Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter

THOMAS M. HAMILL AND JEFFREY S. WHITAKER

NOAA–CIRES Climate Diagnostics Center, Boulder, Colorado

CHRIS SNYDER

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 18 January 2001, in final form 1 May 2001)

ABSTRACT

The usefulness of a distance-dependent reduction of background error covariance estimates in an ensemble Kalman filter is demonstrated. Covariances are reduced by performing an elementwise multiplication of the background error covariance matrix with a correlation function with local support. This reduces noisiness and results in an improved background error covariance estimate, which generates a reduced-error ensemble of model initial conditions.

The benefits of applying the correlation function can be understood in part from examining the characteristics of simple 2×2 covariance matrices generated from random sample vectors with known variances and covariance. These show that noisiness in covariance estimates tends to overwhelm the signal when the ensemble size is small and/or the true covariance between the sample elements is small. Since the true covariance of forecast errors is generally related to the distance between grid points, covariance estimates generally have a higher ratio of noise to signal with increasing distance between grid points. This property is also demonstrated using a two-layer hemispheric primitive equation model and comparing covariance estimates generated by small and large ensembles. Covariances from the large ensemble are assumed to be accurate and are used a reference for measuring errors from covariances estimated from a small ensemble.

The benefits of including distance-dependent reduction of covariance estimates are demonstrated with an ensemble Kalman filter data assimilation scheme. The optimal correlation length scale of the filter function depends on ensemble size; larger correlation lengths are preferable for larger ensembles.

The effects of inflating background error covariance estimates are examined as a way of stabilizing the filter. It was found that more inflation was necessary for smaller ensembles than for larger ensembles.

1. Introduction

Many groups are experimenting with data assimilation schemes for complex numerical weather and oceanographic prediction models where background forecast error covariances are estimated using an ensemble (e.g., Evensen 1994; Evensen and van Leeuwen 1996; Houtekamer and Mitchell 1998, 2001; Burgers et al. 1998; Mitchell and Houtekamer 2000; Lermusiaux and Robinson 1999; van Leeuwen 1999; Anderson and Anderson 1999; Hamill and Snyder 2000; Heemink et al. 2001; Hansen and Smith 2000; Keppenne 2000; Anderson 2001). Much of this experimentation is based on approach is known as the ensemble Kalman filter (EnKF). The EnKF consists of a set (or ensemble) of parallel short-term forecasts and data assimilation cycles. Statistics derived from the short-term forecasts are used to estimate background error covariances during the subsequent data assimilation step.

Though the ensemble-based data assimilation approaches such as the EnKF have not been demonstrated operationally yet, their potential for improving numerical weather prediction has gained them some attention. One potentially large benefit is that by construction, the ensemble forecast and data assimilation steps are unified, so a consistent, reliable, reduced-error ensemble of initial conditions may be available for generating ensemble forecasts (e.g., Hamill et al. 2000; Hamill and Snyder 2000). This would obviate the need to add structured or unstructured noise to a control forecast to generate initial conditions, as is currently done in Europe and the United States (e.g., Molteni et al. 1996; Toth and Kalnay 1993, 1997). Further, ensemble-based techniques have two potential advantages over the tradi-

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

Corresponding author address: Dr. Thomas M. Hamill, NOAA–CIRES Climate Diagnostics Center, R/CDC 1, 325 Broadway, Boulder, CO 80303-3328.
E-mail: hamill@cdc.noaa.gov

© 2001 American Meteorological Society

tional extended Kalman filter (EKF) (Cohn 1997). Compared with the EKF, the cost of ensemble techniques should be significantly less, since covariances are estimated using a limited-size random sample. In the EKF, error covariances for each model component are propagated using linear tangent and adjoints of the fully nonlinear model, an exorbitant computational expense for a high-dimensional model [however, it may be possible to reduce the computations by computing in a reduced-dimensional subspace; see, e.g., Fisher (1998)]. In terms of accuracy, ensemble filters may be more accurate than the EKF since covariances are estimated by propagating model states with a fully nonlinear model rather than under assumptions of linearity.

Ensemble-based data assimilation approaches offer the potential of providing better initial conditions than may be possible using existing methods such as three-dimensional variational data assimilation (3DVAR; Lorenc 1986; Parrish and Derber 1992) or four-dimensional variational data assimilation (4DVAR; Le Dimet and Talagrand 1986, Rabier et al. 1998). Hamill and Snyder (2000) showed that the a hybrid ensemble Kalman filter provided substantial improvements over 3DVAR in a perfect-model context. Anderson (2001) has done some preliminary experiments that suggest that an ensemble-based approach may also be better than 4DVAR.

The relative cost of an ensemble approach versus the current standard, 4DVAR, is hard to estimate; it may depend on whether the analysis scheme sequentially processes the observations (e.g., Houtekamer and Mitchell 2001) or simultaneously processes them (e.g., Hamill and Snyder 2000). As well, it may depend on the size of the ensemble necessary to provide adequate background error statistics and the complexity of forward operators (which convert model states to observations; these operations must be carried out separately for each ensemble member). Further, the extent to which each can be parallelized may be quite different. Ensemble forecasts are of course easily parallelizable; one member forecast can be farmed out to each central processing unit. The data assimilation may also be somewhat parallelizable, perhaps using algorithms such as that suggested by Houtekamer and Mitchell (2001). As a rough guess, however, one might expect the computational expense of an ensemble assimilation approach to be about the same order of magnitude as 4DVAR.

Despite the appeal of ensemble-based data assimilation approaches, there is much yet to be learned before any will be considered for operational use. One substantial problem is caused by using small-sized ensembles to estimate background error covariances. Houtekamer and Mitchell (1998) noted that the EnKF analysis could be improved by excluding observations greatly distant from the grid point being analyzed. They concluded that this was because background error covariance estimates generated from a small ensemble often produced spuriously large magnitude background error covariance estimates between greatly separated grid points; estimates from a larger ensemble showed that the covariances were actually small. Hence, the analysis was more accurate when the observations were excluded than when they were included and assimilated with degraded background error statistics. Houtekamer and Mitchell (2001) have since experimented with filtering covariance estimates produced by the ensemble using a ''Schur product,'' whereby the ensemble-based covariance estimates are multiplied element by element with a distance-dependent correlation function that varies from 1.0 at the observation location to 0.0 at some prespecified radial distance. They have found that the analysis errors are substantially improved when this ''covariance localization'' is incorporated. Also, the analyses were smoother than when observations beyond a specified distance from the observation were excluded, and properties of the covariance matrices such as their conditioning were improved.

This research continues the exploration into the distance-dependent filtering of covariance estimates generated by a finite ensemble. Our goal is to understand why such filtering may be beneficial, how much improvement may be expected from filtering, and how this may change with the size of the ensemble and the observational data density. To this end, section 2 will discuss some of the basic aspects of data assimilation when covariance matrices are estimated using random samples of the model state. We explain how the noise in the estimates of covariances can affect the accuracy of analyses and illustrate how such errors will change depending on the ensemble size and the true correlation structure. This leads us to an understanding of why the ratio of noise to signal in covariance estimates typically varies with distance between the observation location and the grid point being analyzed. Section 3 provides a brief review of the two-layer primitive equation (PE) model and the data assimilation system used in this experiment, as well as a description of how the covariance localization was implemented. Section 4 provides results from a large set of single-observation experiments with the PE model designed to explore how the quality of covariance estimates from an ensemble changes as a function of distance from the observation. Section 5 describes our tests of the covariance localization embedded in the data assimilation system of the PE model. It also discusses the relative improvement that can be gained by inflating the background error covariance estimates produced by the ensemble. We then examine the effects of covariance localization on the eigenvalue spectrum of the covariance matrices. Section 6 discusses these results and concludes.

2. Simple properties of covariance matrices from random samples

Let us start by trying to understand one of the most basic effects that an error in the specification of covariances in the background error will have on data

assimilation. Namely, we consider how an error in the covariance will affect the analysis at a grid point away from the observation location. We consider the simplest system possible, a two-dimensional model state with a single observation and Gaussian statistics. In this section, capital letters will denote continuous random variables, and lowercase letters the actual values, and we will use the nomenclature of Bayesian statistics. Assume we have a random vector $\mathbf{X}^{\mathrm{T}} = (X_1^{\mathrm{T}}, X_2^{\mathrm{T}})$ representing the unknown true state of the model. We have a sample forecast $\mathbf{x}^{\mathrm{b}} = (x_1^{\mathrm{b}}, x_2^{\mathrm{b}})$, denoting the background, or "first-guess" forecast sample of the true state, with background error covariance matrix $\mathbf{P}^{\mathrm{b}}$ defined by

$$\mathbf{P}^{\mathrm{b}} = \begin{pmatrix} \sigma_1^2 & c_{12} \\ c_{12} & \sigma_2^2 \end{pmatrix}. \tag{1}$$

Thus, in the absence of new observations, we have a prior probability distribution $\pi(\mathbf{X}^{\mathrm{T}}) \sim N(\mathbf{x}^{\mathrm{b}}, \mathbf{P}^{\mathrm{b}})$, where $N(\mathbf{a}, \mathbf{B})$ indicates the distribution is normal with expected value $\mathbf{a}$ and variance/covariance $\mathbf{B}$. Assume a new observation then becomes available. This Y is a scalar random variable denoting the observation, and the actual observation is y, taken at the location of the first component of the state vector. Errors ϵ_0 for the observation are defined by $\epsilon_0 \sim N(0, \sigma_0^2)$.

