A Comparison of Hybrid Ensemble Transform Kalman Filter–Optimum Interpolation and Ensemble Square Root Filter Analysis Schemes

XUGUANG WANG

CIRES Climate Diagnostics Center, University of Colorado, and Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado

THOMAS M. HAMILL AND JEFFREY S. WHITAKER

Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado

CRAIG H. BISHOP

Naval Research Laboratory, Monterey, California

ABSTRACT

A hybrid ensemble transform Kalman filter (ETKF)–optimum interpolation (OI) analysis scheme is described and compared with an ensemble square root filter (EnSRF) analysis scheme. A two-layer primitive equation model was used under perfect-model assumptions. A simplified observation network was used, and the OI method utilized a static background error covariance constructed from a large inventory of historical forecast errors. The hybrid scheme updated the ensemble mean using a hybridized ensemble and static background-error covariance. The ensemble perturbations in the hybrid scheme were updated by the ETKF scheme. The EnSRF ran parallel data assimilation cycles for each member and serially assimilated the observations. The EnSRF background-error covariance was estimated fully from the ensemble.

For 50-member ensembles, the analyses from the hybrid scheme were as accurate or nearly as accurate as those from the EnSRF, depending on the norm. For 20-member ensembles, the analyses from the hybrid scheme were more accurate than analyses from the EnSRF under certain norms. Both hybrid and EnSRF analyses were more accurate than the analyses from the OI. Further reducing the ensemble size to five members, the EnSRF exhibited filter divergence, whereas the analyses from the hybrid scheme were still better than those updated by the OI. Additionally, the hybrid scheme was less prone to spurious gravity wave activity than the EnSRF, especially when the ensemble size was small. Maximal growth in the ETKF ensemble perturbation space exceeded that in the EnSRF ensemble perturbation space. The relationship of the ETKF ensemble variance to the analysis error variance, a measure of a spread–skill relationship, was similar to that of the EnSRF ensemble. The hybrid scheme can be implemented in a reasonably straightforward manner in the operational variational frameworks, and the computational cost of the hybrid is expected to be much less than the EnSRF in the operational settings.

1. Introduction

Variational techniques are now almost universally used for operational atmospheric data assimilation, either in its three-dimensional form (3DVAR; e.g., Parrish and Derber 1992; Courtier et al. 1998; Gauthier et al. 1998; Cohn et al. 1998) or in its four-dimensional form (4DVAR, e.g., Courtier et al. 1994; Rabier et al. 1998, 2000). Both 3DVAR and 4DVAR assimilations generally utilize a nearly homogeneous, isotropic, and stationary background error covariance model, which only provides a crude estimate of the actual flow-dependent forecast-error structure. Currently, 3DVAR is still utilized in many operational centers because of its computational efficiency and algorithmic simplicity.

Recently, a variety of techniques have been explored to relax the restrictions on the background-error covariance model in data assimilations. For 3DVAR, techniques are being developed that make it possible to

Corresponding author address: Dr. Xuguang Wang, Physical Sciences Division, NOAA/Earth System Research Laboratory, 325 Broadway, R/PSD1 Boulder, CO 80305-3328.
E-mail: xuguang.wang@noaa.gov

include some spatial inhomogeneity, nonstationarity, and anisotropy (e.g., Riishøjgaard 1998; Wu et al. 2002; Purser et al. 2003; Liu et al. 2005). A different approach is to use ensemble-based data assimilation techniques, where the background-error covariances are estimated from an ensemble of short-term forecasts (e.g., Evensen 1994; Burgers et al. 1998; Houtekamer and Mitchell 1998, 2001, 2005, hereafter HM05; Anderson 2001; Whitaker and Hamill 2002; Tippett et al. 2003; Snyder and Zhang 2003; Ott et al. 2004; Hamill and Whitaker 2005; Szunyogh et al. 2005; Tong and Xue 2005; Hanea et al. 2007). For reviews on ensemble-based techniques, see Evensen (2003), Lorenc (2003), and Hamill (2006).

The presumed benefit of utilizing these ensemble-based techniques is their ability to provide a flow-dependent estimate of the background-error covariances so that the relative contribution of the background and the observations are more appropriately weighted. Another potential benefit is that the processes of ensemble forecasting and data assimilation are unified. In many controlled tests with simple models and simulated observations, ensemble-based methods have demonstrated dramatically improved analysis skill. However, such dramatic improvements have not yet been seen in the tests of ensemble-based methods in realistic operational environments. Recent global, real-data experiments show that the ensemble Kalman filter (EnKF; Houtekamer et al. 2005; HM05) and the ensemble square root filter (EnSRF; Whitaker et al. 2006, manuscript submitted to *Mon. Wea. Rev.*), provide comparable or slightly better results to operational 3DVAR algorithms.

Many of the current ensemble-based data assimilation techniques serially process the observations, so the computational expense of the algorithm scales linearly with the number of observations. Attractively, the operational 3DVAR at the National Centers for Environmental Prediction (NCEP) enables many more observations to be added without dramatically increasing computational cost (J. Derber 2005, personal communication). The computational expense of such ensemble-based algorithms may make them unattractive for use in operational numerical weather prediction centers, where the numbers of available observations are increasing rapidly each year.

Is there a method that can take advantage of the computational efficiency of the 3DVAR while also benefiting from the ensemble-estimated error covariance? Recently, Hamill and Snyder (2000, henceforth HS00) proposed a hybrid EnKF–3DVAR method. In this scheme, the background error covariance was obtained by linearly combining a flow-dependent background error covariance from the ensemble with the standard error covariance estimate from 3DVAR. Each ensemble member was then updated variationally with perturbed observations. Subsequently, Lorenc (2003) and Buehner (2005) discussed how an ensemble-based covariance model could be adapted conveniently to the variational framework by extending the control variables. A proof of the equivalence of the hybrid method realized by the augmented control variables and by directly weighting the covariances was recently provided in Wang et al. (2007).

More recently, Etherton and Bishop (2004, hereafter EB04) tested a hybrid scheme in a two-dimensional turbulence model. In EB04, the ensemble covariance was estimated from the ensemble transform Kalman filter (ETKF) ensemble, and the static background error covariance was formulated in a similar way to the optimal interpolation (OI) scheme (Daley 1985; Lorenc 1981; Daley 1991). Unlike the HS00 scheme where K parallel data assimilation cycles for the K members were required, in the EB04 scheme, a single update of the mean was performed, while the ETKF transformed the background perturbations into analysis perturbations in a computationally efficient manner. EB04 found that the performance of the hybrid scheme using the ETKF ensemble was comparable to that obtained using considerably more expensive ensemble generation schemes.[1]

The purpose of this study is to further explore the potential skill of the hybrid scheme proposed in EB04 by comparing it with the EnSRF, one of the more well-tested ensemble-based assimilation schemes (Whitaker and Hamill 2002; Whitaker et al. 2006, manuscript submitted to *Mon. Wea. Rev.*; Snyder and Zhang 2003; Tong and Xue 2005; Zhang et al. 2004). Note that in this paper, the proposed hybrid scheme will update the mean state by using the classic statistical, or OI formula (Schlatter 1975), which is the same method of solution as the Kalman filter state update equation (Daley 1991). We thus call it the hybrid ETKF–OI scheme. However, under our experiment design, it will provide the same solution as if we had updated the mean state in a 3DVAR framework using the same background-error covariance model (Wang et al. 2007; Daley 1991).

---

[1] The ETKF was proposed by Bishop et al. (2001). It has been demonstrated to be a useful tool for targeted observations (Majumdar et al. 2001, 2002a,b). Wang and Bishop (2003) and Wang et al. (2004) also showed that the ETKF provided an inexpensive yet demonstrably superior scheme to the breeding method (Toth and Kalnay 1993, 1997) for generating perturbed initial conditions for ensemble forecasts.
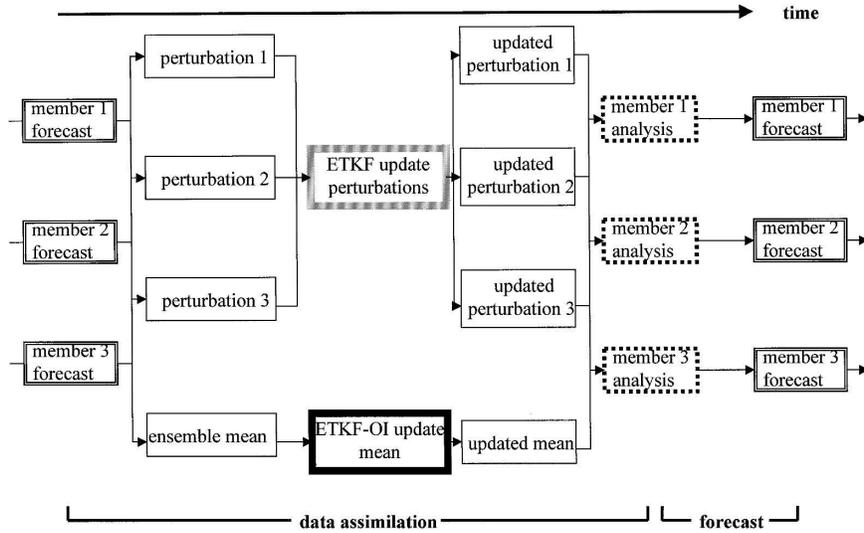
FIG. 1. Illustration of the hybrid ETKF–OI analysis and ensemble generation cycle for a hypothetical three-member ensemble. The ensemble mean is updated with the background-error covariance estimated by the linear combination of the ETKF ensemble covariance and the static covariance. The ensemble perturbations are updated by the ETKF transformation matrix, with innovation-based inflation applied.

The motivating factors for exploring the hybrid ETKF–OI scheme in more depth include the following: (i) the hybrid can be incorporated into the operational variational schemes through minor extensions of the current variational framework (Lorenc 2003; Buehner 2005; Wang et al. 2007); (ii) it can perform no worse than 3DVAR, for its blend of covariances can always be adjusted to exclude the contribution of the ensemble; and (iii) the hybrid is computationally less expensive than the EnKF and the EnSRF. Hence, if the hybrid scheme controls error as well as the EnSRF in the simple-model experiments considered here, the motivation for testing the hybrid at operational centers would increase.

As an initial attempt to investigate the potential skill of the hybrid ETKF–OI relative to the EnSRF, we conduct experiments using a two-layer primitive equation model under the perfect-model assumption. The simple-model tests considered here allow us to directly compare results from the hybrid, EnSRF, and OI methods. In imperfect-model simulations or real-data experiments, differences in quality would be affected by many other factors, such as the parameterization of model errors or the quality control of the observations. Future work by the coauthors will extend this research to compare the methods in simulations including model error.

The rest of the article is organized as follows. In section 2, we briefly review the hybrid ETKF–OI and the EnSRF analysis schemes. Section 3 describes the ex-

perimental design. Results of comparing the two schemes are reported in section 4. Section 5 provides a discussion and conclusions.

2. The hybrid ETKF–OI and the EnSRF analysis schemes

a. The hybrid ETKF–OI scheme

Figure 1 illustrates how the hybrid ETKF–OI data assimilation cycle works. Start with an ensemble of K background forecasts at time t_0. The following four steps are then repeated for each data assimilation cycle. 1) Update the ensemble-mean or a high-resolution control forecast by the hybridized background error covariance. 2) Update the forecast perturbations using the computationally inexpensive ETKF. 3) Add the updated ensemble perturbations to the updated ensemble mean to generate K initial ensemble members. 4) Make K forecasts starting from the K initial ensemble members forward to the next analysis time.

We first consider the update of the mean in the hybrid method. Here, the background-error covariance $\mathbf{P}^b$ is approximated by a linear combination of the sample covariance matrix of the ETKF forecast ensemble $\mathbf{P}^e$ and the static covariance matrix $\mathbf{B}$; that is,

$$\mathbf{P}^b = (1 - \alpha)\mathbf{P}^e + \alpha\mathbf{B}, \tag{1}$$

where α is the weighting coefficient, $0 \leq \alpha \leq 1$, and $\mathbf{P}^e$ is given by

$$\mathbf{P}^e = \frac{\mathbf{X}^b(\mathbf{X}^b)^{\mathrm{T}}}{K-1}, \qquad (2)$$

where the columns of $\mathbf{X}^b$ contain K ensemble perturbations from the mean, $\mathbf{x}_i'^b$, $i = 1, \ldots K$.

