\boldsymbol{\Psi}_j^f})[\mathbf{H}(\boldsymbol{\Psi}_{i,j}^f - \overline{\boldsymbol{\Psi}_j^f})]^T. \quad (15)$$

Although ostensibly of order the total number of observations, the global matrix $\mathbf{HP}_j^f\mathbf{H}^T$ is actually sparse, due to the use of a cutoff radius. Since each computed element of this matrix is likely required for the analysis at many grid columns, this sparse matrix was precomputed and stored. For the analysis at a particular grid column, we retrieve the block of relevant elements of the global matrix $\mathbf{HP}_j^f\mathbf{H}^T$. This block matrix has no negative eigenvalues because it could have been calculated directly from the ensemble of first-guess fields using an operator $\mathbf{H}$ that interpolates only to the selected local observations. Its rank will be less than or equal to the minimum of N and the number of such observations. So, as soon as the number N exceeds the number of selected local observations, $\mathbf{HP}_j^f\mathbf{H}^T$ (and $\mathbf{P}_j^f\mathbf{H}^T$) have full rank. In this way we avoid, or in the case of our smallest ensembles alleviate, a potential rank problem with the ensemble Kalman filter. Evensen and van Leeuwen (1996) have presented a different approach for dealing with the rank problem that occurs in the global solution of (11) if there are considerably fewer ensemble members than observations.

Equation (13) evaluates the covariances between forecast values at analysis points and at observation points, while (15) evaluates the covariances between forecast values at observation points. To evaluate these, we require only the forward interpolation operator $\mathbf{H}$ for each type of observation being used. For the u and v wind components, we define $\mathbf{H}$ as the horizontal derivative of streamfunction evaluated in spectral space, while for the satellite thicknesses $\mathbf{H}$ involves taking the appropriate vertical differences.

Note that the data assimilation algorithm does not use any correlation model. Nor does it require the use of adjoint operators. In fact, one just needs to specify where the analysis grid points are and how to interpolate a model state to the observed quantities. This makes the analysis code almost completely independent of the forecast model.

Finally, we note that the covariances of the analysis error can be computed from the ensemble of analyzed states (Evensen 1994). Our current algorithm does not make use of these covariances. The integration of the $2N$-analyzed states with the nonlinear forecast model closes the assimilation loop.

3. Results

To evaluate the performance of the ensemble Kalman filters, we performed data assimilation cycles extending over a 30-day period. The initial time for these cycles (t_0) is denoted 0000 UTC of day 1 and the final time is 0000 UTC of day 31.

TABLE 1. The rms analysis error (10^6 m^2 s^{-1}) after the first analysis for several configurations described in the text. For each configuration, the spread in each of the two ensembles of the pair is shown. All experiments were performed with $r_{max} = 20°$.

Configuration	N	Ensemble 1	Ensemble 2
EnsKF	16	0.846	0.793
EnsKF	32	0.751	0.761
EnsKF	64	0.727	0.728
EnsKF	128	0.705	0.712
OI	128	0.692	0.694

We first examine the root-mean-square (rms) analysis error for streamfunction after the first analysis (i.e., at 0000 UTC of day 1) for several different configurations. This is presented in Table 1. Each of the configurations denoted EnsKF corresponds to a pair of ensemble Kalman filters, each having N members, configured as in Fig. 2. It can be seen that the rms analysis error decreases as N increases, as expected. In addition as N increases, resulting in reduced sampling error, the difference between the rms analysis error in the two ensembles of each pair also decreases. The configuration denoted OI (optimum interpolation) corresponds to a pair of ensembles where each ensemble member is obtained using statistical (i.e., optimum) interpolation. Each of these OI analyses was performed using the same forecast-error statistics that were used to specify the random perturbation fields in (7) and (8), as discussed in section 2c and the appendix. It follows that these OI analyses are, in fact, using the optimum statistics and should yield optimum analyses, that is, with minimum possible error. The results in the table indicate that this is indeed the case, with the rms error for this configuration yielding an asymptotic value for the ensemble Kalman filter analyses.

Results for six 30-day experiments are presented in Fig. 3. Two measures of error are shown in each panel: the rms difference between the ensemble mean and the true state and the rms spread in the ensemble. Of the two measures, the ensemble spread is seen to behave in a more stable fashion than the more erratic error in the ensemble mean. This is due to the fact that the error in the ensemble mean is much more susceptible to the sampling error associated with the single realization of (6), the generation of the simulated observations, than is the ensemble spread, which is the product of an ensemble of realizations of (9).

The left-hand panels of Fig. 3 show the performance of the filter when only a single ensemble (with 16, 32, and 128 members) is used. Each corresponding right-hand panel shows the performance obtained with a pair of ensemble Kalman filters, configured as in Fig. 2 and with the same *total* number of ensemble members. The results exhibited in each of these right-hand panels are calculated from the first ensemble of the pair only. (Cor-