We seek the posterior probability distribution for the analyzed state conditional on (updated to) the new observation, $\pi(\mathbf{X}^{\mathrm{T}} = \mathbf{x} \,|\, Y = y) = \mathbf{X}^{\mathrm{a}} = (X_1^{\mathrm{a}}, X_2^{\mathrm{a}})$. It can be shown that $\mathbf{X}^{\mathrm{a}} \sim N(\mathbf{x}^{\mathrm{a}}, \mathbf{P}^{\mathrm{a}})$, where $\mathbf{x}^{\mathrm{a}} = (x_1^{\mathrm{a}}, x_2^{\mathrm{a}})$ and

$$\mathbf{P}^{\mathrm{a}} = \begin{pmatrix} \mathrm{Var}(X_1^{\mathrm{a}}) & \mathrm{Cov}(X_1^{\mathrm{a}}, X_2^{\mathrm{a}}) \\ \mathrm{Cov}(X_1^{\mathrm{a}}, X_2^{\mathrm{a}}) & \mathrm{Var}(X_2^{\mathrm{a}}) \end{pmatrix}. \tag{2}$$

Thus, the analyzed values and the expected analysis error variance obtained by updating the background are

$$x_1^{\mathrm{a}} = x_1^{\mathrm{b}} + \frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2}(y - x_1^{\mathrm{b}}),$$

$$\mathrm{Var}(X_1^{\mathrm{a}}) = \sigma_1^2\left(1 - \frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2}\right)$$

$$x_2^{\mathrm{a}} = x_2^{\mathrm{b}} + \frac{c_{12}}{\sigma_0^2 + \sigma_1^2}(y - x_1^{\mathrm{b}}),$$

$$\mathrm{Var}(X_2^{\mathrm{a}}) = \sigma_2^2 - \frac{c_{12}^2}{\sigma_0^2 + \sigma_1^2}$$

$$\mathrm{Cov}(X_1^{\mathrm{a}}, X_2^{\mathrm{a}}) = \frac{c_{12}\sigma_0^2}{\sigma_0^2 + \sigma_1^2} \tag{3}$$

(similar derivations are provided in Daley 1991). Now, suppose we have an inaccurate estimate $\hat{\mathbf{P}}^{\mathrm{b}}$ of the covariance matrix $\mathbf{P}^{\mathrm{b}}$, where variances are correctly specified but the covariance has an error, or "noise" $\epsilon_c \sim N(0, \tau_{12})$:

$$\hat{\mathbf{P}}^{\mathrm{b}} = \begin{pmatrix} \sigma_1^2 & c_{12} + \epsilon_c \\ c_{12} + \epsilon_c & \sigma_2^2 \end{pmatrix}. \tag{4}$$

We seek to understand the effect on the quality of the analysis for x_2^{a}. If the error ϵ_c is uncorrelated with errors in y, x_1^{b}, and x_2^{b}, it can be shown that

$$\begin{aligned} \mathrm{Var}(X_2^{\mathrm{a}}) &= \sigma_2^2 - \frac{c_{12}^2}{\sigma_0^2 + \sigma_1^2} + \frac{\tau_{12}^2}{\sigma_0^2 + \sigma_1^2} \\ &= \sigma_2^2 - \frac{c_{12}^2}{\sigma_0^2 + \sigma_1^2}\left[1 - \left(\frac{\tau_{12}}{c_{12}}\right)^2\right]. \end{aligned} \tag{5}$$

Let us denote τ_{12}/c_{12} the "relative error" in the covariance, a measure of noise relative to signal. Notice that when the relative error is greater than 1.0, the analysis of x_2^{a} is typically *degraded* by assimilating the observation y. Notice also that the amount of improvement or degradation will be proportional to the square of the covariance c_{12}. That is, for a given relative error >1.0, the degradation will be worse for larger covariances.

Given that large relative errors in the magnitude of background error covariances may degrade the analysis, we shift focus to understand what can cause such errors when they are estimated from an ensemble. We examine this question through some simple experiments with 2×2 sample covariance matrices. Again, assume we have an ensemble of vector background values $\mathbf{x}^{\mathrm{b}} = (x_1^{\mathrm{b}}, x_2^{\mathrm{b}})$ sampled from $\mathbf{X}^{\mathrm{T}}$. Here x_1^{b} represents the value at the observation location, and x_2^{b} is the value at some distance from the observation.

It can be shown that given the true covariance matrix $\mathbf{P}^{\mathrm{b}}$ with variances $\sigma_1^2 = 1$, $\sigma_2^2 = \eta^2$, true correlation $\rho = \mathrm{Corr}(X_1^{\mathrm{T}}, X_2^{\mathrm{T}})$ (and hence true covariance $c_{12} = \rho\sqrt{\sigma_1^2}\sqrt{\sigma_2^2} = \rho\eta$), then the variance τ_{12} of the error in the calculation of the covariance from a sample ensemble of $\mathbf{x}^{\mathrm{b}}$ is approximately

$$\mathrm{Var}(\epsilon_c) = \tau_{12} \simeq \frac{1}{n}(1 + \rho^2)\eta^2 \tag{6}$$

for large enough sample sizes n. For brevity, the full derivation is excluded here. The derivation assumes that n random vectors $(X_1^{\mathrm{b}}, X_2^{\mathrm{b}})_1, \ldots, (X_1^{\mathrm{b}}, X_2^{\mathrm{b}})_n$ are sampled from a $N(0, \mathbf{P}^{\mathrm{b}})$ distribution. We are interested in the variance of the estimator $\theta = 1/n\, \Sigma_{i=1}^n\, (X_1^{\mathrm{b}} - \overline{X_1^{\mathrm{b}}})_i\, (X_2^{\mathrm{b}} - \overline{X_2^{\mathrm{b}}})_i$, where the overline indicates an expected value. This would be tedious to calculate, but for large n, $\overline{X_1^{\mathrm{b}}} \simeq 0$ and $\overline{X_2^{\mathrm{b}}} \simeq 0$ so $\theta \simeq 1/n\, \Sigma_{i=1}^n\, (X_1^{\mathrm{b}})_i (X_2^{\mathrm{b}})_i$. The derivation is further simplified by transforming using two independent variables U and $V \sim N(0, 1)$, where $(X_1^{\mathrm{b}})_i = U$ and $(X_2^{\mathrm{b}})_i = [\rho U + (1 - \rho^2)^{1/2} V]\eta$. In this derivation, it is not necessarily appropriate to assume as in (3) that the variances are without error. In fact, if we denote the error in the background state at the observation location as ϵ_1, $\mathrm{Var}(\epsilon_1) \simeq 2/n$.

How do errors change as the true correlation and the ensemble size changes? Figure 1 shows the corresponding relative error of the covariance, τ_{12}/c_{12}. Relative er-

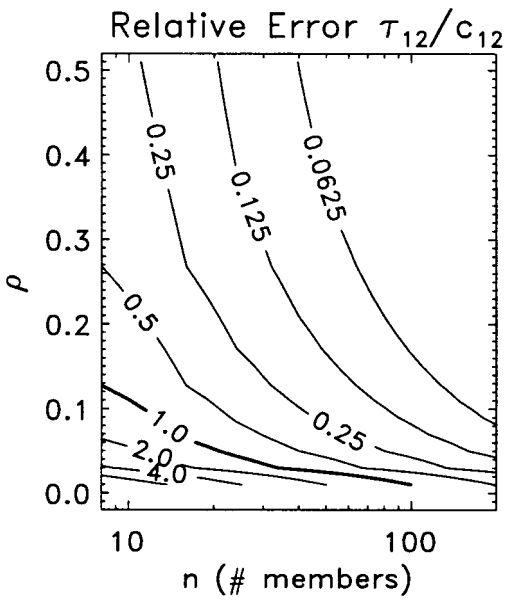

FIG. 1. Relative errors of ensemble of size n and correlation ρ.

ror increases greatly as ρ decreases and as the ensemble size decreases. Since ρ typically decreases with increasing distance from the observation, in a numerical weather prediction model, the noise-to-signal ratio would thus be expected to typically increase with increasing distance from the observation. (This is, on average, the case; sometimes there may be large magnitude true correlations over long distances. Section 4 provides some evidence that this is quite uncommon, however.)

3. Design of the experiment

The experiments conducted here will assume that the forecast model is perfect. A long reference integration of the forecast model provides the true state; the assimilation experiments then use that same model, assimilating imperfect observations generated by adding noise to the true state.

We conducted two general sets of tests, a set of single observation experiments designed to illuminate the characteristics of signal and noise in the ensemble, and a test of analysis accuracy for different observational networks, different sized ensembles, and different filter characteristics. For both, a 90-day set of analyses were computed, updated with new observations every 12 h.

a. Forecast model

Results in the rest of the paper will be based on a dry two-layer PE model. The forecast model was described in Zou et al. (1993). The model state vector consists of vorticity and divergence spectra at two levels as well as Exner function π at the lower surface and at an interface. The model is spectral with a T31 triangular truncation. There is a simple, wavenumber-2 terrain, but there are no land–water interfaces. A fourth-order Runge–Kutta scheme is used for the numerical integration, there is ∇^8 diffusion, and the model is forced by damping the interface π toward an equilibrium state. Error doubling times are somewhat slow in this model, slightly greater than 4 days.