The ensemble-mean forecast $\bar{\mathbf{x}}^b$ is then updated by the new observations $\mathbf{y}$ to obtain the ensemble-mean analysis $\bar{\mathbf{x}}^a$ using

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{P}^b\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^b), \qquad (3)$$

where $\mathbf{H}$ is the observation operator mapping from the model state variables to the observed variables, here presumed linear, $\mathbf{R}$ is the observation-error covariance matrix, and $\mathbf{P}^b$ is given by Eq. (1). Note as in EnKF framework (Evensen 1994; Houtekamer and Mitchell 1998), there is no need to compute and store the full matrix $\mathbf{P}^b$. Instead we first form $\mathbf{BH}^{\mathrm{T}}$ and $\mathbf{HBH}^{\mathrm{T}}$ (see section 3b for details), and calculate $\mathbf{P}^e\mathbf{H}^{\mathrm{T}}$ and $\mathbf{HP}^e\mathbf{H}^{\mathrm{T}}$ from the ensemble. Then we use the linear coefficient α as in Eq. (1) to form $\mathbf{P}^b\mathbf{H}^{\mathrm{T}}$ and $\mathbf{HP}^b\mathbf{H}^{\mathrm{T}}$; that is,

$$\mathbf{HP}^b\mathbf{H}^{\mathrm{T}} = (1 - \alpha)\mathbf{HP}^e\mathbf{H}^{\mathrm{T}} + \alpha\mathbf{HBH}^{\mathrm{T}}, \qquad (4)$$

$$\mathbf{P}^b\mathbf{H}^{\mathrm{T}} = (1 - \alpha)\mathbf{P}^e\mathbf{H}^{\mathrm{T}} + \alpha\mathbf{BH}^{\mathrm{T}}. \qquad (5)$$

When $\alpha = 1$, the background error covariance is fully estimated from the static background error covariance.

A now-common adjustment to error covariances estimated directly from the ensemble is to apply a "localization." Covariance localization implements an elementwise multiplication of the background error covariance matrix with a correlation function with local support (Houtekamer and Mitchell 2001; Hamill et al. 2001). This typically stabilizes the filter by damping out spurious covariances between distant locations, an inevitable consequence of estimating error covariances with a limited-size ensemble. As will be shown in the results, covariance localization is not strictly required in the hybrid scheme when updating the mean, for the addition of a higher-dimensional static background-error covariance itself stabilizes the filter. However, the mitigating effects of covariance localization will be tested when updating the ensemble-mean state. To apply covariance localization, we replace $\mathbf{HP}^e\mathbf{H}^{\mathrm{T}}$ in Eq. (4) and $\mathbf{P}^e\mathbf{H}^{\mathrm{T}}$ in Eq. (5) by $\boldsymbol{\rho}_s \circ \mathbf{HP}^e\mathbf{H}^{\mathrm{T}}$ and $\boldsymbol{\rho}_s \circ \mathbf{P}^e\mathbf{H}^{\mathrm{T}}$, respectively, where the operation $\boldsymbol{\rho}_s \circ$ denotes a Schur product of a correlation matrix $\boldsymbol{\rho}_s$ with the covariance matrix generate by the ensemble. For horizontal localization, one such correlation matrix can be constructed using Gaspari and Cohn's (1999) approximately Gaussian-shaped function with local support, which will be used here. Henceforth we shall refer to a "scale" of the localization. A scale of 3000 km indicates that the Gas-

pari and Cohn correlation function tapers to zero at this distance.

As mentioned in the previous section and section 5, previous work (Lorenc 2003; Buehner 2005; Wang et al. 2006) has shown that hybridizing the ensemble covariance in the existing operational 3DVAR framework should be straightforward. In this study, we updated the mean using Eqs. (3)–(5). The solution given by Eqs. (3)–(5) is the same as the solution given by the variational framework proposed by Lorenc (2003) and Buehner (2005), under the assumptions of the normality of the errors.

We now consider the method for updating perturbations around the mean state. In the proposed hybrid scheme, the ensemble perturbations are updated by the ETKF. The ETKF transforms the matrix of forecast perturbations $\mathbf{X}^b$ into a matrix of analysis perturbations $\mathbf{X}^a$, whose columns contain K analysis perturbations, $\mathbf{x}_k'^a$, $k = 1, \ldots, K$. The transformation happens through the postmultiplication by the matrix $\mathbf{T}$, that is,

$$\mathbf{X}^a = \mathbf{X}^b\mathbf{T}. \qquad (6)$$

The transformation matrix is chosen to ensure that the analysis-error covariance formed from the outer product of the transformed perturbations will be precisely equal to the true analysis-error covariance, assuming that Eq. (2) denotes the true forecast-error covariance, all errors are normally distributed, and $\mathbf{H}$ is linear. As shown in Bishop et al. (2001), Wang and Bishop (2003), and Wang et al. (2004), a precise spherical simplex solution of $\mathbf{T}$ is

$$\mathbf{T} = \mathbf{C}(\boldsymbol{\Gamma} + \mathbf{I})^{-1/2}\mathbf{C}^{\mathrm{T}}, \qquad (7)$$

where $\mathbf{C}$ contains the eigenvectors and $\boldsymbol{\Gamma}$ the eigenvalues of the $K \times K$ matrix $(\mathbf{X}^b)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{X}^b$ and $\mathbf{I}$ is the identity matrix. For the ensemble size K of 100 or less, the computation of Eq. (7) is relatively inexpensive.

Recent research has shown that when K is significantly smaller than the rank r of the true forecast-error covariance, Eq. (7) systematically underestimates the analysis-error variance. In Wang and Bishop (2003), the ensemble of analyzed deviations $\mathbf{X}^a$ were inflated by large factors to compensate for the ETKF's underestimate of the analysis-error variance (Wang and Bishop 2003). An alternative formulation of the ETKF (appendix A) significantly ameliorated this bias by accounting for (i) the fact that the sample covariance of K forecast trials systematically overestimates the true error variance within the ensemble subspace when $K \ll r$, and (ii) the expected difference in angle subtended between ensemble-based eigenvectors and true eigenvectors. Based on these arguments, the ETKF transformation matrix $\mathbf{T}$ as applied here is

$$\mathbf{T} = \mathbf{C}(\rho\mathbf{\Gamma} + \mathbf{I})^{-1/2}\mathbf{C}^{\mathrm{T}}, \qquad (8)$$

where the scalar factor ρ is the fraction of the forecast error variance projected onto the ensemble subspace. It is estimated by

$$\rho = \frac{\overline{(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\bar{\mathbf{x}}^b)^{\mathrm{T}}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\bar{\mathbf{x}}^b)} - (K - 1)}{\overline{(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\bar{\mathbf{x}}^b)^{\mathrm{T}}(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\bar{\mathbf{x}}^b)} - p}, \qquad (9)$$

where p is the number of observations, $\tilde{\mathbf{H}}$ is the normalized observation operator $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}$, and the columns of $\mathbf{E}$ contain the eigenvectors of the ensemble covariance in normalized observation space. As shown in Eq. (12) of Bishop et al. (2001),

$$\mathbf{E} = \tilde{\mathbf{H}}\mathbf{X}^b\mathbf{C}\mathbf{\Gamma}^{-1/2}\big/\sqrt{K - 1}. \qquad (10)$$

The overbar in Eq. (9) represents the average over some independent samples. In this experiment it is the average over two weeks' computations prior to each assimilation time. For derivations of Eqs. (8) and (9), please refer to appendices A and B, respectively.

Note that the computational efficiency of the ETKF is realized by solving the transformation matrix in ensemble perturbation subspace (Bishop et al. 2001). Thus, to maintain the efficiency of the ETKF, the ensemble update is performed using only the ensemble-based covariance, not a hybrid covariance. Also, no covariance localization via either a smooth correlation function (e.g., HM05; Whitaker and Hamill 2002) or via domain decomposition (Ott et al. 2004; Szunyogh et al. 2005) was applied when the ETKF updated the ensemble perturbations. The transformation was done for the global perturbations to take advantage of the computational efficiency of the ETKF.

Because the new ETKF formulation, Eq. (8), can only partly ameliorate the old formulation's underestimate of the analysis-error variance, we still apply an inflation to increase the ensemble covariance. In this study, the maximum-likelihood inflation method (Wang and Bishop 2003) is applied to the analysis perturbations. The idea is to multiply the initial perturbations obtained at time t_i by an estimated inflation factor Π_i; that is,

$$\mathbf{X}_i^a = \mathbf{X}_i^f\mathbf{T}_i\Pi_i. \qquad (11)$$

The purpose of this is to ensure that at time t_{i+1} the background ensemble forecast variance is consistent with the ensemble-mean background-error variance over global observation sites. Specifically, define $\tilde{\mathbf{d}}_i$ as the innovation vector at t_i, normalized by the square root of the observation error covariance matrix, that is,

$\tilde{\mathbf{d}}_i = \mathbf{R}^{-1/2}(\mathbf{y}_i - \mathbf{H}\bar{\mathbf{x}}_i^b)$, where $\mathbf{y}_i$ is the observation vector at t_i and $\mathbf{H}\bar{\mathbf{x}}_i^b$ is the ensemble mean background forecast valid at the time t_i mapped into observation space by the observation operator $\mathbf{H}$. Given that the inflation factor at t_{i-1} was Π_{i-1}, the inflation factor for the transformed perturbation at t_i is obtained by first checking if $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i$ is equal to $\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}_i^e\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})$, where Tr denotes the trace. If not, we need to introduce a parameter c_i so that

$$\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i = \mathrm{Tr}(\tilde{\mathbf{H}}c_i\mathbf{P}_i^e\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I}). \qquad (12)$$

Then the inflation factor Π_i is defined as

$$\Pi_i = \Pi_{i-1}\sqrt{c_i}. \qquad (13)$$

This rescaling of the initial perturbations by Eq. (13) attempts to correct the spread of the set of forecast ensemble perturbations at time t_{i+1} by using the rescaling factor that would have produced a proper forecast ensemble spread at t_i if it had been applied to the transformed perturbations at t_{i-1}. From Eq. (12),

$$c_i = \frac{\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i - p}{\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}_i^e\tilde{\mathbf{H}}^{\mathrm{T}})}, \qquad (14)$$

where p is the number of observations. From Eq. (13), Π_i is a product of these c parameters from the first forecast at time t_1 to that at time t_i, that is,

$$\Pi_i = \sqrt{c_1c_2\cdots c_i}. \qquad (15)$$

A lower bound of 1.0 is set for Π_i. Implicitly in Eq. (12), we assume $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i = \mathrm{Tr}\langle\tilde{\mathbf{d}}_i\tilde{\mathbf{d}}_i^{\mathrm{T}}\rangle$, which requires the number of independent elements in the innovation vector $\tilde{\mathbf{d}}_i$ to be large. The real-time global observational network meets this assumption well (Dee 1995). Because the number of observations in our experiment is rather limited, we replace $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i$ in Eqs. (12) and (14) by using the average of squared innovation vectors two weeks prior to time t_i, denoted as $\overline{\tilde{\mathbf{d}}^{\mathrm{T}}\tilde{\mathbf{d}}}_{\mathrm{prior}\,t_i}$ (see discussion in appendix B also). Thus Eq. (14) becomes

$$c_i = \frac{\overline{\tilde{\mathbf{d}}^{\mathrm{T}}\tilde{\mathbf{d}}}_{\mathrm{prior}\,t_i} - p}{\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}_i^e\tilde{\mathbf{H}}^{\mathrm{T}})}. \qquad (16)$$

In the following experiment, the averaged value and standard deviation of Π_i decreased with increasing ensemble size.

b. The EnSRF analysis scheme

The EnSRF was fully described in Whitaker and Hamill (2002). It is one of the simpler implementations of a class of ensemble square root filters, which includes the ETKF (Tippett et al. 2003). Whereas the primary advantage of the hybrid scheme is its potential low computational expense, the comparative advantage of the EnSRF is its relative algorithmic simplicity and its

full exploitation of ensemble-based background error covariances. In the EnSRF, covariance localization is required to avoid filter divergence (Houtekamer and Mitchell 2001; Hamill et al. 2001).

