A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes

XUGUANG WANG

The Pennsylvania State University, University Park, Pennsylvania

CRAIG H. BISHOP

UCAR, Boulder, Colorado, and Naval Research Laboratory, Monterey, California

(Manuscript received 28 May 2002, in final form 1 November 2002)

ABSTRACT

The ensemble transform Kalman filter (ETKF) ensemble forecast scheme is introduced and compared with both a simple and a masked breeding scheme. Instead of directly multiplying each forecast perturbation with a constant or regional rescaling factor as in the simple form of breeding and the masked breeding schemes, the ETKF transforms forecast perturbations into analysis perturbations by multiplying by a transformation matrix. This matrix is chosen to ensure that the ensemble-based analysis error covariance matrix would be equal to the true analysis error covariance if the covariance matrix of the raw forecast perturbations were equal to the true forecast error covariance matrix and the data assimilation scheme were optimal. For small ensembles ($\sim$100), the computational expense of the ETKF ensemble generation is only slightly greater than that of the masked breeding scheme.

Version 3 of the Community Climate Model (CCM3) developed at National Center for Atmospheric Research (NCAR) is used to test and compare these ensemble generation schemes. The NCEP–NCAR reanalysis data for the boreal summer in 2000 are used for the initialization of the control forecast and the verifications of the ensemble forecasts. The ETKF and masked breeding ensemble variances at the analysis time show reasonable correspondences between variance and observational density. Examination of eigenvalue spectra of ensemble covariance matrices demonstrates that while the ETKF maintains comparable amounts of variance in all orthogonal and uncorrelated directions spanning its ensemble perturbation subspace, both breeding techniques maintain variance in few directions. The growth of the linear combination of ensemble perturbations that maximizes energy growth is computed for each of the ensemble subspaces. The ETKF maximal amplification is found to significantly exceed that of the breeding techniques. The ETKF ensemble mean has lower root-mean-square errors than the mean of the breeding ensemble. New methods to measure the precision of the ensemble-estimated forecast error variance are presented. All of the methods indicate that the ETKF estimates of forecast error variance are considerably more accurate than those of the breeding techniques.

1. Introduction

Since ensemble forecasting was recognized as a practical way for providing probabilistic forecasts in the early 1970s (Leith 1974), ensemble generation schemes have been developed and used in weather prediction centers, for example, the breeding method (Toth and Kalnay 1993, 1997) used at the U.S. National Centers for Environmental Prediction (NCEP), the singular vector method (Buizza and Palmer 1995; Molteni et al. 1996) applied at the European Centre for Medium-Range Weather Forecasts (ECMWF), and the system simulation method (Houtekamer et al. 1996) used at the Canadian Meteorological Centre (CMC).

Previous comparisons of the skills of these ensemble generation schemes include that of Houtekamer and Derome (1995) who found that although these three ensemble generation schemes are fundamentally different, their medium-range ensemble mean error reduction skills are quite comparable given that all initial perturbations are centered around the control analysis. Hamill et al. (2000) further studied the ensemble spread skill of these three ensemble generation schemes. They found that the singular vector and breeding methods produce less skillful probabilistic forecasts than a variant of the system simulation method known as the perturbed observation method.[1] This is partly due to a stronger spread–skill relationship in the perturbed observation ensemble.

Corresponding author address: Xuguang Wang, Department of Meteorology, The Pennsylvania State University, 503 Walker Building, University Park, PA 16802.
E-mail: xuguang@essc.psu.edu

[1] In the system simulation method uncertainties in both model physics and the initial condition error are accounted for.

Compared to a single control forecast, an ensemble forecast not only provides a more accurate estimate of the first moment (the mean) of the probability density function (PDF) of future atmospheric states, but also provides higher-order moment estimations such as the forecast error variance. Studies by Toth et al. (2001) argued that the major extra information provided by ensemble forecasts relative to a single control forecast is that ensemble forecasts can estimate the case to case variations in forecast uncertainty. Zhu et al. (2002) also attributes much of the economic value of an ensemble forecast to its ability to forecast regional and temporal variations in forecast error variance. Thus, measuring error variance prediction accuracy is an important criterion in evaluating an ensemble forecast system.

The basic idea behind an ensemble generation scheme is to generate independently perturbed initial conditions such that the covariance of the ensemble perturbations approximates the analysis error covariance matrix at the initial time and thus the forecast error covariance matrix at the forecast time. For an optimal data assimilation scheme, the analysis error covariance matrix $\mathbf{P}^a$ and the forecast error covariance $\mathbf{P}^f$ are related by

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f, \quad (1)$$

where the matrix $\mathbf{H}$ is the linear observation operator that maps model variables to observed variables and the matrix $\mathbf{R}$ is the observation error covariance matrix. The notational convention in this paper is roughly consistent with Ide et al. (1997). There are two major characteristics in Eq. (1). First, the error variance reduction varies with geographical variations of observation density and accuracy. Second, error variance in directions with large forecast error variance is reduced by a larger factor than error variance in directions with small forecast error variance is reduced (see appendix and/or Daley 1991); in other words, data assimilation schemes use observations to filter out the most uncertain or ''noisy'' aspects of the forecast first-guess field. To provide accurate estimate of $\mathbf{P}^a$ and $\mathbf{P}^f$, an ensemble generation scheme should be able to reflect these two characteristics.