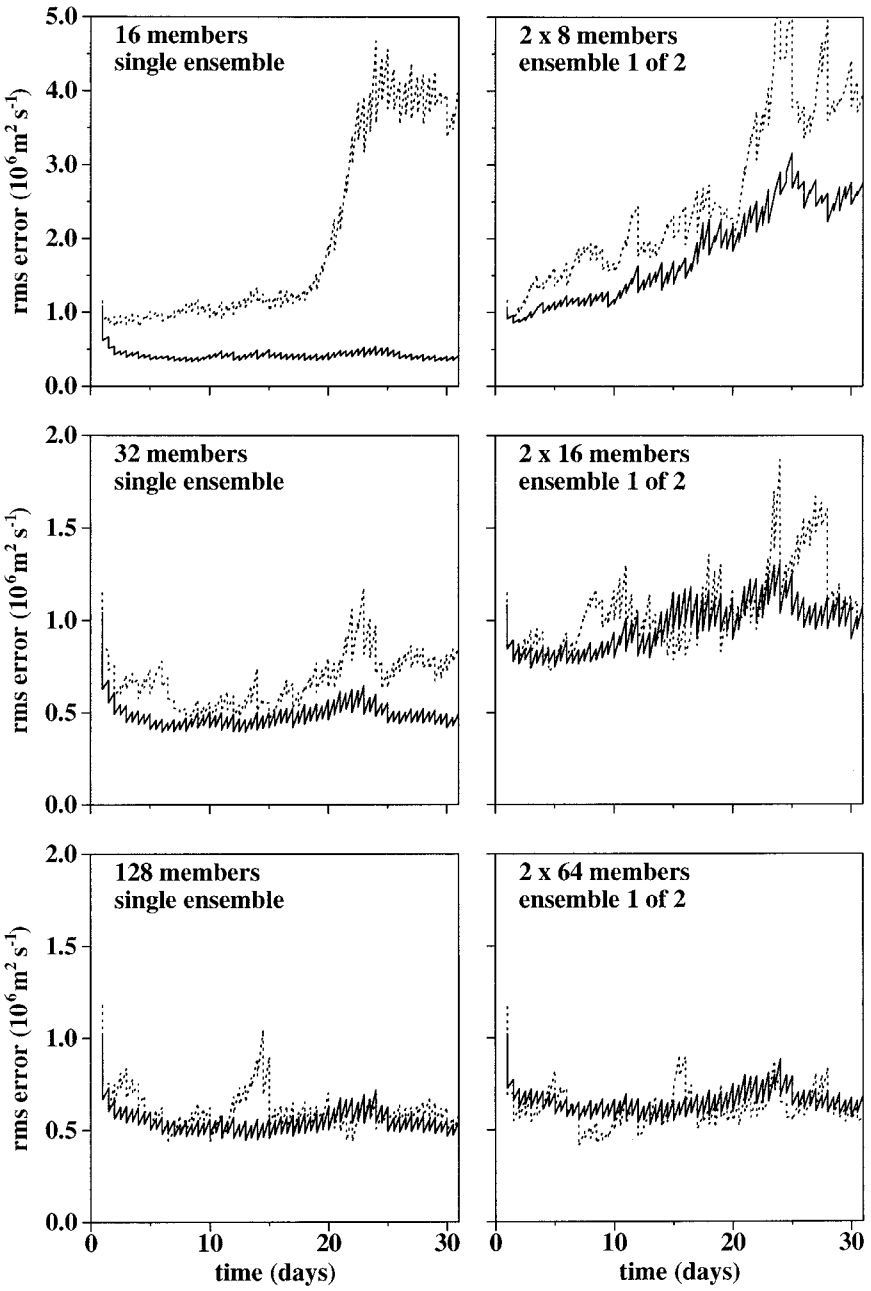

FIG. 3. Forecast and analysis error every 12 h during 30-day assimilation cycles. The solid line is for the rms spread in the ensemble. The dashed line gives the rms error of the ensemble mean. The left-hand panels are for the setup that uses a single ensemble. The right-hand panels are for the first ensemble of a pair configured as in Fig. 2. The upper, middle, and lower panels are for configurations having a total of 16, 32, and 128 members, respectively. In all cases $r_{max} = 20°$. Note the difference in scale for the rms error between the upper panels and the other panels.

responding results calculated from the second ensemble of two of the pairs will be exhibited in Fig. 4 to permit an impression of the degree of similarity between the two ensembles.)

In the case of the pair of ensemble Kalman filters, one could also combine the two ensembles together and then calculate the ensemble mean error and the ensemble spread from this larger ensemble. The effect of this, when averaged over the 30-day assimilation period, was found to be a *reduction* in the ensemble mean error (by approximately 25% for the pair with $N = 8$, 17% for the pair with $N = 16$, and 5% for the pair with $N = 64$) and an *increase* in the ensemble spread by similar percentages in the three cases. It is important to note

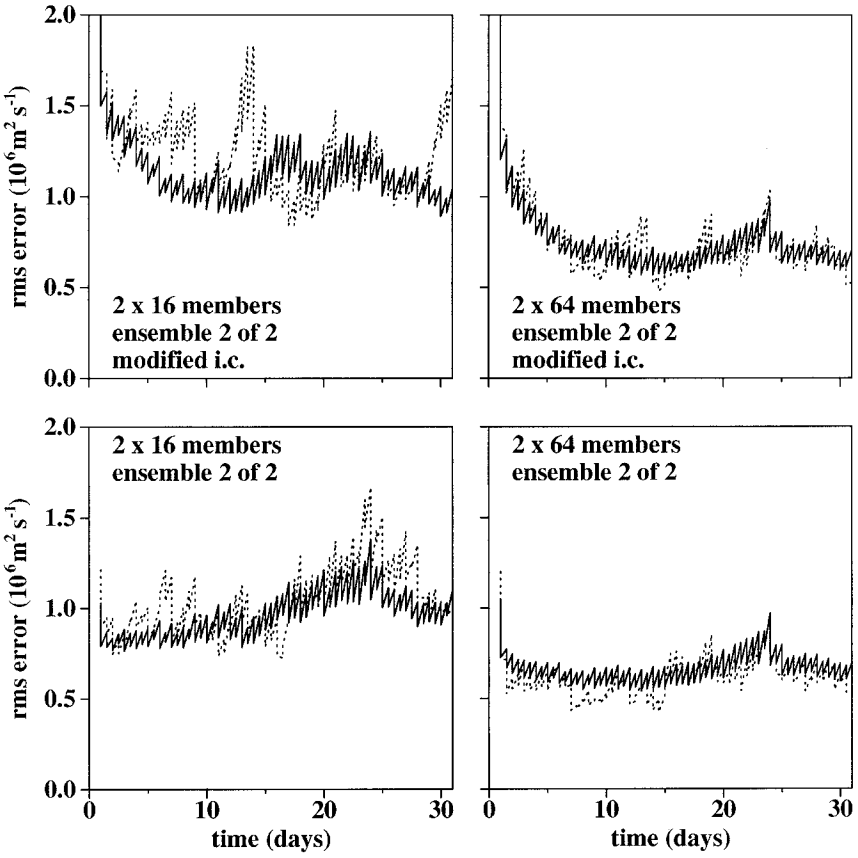

FIG. 4. Forecast and analysis error every 12 h during 30-day assimilation cycles using the setup with a pair of ensemble Kalman filters. The upper panels are with initial conditions modified as described in the text, while the lower panels are with the standard initial conditions. All results are for the second ensemble of the pair. The left-hand panels are for $N = 16$, while the right-hand panels are for $N = 64$. The plotting convention and value of r_{max} are as in Fig. 3.

that such a combined ensemble does not retain its representativeness properties. We note that, due to the sampling error associated with (6), caution should be exercised in judging the relative merits of the ensemble mean in the right- versus left-hand panels. To obtain stronger conclusions, we might have used multiple realizations of (6), as in Houtekamer and Derome (1995).