The EnSRF serially assimilates observations. The ensemble of perturbations updated by the previous observations is used to model the background-error covariance for assimilating the next observation (for details see Whitaker and Hamill 2002). Similarly, the updated mean from the assimilation of the previous observation is used as the prior state for the assimilation of the next observation. The EnSRF update equations are as follows:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^b), \qquad (17)$$

$$\mathbf{x}_i'^a = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H})\mathbf{x}_i'^b. \qquad (18)$$

Here $\mathbf{K}$ is the Kalman gain modified by the covariance localization,

$$\mathbf{K} = (\boldsymbol{\rho}_s \circ \mathbf{P}^e\mathbf{H}^T)(\boldsymbol{\rho}_s \circ \mathbf{H}\mathbf{P}^e\mathbf{H}^T + \mathbf{R})^{-1}, \qquad (19)$$

where operation $\boldsymbol{\rho}_s \circ$ denotes a Schur product of a correlation matrix $\boldsymbol{\rho}_s$ with the covariance matrix generated by the ensemble. Horizontal localization by using Gaspari and Cohn's (1999) approximately Gaussian-shaped function is used here; $\tilde{\mathbf{K}}$ in Eq. (18) is called the "reduced" Kalman gain matrix (Whitaker and Hamill 2002). When sequentially processing independent observations, $\mathbf{K}$, $\tilde{\mathbf{K}}$, and $\mathbf{P}^e\mathbf{H}^T$ are all vectors with the same number of elements as the model-state vector, and $\mathbf{H}\mathbf{P}^e\mathbf{H}^T$ and $\mathbf{R}$ are scalars. The reduced gain matrix becomes

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^e\mathbf{H}^T + \mathbf{R}}}\right)^{-1} \mathbf{K}. \qquad (20)$$

Inflation factors were computed and applied to the EnSRF using the same method as that used for the ETKF ensemble in Eqs. (15)–(16).

3. Experiment design

a. Model, observations, and ensemble configuration

In this study, we ran a dry, global, two-layer primitive equation model (Zou et al. 1993). It was previously used in Hamill et al. (2001) and Hamill and Whitaker (2005) for ensemble data assimilation experiments in both a perfect-model and imperfect-model contexts. The model is spectral, and the model-state vector includes coefficients of vorticity and divergence at two levels and coefficients of two-layer thicknesses $\Delta\pi_1$ and $\Delta\pi_2$, where π is the Exner function. There is a simple, zonal wavenumber-2 terrain. The model is forced by
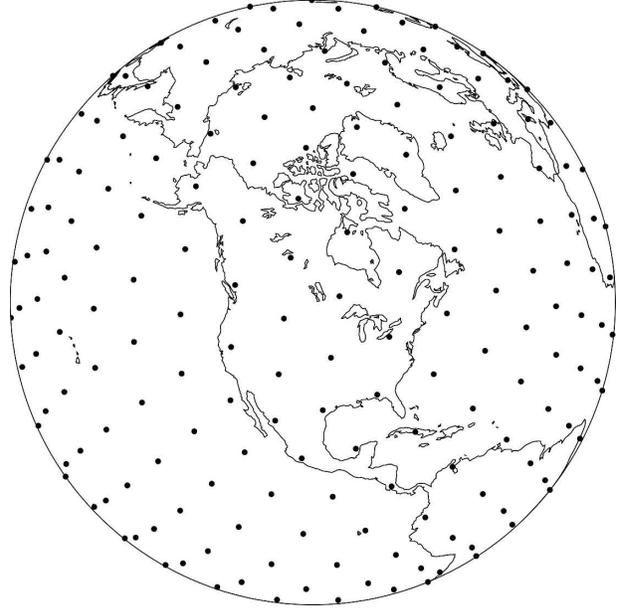

FIG. 2. Observation sites in the simulated observational network (dots). There are 362 sites nearly uniformly distributed on a spherical geodesic grid.

Newtonian relaxation to a prescribed interface Exner function. A fourth-order Runge–Kutta scheme is used for numerical integration, and ∇^8 hyperdiffusion is used. The parameters chosen are the same as in Hamill and Whitaker (2005). The model was run at T31 resolution. The error-doubling time of the model at T31 is 3.78 days. The perfect-model assumption was made in the following experiment. The model was shown by previous studies to realistically represent the nonlinear baroclinic wave cycles (Zou et al. 1993). For more detailed characteristics of the model including the kinetic energy power spectra and error growth characteristics, see Hamill and Whitaker (2005).

Observations of interface π and surface π were taken at a set of nearly equally spaced locations on a spherical geodesic grid (Fig. 2). The 362 observations of each consisted of the T31 true state plus errors drawn from a distribution with zero mean and standard deviation of 8.75 J kg^{-1} K^{-1} for interface π and 0.875 J kg^{-1} K^{-1} for surface π, respectively. The numbers chosen were about one quarter of the globally averaged climatological spread of the nature runs. Observation errors were constructed to be independent spatially and temporally, and observations were assimilated every 24 h.

Ensembles of size 50 to 5 were examined. The ensemble was initialized with random draws from the model climatology. The data assimilation was conducted for 250-day period, and the statistics were accumulated over the last 150 days.

b. Formulation of the static background error covariance

In this experiment, we used a technique for developing a static background error covariance model similar to those described by Evensen (2003), Zhang and Anderson (2003), and Hamill and Whitaker (2005). The static background-error covariance was formed from a large inventory of historical forecast errors over many separate times.

To optimize the static background error covariance model, we performed the following steps: 1) a first estimate of the static background-error covariance matrix was constructed from the estimated covariance based on 250 samples of historical 24-h forecast errors. Because 250 samples were much smaller than the dimension of the model,[2] a Gaspari and Cohn (1999) horizontal localization was used to reduce the sampling error in this first estimate of the static background-error covariance. A few localization scales were tried, and we chose to apply the 15 000-km scale as it produced the smallest analysis errors.[3] 2) Then 7000 data assimilation cycles were run with the background-error covariance approximated by the first estimate of the static background error covariance given in step 1. Equations (3)–(5) were used with $\alpha = 1$, and at each assimilation time the first-guess forecast was the single 24-h control forecast valid at this time. From these 7000 cycles, we were able to collect 7000 samples of 24-h forecast errors. Note the number of forecast error samples, 7000, was greater than the number of variables of the model used in this study. 3) The static background error covariance matrix was constructed by calculating the covariance of this large inventory of the forecast error samples. No covariance localization was applied. 4) With the static background-error covariance model constructed in step 3, we ran another 7000 cycles of data assimilation. Steps 3 and 4 were then iterated until the error did not decrease significantly. Figure 3 shows the root-mean-square (rms) analysis errors measured with three different norms corresponding to the first estimate of the static covariance, and the first, second, and third iteratively constructed static covariances. Figure 4 shows the absolute value of the hourly surface Exner function tendency, an analogy to surface pressure tendency, averaged globally and over the 24-h forecast periods. After each iteration, the rms analysis errors were reduced (Fig. 3) and the analyses became more balanced (Fig. 4). After repeating steps 3 and 4 three times, the decrease in rms analysis error and imbalance was statistically insignificant. We thus used the static background-error covariance model at the last iteration, denoted as $\mathbf{B}_3$, for the rest of the experiments.

Compared with the classic static background-error covariance in OI and 3DVAR, the static background-error covariance matrix constructed from a large inventory of the forecast errors made far fewer assumptions on the form of the covariances. For example, it did not presuppose the error covariances were homogeneous or isotropic, as typically assumed in OI and 3DVAR. Because the covariances were estimated directly from model states, whatever mass-wind balance properties possessed by these model states were preserved in the covariances. It is thus expected that the static background-error covariance model constructed in this simulation experiment is actually better than background-error covariances formulated for operational OI and 3DVAR, where the true state can never be known.

In this paper, following EB04, an online-estimated rescaling factor was used to rescale the static covariance matrix so that the total variance of the rescaled covariance matrix more appropriately estimated the total forecast-error variance, which varied with time. To be specific, during each assimilation cycle we rescaled the static covariance models so that in the normalized observation space, the total variance of the rescaled covariance model was consistent with the total forecast error variance estimated by the innovations. Mathematically, at data assimilation time t_i, we sought a rescaling factor f_i that satisfied

$$\overline{\tilde{\mathbf{d}}^T\tilde{\mathbf{d}}}_{\text{prior } t_i} = \text{Tr}(\tilde{\mathbf{H}}f_i\mathbf{B}_3\tilde{\mathbf{H}}^T + \mathbf{I}), \qquad (21)$$

where $\overline{\tilde{\mathbf{d}}^T\tilde{\mathbf{d}}}_{\text{prior } t_i}$ was defined the same as in Eq. (16). The rescaled gain components at t_i were $\mathbf{H}\mathbf{B}_i\mathbf{H}^T = \mathbf{H}f_i\mathbf{B}_3\mathbf{H}^T$ and $\mathbf{B}_i\mathbf{H}^T = f_i\mathbf{B}_3\mathbf{H}^T$. For more detailed discussions, please refer to EB04. As shown in Eq. (21), the value of the rescaling factor f_i varies depending on the magnitude of the innovation vectors. In the following experiment running the OI data assimilation cycles, f_i varied from 0.5 to 1.4. Note that the rescaling strategies adopted made traces of the ensemble covariance and the rescaled static error covariance approximately

---

[2] The number of complex spectral coefficients predicted by the model is 3168.

[3] The first estimate of the static covariance model was also constructed iteratively. 1) The first 250 samples of the forecast errors were collected from the ensemble mean forecast errors of the EnSRF. 2) Different covariance localizations were tried on the sample covariance of the 250 samples and data assimilation cycles were conducted with these different trials. 3) The one that produced the smallest rms errors were picked and its 250 forecast errors samples were collected. Steps 2) and 3) were repeated until the rms analysis errors reached the lowest value and saturated. The final static covariance was the first estimate B0 in Fig. 3 and Fig. 4.

Analysis errors in kinetic energy

Analysis errors in $\Delta \pi 2$

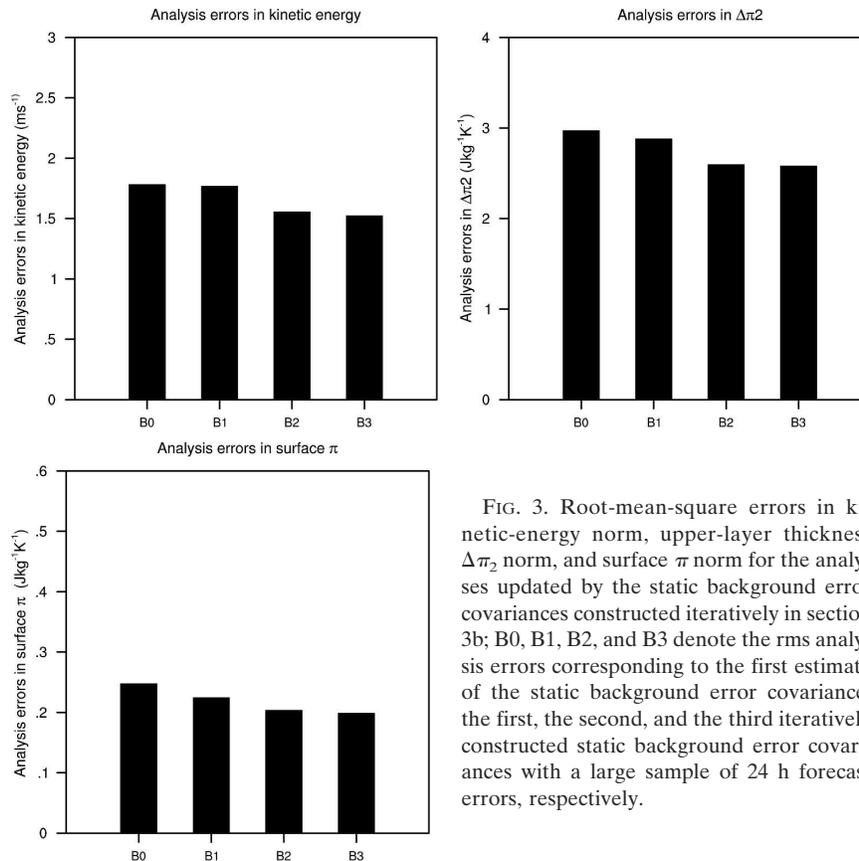Analysis errors in surface π

Fig. 3. Root-mean-square errors in kinetic-energy norm, upper-layer thickness $\Delta \pi_2$ norm, and surface π norm for the analyses updated by the static background error covariances constructed iteratively in section 3b; B0, B1, B2, and B3 denote the rms analysis errors corresponding to the first estimate of the static background error covariance, the first, the second, and the third iteratively constructed static background error covariances with a large sample of 24 h forecast errors, respectively.

equal. Thus, the coefficient α in Eqs. (1), (4), and (5) gave a clearer indication of the relative weights assigned on the error correlations given by the ensemble and the static covariance model.