Of all ensemble generation schemes, the breeding method is the most computationally inexpensive. A hypothesis of the breeding method is that the important part of analysis errors is the dynamically constrained part contributed by errors of the forecast background. In the simple form of the breeding method (Toth and Kalnay 1993), forecast perturbations are transformed into analysis perturbations by multiplying a globally constant factor whose magnitude is less than one so that the size of the scaled forecast perturbations is consistent with the empirically estimated global analysis uncertainty. Consequently, perturbation amplitude does not reflect geographical variations in the observational network density and accuracy nor does it reflect the fact that data assimilation schemes reduce error variance in directions corresponding to large error variance by a larger factor than error variance in directions corresponding to small error variance. To allow bred-perturbation amplitude to reflect geographical variations in the observational network, Toth and Kalnay (1997) introduced a regional rescaling method where perturbations at all levels are multiplied with a geographically dependent rescaling factor. We shall refer to this more sophisticated breeding scheme as ''masked breeding.'' While masking enables perturbation amplitude to reflect geographical variations in observational density and accuracy, it has a severely limited ability to make perturbation amplitudes reflect the fact that data assimilation schemes attenuate error variance in directions corresponding to large forecast error variance more than in directions corresponding to small forecast error variance.

The ensemble transform Kalman filter (ETKF) theory was first introduced as an adaptive sampling method (Bishop et al. 2001). It produces initial perturbations consistent with the error covariance update equation (1) within the vector subspace of ensemble perturbations. Thus, one would expect the aforementioned two problems from the breeding scheme would be ameliorated (see section 2c in detail). Different from other ensemble-based Kalman filter (EnKF) theories that have been explored for data assimilation (Houtekamer and Mitchell 1998, 2001; Hamill and Snyder 2000; Hamill et al. 2001; Whitaker and Hamill 2002; Anderson 2001), the ETKF ensemble is used to estimate $\mathbf{P}^f$ only for predicting $\mathbf{P}^a$, not for updating the mean state. Thus, the control analysis used in the ETKF is potentially not as accurate as the control analysis created in the EnKF data assimilation schemes. However, the computational expense of the ETKF ensemble generation is considerably less than the EnKF ensemble.

To gain insight into the differences between the breeding and ETKF ensemble generation schemes, it is helpful to first consider the types of perturbations each scheme would produce in an idealized system for a time-invariant linear dynamics, fixed observational network, fixed observation error statistics, and no model error. In such an system, forecast covariances at time t_k is related to analysis covariances at time t_{k-1} by

$$\mathbf{P}^f_k = \mathbf{M}\mathbf{P}^a_{k-1}\mathbf{M}^{\mathrm{T}}, \quad (2)$$

where $\mathbf{M}$ is the linear dynamics operator. According to Cohn and Dee (1988), in such a system the Kalman filter asymptotically approaches a steady state when the system is observable. Closed form solutions for the exact error covariances that would be produced by infinite time Kalman filter scheme for such a system are given in Bishop et al. (2003, hereafter BRT). It demonstrates that *all* nondecaying eigenvectors (or normal modes) of $\mathbf{M}$ are required to precisely describe the forecast and analysis error covariances in such a system with Kalman filter for an infinite time. With a fixed dynamics propagator, the simple breeding method is equivalent to the power method for finding the eigenvector corresponding

to the largest eigenvalue of **M**. Consequently, perturbations produced by the simple breeding method would all eventually point in the direction of the leading eigenvector (or normal mode) of **M**. Thus, for the simple breeding ensemble, the rank of the ensemble-based sample covariance matrix would be limited to 1.[2] From Cohn and Dee (1988) and BRT, in such an idealized system, error covariances started from *any* initial (at t_0) estimate, updated by Eq. (1), and propagated by Eq. (2), will eventually (at t_∞) converge toward the error covariances for an infinite time optimal scheme if the initial estimate covers all amplifying normal modes that the true initial analysis/forecast error covariance would cover. Since the ETKF perturbations satisfy (1) and (2) in such an idealized system, they would eventually (at t_∞) provide a precise description of the forecast and analysis error covariances that an infinite time optimal data assimilation scheme would have if the ETKF initial ensemble perturbations span the vector space of unstable normal modes of the dynamics operator.[3] Thus, the ETKF ensemble would be expected to markedly outperform the breeding ensembles in estimating the forecast/analysis error covariances in such an idealized system. A comparison of the performance of the ETKF, the simple breeding, and the masked breeding ensembles in an imperfect system where the dynamics propagator is imperfect and time-dependent and the number of ensemble members is less than the number of growing normal modes of the dynamics operator is the primary aim of this paper.