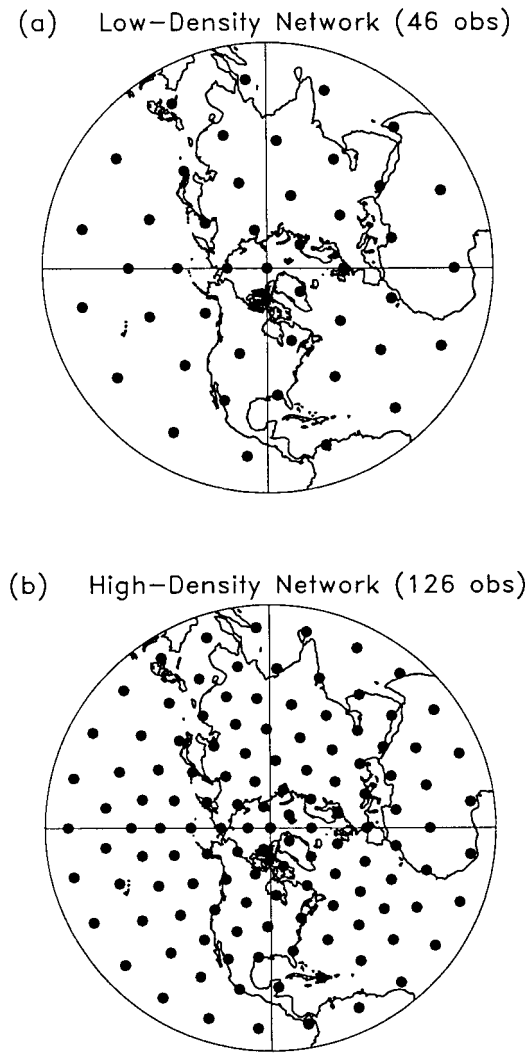

FIG. 2. Observation locations for two network configurations. (a) Lower-density network, (b) higher-density network.

b. Observations

Two observational networks with approximately uniform data density were tested (Fig. 2). We observed u and v components of the wind at both model levels and π at the lower surface and interface. Observations have uncorrelated errors. Wind component error variances were assumed to be 9 $m^2\ s^{-2}$. Lower boundary π variances are assumed to be 0.09 $J^2\ kg^{-2}\ K^{-2}$, or about 1 hPa^2 pressure error variance. Interface π variances were

set to 9.0 $J^2\ kg^{-2}\ K^{-2}$, which corresponds to about 1 K^2 temperature error variance. These same observation error variances were used both to generate random observation errors and were those assumed by the data assimilation scheme. Observations and new analyses were generated every 12 h, followed by a 12-h forecast with the PE model that served as background at the next analysis time.

c. Ensemble Kalman filter data assimilation system

Notational convention in this section roughly follows that suggested in Ide et al. (1997). Let $\mathbf{X}$ be a random vector denoting the model state vector, here converted from spectral components to gridded u and v wind components at the two model levels, as well as lower surface and interface π. Given a set of control observations $\mathbf{y}^o$ of dimension n_o and a background forecast $\mathbf{x}^b$, we seek the specific analysis state $\mathbf{x}^a$, which is the $\mathbf{X}$ that minimizes

$$J(\mathbf{X}) = \frac{1}{2}[(\mathbf{X} - \mathbf{x}^b)^T\mathbf{P}^{b^{-1}}(\mathbf{X} - \mathbf{x}^b) + (\mathbf{y}^o - \mathbf{H}\mathbf{X})^T\mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\mathbf{X})] \quad (7)$$

(Lorenc 1986). Here $\mathbf{P}^b$ represents the background error covariance, and $\mathbf{H}$ (here assumed linear) is an operator that converts the model state to the observation type and location. Further, $\mathbf{R}$ is the $n_o \times n_o$ measurement error covariance matrix.

As in Lorenc (1986), it can be shown that the analysis state that minimizes this functional can be expressed as

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{P}^b\mathbf{H}^T[\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}]^{-1}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b). \quad (8)$$

One of the greatest challenges in data assimilation is formulating a reasonably accurate model of background error covariances $\mathbf{P}^b$. The EnKF presupposes that an ensemble of background states are available to generate background covariance estimates. Ideally, this ensemble approximates a random sample from the probability distribution of plausible background states given all previously and currently available observations. To this end, we used a Monte Carlo procedure similar to that of Houtekamer and Mitchell (1998). We started with an ensemble of n analyses at some time t_0 generated in the manner described in Hamill and Snyder (2000). These perturbed analyses were generated by adding random spatially correlated noise to the an estimate of the truth. We then repeated the following three-step process for each data assimilation cycle: 1) Make n forecasts to the next analysis time, here, 12 h hence. These forecasts will be used as background fields for n parallel analyses. 2) Given the already imperfect observations at this next analysis time (hereafter called the ''control'' observations), generate $i = 1, \ldots, n$ independent sets of *perturbed* observations $\mathbf{y}_i^o$ by adding random noise to the control observations $\mathbf{y}^o$. The noise is drawn from the same distributions as the observation errors (see section 3b), and the noise is constructed to ensure that the mean of the perturbed observations is equal to the control observation. 3) Perform n objective analyses via (9) below, updating each of the n background forecasts using the associated set of perturbed observations. The analysis equation for the ith member is

$$\mathbf{x}_i^a = \mathbf{x}_i^b + \hat{\mathbf{P}}^b\mathbf{H}^T[\mathbf{H}\hat{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R}]^{-1}(\mathbf{y}_i^o - \mathbf{H}\mathbf{x}_i^b). \quad (9)$$

Here $\mathbf{x}_i^b$ is the m-dimensional model state vector for the ith member background forecast of an n-member ensemble, and $\mathbf{x}_i^a$ is the subsequently analyzed state for the ith member. Then $\hat{\mathbf{P}}^b$ is now an approximation of the background error covariances generated from the collection of background forecasts. In its most simple form, $\hat{\mathbf{P}}^b$ is approximated by

$$\hat{\mathbf{P}}^b = \left[\frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i^b - \overline{\mathbf{x}^b})(\mathbf{x}_i^b - \overline{\mathbf{x}^b})^T\right], \quad (10)$$

where $\overline{\mathbf{x}^b} = 1/n \sum_{i=1}^{n} \mathbf{x}_i^b$ is the ensemble mean.

Some additional complexity will be introduced to the standard EnKF design to deal with the detrimental process known as filter divergence (e.g., Houtekamer and Mitchell 1998; van Leeuwen 1999). In this process, the ensemble progressively ignores observational data more and more in successive cycles, leading to a useless ensemble. For the EnKF, much of this problem is a consequence of using the ensemble to produce a reduced-rank representation of background error statistics.

Two potential sources of filter divergence are illustrated in Fig. 3. First, the background at the observation location is adjusted toward the observation only to an extent consistent with the ratio of background (prior) and observational covariances. Figure 3a illustrates a hypothetical posterior probability distribution when background error covariances are estimated correctly (in this example, the covariance between the two components is zero). If background errors are underestimated, the observation is comparatively ignored (Fig. 3b) and the posterior distribution unduly resembles the prior. If there are directions in phase space where the ensemble underestimates the true background covariances because of sampling errors, or at its worst assumes no variance at all because of the limited span of a finite number of ensemble members, then the background is not sufficiently corrected back toward the observation in these directions. Similarly, if the magnitude of background error covariances between an observation location and a far-removed grid point are overestimated due to sampling errors, the posterior probability distribution at this far-removed grid point will be adjusted too much (Fig. 3c). This can generate a posterior probability distribution that is biased and/or has too little variance. In probabilistic terms, the posterior distribution has insufficient probability in the region in phase space near to the true state.