4. Results

a. Analysis errors

In this section, we evaluate the characteristics of the ensemble-mean analysis errors for the hybrid and the EnSRF analysis schemes. Figure 5 shows the rms analysis error in the norms of kinetic energy, upper-layer thickness $\Delta \pi_2$, and the surface pressure π for the 50-member ensemble. Definitions of these norms are provided in Hamill and Whitaker (2005). The gray bars correspond to the rms analysis errors of the EnSRF with respect to different covariance localization length scales of 3000, 5000, 15 000, 25 000, and 35 000 km. The EnSRF exhibited filter divergence if no localization was applied. The black bars correspond to the results of the hybrid scheme as a function of the weighting coefficient α. Different covariance localization scales were tried on the ensemble covariance at each α for the hybrid
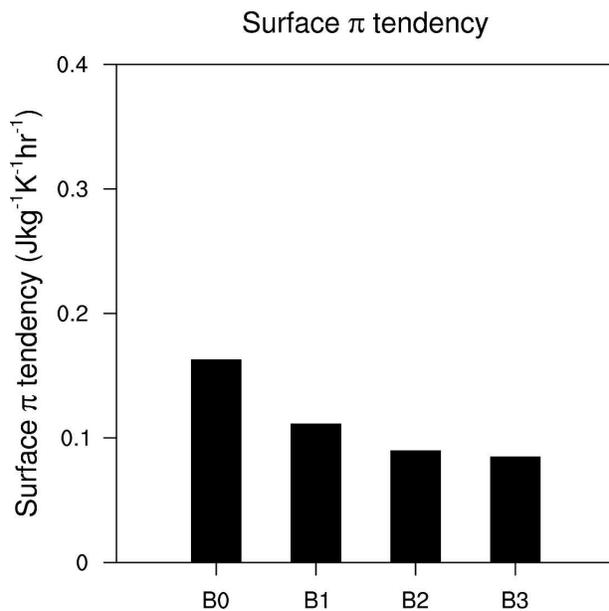
Surface π tendency

Fig. 4. Mean absolute surface π tendency (J kg^{-1} K^{-1} h^{-1}) averaged globally, over the subsequent twenty-three 1-h periods for forecasts initialized with the analyses updated by the static background error covariance constructed in section 3b.
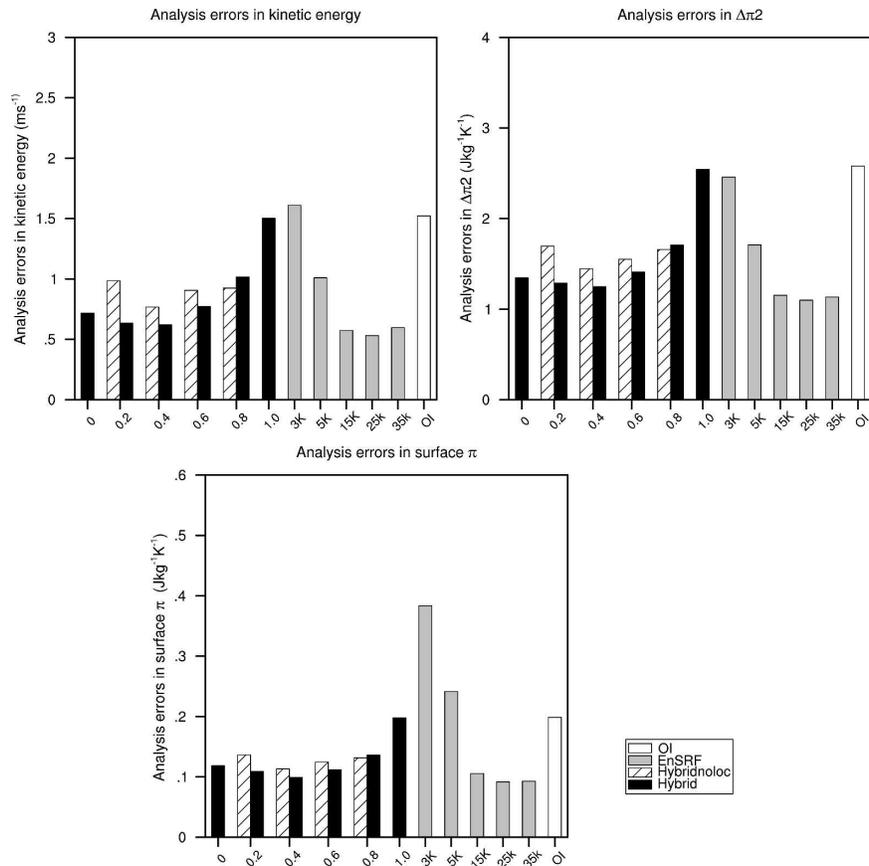
FIG. 5. Root-mean-square analysis error for kinetic-energy norm, upper-layer thickness $\Delta\pi_2$ norm, and surface π norm for 50-member ensembles. The black bars are results for the hybrid ETKF–OI scheme with localization. The hatched bars are results for the hybrid ETKF-OI with no localization. The gray bars are results for the EnSRF. The white bar is for the OI experiment.

scheme when updating the mean. Each black bar corresponds to the localization scale that produced the smallest rms errors at that α. Specifically, for $\alpha = 0.0$, 0.2, 0.4, 0.6, and 0.8, the optimal covariance localization scales for the black bars were 25 000, 25 000, 35 000, 45 000, and 45 000 km. The hatched bars at $\alpha = 0.2$, 0.4, 0.6, 0.8 are for the hybrid with no covariance localization applied on the ensemble when updating the mean. Filter divergence occurred for the hatched bar at $\alpha = 0.0$, and hence no hatched bar is shown at $\alpha = 0.0$. The white bar is the result for the experiments where control forecast was updated using the static background error covariance model. We denote in this subsection 4b the results for the white bar as OI. Note that the difference between the OI and the $\alpha = 1.0$ experiments is that the background forecast for the former was from the single control forecast whereas for the latter it was from the ETKF ensemble-mean forecast.

Consider first the characteristics of the 50-member

hybrid in Fig. 5 with localization (black bars) and without (hatched). The optimal α for the hybrid scheme was 0.4 for all three norms, both with and without covariance localization. When the background error covariance was purely from the static covariance model; that is, $\alpha = 1.0$, the analysis error was less than the error of the OI control simulation, which presumably was because the background forecast from the ensemble mean was more accurate than the control run; a similar result was shown in the perturbed-observation experiments in Hamill et al. (2000). The covariance localization applied when updating the mean improved the analysis of the hybrid scheme when $\alpha < 0.8$.

Now consider the EnSRF and the hybrid's relative performance. The localization scale that produced the smallest rms error for the EnSRF in this experiment was 25 000 km. Depending on the norm, the best performance of the hybrid scheme was slightly worse than or similar to that of the EnSRF. Typically, more than
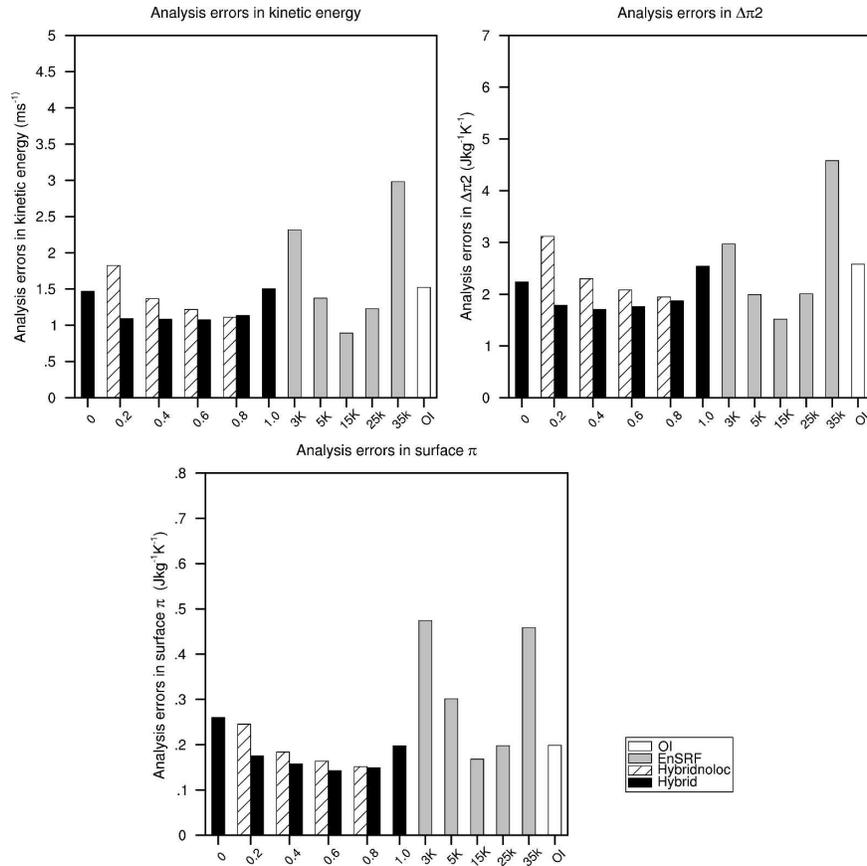
FIG. 6. As in Fig. 5 but for the 20-member ensembles.

90% of the improvement of the EnSRF over the OI was achieved by the hybrid scheme. Less severe localization was needed in the hybrid than in the EnSRF to achieve the best performance of the analysis. This was because the hybridization with the static covariance model already ameliorated the sampling error of the ensemble covariance. The hybridized static covariance itself (with no help from covariance localization on the ensemble covariance) can avoid filter divergence for the hybrid scheme.

Now consider results from the smaller 20-member ensemble (Fig. 6). Compared with the 50-member results in Fig. 5, both the hybrid and the EnSRF were less accurate. The best-performing hybrid and the best-performing EnSRF were still more accurate than the OI. The best-performing hybrid still achieved a large portion of the improved accuracy of the best-performing EnSRF over OI; 71% and 82% of the improvement between OI and the EnSRF was achieved by the hybrid for the kinetic energy and for the second-layer thickness norms, respectively. In the surface π norm, the best-performing hybrid was better than the best-performing EnSRF. The optimal localization scale

for the EnSRF was reduced to 15 000 km compared with 25 000 km for the 50-member runs. Compared with the 50-member runs, a tighter localization was also used for the hybrid scheme. The covariance localization scales for the localized hybrid were 15 000, 15 000, 15 000, 35 000, and 45 000 km for $\alpha = 0.0, 0.2, 0.4, 0.6,$ and 0.8.

When the ensemble size was further reduced to 10 members, the EnSRF experienced filter divergence for localization scales greater than 5000 km (not shown). The analyses for the EnSRF at 5000-km localization were better than at 3000 km, but they were worse than the optimal hybrid scheme. For the surface π norm, the EnSRF was even worse than the OI. The hybrid still performed better than the OI.

Figure 7 shows the results when the ensemble size was further reduced to five members. The EnSRF experienced filter divergence for all localization scales tried. However, the hybrid scheme with localization still performed slightly better than the OI (Fig. 7). In Fig. 7, the localizations were 15 000, 15 000, 25 000, and 35 000 km for $\alpha = 0.2, 0.4, 0.6,$ and 0.8 respectively. The hybrid with no covariance localization did no better than at
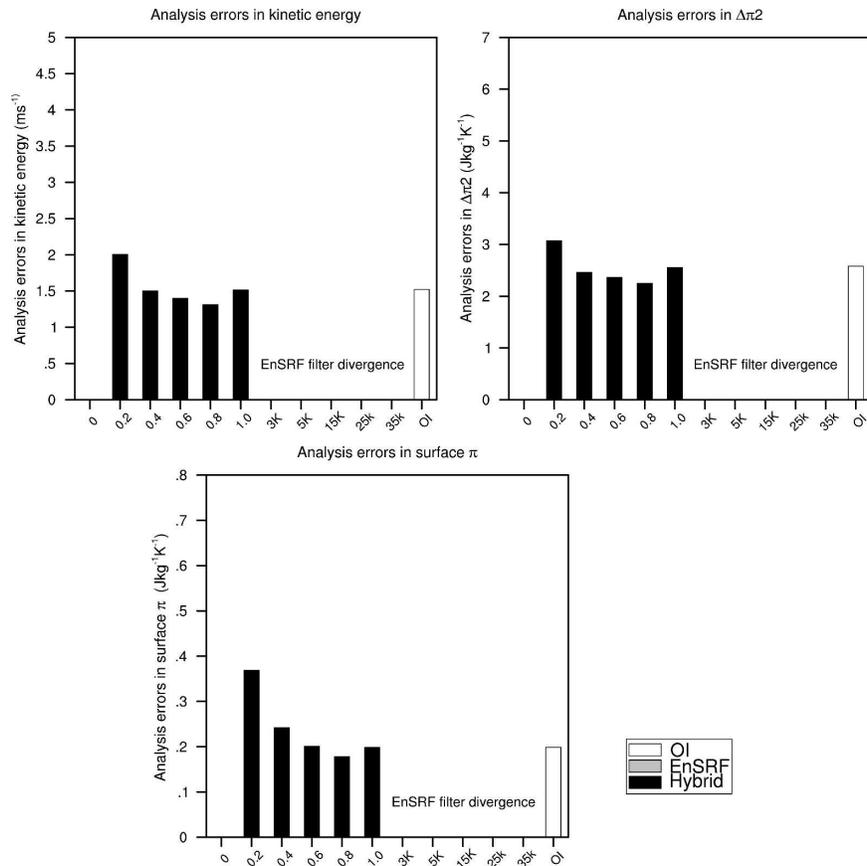
FIG. 7. As in Fig. 5 but for the 5-member ensembles.