Looking first at the upper left-hand panel of Fig. 3, it can be seen that while the spread between the ensemble members decreases initially and then remains at a fairly low level, this in no way reflects the error in the ensemble mean. The latter increases, slowly at first, and then dramatically. By way of contrast, the upper right-hand panel shows the performance when a pair of ensemble Kalman filters, configured as in Fig. 2 and with the same total number of ensemble members (i.e., 16), is used. While the error in the ensemble mean initially grows more quickly than in the left-hand panel, the spread in the ensemble now grows as well and reflects, albeit by underestimation, the error in the ensemble mean.

The very substantial overall decrease in error that occurs with both a single or a pair of ensembles upon doubling the ensemble sizes can be seen by examining the two middle panels. (Note the difference in scale for the rms error between the upper and middle panels.) With regard to the representativeness of the ensemble in this case, while the ensemble spread consistently underestimates the true error in the left-hand panel, the ensemble spread much more nearly represents the mean error when a pair of filters is used (right-hand panel). An exception to the general decrease in error from the upper to the middle panels is exhibited in the case of a single ensemble by the ensemble spread, which can be seen to increase slightly in conjunction with the doubling in the ensemble size. This undesirable behavior is consistent with the discussion in section 2e about underestimation of the spread in the analyses in the case of a single ensemble. It is clear from that discussion that increasing the ensemble size alleviates the underestimation of the ensemble spread.

The effect of further quadrupling the number of ensemble members is shown in the bottom two panels. In the case of a single ensemble Kalman filter (left-hand

panel), this results in an approximately 20% decrease in the ensemble mean error in conjunction with a further small increase in the ensemble spread. The result is a fairly good agreement between the two error estimates, with only a lingering tendency for the ensemble spread to underestimate the error in the ensemble mean. In contrast, when a pair of filters is used (right-hand panel), both error estimates decrease substantially and there is no evidence that the ensemble spread underestimates the ensemble mean error.

In summary, Fig. 3 demonstrates how the configuration of Fig. 2 permits ensemble representativeness to be maintained with much smaller ensembles than those required when using a single ensemble. Therefore, we adopt the configuration of Fig. 2 for all further experiments.

To examine the sensitivity of our results to the initial conditions, two of the cycles of Fig. 3 were repeated with substantially modified ensembles of initial first-guess fields. To generate these modified ensembles, we quadrupled all error covariances on the right-hand side of (4), which results in a doubling of the initial rms error, and used different realizations of the random fields in (7) and (8) than had been used before. The results for the two cycles are presented in the upper panels of Fig. 4 and can be compared to the two standard cycles in the lower panels. (Note that we present results for the *second* ensemble of each pair. Comparing each of the lower panels with the appropriate panel of Fig. 3 allows an impression of the similarities and differences between the two ensembles of a pair.)

An examination of Fig. 4 shows that both for the configurations with $N = 16$ and $N = 64$, the rms error of the cycle with modified initial background fields drops to the error level of the corresponding standard cycle within about 10 days. However, the effect of the modified initial conditions on the behavior in the latter part of the 30-day period is rather different: the two ensembles with $N = 16$ (left-hand panels) exhibit rather different behavior right up until day 31, while the two ensembles with $N = 64$ (right-hand panels) behave similarly for the last 14 or so days. This difference in behavior was confirmed by extending the integrations for a further 30-day period (not shown). These results indicate that ensemble size plays an important role in determining the degree of asymptotic stability of the ensemble Kalman filter.

We wished to examine the effect of r_{max}, the cutoff radius used for data selection, on filter performance. To do this, a series of 30-day data assimilation cycles was performed for a range of values of r_{max} for several different values of N. The results are summarized in Fig. 5 in terms of the rms spread in the first ensemble of analyses of each pair at the end of each assimilation cycle, that is, at 0000 UTC day 31. (The corresponding results for the second ensemble of each pair are very similar and are not shown.)

First of all, Fig. 5 confirms the benefits of larger

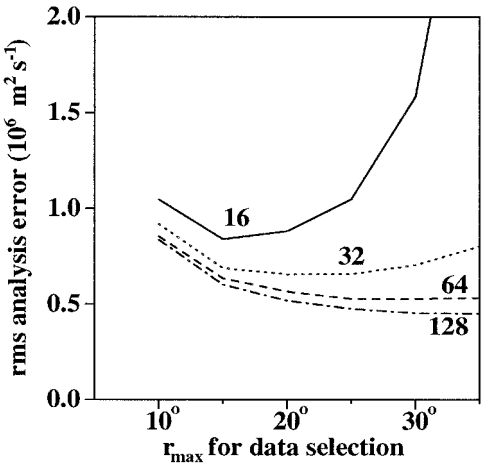

FIG. 5. Performance of the ensemble Kalman filters evaluated at 0000 UTC day 31 as a function of the cutoff radius r_{max} for different values of N.

ensemble sizes, with the greatest impact of doubling the ensemble size occurring for N small. In addition, the figure indicates that (i) for each value of N there is an optimal value of r_{max} and (ii) this value increases as N increases. To investigate this behavior, we examine global correlations calculated from the ensembles of background fields valid at 0000 UTC day 31.

To illustrate, we show in Fig. 6a the 50-kPa global correlation field with respect to a point off the west coast of North America. These correlations were calculated from the first ensemble of the pair with $N = 32$ and $r_{max} = 20°$. It can be seen that the correlation maximum tends to be pear shaped with a pronounced extension toward the northwest. Large negative values are evident to the east-northeast and southwest. Note also the many centers with correlations exceeding ± 0.25 scattered around the globe, and as far away as the eastern Mediterranean and the coast of Antarctica. An indication of the accuracy of these correlation features can be obtained by comparing this field with the corresponding field calculated from the second ensemble of the pair, shown in Fig. 6b. It can be seen that whereas the features in the eastern part of the North Pacific are confirmed by the second ensemble, this is not the case for the correlations at larger distances.

The corresponding global correlation fields, calculated from the pair of ensembles with $N = 128$ and $r_{max} = 35°$, are shown in Figs. 6c and 6d. These two fields confirm many of the main features noted in Figs. 6a and 6b. In addition, it can be seen that there is now much better agreement between the two correlation fields of the pair and a marked reduction in the correlations at large distances.