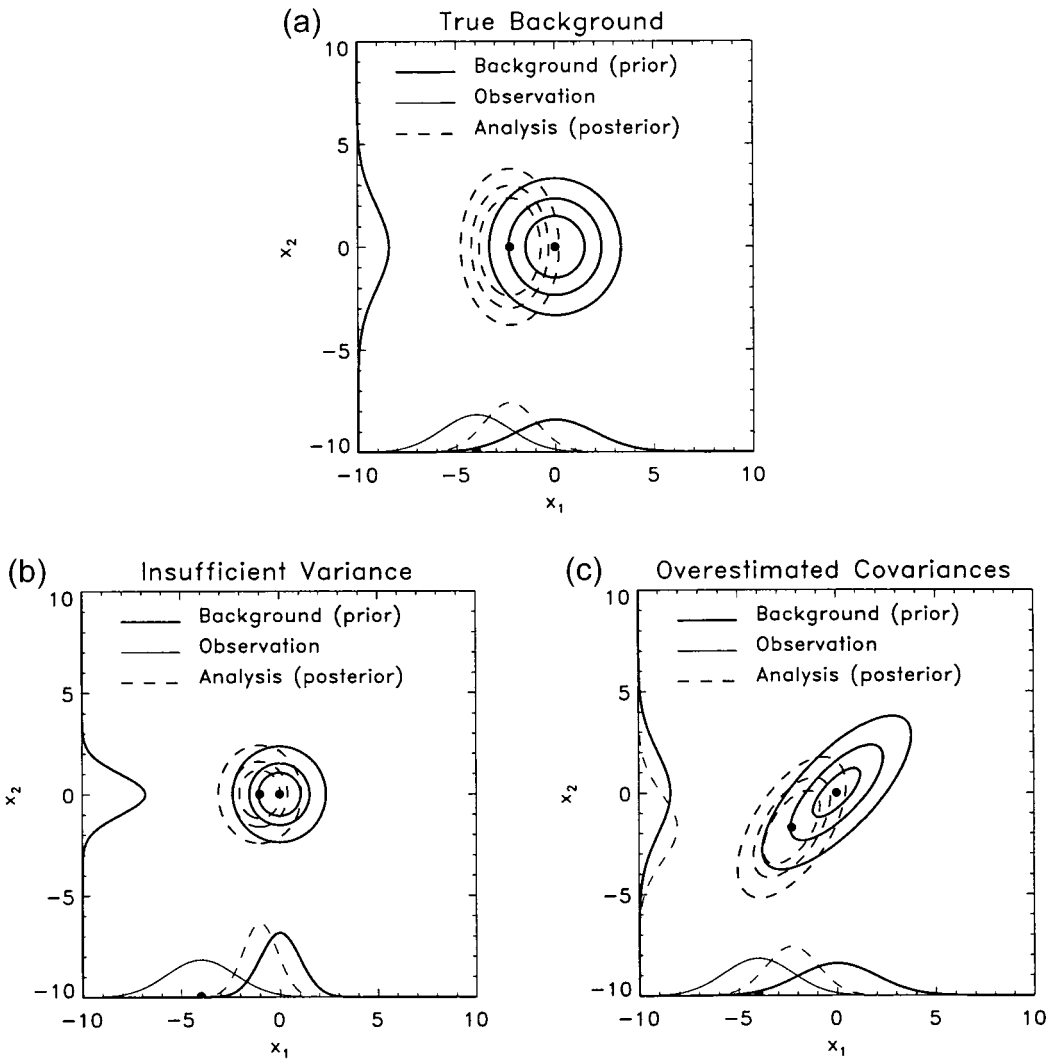

FIG. 3. (a) Hypothetical data assimilation for two-dimensional state vector with an observation in only the x_1 component. Heavy lines denote the true background error distribution, or prior (marginal distributions plotted along each axis). Light solid line denotes marginal distribution for observation. Dot on x_1 axis denotes value of observation. Dashed line denotes distribution of the analysis (posterior). (b) As in (a), but assuming the the background error distribution is underestimated in magnitude. Note that the posterior is shifted very little from the prior. (c) As in (a), but where correlations between the two components are overestimated, so the posterior of x_2 is inappropriately shifted.

In the context of ensemble forecasting, the prior and posterior are represented by a sample of model states. An initial error in the covariances caused by estimating them from a small sample can thus create a ensemble of analyses with a biased mean state and insufficient variance. During the next forecast step, chaotic dynamics may cause member forecasts to drift yet farther from the truth. During the subsequent assimilation cycle, the variance-deficient ensemble thus further underestimates the background error statistics, disregarding even more the influence of the new observations. This problem can progressively worsen, resulting in a useless ensemble of forecasts.

Many approaches have been suggested to lessen or prevent the tendency toward filter divergence. One approach is to localize background error covariances by applying a Schur product with a correlation function, as discussed in Houtekamer and Mitchell (2001) and in greater depth later in this paper. The product of these two covariance models reduces spurious noise in the covariances and the resulting tendency toward introducing unrealistically large analysis increments far from the observation (adjusting ensemble members toward the observations at grid points where it is not appropriate contributes to an inappropriate reduction of the analysis variance).

Another approach that can ameliorate the tendency toward filter divergence is to use a ''double'' EnKF

(Houtekamer and Mitchell 1998), whereby ensemble members are kept in two separate batches; the covariance model from one batch is used in the assmilation of the other. This can help prevent the feedback cycle toward smaller and smaller background error covariances. Hamill and Snyder (2000) suggested a hybrid EnKF, whereby covariances are modeled as a combination of covariances from the ensemble and from a stationary model like 3DVAR. Neither the double EnKF nor the hybrid approach are used in this experiment.

Anderson and Anderson (1999) suggested increasing background error covariances somewhat by inflating the deviation of background members with respect to their mean by a small amount. This is one approach we shall follow here. Before the first observation is assimilated in a new cycle, background forecasts deviation from the mean are inflated by an amount r, slightly greater than 1.0:

$$\mathbf{x}_i^{\mathrm{b}} \leftarrow r(\mathbf{x}_i^{\mathrm{b}} - \overline{\mathbf{x}}^{\mathrm{b}}) + \overline{\mathbf{x}}^{\mathrm{b}}. \tag{11}$$

Here, the operation $\leftarrow$ denotes a replacement of the previous value of $\mathbf{x}_i^{\mathrm{b}}$. Unless noted otherwise, hereafter $r = 1.01$ (1 percent inflation each cycle).

As in Evensen (1994) and Houtekamer and Mitchell (1998, 2001), $\hat{\mathbf{P}}^{\mathrm{b}}$ is not computed explicitly by itself. Rather, for computational efficiency, the matrix operations $\hat{\mathbf{P}}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ and $\mathbf{H}\hat{\mathbf{P}}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ in (9) are computed together using data from the ensemble of background states. Define

$$\overline{\mathbf{Hx}}^{\mathrm{b}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Hx}_i^{\mathrm{b}},$$

which represents the mean of the estimate of the observation generated from the background forecasts. Then

$$\hat{\mathbf{P}}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} = \rho_S \circ \left[\frac{1}{n-1}\sum_{i=1}^{n} (\mathbf{x}_i^{\mathrm{b}} - \overline{\mathbf{x}}^{\mathrm{b}})(\mathbf{Hx}_i^{\mathrm{b}} - \overline{\mathbf{Hx}}^{\mathrm{b}})^{\mathrm{T}}\right], \tag{12}$$

and

$$\mathbf{H}\hat{\mathbf{P}}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} = \frac{1}{n-1}\sum_{i=1}^{n} (\mathbf{Hx}_i^{\mathrm{b}} - \overline{\mathbf{Hx}}^{\mathrm{B}})(\mathbf{Hx}_i^{\mathrm{b}} - \overline{\mathbf{Hx}}^{\mathrm{B}})^{\mathrm{T}}. \tag{13}$$

The operation $\rho_S \circ$ in (12) denotes a Schur product (an element-by-element multiplication) of a correlation matrix $\mathbf{S}$ with the covariance model generated by the ensemble. The Schur product of matrices $\mathbf{A}$ and $\mathbf{B}$ is a matrix $\mathbf{C}$ of the same dimension, where $C_{ij} = A_{ij}B_{ij}$. For sequential data assimilation, the function $\mathbf{S}$ depends upon the observation location; it is a maximum of 1.0 at the observation location and typically decreases monotonically to zero at some finite distance from the observation. As noted in Houtekamer and Mitchell (2001) and references therein, the product of a covariance matrix and correlation function is also a covariance function. Note that because of our use of a simple $\mathbf{H}$, involving only grid points near the observation, the Schur product is not included in (13) (this is a minor approximation since all values of the correlation function at these stencil points used in $\mathbf{H}$ are $\sim$1.0).

To define the correlation matrix $\mathbf{S}$, we used a fifth-order function of Gaspari and Cohn (1999), which is similar to a Gaussian function in shape but compactly supported, that is, correlations decreased to zero at a finite radius. Define a length scale l_c, and let $F_c = \sqrt{10/3}\ l_c$. Define $\|D_{ij}\|$ to be the Euclidean distance between grid point (i, j) and the observation location. Then a correlation matrix $\mathbf{S}$ is defined for every grid point (i, j) in the domain according to $S(i, j) = \Omega(F_c, \|D_{ij}\|)$. Let $a = F_c$ and $b = \|D_{ij}\|$. Then

$$\Omega(a, b) = \begin{cases} -\frac{1}{4}\left(\frac{b}{a}\right)^5 + \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 - \frac{5}{3}\left(\frac{b}{a}\right)^2, & 0 \le b \le a; \\ \frac{1}{12}\left(\frac{b}{a}\right)^5 - \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 + \frac{5}{3}\left(\frac{b}{a}\right)^2 - 5\left(\frac{b}{a}\right) + 4 - \frac{2}{3}\left(\frac{b}{a}\right)^{-1}, & a < b \le 2a; \\ 0, & b > 2a. \end{cases} \tag{14}$$

Because observations were constructed under the assumption of independence of errors, the analysis produced by the sequential assimilation of observations should be identical to the analysis produced by assimilating all simultaneously. [We note that this is strictly true in the context of an extended Kalman filter (e.g., Anderson and Moore 1979), but this is an approximation in the EnKF. Whitaker and Hamill (2001, manuscript submitted to *Mon. Wea. Rev.*) explores this in more depth.] In any case, we assumed this approximation was acceptable here. Thus, each individual observation of u, v, and π at each location were assimilated sequentially. This simplification was attractive; it reduced the rank of $[\mathbf{H}\hat{\mathbf{P}}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}]$ to 1, so computation of its inverse was trivial. The sequential processing of observations also makes application of a correlation function much simpler, as described below. We note that this manner of computation is simple and useful for state vectors of relatively limited dimension and/or a small number of observations, but for realistic numerical prediction ap-

plications, many modifications will be necessary (see Houtekamer and Mitchell 2001 for one possible algorithm).