$\alpha = 1.0$ (not shown). The analyses at $\alpha = 1.0$ were still slightly more accurate than the OI due to the use of the ensemble mean forecast instead of the single control forecast as the background forecast.

The results for 5- and 10-member ensembles indicate that the hybrid scheme was more robust than the EnSRF when the ensemble size was small. Note also that for 5-member ensembles, the hybrid did not experience filter divergence for $\alpha \geq 0.2$, even with no covariance localization applied. This indicates that the static background-error covariance stabilized the filter. Figures 5–7 also show that with decreasing ensemble size, the improvement of the hybrid and EnSRF over OI became smaller, as expected due to the decreasing quality of the ensemble-based covariances.

As discussed above, the optimal localization scales of the EnSRF were shorter for smaller ensembles. The optimal localization scale of the EnSRF was larger than what may have been expected by other studies with the same ensemble size (e.g., Hamill et al. 2001; Houteka-mer et al. 2005). The possible reasons are (i) the model resolution in this experiment was coarse, T31, and the covariance localization scale for a coarse-resolution

model may be larger than that for a fine-resolution model that can resolve small-scale features; (ii) the best localization scale for the EnSRF is longer than that for the EnKF, because the EnKF introduces noise through the perturbation of observations (Whitaker and Hamill 2002); and (iii) the two-layer primitive equation model in this experiment did not include an initialization step, to efficiently damp the gravity wave noise generated by the imbalance due to a short localization scale, and the explicit fourth-order Runge–Kutta time integration scheme also failed to damp these fast-propagating wave motions. Thus an imbalanced background inherited from the previous assimilation cycle together with the imbalance introduced by the EnSRF assimilation may have accumulated during the assimilation cycles, more acutely so with a more severe localization. Perhaps this aspect could have been better controlled with a different time integration scheme or an explicit initialization.

b. Similarity of hybrid and EnSRF covariance models

To compare the background-error covariance models from the hybrid and the EnSRF, we performed single-

observation experiments and plotted the analysis increment associated after assimilating the observation. For illustration, we conducted the experiment based on the background ensemble at the 100th data assimilation cycle from the experiment above. Figure 8 shows the results for the 50-member hybrid scheme with the best coefficient $\alpha = 0.4$, and a covariance localization scale of 35 000 km. Figure 9 shows the results for the 50-member EnSRF with its best covariance localization scale of 25 000 km. The contours are the background ensemble mean of upper-layer zonal wind, meridional wind, and thickness, at the 100th cycle. The colors are the analysis increment from the assimilation of one $\Delta \pi_2$ observation with the value of 3 J kg^{-1} K^{-1} smaller than the mean background at 47°N, 108°W. Both increments from the hybrid and the EnSRF were aligned along the flow pattern, and the wind increment appeared to be dynamically consistent with the layer thickness increment, for example, a cyclonic wind was associated with a decrease of layer thickness around the observation site. The patterns of the increments from the hybrid and the EnSRF were similar. Similar plots (not shown) at the 200th data assimilation cycle showed the pattern of the analysis increments was oriented along the background flow of the 200th day. Hence, both background error covariance models from the hybrid and EnSRF were flow dependent and similar to each other. As expected, the increment when the background error covariance was estimated purely from the static background error covariance did not align with the flow pattern (not shown), which is consistent with previous studies (e.g., HS00; Parrish and Derber 1992). The OI increment of the second-layer thickness around the observation at 47°N, 108°W, was elliptical. Unlike the increments of the hybrid (Fig. 8c) and the EnSRF (Fig. 9c), the OI increment was not parallel to the southwest–northeast oriented jet; that is, the "flow of the day," but rather elongated in the west–east direction. The OI wind increments, although dynamically consistent with the layer thickness increment, were not oriented along the flow of the day either.

c. Comparison of maximal perturbation growth rates

A desirable property of an ensemble of initial conditions is an appropriately rapid growth of the subsequent forecast perturbation. One useful metric for this is the fastest growth within the ensemble subspace during a short lead time (Palmer et al. 1998). This can be estimated by the following method. First, we assumed for a short-term forecast that the dynamic operator **M** was linear, that is,

$$\mathbf{X}^b = \mathbf{M}\mathbf{X}^a. \tag{22}$$

We then identified the direction in the initial ensemble perturbation subspace where the subsequent amplification was maximized in a chosen norm. In other words, we found the vector of linear combination coefficients **b** to apply to the initial perturbation such that

$$\max \left[\frac{\mathbf{b}^{\mathrm{T}}(\mathbf{X}^b)^{\mathrm{T}}\mathbf{S}\mathbf{X}^b\mathbf{b}}{\mathbf{b}^{\mathrm{T}}(\mathbf{X}^a)^{\mathrm{T}}\mathbf{S}\mathbf{X}^a\mathbf{b}} \right]. \tag{23}$$

Here **S** defines a particular norm. In this case we chose a global kinetic-energy norm and **S** was a diagonal matrix. As shown in Bishop and Toth (1999), **b** actually was the leading eigenvector of $\boldsymbol{\beta}^{-1/2}\mathbf{D}^{\mathrm{T}}(\mathbf{X}^b)\mathbf{S}\mathbf{X}^b\mathbf{D}\boldsymbol{\beta}^{-1/2}$ where $\boldsymbol{\beta}$ and **D** were the eigenvalue and eigenvector matrices of $(\mathbf{X}^a)^{\mathrm{T}}\mathbf{S}\mathbf{X}^a$, and the corresponding eigenvalue was the maximal growth.

Figure 10 shows the averaged 24-h maximum growth in the global kinetic-energy norm within the ETKF and EnSRF ensemble subspaces for the 50-member, 20-member, and 10-member runs.[4] Each bar corresponds to the covariance localization scale and weighting coefficient α that produced the smallest rms analysis errors. The maximum growth in the ETKF ensemble subspace was larger than that in the EnSRF ensemble subspace. We speculate the slower maximal growth of the EnSRF may be due to the imbalanced perturbations induced by the covariance localization (Mitchell et al. 2002; Lorenc 2003; HM05). Note that no covariance localization was applied when the ETKF updated the ensemble perturbations (section 2a).

d. Initial-condition balance

In an operational data assimilation system, an explicit or implicit initialization is typically utilized, otherwise imbalances between the mass, momentum, and diabatic heating in the analysis can produce large-amplitude gravity waves. However, initialization requires extra computational cost. Hence a data assimilation technique that itself can produce balanced initial condition is desirable.

The mean absolute tendency of surface pressure (Lynch and Huang 1992) is a useful diagnostic of the amount of imbalance for an analysis generated by a data assimilation scheme. Surface π is the quantity analogous to the surface pressure in this two-layer primitive equation model. To examine surface π tendencies, we reran forecasts from the analysis ensembles to 24-h lead, producing output every hour. We then

---

[4] For the 5-member ensemble, the EnSRF experienced filter divergence, so no calculations of Eq. (23) were performed.
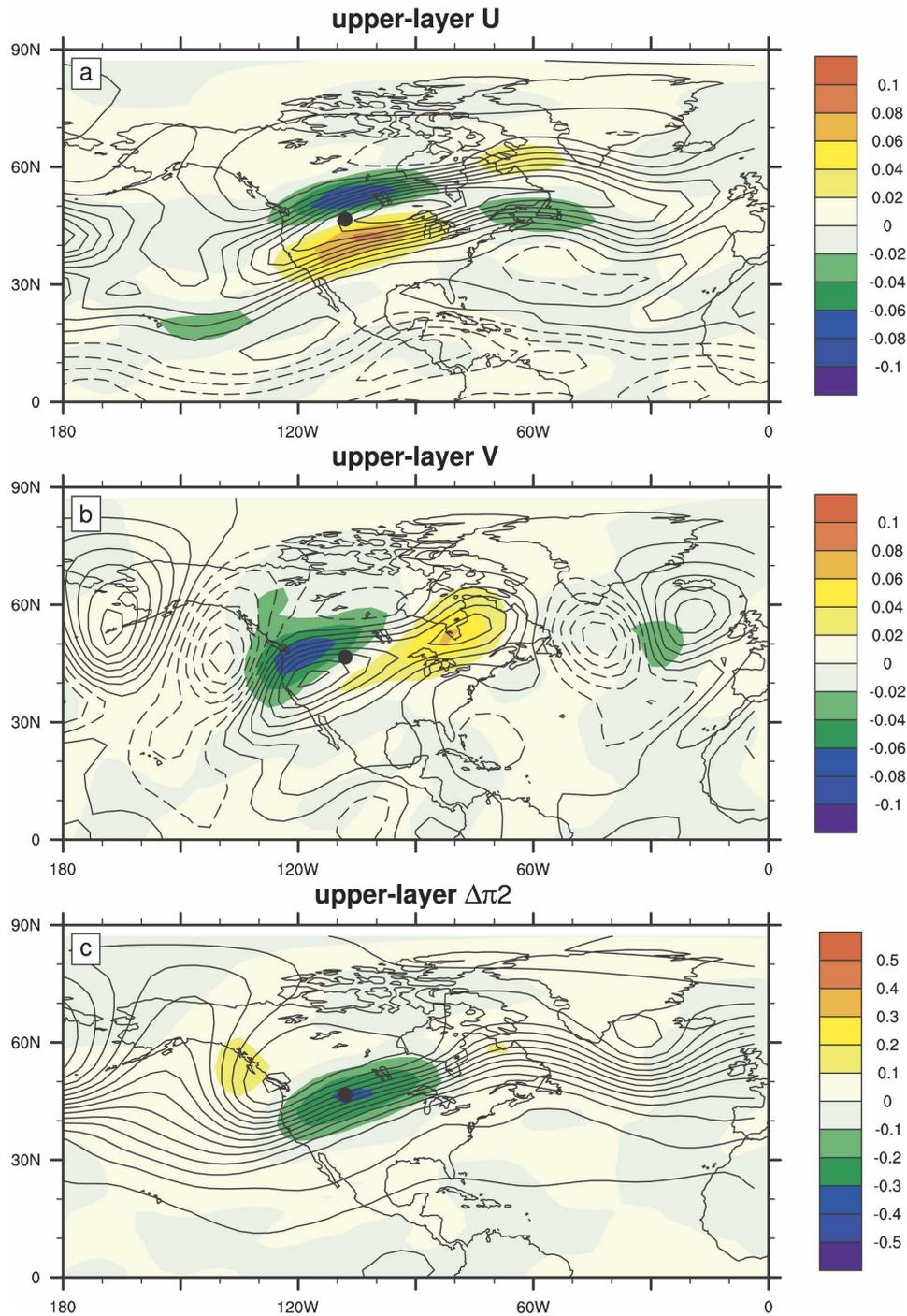
Hybrid 35000km localization α=0.4 analysis increment

upper-layer U

upper-layer V

upper-layer $\Delta\pi2$

FIG. 8. A snapshot (at the 100th analysis cycle) of the ensemble mean upper-layer wind (U, V) and thickness ($\Delta\pi_2$) increments for a single -3 J kg^{-1} K^{-1} $\Delta\pi_2$ observation increment for the hybrid scheme with $\alpha = 0.4$ and a 35 000-km localization scale. The black dot is the observation location. The contours and color shades are the background ensemble mean and the analysis increment at the 100th cycle for (a) upper-layer U wind (m s^{-1}), (b) upper-layer V wind (m s^{-1}), and (c) upper-layer thickness $\Delta\pi_2$ (J kg^{-1} K^{-1}). The contour intervals for the background mean are (a) from -20 to 45 m s^{-1} with a 5 m s^{-1} interval, (b) from -24 to 28 m s^{-1} with a 4 m s^{-1} interval, and (c) from 75 to 450 J kg^{-1} K^{-1} with a 25 J kg^{-1} K^{-1} interval.