Comparisons such as these permit the behavior noted in Fig. 5 to be explained in terms of (12). If N is small, the accuracy with which covariance fields can be computed from the ensemble is relatively poor. As N in-

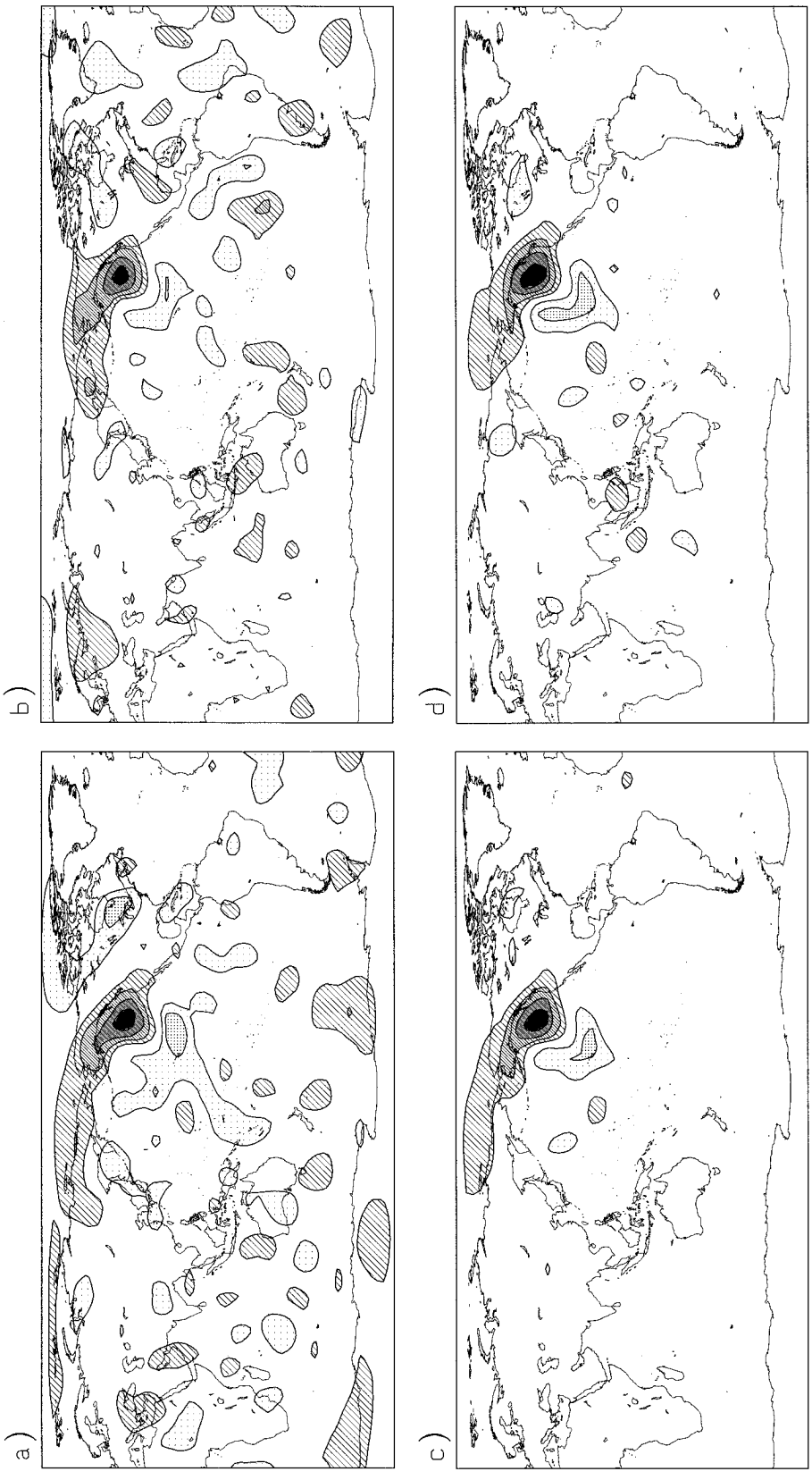

FIG. 6. Global correlation fields with respect to a point off the west coast of North America at 0000 UTC day 31. Panels (a) and (b) show the 50-kPa correlation fields for ensembles 1 and 2 when $N = 32$ and $r_{max} = 20°$. Panels (c) and (d) show the corresponding fields when $N = 128$ and $r_{max} = 35°$. The contours are at ± 0.25, ± 0.50, ± 0.75, and ± 0.90. Shading with dotted (hatched) patterns indicates negative (positive) correlations. The intensity of the shading increases as the magnitude of the correlation increases. Regions with correlations between -0.25 and $+0.25$ are not shaded.

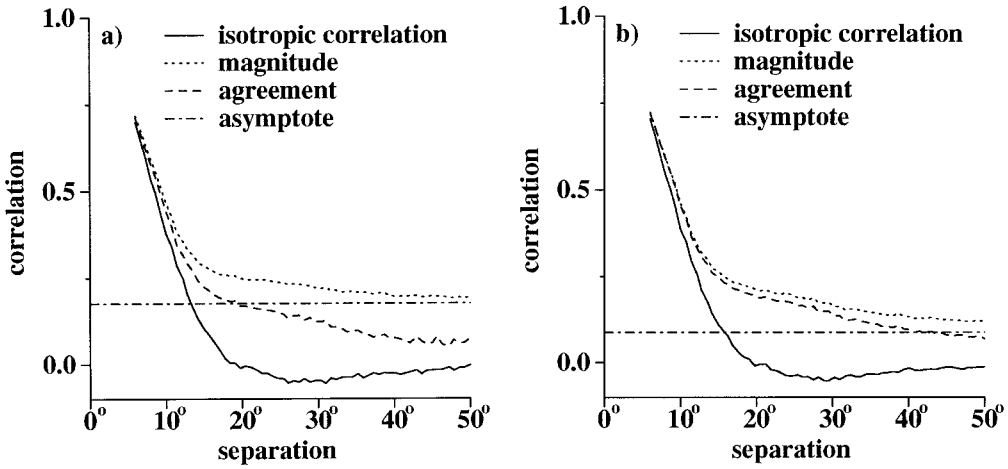

FIG. 7. The solid curve in each panel shows the mean isotropic correlation computed over the entire sphere for the ensembles of first-guess fields valid at 0000 UTC day 31. The dotted curve is the square root of the corresponding mean squared correlations. The dashed–dotted line indicates the noise level for the zero correlation estimates, i.e., $1/N^{0.5}$. The dashed curve shows the agreement between the two ensembles of the pair, i.e., the square root of the mean product of corresponding correlations computed from each ensemble. In (a) $N = 32$ and $r_{max} = 20°$. In (b) $N = 128$ and $r_{max} = 20°$.

creases, not only does accuracy improve in general, but the distance over which it becomes possible to accurately compute covariances also increases. Now it is advantageous to utilize all the data for which accurate forecast-error covariances can be computed but detrimental to use data for which forecast-error covariances are inaccurate. Since the distance between two points gives a rough measure of the accuracy with which the covariance between them can be computed, there is a value of r_{max} that is approximately optimal for a given value of N.