4. Signal and noise estimates from ensembles

We would like to develop some intuition about whether errors in covariance estimates from small ensembles are especially problematic, and how these errors depend on distance from an observation. To explore this, we first generated an ensemble of assimilations over a 90-day period, assimilating observations every 12 h from the network shown in Fig 2a. To keep the EnKF from diverging, covariances were inflated by the factor $r = 1.01$, and a broad correlation function $\mathbf{S}$ was applied, with $l_c = 4500$ km.

At several points during the data assimilation cycle, background forecasts were used to generate covariances for a set of single-observation experiments. Covariance estimates from a large ensemble ($n = 400$) were assumed accurate and then used to evaluate the noise properties of covariance estimates from a smaller ensemble ($n = 25$). Covariance estimates for these single-observation experiments were then generated as a direct outer product of member deviations about the ensemble mean, that is, $\rho_S \circ$ was not applied in (12) when assimilating the observation. For the assimilation of a single observation with known observation variance, the analysis increment ($\mathbf{x}^a - \mathbf{x}^b$) is then directly proportional to the background error covariance. Hence, a map of analysis increments from the 400-member ensemble will be assumed to be related to the ''true'' covariances, that is, the error is small enough to permit its use for evaluating the accuracy of covariance estimates from a smaller ($n = 25$) member ensemble sampled from this larger ensemble. As demonstrated in section 2, the rms errors [square root of Eq. (5)] in covariance estimates should scale approximately as $1/\sqrt{n}$, so the rms covariance errors from a 400-member ensemble should actually be about 1/4 that of a 25-member ensemble.

For these experiments, the normal data assimilation cycle was interrupted on 25 different case days starting 30 days into the assimilation and with 2½ days between cases. For each case day, 26 independent single-observation experiments were performed, each at a different observation location in the domain above 30°N latitude; this minimum latitude was chosen so that sample points would be affected by extratropical dynamics and thus have substantial background errors. The 26 observations on 25 days thus produced a total of 650 single-observation experiments. In each experiment, an observation increment ($\mathbf{y}^o - \mathbf{H}\mathbf{x}$) of $+3$ J kg^{-1} K^{-1} in π was induced at the interface. This corresponds to an increment of approximately 1 K. We then kept track of the analysis increments ($\mathbf{x}^a - \mathbf{x}^b$). A sample of these increments from each sized ensemble is shown in Fig. 4; note generally larger increments away from the observation location for the 25-member ensemble.

Let $f_{25} = f_{25}(r, \theta, i)$ represent the interface π analysis increment from the 25-member ensemble generated for the ith observation in series of single-observation experiments. Here r denotes the distance from the grid point to the observation location and θ the angle in polar coordinates. Similarly, let f_{400} represent the analysis increment from a 400-member ensemble. We assume the noise N in the 25-member estimate is represented by $N = |f_{400} - f_{25}|$ and the true response, or signal S is $S = |f400|$. We will keep track of the 5th, 50th, and 95th percentiles of N, S, and N/S as a function of r over the 650 replications.

Figure 5 shows these quantiles of S, N, and N/S as a function of the distance from the observations. All quantiles of the typical signal drops off rapidly with increasing distance; the noise is slightly larger near the observation, but its decrease with distance is much less pronounced than the signal. Of particular importance is the ratio of noise to signal, since as indicated in (4), when this ratio is greater than 1.0, assimilating the observation does more harm than good. By ~5000 km, the median N/S is around 1, and beyond this distance the ratio asymptotically approaches a value near 1.5.

Since N/S on average increases monotonically away from the observation, application of a distance-dependent correlation function that decreases covariance estimates monotonically as outlined in section 3c seems to be a plausible choice for improving background covariance estimates and hence the quality of analyses. There may be some occasional circumstances where there are truly large covariances between widely separated grid points that are anomalously damped by the localization. The amount of filtering (i.e, the length scale) should be tuned so that on average correctly large covariances are not excessively damped while anomalous ones are damped.

5. Analysis errors with filtered covariances

We examined the accuracy when the filter length scale was varied while holding the inflation factor fixed. Similarly, we examined the accuracy as a function of the inflation factor while the filter length scale was held fixed. Ensembles with 25, 100, and 400 members were tested. Forecasts and analyses were cycled for 90 days, with updates every 12 h. We examined the analysis error characteristics of interface π averaged over the last 60 days of the integration (errors in other norms were qualitatively similar). The filtering used a fifth-order function in Gaspari and Cohn (1999) as discussed in section 3c.

a. Analysis errors as function of filter length scale

Figures 6a,b present the time average ensemble mean error for the sparse network (46 observation locations; Fig. 2a) and the denser network (126 observation locations; Fig. 2b). The inflation factor r is fixed at 1%.

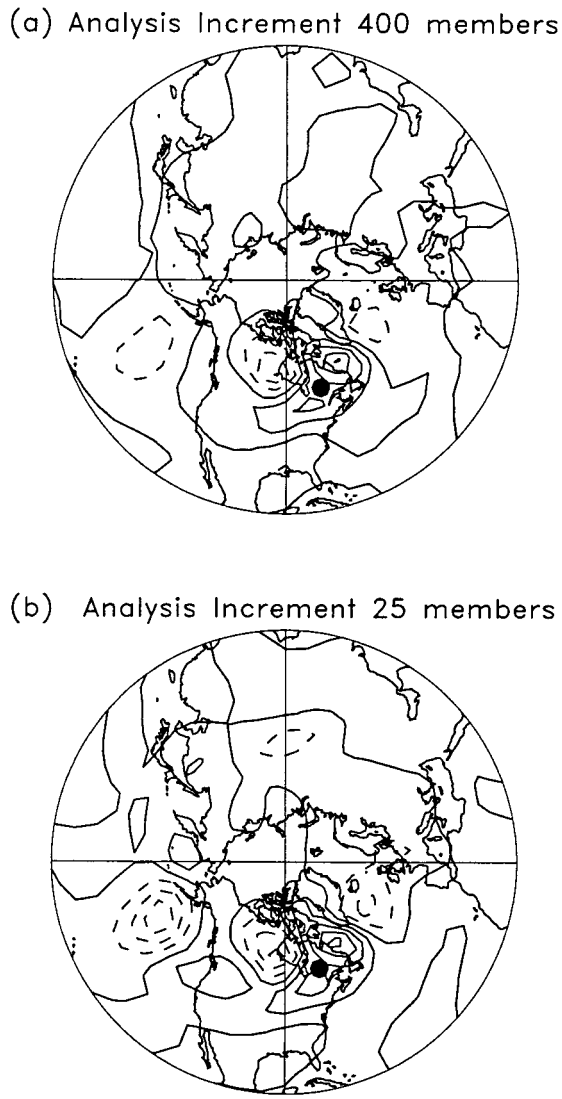

FIG. 4. Analysis increments from single observation experiment, where $+3$ J kg^{-1} K^{-1} observation increment in π was induced at location designated by dot. Contours every 0.01 J kg^{-1} K^{-1}. Negative increments dashed. (a) 400-member ensemble, (b) 25-member ensemble.

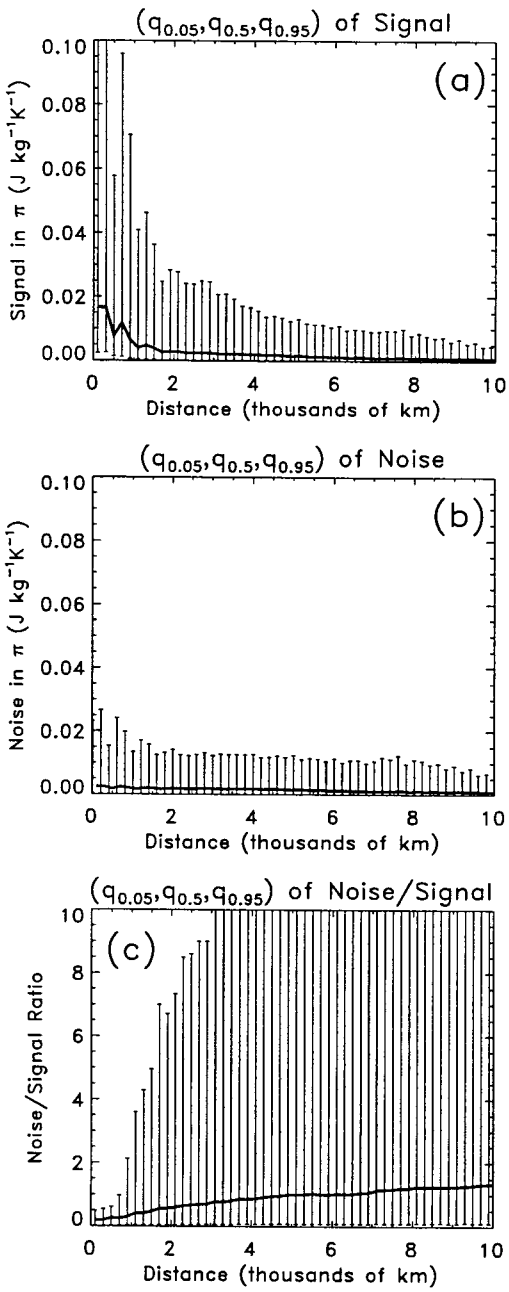

FIG. 5. (a) 5th, 50th (solid line), and 95th percentiles of reference analysis increment (the ''signal'') as function of distance from the observation generated from 400-member ensemble. Original observation increment is $+3$ π at interface (around 1 K). (b) As in (a), but for the increment error (''noise'') of a 25-member ensemble as a function of distance from observation (error is relative to signal from 400-member ensemble). (c) As in (a), but for ratio of noise to signal.