EnSRF 25000km localization analysis increment

upper-layer U

upper-layer V
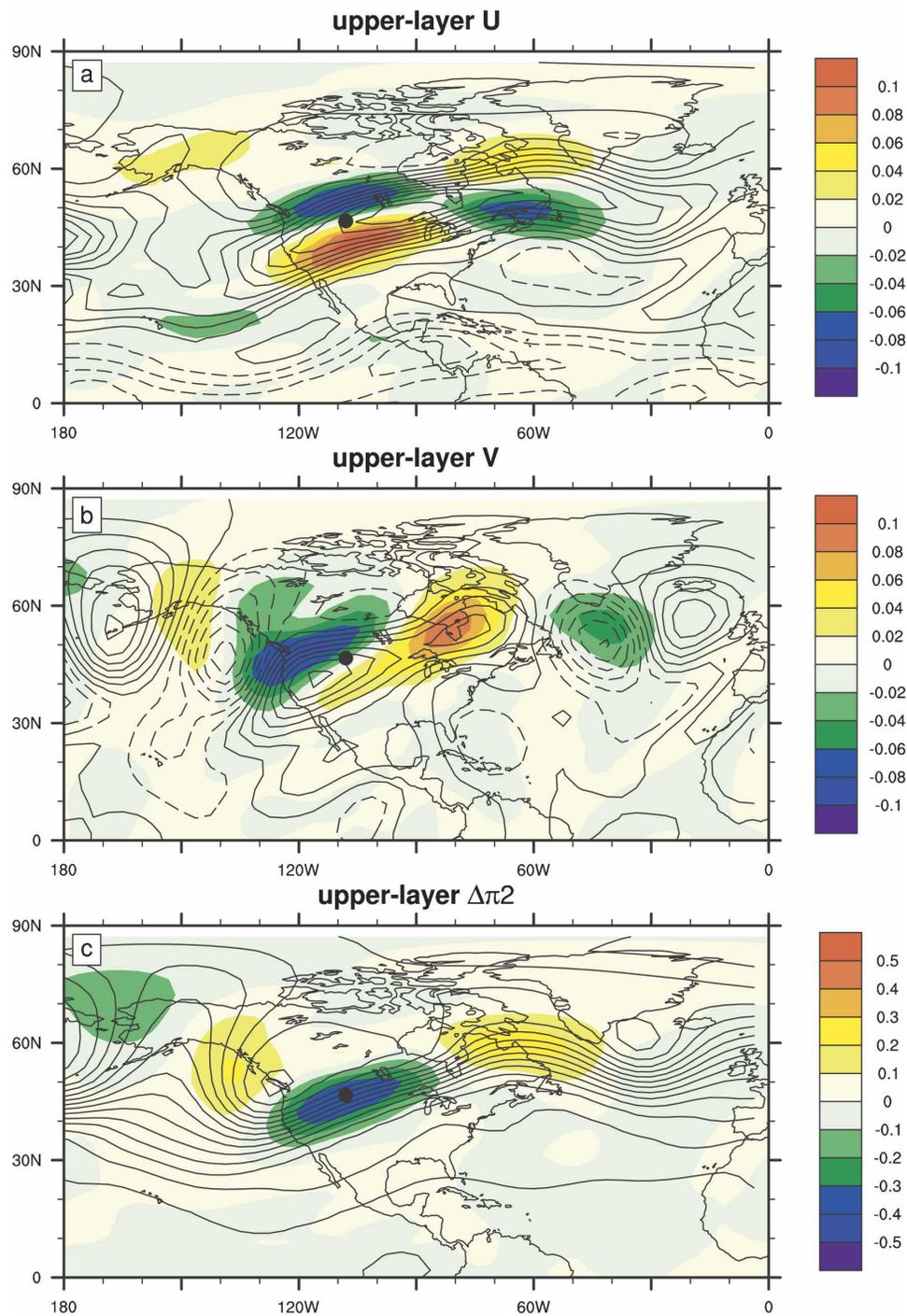

upper-layer $\Delta\pi 2$

FIG. 9. Same as Fig. 8, but for the EnSRF with localization scale of 25 000 km.

calculated the hourly surface Exner function tendencies. Figure 11 shows the mean absolute tendency of the hourly surface π averaged over global grids, all ensemble members, all times, and the 23-hourly tendency snapshots from the 50-member ensemble forecasts. As expected, the EnSRF tendencies were larger than the truth run tendency, and this discrepancy was greater when the localization length scale was shorter. For the
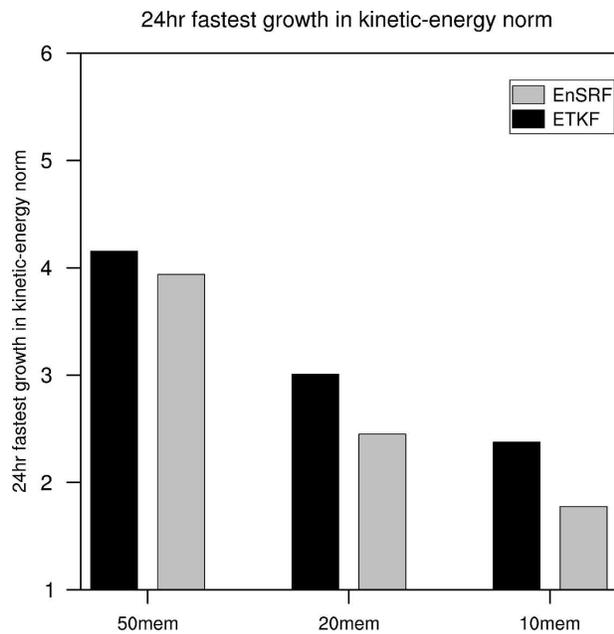
24hr fastest growth in kinetic-energy norm

FIG. 10. Averaged maximum 24-h perturbation growth in global kinetic-energy norm within the ensemble perturbation subspace of the ETKF (black bars), and the EnSRF (gray bars) for 50-member, 20-member, and 10-member ensembles. Each bar corresponds to the configuration, i.e., localization scale and α, that produced the most accurate analyses.

Surface π tendency

FIG. 11. Mean absolute surface π tendency (J kg^{-1} K^{-1} h^{-1}) averaged globally, over the subsequent twenty-three 1-h forecast periods and over all ensemble members for 50-member ensembles. Bars are defined the same as in Fig. 5, but the white bar is for the model nature run.

hybrid, covariance localizations applied when updating the mean degraded the balance. At small αs, where more severe localizations were applied, the analyses were more imbalanced. If no covariance localization was applied for the hybrid, the analyses were more balanced, with less weight placed on the static background error covariance. Note the static covariance constructed in this study only ensured climatological balance, which was not flow-dependent. For the 50-member ensembles, the optimal analyses of the hybrid were slightly more balanced than that of the EnSRF. For small ensemble size (not shown), it was apparent that the hybrid analysis was more balanced than the EnSRF. This result suggests that relative to the hybrid, the EnSRF may be more in need of an explicit initialization step or to use a time integration scheme with extra numerical damping of gravity waves.

e. Spread–skill relationships

To measure a spread–skill relationship, we used a method similar to that used in Wang and Bishop (2003) and Majumdar et al. (2002a). We first produced a scatterplot where the ordinate and abscissa of each point were the squared analysis error and the analysis ensemble variance for a particular variable of interest,
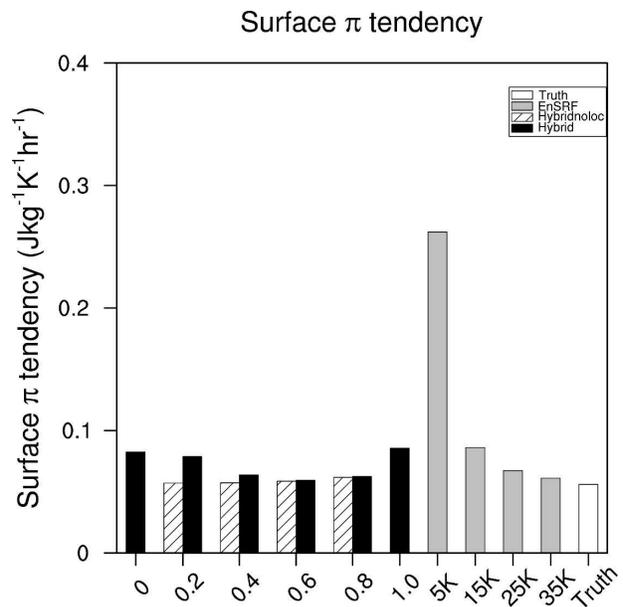
respectively. We collected samples over all grid points and all analysis cycles. We then divided these points into four equally populated bins, arranged in order of increasing ensemble variance. Next, we averaged the squared analysis error and analysis ensemble variance separately for each bin. We then plotted the square roots of the averaged squared analysis error against the square roots of the averaged analysis ensemble variance. The connecting curve described the relationship between the analysis ensemble spread and the rms analysis error. Figure 12 shows an example of such curves of surface π for the hybrid with $\alpha = 0.6$ and localization scale of 35 000 km, and the EnSRF with localization length scale equal to 15 000 km for the 20-member ensembles. Both correspond to the configuration that produced the smallest rms errors in Fig. 6. There were two aspects of the curve that we were interested to examine. First, after further averaging the values of the four points, we found that both schemes' ensemble spread was approximately equal to their rms analysis error, which means the ensemble spreads were reliable overall. Second, we examined whether spread was an accurate predictor of the error for the four individual subsets. Ideally, such spread–skill curve should follow the 45° reference line. In Fig. 12, the ensemble spreads of both schemes were suboptimal; they were negatively biased when the analysis error was small and positively biased when the analysis error was large. The
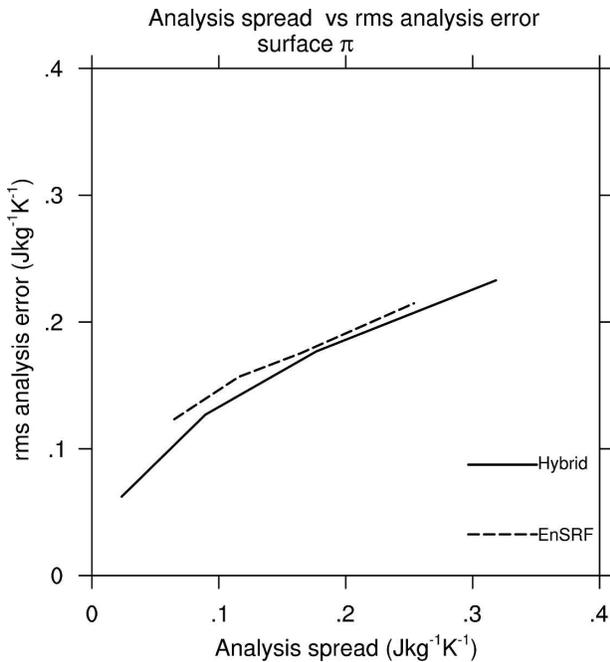
Analysis spread vs rms analysis error
surface π

FIG. 12. The relationship between the surface π analysis spread and rms analysis error for the 20-member hybrid (solid) and the 20-member EnSRF (dashed).

curves for the hybrid and the EnSRF were approximately parallel under the surface π norm (Fig. 12), the wind norm and layer thickness norm (not shown). This indicated that the ETKF and the EnSRF ensembles had similar skill in estimating the error variance. Note the range of the hybrid analysis error variance in Fig. 12 was larger than that of the EnSRF, which indicated a larger spatial and/or temporal variability in the magnitude of the hybrid analysis errors. The biases of the spread–skill curves were due to the deficiencies of the ensembles generated in each system. In the perfect model experiment conducted here, it was presumably because of the sampling error and the approximations made in estimating the error covariances.

5. Conclusions and discussion

In this paper, we compared the skill of the hybrid ETKF–OI and the EnSRF analysis schemes in an observing-system simulation experiment. A perfect model and simplified observation network were assumed, and a two-layer primitive equation model was used. Static background error covariances for the OI were constructed from a large inventory of historical forecast errors. The EnSRF analyses provided a reference for the expected analysis accuracy that may be obtained from a state-of-the-art ensemble-based data assimilation method. However, the EnSRF assimilation method

is expected to be computationally expensive in operational usage, as it requires parallel assimilation cycles for each member and costs scale linearly with the number of observations, which increase each year. In the hybrid scheme, the ensemble mean was updated utilizing a hybridized ETKF ensemble covariance and static covariance models. According to previous work (e.g., Lorenc 2003; Buehner 2005; Wang et al. 2007), this can be conveniently adapted to the existing operational variational codes, which handle large numbers of observations more efficiently. The ensemble perturbations in the hybrid were updated by the computationally efficient ETKF scheme.

Results from this paper demonstrated that for 50-member ensembles, the analyses of the hybrid scheme were almost as accurate as those from the EnSRF. A large fraction of the EnSRF's improvement over OI was achieved by the hybrid scheme. For 20-member runs, under certain norms, the analyses of the hybrid were more accurate than the EnSRF. Further reducing the ensemble size to five members, the EnSRF exhibited filter divergence whereas the analyses of the hybrid were still better than that of the OI analyses. The ability of the ETKF ensemble spread to predict the ensemble-mean analysis error was similar to the EnSRF. Further, the initial conditions of the hybrid were more balanced than those of the EnSRF, especially when the ensemble size was small. The maximum growth in the ETKF ensemble perturbation subspace was larger than that in the EnSRF ensemble perturbation space. In summary, the hybrid is nearly as accurate as the full EnSRF for large ensembles, it is more robust for smaller ensembles, and it may be more attractive operationally because of its reduced computational expense.