To quantify the above observations somewhat more, we have taken a very large number of randomly chosen pairs of points, (l, k), on the sphere and, using both ensembles of a pair together, computed correlations, $\rho_{l,k}$. Sorting these by great-circle distance and averaging these in distance bins yields mean isotropic correlations, $\overline{\rho(r)}$. The solid curve in Fig. 7a shows how these mean correlations fall off with distance for the case $N = 32$ and $r_{max} = 20°$. It can be seen that the mean correlations become zero, or even slightly negative, beyond a distance of about 20°. This suggests that traditional statistical interpolation schemes, applied here, could not make any use of observations beyond this distance. However, it may be that for certain points significant, possibly nonisotropic, correlations extend to much farther distances. The pair of ensemble Kalman filters could then exploit observations at distances beyond 20°. To quantify this, we took the same pairs of points as above but we now squared each of the correlations to obtain $\rho^2_{l,k}$. Sorting these squared correlations by great-circle distance and averaging in distance bins as before, we then took the square root for each bin to obtain $\sqrt{\overline{\rho^2(r)}}$. This is plotted as the dotted line. We see that values exceeding 0.2 are found even out to distances of 50°. Unfortunately, as we shall now argue, these correlations are not very significant in the present case.

Setting $N = 32$ and $\rho = 0$ in (12), we obtain an rms error for the estimated correlation of $1/\sqrt{N} \approx 0.18$, as indicated by the dashed–dotted line. The dotted curve asymptotes to this noise level rather than to zero. Finally, the dashed line in Fig. 7a shows the agreement between the two ensembles of the pair. It was obtained by calculating the product of corresponding correlations computed from each ensemble, computing mean values for each distance bin, and then taking the square root. Now at large distances, we expect the two estimates to be independent values near zero; therefore, the agreement should asymptote to zero at large distance, as is observed. Where the agreement between the two sets becomes smaller than the noise level, we do not expect our correlations to be of much use for the purposes of data assimilation. In this way we would roughly estimate an optimal value of $r_{max} = 20°$ for $N = 32$ from Fig. 7a. In Fig. 7b we show the corresponding curves for the case with $N = 128$ and $r_{max} = 20°$. Due to the fourfold increase in the size of the ensembles, the noise level is reduced by a factor of 2 and the estimates of the correlations are now relatively significant for distances of up to $r_{max} = 40°$. These estimates are in agreement with the results shown in Fig. 5. As the ensemble size becomes bigger, the ensemble Kalman filters can benefit from a larger cutoff radius r_{max}.

An overall impression of the nonhomogeneous nature of the statistics produced by the ensemble Kalman filters can be obtained from Fig. 8. This figure shows the global rms forecast-error field from the first ensemble of the pair with $N = 128$ and $r_{max} = 35°$. (Corresponding values from the second ensemble of the pair were generally within 10% of these values and are not shown.) It can

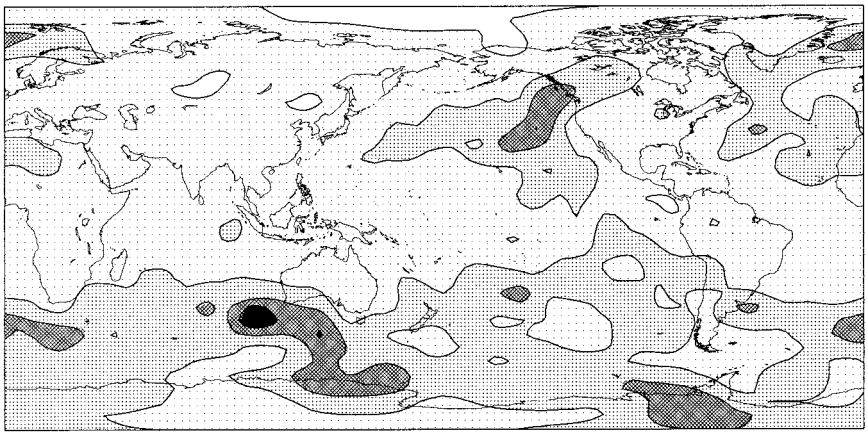

FIG. 8. The 50-kPa rms forecast-error field (10^6 m^2 s^{-1}) at 0000 UTC day 31 computed from one ensemble of the pair with $N = 128$, $r_{max} = 35°$. The contours are at 0.3, 0.6, 0.9, and 1.2. The intensity of the shading increases as the magnitude of the error increases.

be seen that the forecast-error magnitude exhibits considerable inhomogeneity. This inhomogeneity is consistent with the observation network of Fig. 1, with relatively large errors occurring over most of the Southern Hemisphere oceans and over the North Pacific and Atlantic Oceans. However, it is clear that the observation network is not the only determining factor and that the dynamics is playing an important role as well. For example, the maxima in the Northern Hemisphere occur in the *eastern* North Pacific and Atlantic Oceans, evidence of advection by the prevailing westerlies [as in Fig. 5 of Daley (1992b), which shows the error due to a data void in the case of a constant advecting velocity].

To convey an impression of the nonhomogeneous nature of the horizontal correlations, we present in Fig. 9 two further examples of 50-kPa global correlations computed from ensemble 1 of the pair with $N = 128$ and $r_{max} = 35°$. The correlations in Fig. 9a are computed with respect to a point off the east coast of North America and those in Fig. 9b are with respect to a point in the vicinity of Lake Superior. Both points are at the same latitude as the base point for Fig. 6. The correlations in Fig. 9a exhibit a highly anisotropic structure, like those in Fig. 6, while the correlation structure in Fig. 9b is much more isotropic. It seems unlikely that a simple correlation model based on a small number of free parameters could faithfully represent these highly anisotropic correlation structures.