To the right of the dots plotted for a given correlation length scale in Fig. 6, filter divergence occurred for tested larger length scales, and the analyses were useless.

Figure 6 suggests some interesting characteristics of the EnKF coupled with the localization of covariances. First, as expected, the analyses were significantly improved by using more observations. Also note that the optimal length scale is a function of the size of the ensemble. Smaller ensembles had a smaller optimal length scale than for larger ensembles, indicating that noise in the covariance estimates overwhelms signal at relatively short distances from the observations when the ensemble size is small, but for larger ensembles, noise does not overwelm signal until much farther from the observation. This is similar to a result Houtekamer and Mitchell (1998) found using a cutoff radius to eliminate observations.

We also generated rank histograms (Hamill 2001 and

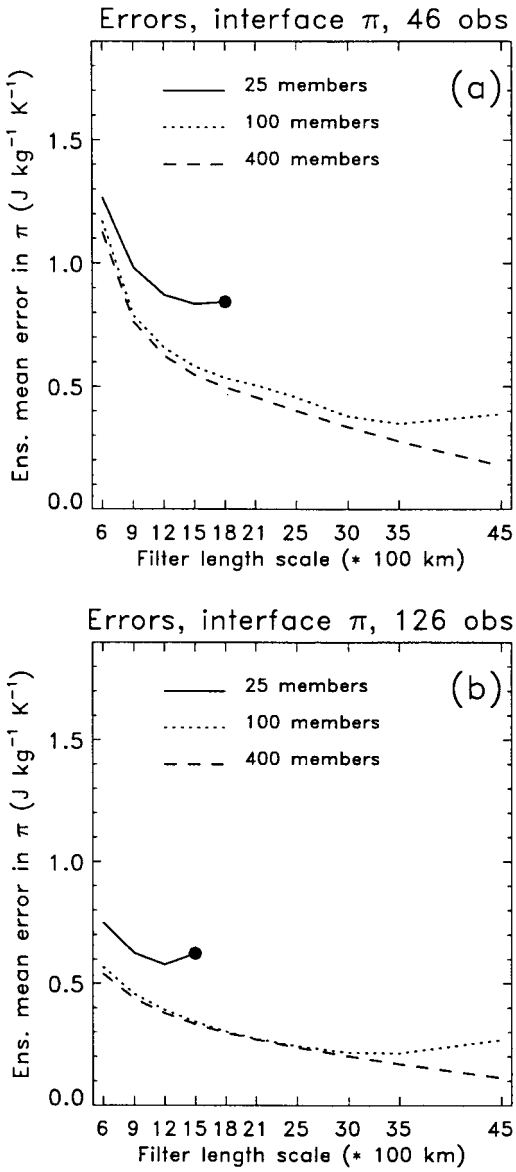

FIG. 6. (a) Time averaged ensemble mean error in interface π for 46-observation network as function of correlation length scale of the filter. (b) As in (a), but for the 126-observation network.

references therein) as a way of measuring the reliability of the ensemble. Ideally, a sample of forecast values from the ensemble and the true state ought to be able to be considered random samples from the same probability distribution. If this is true, then when the rank of the true state is compared to an n-member ensemble sorted from lowest to highest, the rank of the true state should be equally likely to occur in any of the $n + 1$ possible ranks. A histogram of the rank of the truth tallied over many points provides evidence of the reliability of the ensemble. A U-shaped rank histogram (excessive population at the lowest and highest ranks) indicates insufficient spread or bias in the ensemble. An excess population at the middle ranks indicates too much spread.

Figures 7a,b show rank histograms for the 46- and 126-observation networks, respectively. Rank histograms for the 100- and 400-member ensembles were generated by taking a subset of 25 of the members, so that comparisons with the 25-member ensemble could be facilitated. For the 25-member ensembles, at all but the shortest tested length scale, rank histograms are overpopulated at the extreme ranks. This result suggests that the small ensemble may not be able to correctly specify error variances over the full range of growing directions in the ensemble. We will examine this more in section 5c.

For the 25- and 100-member ensembles, there is a trend toward more population at the extreme ranks as the filter length scale increases. This change from underpopulation to overpopulation as filter length increases is a primarily reflection of differing amounts of variance reduction associated with different filter lengths. With strong filtering (a short l_c), only grid points very near the observations are adjusted during the assimilation, and at the rest, the original variance in the background is preserved in the ensemble of analyses and propagated forward to the next cycle. Thus, when filter length is shorter than appropriate for a given sized ensemble, the background covariances estimated from the ensemble are reduced too much in magnitude, undercorrecting the analysis far away from observation locations.

As the length scale of the filter increases, the more background error covariances from the ensemble are trusted far from the observation; hence more and bigger corrections to the analysis are possible. If the covariances are very noisy, though, as shown before, the corrections are inappropriate, and the result is an overly adjusted, variance-deficient ensemble. In the extreme, for very long correlation lengths, this can induce filter divergence. This can be noticed in the rank histograms, which become increasingly U-shaped as correlation length is increased.

b. Analysis errors as function of inflation factor

In all of the experiments described above, background forecast deviations were inflated about their mean by 1% before the data assimilation. It is possible that 1% is not an optimal factor for all ensemble sizes and observation densities. It was too computationally expensive to try a range of inflation factors for all of the correlation length scales. However, we did test a range of inflation factors for the correlation length scale of 1200 km. The corresponding ensemble mean errors are shown in Figs. 8a,b. The optimal inflation factor is a function of the ensemble size. For example, with the 46-observation network and the 25-member ensemble, the analysis can be improved by inflating covariances by ~2%–4%; for the 100-member ensemble, 1% or 2% appears optimal, and for the 400-member ensemble, the 0.25%–1.0% inflations produced the best results of those tested. The minima are less pronounced for the 126-