The hybrid can also be adapted into existing variational schemes relatively easily. Variational analysis schemes typically perform a preconditioning with respect to the background terms in order to speed up the convergence rate of minimization. As discussed by Lorenc (2003), Buehner (2005), and Wang et al. (2007), to extend this to a hybrid scheme, in the cost function one can add another background term with extended control variables preconditioned on the square root of the ensemble covariance. The resulting analysis increment is then equal to the weighted sum of the standard 3DVAR increment and that associated with the ensemble-related extended control variables. Experiments with the framework proposed by Lorenc (2003), Barker (1998), and D. Barker (2005, personal communication) suggested that only modest extra cost relative to the standard 3DVAR would be needed. Observation-space preconditioning, such as the Naval Research Laboratory Atmospheric Variational Data Assimila-

A

tion System (NAVDAS; Daley and Barker 2001), can also incorporate the ensemble covariance easily by linearly combining the ensemble covariance with the standard static covariance.

The idea of combining the ETKF ensemble covariance and the static background error covariance, may be extended to the 4DVAR framework also. The incorporation of the ensemble covariance may improve its initial background-error covariance estimate and thus improve the 4DVAR analysis. In this case, the ETKF transformation matrix is calculated with observations distributed in both space and time.

We are now comparing the EnSRF and the hybrid scheme within an imperfect-model environment where, in particular, we consider the model errors due to truncation. Preliminary results show that the analyses of the hybrid scheme are still approximately as accurate as those from the EnSRF. The hybrid and the EnSRF were still more accurate than the OI, and the improvement of both was less than in the perfect model experiment. In these experiments, the structure of the model error were known and relatively well represented in the ensemble. In operational forecasting, the structure and amplitude of the error that is because model error is largely unknown. In this situation the hybrid may be more compelling than the full ensemble filters because the static covariance can serve as a plausible catch-all for unknown model errors. EB04's simple model results supported this argument.

The simulated observational network in the current study (Fig. 2) is of course much simpler and more uniform than the real observing network. Discussions about the advantage of the hybrid over the 3DVAR with respect to different observational networks can be found in HS00. In general, they found less improvement of the hybrid over 3DVAR for data-rich and uniform networks than for data-poor and nonuniform networks. Comparison of the hybrid and the EnSRF with different configurations of the observational network may be studied in future work. In general, the best linear combination coefficient α for the hybrid and the best covariance localization scale of the EnSRF can be expected to vary depending on the ratio of the number of observations and the ensemble size.

The encouraging results of the hybrid in our simple system and the fact that the hybrid is straightforward to implement in an operational variational system strongly motivate further exploration of the hybrid scheme in operational settings.

APPENDIX A

Derivation of an Improved ETKF Formulation under the White Noise Assumption

Here, we identify gross bias in the ETKF formulation [Eq. (7)] given in Bishop et al. (2001) and Wang et al. (2004), which occurs when the ensemble size K is significantly less than the rank r of the true forecast error covariance matrix $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$. We shall refer to this version of the ETKF as the "old ETKF". We believe that the formulation given by Eq. (8) is superior to that given by Eq. (7) because (i) extensive 1D simple-model experiments showed that the analysis error variance estimate by Eq. (8) is closer to an optimal scheme and (ii) experiments in this paper also showed that formula Eq. (8) delivered superior performance to Eq. (7). Unfortunately, Eq. (8) was arrived at by a series of guesses, and not by a rigorous analysis of the state estimation equations and how they might be approximated. As such, the purposes of this appendix are modest. We merely wish to highlight the bias in Eq. (7) and indicate aspects of the train of thought that motivated the series of guesses that led to Eq. (8).

The sample background forecast-error covariance of a K-member ensemble in the normalized observation space $\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T$ is given by

$$\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T = \frac{\sum_{k=1}^{K}(\tilde{\mathbf{H}}\mathbf{x}_k'^b)(\tilde{\mathbf{H}}\mathbf{x}_k'^b)^T}{K-1}, \qquad (A1)$$

where $\tilde{\mathbf{H}}$ is the normalized observation operator; that is, $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}$, and $\mathbf{x}_k'^b$ is the deviation of the kth member from the ensemble mean. With the innovation-based inflation factor applied (Wang and Bishop 2003), the sample covariance in Eq. (A1) provides a reasonable estimate of the total forecast error variance on the entire normalized observation space; that is,

$$\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T) = \mathrm{Tr}(\mathbf{E}\mathbf{\Gamma}\mathbf{E}^T) = \mathrm{Tr}(\mathbf{\Gamma}) = \sum_{j=1}^{K-1}\gamma_j$$

$$\approx \mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T), \qquad (A2)$$

where $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_j, \ldots, \mathbf{e}_{K-1}]$ lists the $K-1$ orthonormal eigenvectors and the diagonal matrix $\mathbf{\Gamma} = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_j, \ldots, \gamma_{K-1}\}$ lists the corresponding nonzero eigenvalues of $\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T$; and $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$ is the true forecast error covariance in normalized observation space.

Note that the dimensions of $\mathbf{E}$, $\boldsymbol{\Gamma}$, $\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T$, $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$ are $p \times (K-1)$, $(K-1) \times (K-1)$, $p \times p$, and $p \times p$, respectively, where p is the number of observations.

When K is significantly less than the rank r of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$, the sample covariance [Eq. (A1)] will generally overestimate the forecast error variance within the vector space spanned by the ensemble perturbations. To see this, we consider the special case where the true forecast error covariance matrix in normalized observation space $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$ has r nonzero eigenvalues all of which have the same value λ; that is

$$\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T = \mathbf{V}\lambda\mathbf{I}\mathbf{V}^T \quad \text{and} \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}, \quad (\text{A3})$$

where $\mathbf{V}$ is a $p \times r$ matrix and $\mathbf{I}$ is the $r \times r$ identity matrix. Equation (A3) defines what we refer to as the white noise assumption. From Eq. (A3), $\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T) = r\lambda$ and thus from Eq. (A2)

$$\sum_{j=1}^{K-1} \gamma_j \approx r\lambda. \quad (\text{A4})$$

For simplicity, we will assume that all ensemble perturbations can be precisely represented as a linear combination of the true eigenvectors $\mathbf{V}$ (This is always possible if the ensemble perturbations are random draws from a distribution with mean equal to zero and covariance equal to the true error covariance). Equation (A4) shows that the sum of all of the r nonzero eigenvalues of the entire system is equal to the sum of the $K-1$ eigenvalues γ_j corresponding to the $K-1$ orthogonal eigenvectors of $\tilde{\mathbf{H}}\mathbf{P}^e\tilde{\mathbf{H}}^T$. Note that, in general, $r \gg K$. Note also that the true error variance in one of the sample covariance eigenvector directions $\mathbf{e}_j$ is given by

$$\mathbf{e}_j^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T\mathbf{e}_j = \lambda. \quad (\text{A5})$$

Further from Eq. (A4),

$$\lambda \approx \frac{\sum_{j=1}^{K-1} \gamma_j}{r} = \frac{(K-1)}{r}\overline{\gamma}, \quad (\text{A6})$$

where $\overline{\gamma} = [1/(K-1)] \Sigma_{j=1}^{K-1} \gamma_j$ gives the mean of the sample covariance eigenvalues. Comparison of Eqs. (A5) and (A6) shows that the eigenvalue γ_i overestimates the error variance in the direction associated with its eigenvector by the factor $r/(K-1)$ on average.[5]

Noting that the sample analysis error covariance by

_____

[5] This type of property was obtained and discussed in the context of both the background and observation error covariances by Kepert (2004).

the old ETKF (Bishop et al. 2001; Wang and Bishop 2003; Wang et al. 2004) is

$$\tilde{\mathbf{H}}\mathbf{S}^a\tilde{\mathbf{H}}^T = \mathbf{E}\boldsymbol{\Gamma}(\boldsymbol{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T = \sum_{i=1}^{K-1} \mathbf{e}_i \frac{\gamma_i}{\gamma_i + 1} \mathbf{e}_i^T. \quad (\text{A7})$$

It is evident from Eq. (A7) that the filtering properties of the ETKF error covariance update equation are sensitive to the eigenvalues γ_i. Thus, it seems appropriate to replace these eigenvalues by a set of eigenvalues that better reflect the error variance within the ensemble subspace. In other words, it suggests that we replace Eq. (A7) with

$$\tilde{\mathbf{H}}\mathbf{S}^a\tilde{\mathbf{H}}^T = a\mathbf{E}\rho\boldsymbol{\Gamma}(\rho\boldsymbol{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T = a \sum_{i=1}^{K-1} \mathbf{e}_i \frac{\rho\gamma_i}{\rho\gamma_i + 1} \mathbf{e}_i^T. \quad (\text{A8})$$

The factor ρ in Eq. (A8) gives the percentage of total variance that projects onto the ensemble subspace; that is,

$$\rho = \frac{\mathrm{Tr}(\mathbf{E}^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T\mathbf{E})}{\mathrm{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T)}. \quad (\text{A9})$$

This choice is motivated by the fact that Eq. (A9) gives

$$\rho = \frac{K-1}{r}, \quad (\text{A10})$$

which according to the previous discussion is precisely what is required under the white noise assumption. Please refer to appendix B to see how to estimate ρ from Eq. (A9).

The factor a in Eq. (A8) is motivated by the following arguments. While $\mathbf{E}\rho\boldsymbol{\Gamma}(\rho\boldsymbol{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T$ would improve the filtering properties for directions that lie within the ensemble subspace, it does not improve the filtering properties in directions that lie outside the ensemble subspace. We hypothesize that a coefficient a will be required in order to ensure relatively unbiased filtering properties in directions that do not lie entirely within the ensemble perturbation vector subspace. In other words, we search for an a such that the expected sample analysis error covariance will give a relatively unbiased estimate of the true analysis error covariance, $\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T$; that is,

$$\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T = \langle\tilde{\mathbf{H}}\mathbf{S}^a\tilde{\mathbf{H}}^T\rangle = a\left\langle \sum_{j=1}^{K-1} \mathbf{e}_j \frac{\rho\gamma_j}{\rho\gamma_j + 1} \mathbf{e}_j^T \right\rangle. \quad (\text{A11})$$

In Eq. (A11) the angle bracket on the far right side represents the average over an infinite number of independent calculations of $\mathbf{E}\rho\boldsymbol{\Gamma}(\rho\boldsymbol{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T$ obtained

from a corresponding infinite number of independent K-member ensembles.

Next we choose a so that it satisfies Eq. (A11) under the white noise assumption. Under this assumption, the true analysis error covariance is (Wang and Bishop 2003)

$$\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^{\mathrm{T}} = \mathbf{V}\lambda/(\lambda + 1)\mathbf{I}\mathbf{V}^{\mathrm{T}} = \sum_{i=1}^{r} \mathbf{v}_i \frac{\lambda}{\lambda + 1} \mathbf{v}_i^{\mathrm{T}}.$$

(A12)

Substituting Eq. (A12) into Eq. (A11), premultiplying by $\mathbf{v}_i^{\mathrm{T}}$ and postmultiplying by $\mathbf{v}_i$ one obtains

$$\frac{\lambda}{\lambda + 1} = a\left\langle \mathbf{v}_i^{\mathrm{T}} \left(\sum_{j=1}^{K-1} \mathbf{e}_j \frac{\rho\gamma_j}{\rho\gamma_j + 1} \mathbf{e}_j^{\mathrm{T}} \right) \mathbf{v}_i \right\rangle$$

$$= a\left\langle \sum_{j=1}^{K-1} \left(\frac{\rho\gamma_j}{\rho\gamma_j + 1} \right) \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \right\rangle. \quad \text{(A13)}$$

Note that the term $\mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i$ represents the square of the projection of the ith true eigenvector on the jth sample covariance eigenvector. Under the white noise approximation, the error variance associated with $\mathbf{v}_i$ is the same for all i. Hence, if the ensemble perturbations represented a random draw from the true forecast error covariance matrix, the eigenvalues γ_j of the sample covariance matrix would be entirely unrelated to and uncorrelated with $\mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i$. With this assumption Eq. (A13) can be rewritten in the form,

$$\frac{\lambda}{\lambda + 1} = a \sum_{j=1}^{K-1} \left\langle \left(\frac{\rho\gamma_j}{\rho\gamma_j + 1} \right) \right\rangle \langle \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \rangle. \quad \text{(A14)}$$

Also, the direction $\mathbf{e}_m$ is statistically interchangeable with the direction $\mathbf{e}_n$ under the white noise assumption. Hence, $\langle \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \rangle$ and $\langle (\rho\gamma_j/\rho\gamma_j + 1) \rangle$ have the same value for all j. Furthermore, note that we could augment the $K - 1$ orthonormal basis $\mathbf{E}$ using (for example) Gramm–Schmidt orthogonalization to obtain an expanded r-dimensional subspace $\mathbf{E}^E$ that precisely spanned the same space of the true eigenvectors $\mathbf{V}$. Hence,

$$\langle \mathbf{v}_i^{\mathrm{T}}\mathbf{E}^E(\mathbf{E}^E)^{\mathrm{T}}\mathbf{v}_i \rangle = 1 = \sum_{j=1}^{r} \langle \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \rangle = r\langle \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \rangle.$$

(A15)

Equation (A15) implies that

$$\langle \mathbf{v}_i^{\mathrm{T}}\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\mathbf{v}_i \rangle = \frac{1}{r}. \quad \text{(A16)}$$

Furthermore, under the white noise assumption, we assume that γ_j approximates $\bar{\gamma} = [1/(K-1)] \Sigma_{i=1}^{K-1} \gamma_i$.