The vertical structures of the correlations computed from the ensembles are also of interest. The synoptic situation, and vertical structure, at the three points whose horizontal correlations we have examined differ at 0000 UTC on day 31. The two points off the coast of North America are located in baroclinic zones in regions where the profiles slope westward with height, while the structure near Lake Superior is much more vertical. Figure 10 shows the correlations at each of the three points in a vertical plane oriented in the zonal direction. In each case the correlations are computed with respect to the central grid point, for which the correlation is therefore 100%. Indeed it can be seen from the figure that the correlations with respect to the points off the east and west coasts of North America (Fig. 10a and 10c, respectively) exhibit a clear westward tilt with height, while the correlation with respect to the point near Lake Superior (Fig. 10b) is almost vertical.

4. Summary and concluding discussion

One of the most serious approximations in current NWP practice is that the statistics used for assimilating data are largely homogeneous, isotropic, and independent of the flow. This shortcoming, essentially what distinguishes 3D from 4D data assimilation, has led to an intense effort to develop 4D data assimilation techniques over the past decade. Attention has focused on two techniques: the Kalman filter (e.g., Cohn and Parrish 1991) and the 4D variational algorithm (e.g., Courtier et al. 1994). Due largely to the enormous computational burden associated with the Kalman filter, the cheaper 4D variational approach has benefited from more focused development and now seems closer to practical reality.

Recently, Evensen (1994) has suggested that using the statistics provided by an ensemble of perturbed short-range forecasts may lead to an alternative 4D data assimilation system. The cycling of covariance information from one assimilation cycle to the next, a big advantage of a Kalman filter over the 4D variational approach, is performed using an ensemble of, say, 6-h integrations with the nonlinear forecast model. The purpose of the present study has been to examine this technique, termed an ensemble Kalman filter, in an idealized environment.

Using a three-level, quasigeostrophic, T21 model and

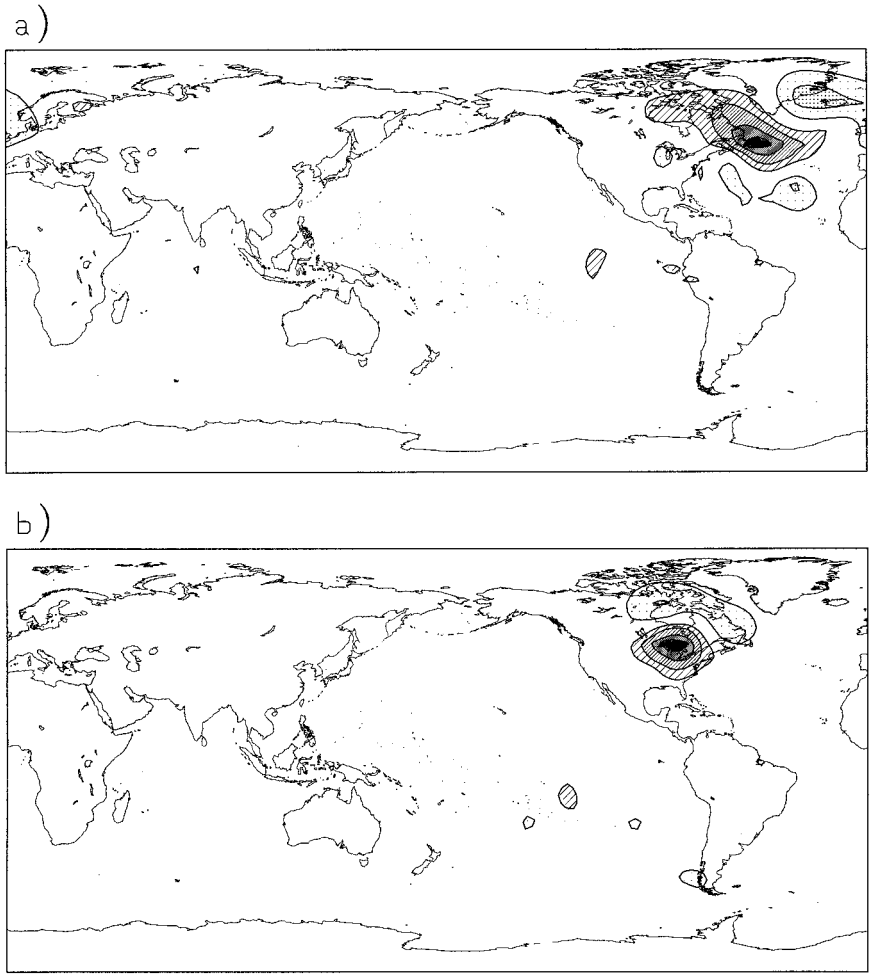

FIG. 9. As in Fig. 6c except (a) shows the global correlation field for ensemble 1 with respect to a point off the east coast of North America. (b) As in (a) except with respect to a point in the vicinity of Lake Superior.

simulated observations, experiments have been performed based on a perfect-model assumption. It was found that the naive approach of using the forecast-error covariances computed from an ensemble of short-range forecasts to calculate weights for the assimilation of data using this same ensemble as background fields, gave rise to an inbreeding problem. In order to avoid this problem and maintain a representative spread between the ensemble members, we have employed a technique that uses a pair of ensemble Kalman filters, configured as shown in Fig. 2. Even with this technique, our smallest ensembles exhibit some underestimation of the error (Fig. 3). To address this further, we might use three (or more) coupled ensembles. However, for a fixed total number of ensemble members, the resulting smaller ensembles would likely incur larger ensemble mean errors. Alternatively, for each ensemble member, one might use a gain matrix computed using all of the other ensemble members. This would mean accepting the cost of computing many gain matrices.

In a series of 30-day data assimilation experiments with ensembles of different sizes, it was found that, the rms analysis error decreased as the size of the ensembles increased, as expected, and ensembles having of the order of 100 members are sufficient to accurately describe local anisotropic, baroclinic correlation structures. The estimation of small correlations, associated with remote observations, is much more difficult and may require very large ensembles, as expected from (12), a standard result from statistical theory.