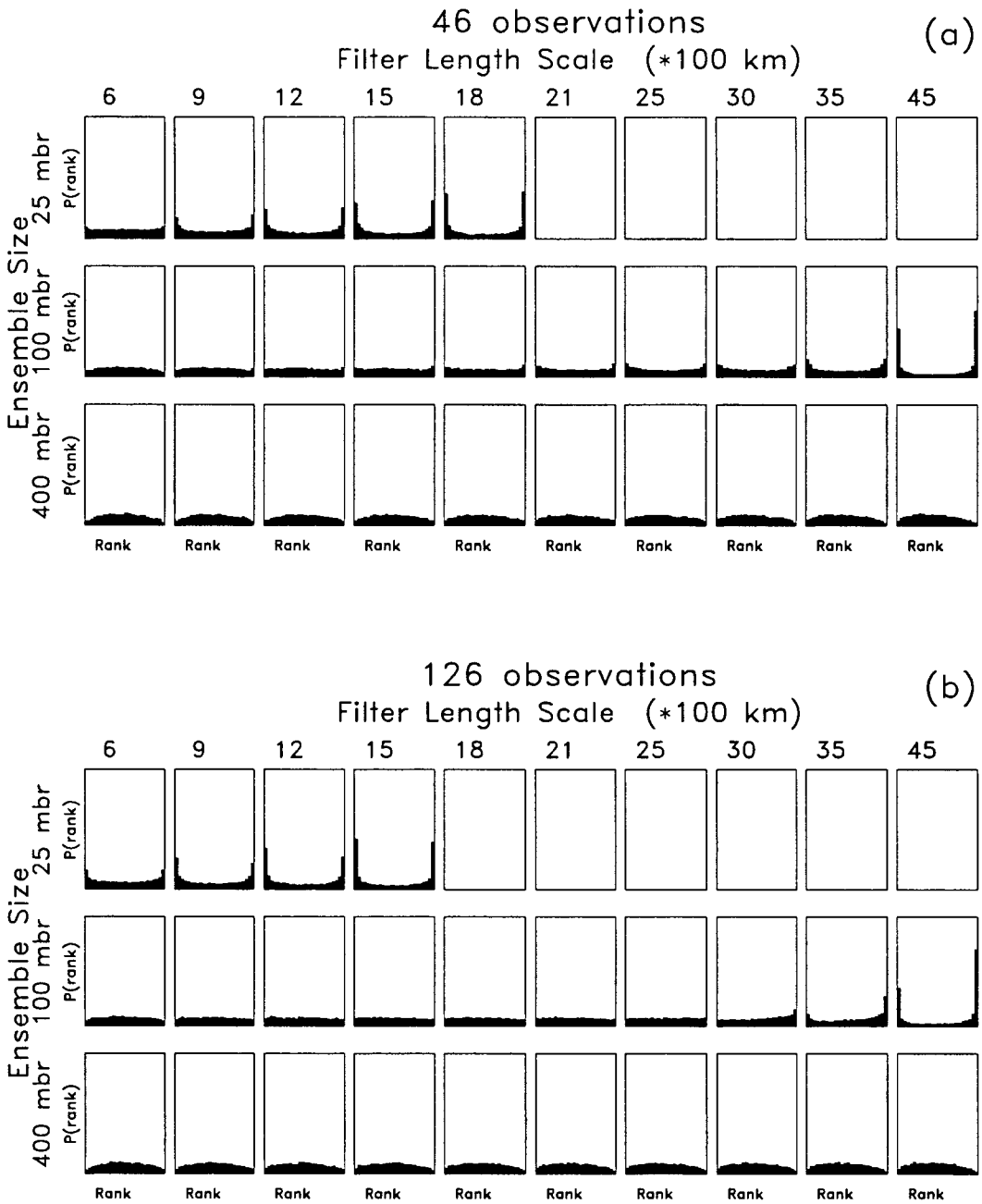

FIG. 7. (a) Rank histograms for the 46-observation network as a function of the ensemble size and the filter correlation length. Where rank histograms are not plotted, filter divergence occurred. (b) As in (a), but for the 126-observation network.

observation ensembles, but the patterns are similar. The rank histograms (Figs. 9a,b) show that, as expected, the larger inflation factors increase the spread in the ensemble, producing less population at the extreme ranks. Note that the 1% inflation factor used for producing Figs. 8 and 9 is nearly optimal for all ensembles; errors can be decreased by a few percent by choosing a different inflation factor, much less of an improvement than can be obtained by adjusting the correlation length.

Given that for all filter length scales in Fig. 7, the rank histograms for the 400-member ensemble were *underpopulated* at the extreme ranks, this suggests that perhaps the inflation factor is too high. We reran all 400-member ensemble forecasts with a 0.25% inflation factor for comparison. The 0.25% inflated ensemble had flat rank histograms and generally slightly lower errors, especially for the longer length scales (not shown).

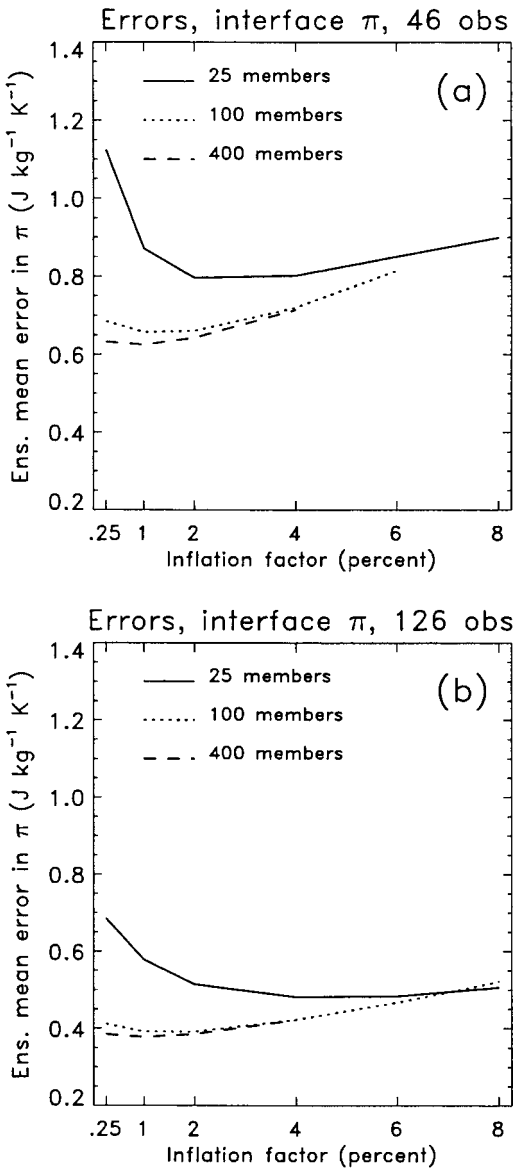

FIG. 8. (a) Time averaged ensemble mean error in interface π for 46-observation network as function of inflation factor and ensemble size. Filter correlation length held fixed at 1200 km. (b) As in (a), but for the 126-observation network.

c. *Eigenvalue spectra of background error covariance matrices*

We would like to develop a qualitative understanding of the why certain inflation factors and correlation length scales are optimal. Some evidence of the deficiencies of small ensembles can be understood from an examination of the eigenspectrum of background error covariance estimates and how they change as a function of ensemble size (P. Houtekamer 1999, personal communication). For a small ensemble, the spectrum of eigenvalues associated with the leading eigenvectors is too steep, indicating that there is insufficient projection upon many of the smaller eigenvectors of the background error covariance matrix; as well, these eigenvectors may increasingly be in the wrong direction. We illustrate this problem here by generating covariance matrices of interface π for ensembles of size 25, 100, and 400, all taken from the experiment with the 400-member ensemble, a 1.01 inflation factor, and 1200-km correlation length scale. For each of 24 sample times (see section 4), a covariance matrix for interface π was calculated from background forecasts for each ensemble size *without covariance localization,* that is, using Eq. (10). The average spectrum of eigenvalues of these covariances is plotted in Fig. 10a. Let us assume that the larger, 400-member ensemble provides a reasonably accurate estimate of the true eigenvalue spectrum. Then the spectra of the 25- and 100-member ensembles can be evaluated. Especially for the 25-member ensemble, there is an excess of variance at the leading eigenvalues, less variance at lower eigenvalues, and of course zero variance beyond eigenvalue 24. One consequence of deficient, reduced-rank approximation was illustrated in Fig. 4, namely, that spurious covariances induce unreasonable corrections to the analysis far distant from the observation location.

When a localization is applied to the covariances, the result is an eigenvalue spectrum that is much flatter (Fig. 10b). The shorter the correlation length, the flatter the spectrum. Why this is so can best be understood by considering the localization in its logical limit, a delta function. This forces all covariances to zero, leaving a diagonal matrix of variances. The rank of this matrix would increase to the dimension of the state vector, and the eigenvalues would be bounded by the largest and smallest variances. Consider also the effect of localization on the analysis increments. Without localization, the correction of the ensemble at two distant observation locations still occurs within the same reduced subspace of the ensemble. With localization, corrections depend on the observation location, introducing extra degrees of freedom.

6. Discussion and conclusions

This paper provided a statistically based rationale for the localization of background error covariances, as proposed by Houtekamer and Mitchell (2001). We demonstrated that the analysis is worsened when the noise (the error) in a covariance estimate is larger than the signal (the true magnitude of the covariance). This ratio of noise to signal is a function of the size of the ensemble; there is less noise with larger ensembles. The ratio of noise to signal is also a function of the magnitude of true correlations between grid points (larger ratios for smaller correlations). Since the correlation is typically a decreasing function of increasing distance, covariances between more distant locations can be expected to have higher ratios of noise to signal.