Simple model experiments have shown that the accuracy of this approximation increases as the ratio (K/r) decreases. Thus from Eqs. (A6) and (A10), $\rho\gamma_j \approx \lambda$, which leads to

$$\frac{\lambda}{\lambda + 1} \approx \left\langle \left(\frac{\rho\gamma_j}{\rho\gamma_j + 1} \right) \right\rangle. \quad \text{(A17)}$$

Using Eqs. (A17) and (A16) in Eq. (A15) then gives

$$a \approx \frac{r}{K-1} = \frac{1}{\rho}. \quad \text{(A18)}$$

Note from Eq. (A16), $(1/a) \approx (K - 1)/r = \rho$ is the expected squared projection of a true eigenvector onto $K - 1$ ensemble-based eigenvectors under the white noise approximation. It is directly related to the expected angle subtended between ensemble eigenvectors and true eigenvectors. Using Eq. (A18) in Eq. (A11) suggests that if we define the sample analysis error covariance as

$$\tilde{\mathbf{H}}\mathbf{S}^a\tilde{\mathbf{H}}^{\mathrm{T}} = \mathbf{E}\mathbf{\Gamma}(\rho\mathbf{\Gamma} + \mathbf{I})^{-1}\mathbf{E} = \sum_{j=1}^{K-1} \mathbf{e}_j \frac{\gamma_j}{\rho\gamma_j + 1} \mathbf{e}_j^{\mathrm{T}},$$

(A19)

the expected value of it will provide a less-biased estimate of the true analysis error covariance $\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^{\mathrm{T}}$ than the old ETKF. From Eq. (10), the sample analysis-error covariance in state space that is consistent with Eq. (A19) is

$$\mathbf{S}^a = \mathbf{X}^b\mathbf{C}(\rho\mathbf{\Gamma} + \mathbf{I})^{-1}\mathbf{C}^{\mathrm{T}}(\mathbf{X}^b)^{\mathrm{T}}/(K - 1). \quad \text{(A20)}$$

Thus, the new transformation matrix given by Eq. (8) is recovered.

While much of the argument given above relies on the white noise approximation, simple model experiments have been performed and have shown that the superiority of the new ensemble update Eq. (8) over the old update Eq. (7) gradually diminishes as the white noise assumption is relaxed. As the true error distribution becomes more red, the true error variance of the forecast is better spanned by the ensemble subspace, the parameter ρ tends toward unity and the differences between the new and the old formulations diminish. In the limit of a true eigenvalue spectrum that contains fewer nonzero eigenvalues than the number of ensemble members, $\rho = 1$ and the differences between Eqs. (8) and (7) vanish.

APPENDIX B

Estimating the ρ Parameter in Eq. (A9)

To estimate the ρ factor in Eq. (A9) we need to estimate the total forecast error variance in normalized

observation space $\text{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T)$ and its projection onto the ensemble subspace $\text{Tr}(\mathbf{E}^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T\mathbf{E})$. The difference between observation and forecast vectors; that is, the innovations, enables both of these tasks.

First, note that because

$$\text{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I}) = \text{Tr}\langle[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b][\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T\rangle$$
$$= \langle[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]\rangle,$$
(B1)

it follows that the approximation

$$\text{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T) \approx [\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b] - p,$$
(B2)

where p is the number of observations, gives an unbiased estimate of the trace of the total forecast error variance in normalized observation space. Note the second equality holds in Eq. (B1) because both sides are equal to the sum of the variances of all elements in $(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b)$. To estimate the accuracy of the approximation [Eq. (B2)], recognize that

$$\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b = \sum_{i=1}^{p} \eta_i \mathbf{v}_i, \qquad \text{(B3)}$$

where $\mathbf{v}_i$ is the ith eigenvector of the normalized innovation covariance $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I}$ and η_i is a random variable with mean zero and variance equal to the ith eigenvalue σ_i of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I}$. Equation (B3) implies that

$$\frac{[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]}{\langle[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]\rangle} = \frac{\sum_{i=1}^{p} \eta_i^2}{\sum_{i=1}^{p} \sigma_i} = \frac{\sum_{i=1}^{p} \eta_i^2}{p\overline{\sigma}}.$$
(B4)

Note that if the eigenspectrum of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I}$ were flat (white) with $\sigma_i =$ constant for all i, and η_i is normally distributed, then the numerator of Eq. (B4) would be a χ^2 variable with p degrees of freedom (Ross 1998, p. 267). Hence, in this case, Eq. (B4) tends to unity as p tends to infinity. Tables of χ^2 statistics can be used to determine the reliability of estimates obtained through Eq. (B2).

In typical atmospheric and oceanographic applications, however, little is known about the eigenvalue spectrum of the innovation covariance matrix. In such cases, the only meaningful test of the accuracy of Eq. (B2) is to check the variance of a sample of squared innovations. Because Eq. (B2) is an unbiased estimate, small variance is indicative of estimation accuracy.

Wang and Bishop's (2003) experience in using Eq. (B2) to generate an inflation factor for an old-ETKF ensemble in a low resolution (T42) global circulation model suggests that, for the atmosphere, the innovation associated with the global rawinsonde network has enough degrees of freedom to make Eq. (B2) useful. If the observational network is sparse, assuming that the distributions from which $[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]$ are sampled have a degree of time invariance, then the approximation

$$\text{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T) \approx \overline{(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b)^T(\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b)} - p,$$
(B5)

can be used. The overbar in Eq. (B5) represents the average of $[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]$ computed over a period of time.

Having obtained a plausible estimate of $\text{Tr}(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T)$, one also needs to estimate $\text{Tr}(\mathbf{E}^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}\mathbf{E})$. Replacing $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I}$ by $\mathbf{E}^T(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I})\mathbf{E}$ in Eq. (B1) leads to the approximation

$$\text{Tr}(\mathbf{E}^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T\mathbf{E}) \approx [\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T\mathbf{E}\mathbf{E}^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]$$
$$- (K - 1), \qquad \text{(B6)}$$

and the $\mathbf{E}^T(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I})\mathbf{E}$ counterpart of Eq. (B4) is

$$\frac{[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T\mathbf{E}\mathbf{E}^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]}{\langle[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T\mathbf{E}\mathbf{E}^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]\rangle} = \frac{\sum_{i=1}^{K-1} \xi_i^2}{\sum_{i=1}^{K-1} \upsilon_i}$$
$$= \frac{\sum_{i=1}^{K-1} \xi_i^2}{(K-1)\overline{\upsilon}},$$
(B7)

where υ_i is the ith eigenvalue of the $(K-1) \times (K-1)$ normalized innovation covariance $\mathbf{E}^T(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I})\mathbf{E}$, ξ_i is a random variable with mean zero and variance υ_i. In Eq. (B7) $\overline{\upsilon}$ gives the mean of the eigenvalues.

From our previous discussion of Eqs. (B2) and (B4), large ensemble sizes K would generally be required in order to give Eq. (B6) a high level of accuracy. If only a small ensemble is available, similarly to Eq. (B5), the approximation

$$\text{Tr}(\mathbf{E}^T\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T\mathbf{E}) \approx \overline{[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]^T\mathbf{E}\mathbf{E}^T[\mathbf{R}^{-1/2}\mathbf{y} - \tilde{\mathbf{H}}\overline{\mathbf{x}}^b]}$$
$$- (K - 1) \qquad \text{(B8)}$$

can be used. Given Eqs. (B5) and (B8), the estimated ρ factor is given by Eq. (9). In the current experiment,

the overbars in Eqs. (B5) and (B8) are the average of the two weeks of data previous to the current time.

REFERENCES

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

Barker, D. M., 1998: Var scientific development paper 25: The use of synoptically-dependent error structures in 3DVAR. UK MET Office Tech. Rep., 2 pp.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.,* **56,** 1748–1765.

——, B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.,* **131,** 1013–1043.

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., D. M. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the effects of data selection with the DAO physical space statistical analysis system. *Mon. Wea. Rev.,* **126,** 2913–2926.

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

——, and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.,* **124,** 1783–1807.

Daley, R., 1985: Analysis of synoptic-scale divergence by a statistical interpolation procedure. *Mon. Wea. Rev.,* **113,** 1066–1079.

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

——, and E. Barker, 2001: NAVDAS: Formulation and diagnostics. *Mon. Wea. Rev.,* **129,** 869–883.

Dee, D. P., 1995: Online estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.,* **123,** 1128–1196.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.,* **132,** 1065–1080.

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

——, 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–367.

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

Gauthier, P. C., L. Cherette, L. Fillion, P. Koclas, and S. Laroche, 1998: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmos.–Ocean,* **37,** 103–156.

Hamill, T. M., 2006: Ensemble based atmospheric data assimilation. *Predictability of Weather and Climate,* R. Hagedorn and T. N. Palmer, Eds., Cambridge University Press, 124–156.

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

——, and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.,* **133,** 3132–3147.

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

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

Hanea, R. G., G. J. M. Velders, A. J. Segers, M. Verlaan, and A. W. Heemink, 2007: A hybrid Kalman filter algorithm for large-scale atmospheric chemistry data assimilation. *Mon. Wea. Rev.,* **135,** 140–151.

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

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

——, and ——, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.,* **131,** 3269–3289.

——, ——, G. Pellerin, M. Buehner, and M. Charron, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.,* **133,** 604–620.

Kepert, J. D., 2004: On ensemble representation of the observations-error covariance in the ensemble Kalman filter. *Ocean Dyn.,* **54,** 561–569.

Liu, H., M. Xue, J. R. Purser, and D. F. Parrish, 2005: Retrieval of moisture from GPS slant-path water vapor observations using 3DVAR with isotropic and anisotropic recursive filters. Preprints, *17th Conf. on Numerical Weather Prediction,* Washington, DC, Amer. Meteor. Soc., CD-ROM, 14B.6.

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

——, 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.,* **129,** 3183–3203.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.,* **120,** 1019–1034.

Majumdar, S. J., C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Quart. J. Roy. Meteor. Soc.,* **127,** 2803–2820.

——, ——, ——, and Z. Toth, 2002a: Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Mon. Wea. Rev.,* **130,** 1356–1369.

——, ——, R. Buizza, and R. Gelaro, 2002b: A comparison of ensemble transform Kalman filter targeting guidance with ECMWF and NRL total-energy singular vector guidance. *Quart. J. Roy. Meteor. Soc.,* **128,** 2527–2549.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 2791–2808.

Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus,* **56A,** 415–428.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.,* **55,** 633–653.

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

Purser, J. R., W. S. Wu, D. F. Parrish, and N. M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomoge-

neous and anisotropic general covariances. *Mon. Wea. Rev.,* **131,** 1536–1548.

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.

——, H. Jarvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

Riishøjgaard, L.-P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus,* **50A,** 42–57.

Ross, S., 1998: *A First Course in Probability.* Prentice Hall, 514 pp.

Schlatter, T. W., 1975: Some experiments with a multivariate statistical objective analysis scheme. *Mon. Wea. Rev.,* **103,** 246–257.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.,* **131,** 1663–1677.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. A. York, 2005: Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus,* **57A,** 528–545.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.,* **131,** 1485–1490.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.,* **133,** 1789–1807.

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.

Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.,* **60,** 1140–1158.

——, ——, and S. J. Julier, 2004: Which is better, an ensemble of positive–negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.,* **132,** 1590–1605.

——, C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. *Mon. Wea. Rev.,* **135,** 222–227.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913–1924.

Wu, W. S., J. R. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.,* **130,** 2905–2916.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.,* **132,** 1238–1253.

Zhang, S., and J. L. Anderson, 2003: Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model. *Tellus,* **55A,** 126–147.

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.