To deal with these small correlations at large distances, a cutoff radius beyond which observations were not used was implemented. It was found that the optimal value of the cutoff radius increased as the number of available ensemble members increased. Arguing heuristically, the ensemble Kalman filter is computationally more efficient than the usual Kalman filter at the expense of not precisely estimating correlations and, from (12), especially the small correlations at large distances. Therefore, as the number of ensemble members increas-

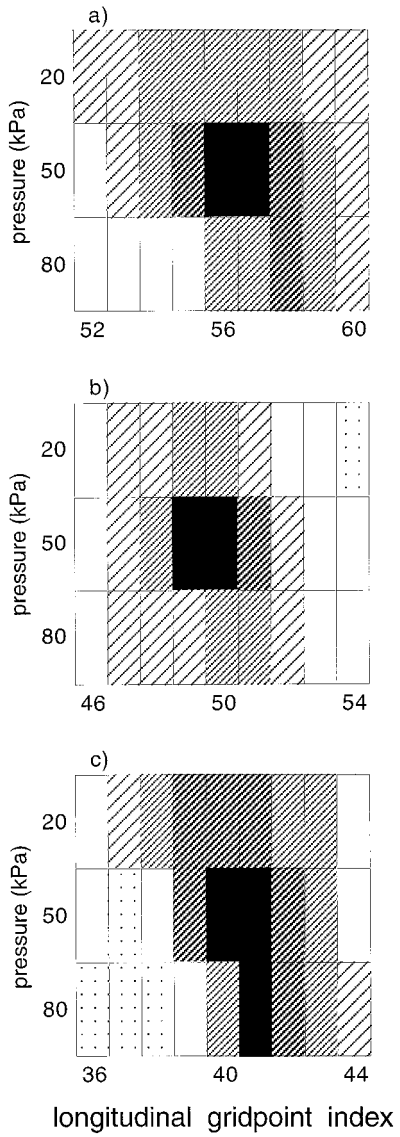

FIG. 10. The vertical correlations with respect to three points located at 50 kPa on the same latitude circle. In all cases, the vertical plane is oriented in the zonal direction and the correlations are computed from the first ensemble of the pair with $N = 128$ and $r_{max} = 35°$. In each case, the longitudinal gridpoint index is indicated along the abscissa. The contour and shading convention is the same as in Figs. 6 and 9. The central points for the three panels are located (a) off the east coast, (b) near Lake Superior, and (c) off the west coast.

es, correlations at larger and larger distances can be accurately estimated and so it becomes advantageous to increase the cutoff radius. Thus, given a certain ensemble size, an appropriate cutoff radius can be specified. One should keep in mind that as observations become very remote from the point being analyzed, their potential positive impact can be expected to be rather small.

A complementary argument in favor of a cutoff radius follows from the fact that, given an N-member ensemble, one can deduce a nonzero forecast error in only $N - 1$ directions. It follows that the observations can only produce a correction in these $N - 1$ directions. If the number of observations is much larger than $N - 1$, this results in a big loss of information; the background can only be corrected in a limited number of directions and consequently the analyses can be expected to diverge from the real state (filter divergence). This rank problem is greatly reduced if a cutoff radius is used to arrive at a large number of small problems, each of which individually has no rank problem.

As is usually done in the 3D and 4D variational approaches, forward interpolation operators from the model variable to the observations were employed in this study to enable the ensemble Kalman filter to make use of nonconventional observations. This implies that some of the work currently being done toward the assimilation of nonconventional data using variational analysis methods could be used for an ensemble Kalman filter. It should be noted that the ensemble Kalman filter does not require the development of the tangent linear model or its adjoint, since it uses the complete nonlinear model to transport the covariances.

The perfect-model assumption must be dropped if this technique is to be successful in an operational setting. We note that the ensemble Kalman filter can easily account for the portion of the model error of known origin, as shown in Houtekamer et al. (1996a) and Houtekamer and Lefaivre (1997) where different versions of the forecast model (for instance, using different convection schemes) were used for different members of the ensemble. It is likely that the remaining portion of the model error will have to be parameterized. The values of the parameters might be estimated from the actual innovations (i.e., differences between the observations and the first-guess values) using the maximum-likelihood method (Dee 1995; Maybeck 1982) or some related adaptive technique (e.g., Blanchet et al. 1997). We intend to examine this problem in a subsequent study. Until model error is properly accounted for, the usefulness of the ensemble Kalman filter for atmospheric data assimilation cannot really be evaluated. One might also want to estimate the error statistics of the observations from the innovations. However, it may be difficult to estimate the covariances of model and observational error simultaneously (e.g., Daley 1992a; Maybeck 1982).

Several operational centers are already doing medium-range ensemble forecasting. Implementation of an ensemble Kalman filter would, of course, provide an ensemble of initial conditions for these forecasts. With regard to the computational feasibility of the ensemble Kalman filter, we note that, for example, nine 10-day ensemble forecasts are being run with a T63 model every day (Houtekamer and Lefaivre 1997) at the Canadian Meteorological Centre (CMC). This corresponds to 90 days of integration. For an equivalent cost, we could run 90 independent analysis cycles. Thus if, as indicated by a naive examination of (12), the accuracy

with which correlations can be computed depends only on the ensemble size, then an ensemble Kalman filter is likely already feasible on the current CMC computers. As pointed out by Evensen (1994), the ensemble Kalman filter approach is embarrassingly parallel.

Our current methodology would become prohibitively expensive if the data were very dense with many analysis points having hundreds of observations lying within the specified cutoff radius. To deal with this, it may be necessary to resort to traditional approaches like superobservation formation (e.g., Lorenc 1981) or a more restrictive data-selection algorithm, although the latter can be an important source of noise and imbalance (da Silva et al. 1995). Alternatively, one could solve each local analysis problem using a variational approach. This would involve solving a number of variational problems for each grid point or volume (as many as there are ensemble members). It has been suggested that it may be possible to improve the preconditioning as we go to successive ensemble members (P. Courtier 1996, personal communication).

The ensemble Kalman filter approach could also be used for data assimilation at high resolution with a regional or mesoscale model. For observations taken at high density, the proper specification of observation-error statistics will likely continue to be a difficult problem. In the case of high-density independent observations and as atmospheric behavior becomes high dimensional (i.e., increasingly unconstrained by simple balance conditions and exhibiting more complicated structure in phase space), larger ensembles *may* be required in order to produce proper high-dimensional forecast-error covariances.

Acknowledgments. This project benefited from the experience gained during the course of a preliminary study that the first author carried out with Luc Fillion using a hemispheric barotropic model. We would also like to thank Cécilien Charette for helpful discussions relating to the programming of the analysis algorithm and Monique Tanguay for her thoughtful review of the manuscript. Further clarifications to the paper resulted from the reviews by Geir Evensen and two anonymous reviewers.