To understand how errors in covariance estimates

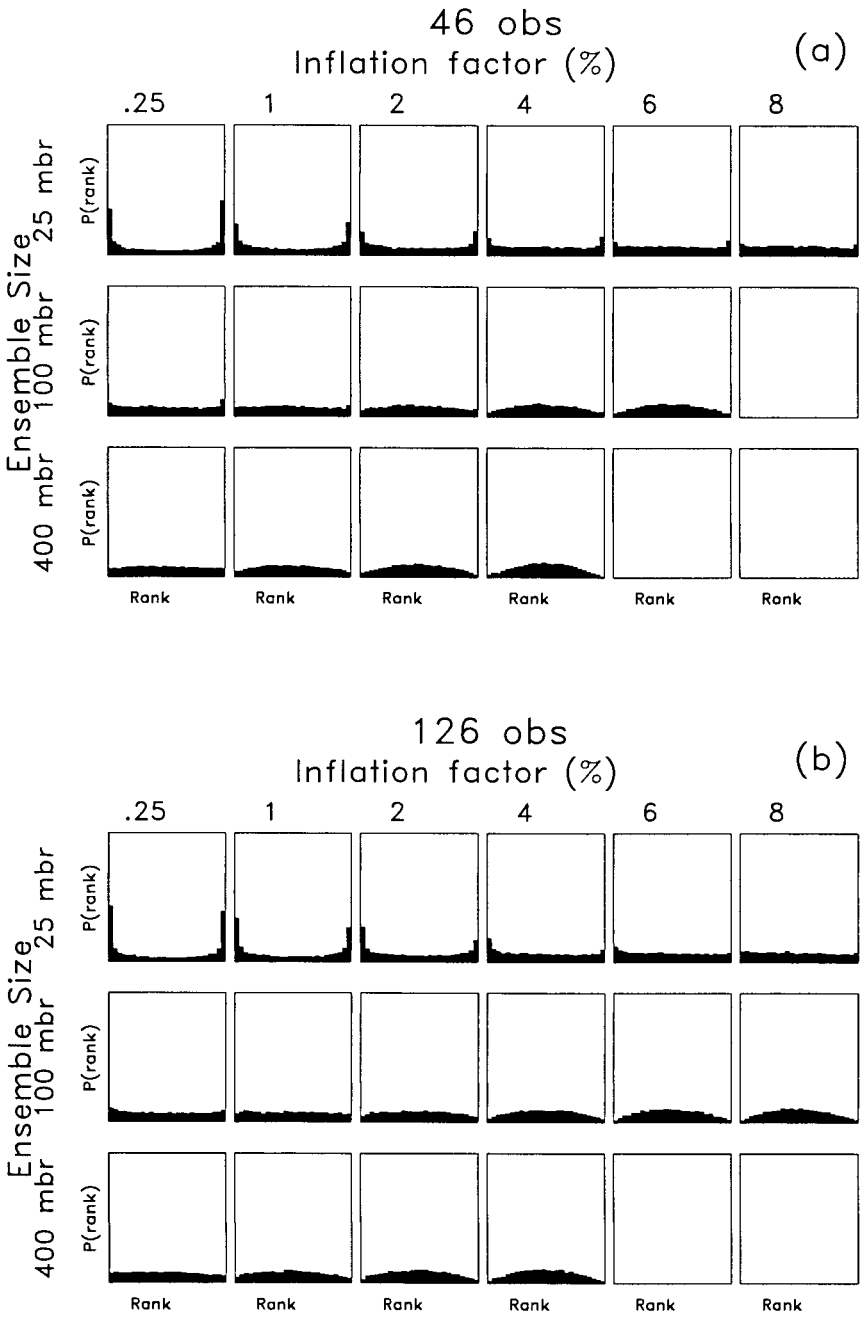

FIG. 9. (a) Rank histograms for 46-observation network as a function of inflation factor and ensemble size. Filter correlation length held fixed at 1200 km. (b) As in (a), but for the 126-observation network.

vary with distance from the observation, covariance estimates were generated from a large ($n = 400$) ensemble and a smaller ($n = 25$) subset of the 400-member ensemble (analysis increments were actually generated, but there is a 1–1 correspondence between the covariances and the increments). These were compared under the assumption that the covariance estimate from the large ensemble could be taken as the true covariance. It was shown that the noise-to-signal ratio for ensemble-based covariances was typically small near the observation and increased to 1.0 at approximately 5000 km from the observation. The ratio of signal to noise continued to increase beyond this distance. This supported the proposition that reducing the magnitude of covariance estimates more as the distance from the observation is increased might beneficially reduce the influence of spurious noise.

An EnKF was tested in which the magnitude of co-

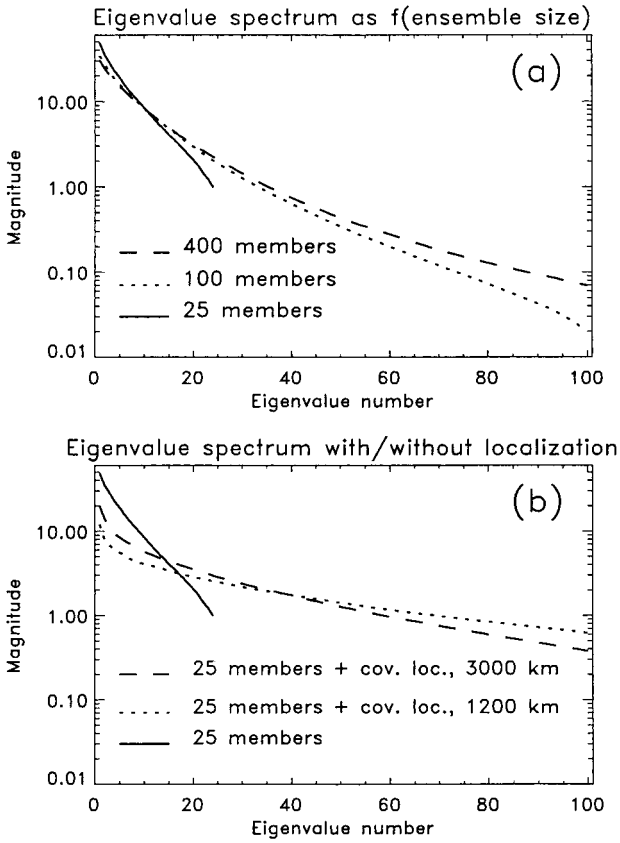

FIG. 10. (a) Average spectrum of eigenvalues of covariance matrix of interface π from 25-, 100-, and 400-member ensembles, all members sampled from the ensemble assimilation test cycle with 400 members, l_c = 1200 km, and r = 1.01. Average determined from 24 sample case days with 2½ days between each sample, starting at the 30th day of the 90-day cycle. (b) As in (a), but for 25-member ensemble with and without covariance localizations of l_c = 1200 and 3000 km applied.

variances were reduced in a distance-dependent manner, with a greater reduction farther from the observation. This was done through a Schur product of ensemble covariances with a correlation function with local support. This localization of covariances provided a notable benefit, especially for the network with few observations. However, the benefit was somewhat smaller for networks with a greater abundance of observations.

We also examined the effects of inflating covariances by increasing the deviation of members around the ensemble mean. We found that the optimal magnitude of the inflation was a function of ensemble size; the smaller the size of the ensemble, the larger the inflation.

An understanding of the dual effects of localization and inflation was gained through a comparison of the eigenvalue spectrum of the background error covariance matrix estimates from small and larger ensembles. It was found that the eigenvalue spectrum of the smaller ensembles was too steep, with an excess of variance associated with the leading eigenvectors and insufficient variance with the trailing eigenvectors. The beneficial aspects of covariance localization and inflation could be partially understood by how they changed the eigenvalue spectrum. Localization tended to increase variance in the tails of the spectrum, while inflation increased the variance associated with all of the resolved eigenvectors.

Overall, the results presented here suggest that a distance-dependent filtering of covariances may provide dramatic improvements to the quality of ensembles from the EnKF or its variants. It is likely that the cost of localizing covariances will be significantly less than the cost of generating a large enough ensemble for the errors to be similar.

The EnKF approach has yet to be tested in an operational environment, though preparation is under way for semioperational testing at the Canadian Meterological Centre (Houtekamer and Mitchell 2001). Though there are many issues yet to be fully understood, our recently submitted work (Whitaker and Hamill 2001, manuscript submitted to *Mon. Wea. Rev.*) addresses two such issues, the effects of noise created by perturbing the observations in the EnKF and the rectitude of serial processing of observations. A tremendous amount of work will be required to determine how best to deal with model error. Regardless, these preliminary results and those of other colleagues demonstrate the potential appeal of ensemble-based data assimilation schemes. We suggest further testing in more complex models, including comparisons with 4D-Var.

Acknowledgments. This research was supported through NCAR's U.S. Weather Research Program. Doug Nychka (NCAR/GSP) is thanked for his assistance with statistical issues. We thank Jean Thiebaux, Jim Purser, and Istvan Szunyogh of NCEP for their advice on an early version of this manuscript. Several library routines were borrowed from *Numerical Recipes* (Press et al. 1992).

This research was conducted partly while the first author was an Advanced Studies Program post-doctoral fellow at NCAR; we thank the NOAA-CIRES Climate Diagnostic Center for allowing us to finish this research.

REFERENCES

Anderson, B. D., and J. B. Moore, 1979: *Optimal Filtering.* Prentice-Hall, 357 pp.

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

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

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

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75** (1B), 257–288.

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

Evensen, G., 1994: Sequential data assimilation with a nonlinear qua-

sigeostrophic 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.

Fisher, M., 1998: Development of a simplified Kalman filter. ECMWF Research Department Tech. Memo. 260, 16 pp. [Available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire, RG2 9AX, United Kingdom.]

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.,* **129,** 550–560.

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

——, ——, and R. E. Morss, 2000: A comparison of probabilistic forecasts from bred, singular vector, and perturbed observation ensembles. *Mon. Wea. Rev.,* **128,** 1835–1851.

Hansen, J. A., and L. A. Smith, 2000: Probabilistic noise reduction. *Tellus,* in press.

Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.,* **129,** 1718–1728.

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

——, and ——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential, and variational. *J. Meteor. Soc. Japan,* **75** (1B), 181–189.

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

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110.

Lermusiaux, P. F. J., and A. R. Robinson, 1999: Data assimilation via error subspace statistical estimation. *Mon. Wea. Rev.,* **127,** 1385–1407.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

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

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119.

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

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran.* 2d ed. Cambridge University Press, 963 pp.

Rabier, F., J.-N. Thepaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quart. J. Roy. Meteor. Soc.,* **124,** 1–39.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74,** 2317–2330.

——, and ——, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

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 T. M. Hamill, 2001: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* submitted.

Zou, X., A. Barcilon, I. M. Navon, J. Whitaker, and D. G. Cacuci, 1993: An adjoint sensitivity study of blocking in a two-layer isentropic model. *Mon. Wea. Rev.,* **121,** 2833–2857.