APPENDIX

Generation of 3D Fields Having a Prescribed Covariance Structure

We give an approximate method for generating realizations of random fields having, on average, a prescribed covariance structure and zero mean. The approximation consists of prescribing the correlation structure only if the horizontal correlation exceeds a certain minimum value, $\rho_{\min}$ (usually taken to be 0.1, to limit the cost of the algorithm). Since Houtekamer et al. (1996a,b) have discussed the generation of 2D fields using the same principles (Epstein 1969; Kendall and Stuart 1979), only a brief description is given here.

First we prescribe a separable 3D covariance structure by specifying (i) a horizontal correlation function, and (ii) a 3×3 vertical covariance matrix. We now generate three independent 2D random fields, each of which will be projected onto an eigenvector of the vertical covariance matrix. Each of these fields is defined on the 64×32 Gaussian grid and is generated as follows.

The points of the horizontal grid are assigned numbers $1, 2, 3, \ldots, N_G$. Since we have no prior information for the first grid point, we assign it a random number drawn from a standard normal distribution. For subsequent grid points i, we do the following in the order $i = 2, \ldots, N_G$.

1) Determine the previously processed points at which, according to the specified horizontal correlation function, the correlation with point i exceeds $\rho_{\min}$.
2) Using Eqs. (27.17) and (27.61) of Kendall and Stuart (1979), determine the mean and standard deviation of the conditional distribution at point i given the values at the points identified in step 1.
3) Draw a random number from a normal distribution having this mean and standard deviation. This number is the value at point i which we seek.

Projecting these three random fields onto the eigenvectors of the vertical covariance matrix yields a single 3D field. Repeating the above procedure using different random numbers yields an ensemble of 3D fields, which has the desired statistical properties.

REFERENCES

Barkmeijer, J., P. L. Houtekamer, and X. Wang, 1993: Validation of a skill prediction method. *Tellus,* **45A,** 424–434.

Bartello, P., and H. L. Mitchell, 1992: A continuous three-dimensional model of short-range forecast error covariances. *Tellus,* **44A,** 217–235.

Blanchet, I., C. Frankignoul, and M. A. Cane, 1997: A comparison of adaptive Kalman filters for a tropical Pacific Ocean model. *Mon. Wea. Rev.,* **125,** 40–58.

Cohn, S. E., and D. P. Dee, 1988: Observability of discretized partial differential equations. *SIAM J. Num. Anal.,* **25,** 586–617.

——, and D. F. Parrish, 1991: The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.,* **119,** 1757–1785.

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

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

——, 1992a: The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.,* **120,** 178–196.

——, 1992b: Forecast-error statistics for homogeneous and inhomogeneous observation networks. *Mon. Wea. Rev.,* **120,** 627–643.

da Silva, A., J. Pfaendtner, J. Guo, M. Sienkiewicz, and S. Cohn, 1995: Assessing the effects of data selection with DAO's physical-space statistical analysis system. *Proc. Second Int. Symp. on Assimilation of Observations in Meteorology and Oceanography,* Tokyo, Japan, World Meteorological Organization, WMO Tech. Doc. 651, 273–278.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.,* **123,** 1128–1145.

Desroziers, G., and J.-P. Lafore, 1993: A coordinate transformation for objective frontal analysis. *Mon. Wea. Rev.,* **121,** 1531–1553.

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus,* **21,** 739–759.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

——, and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.,* **124,** 85–96.

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields* (English translation). Israel Program for Scientific Translations, Gidromet, 242 pp.

Ghil, M., and R. Todling, 1996: Tracking atmospheric instabilities with the Kalman filter. Part II: Two-layer results. *Mon. Wea. Rev.,* **124,** 2340–2352.

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

——, and ——, 1990: Reply. *Mon. Wea. Rev.,* **118,** 1929.

Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123,** 2181–2196.

——, and L. Lefaivre, 1997: Using ensemble forecasts for model validation. *Mon. Wea. Rev.,* **125,** 2416–2426.

——, ——, J. Derome, H. Ritchie, and H. L. Mitchell, 1996a: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124,** 1225–1242.

——, ——, and ——, 1996b: The RPN ensemble prediction system. *Proc. of the ECMWF Seminar on Predictability,* Vol. 2, 121–146. [Available from ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom.]

Jørgensen, A. M. K., 1987: Numerical analysis of the atmospheric state for a Danish fine-mesh numerical weather prediction model. Ph.D. thesis, University of Copenhagen, 157 pp. [Available from Der Naturvidenskabelige Fakultet, University of Copenhagen, Øster Voldgade 3, 1350 Copenhagen K, Denmark.]

Kalnay, E., M. Kanamitsu, and W. E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bull. Amer. Meteor. Soc.,* **71,** 1410–1428.

Kendall, M., and A. Stuart, 1979: *Inference and Relationship.* Vol. 2, *The Advanced Theory of Statistics,* 4th ed. MacMillan, 749 pp.

Lin, H., and J. Derome, 1996: Changes in predictability associated with the PNA pattern. *Tellus,* **48A,** 553–571.

Lönnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus,* **38A,** 137–161.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.,* **109,** 701–721.

Marshall, J., and F. Molteni, 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.,* **50,** 1792–1818.

Maybeck, P. S., 1982: *Stochastic Models, Estimation and Control.* Vol. 2. Academic Press, 423 pp.

Ménard, R., 1994: Kalman filtering of Burgers' equation and its application to atmospheric data assimilation. Ph.D. thesis, McGill University, 211 pp. [Available from Dept. of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke Street W., Montreal, PQ H3A 2K6, Canada.]

Mitchell, H. L., C. Charette, C. Chouinard, and B. Brasnett, 1990: Revised interpolation statistics for the Canadian data assimilation procedure: Their derivation and application. *Mon. Wea. Rev.,* **118,** 1591–1614.

——, ——, S. J. Lambert, J. Hallé, and C. Chouinard, 1993: The Canadian global data assimilation system: Description and evaluation. *Mon. Wea. Rev.,* **121,** 1467–1492.

Molteni, F., and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Polavarapu, S. M., 1995: Divergent wind analyses in the oceanic boundary layer. *Tellus,* **47A,** 221–239.

Rabier, F., and T. McNally, 1993: Evolution of forecast error covariance matrix. ECMWF Tech. Memo. 195, 36 pp. [Available from ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom]

Stuart, A., and J. K. Ord, 1987: *Distribution Theory.* Vol. 1, *Kendall's Advanced Theory of Statistics,* 5th ed., Charles Griffin, 604 pp.