A description of the ETKF and breeding ensemble generation schemes is given in section 2. Section 3 provides a brief introduction of the model, the analysis data, the observational network, and the specific constructions of initial perturbations. Section 4 compares initial perturbations in terms of effective rescaling factors. Section 5 compares the dimension of the subspaces in which ensemble variance is maintained. Section 6 compares the maximal growth within ensemble perturbation subspaces under the total energy norm. Section 7 compares the ensemble forecast skills in terms of ensemble mean. In section 8, we introduce methods to measure the error variance prediction accuracy and perform comparisons between the ETKF and breeding schemes. The computational expense of the techniques is compared in section 9. Section 10 summarizes our results.

2. Ensemble generation methods

a. Simple breeding

Following Toth and Kalnay (1993), a global constant factor is applied to each raw forecast perturbation so that the scaled perturbation has the same size as the empirically determined global analysis root-mean-square (rms) error. Mathematically,

$$\mathbf{x}^a = \mathbf{x}^f \cdot c, \tag{3}$$

where vectors $\mathbf{x}^a$ and $\mathbf{x}^f$ represent an analysis perturbation and a forecast perturbation. Scalar c is a constant globally for each raw perturbation. In our experiment, the value of c for each forecast perturbation is chosen so that 12-h ensemble forecast variance is consistent with the 12-h control forecast error variance on a globally averaged basis. Details of the method for doing this are given in section 2c and section 3d.

b. Masked breeding

The global constant factor c in Eq. (3) cannot reflect the geographically dependent analysis uncertainty. The masked breeding (Toth and Kalnay 1997) ameliorates this problem with a regional rescaling factor. Mathematically,

$$x^a_{ij} = x^f_{ij} c_{ij}, \tag{4}$$

where x^a_{ij} and x^f_{ij} are elements in $\mathbf{x}^a$ and $\mathbf{x}^f$ at latitude i and longitude j. The mask is a function of latitude and longitude whose values bound the vertically averaged and horizontally smoothed ensemble perturbation amplitude. The amplitude for each ensemble perturbation is measured under a user-specified norm that, in our experiment, is the square root of vertically averaged squared wind perturbations. Denote the mask as e_{ij} and the vertically averaged and horizontally smoothed forecast perturbation amplitude as y^f_{ij}. The rescaling factor is determined for each forecast perturbation as

$$c_{ij} = \begin{cases} 1 & y^f_{ij} \le e_{ij} \\ \dfrac{e_{ij}}{y^f_{ij}} & y^f_{ij} > e_{ij}. \end{cases} \tag{5}$$

The rescaling factor c_{ij} is only a function of latitude and longitude. It applies to all variables at all levels. In our experiments, an inflation factor is applied after (5) in order to ensure that 12-h ensemble forecast variance is consistent with the 12-h control forecast error variance on a globally averaged basis (again, see below for details).

c. ETKF

The ensemble transform Kalman filter is a suboptimal Kalman filter (Kalman 1960; Kalman and Bucy 1961; Daley 1991) with the forecast error covariance matrix estimated by the covariance matrix of the ensemble forecast perturbations. As mentioned in the introduction, compared to other ensemble-based Kalman filters, the $\mathbf{P}^f$ estimated by the ETKF is not for updating the mean state but only for estimating $\mathbf{P}^a$.

Different from the breeding method that transforms

[2] In cases where there were $n > 1$ eigenvectors corresponding to the same largest eigenvalue, the rank would be limited by n.

[3] In case not all amplifying normal modes of the linear dynamics operator have errors initially, the statement is true when the ETKF initial perturbations exclude these error-free amplifying normal modes.

forecast perturbations into analysis perturbations by multiplying each forecast perturbation by a constant or regional rescaling factor, the ETKF method transforms forecast perturbations into analysis perturbations by a transformation matrix $\mathbf{T}$, that is,

$$\mathbf{X}^a = \mathbf{X}^f\mathbf{T}, \tag{6}$$

where forecast perturbations are listed as columns in the matrix $\mathbf{X}^f$ and analysis perturbations are listed as columns in the matrix $\mathbf{X}^a$. In our experiment, since we use one-sided perturbations, ensemble mean at the initial time is degraded compared to the control analysis. So in our one-sided ensemble system, we define analysis and forecast perturbations as deviations from the control, that is, $\mathbf{X}^a = (\mathbf{x}_2^a - \mathbf{x}_1^a, \mathbf{x}_3^a - \mathbf{x}_1^a, \ldots, \mathbf{x}_K^a - \mathbf{x}_1^a)$ and $\mathbf{X}^f = (\mathbf{x}_2^f - \mathbf{x}_1^f, \mathbf{x}_3^f - \mathbf{x}_1^f, \ldots, \mathbf{x}_K^f - \mathbf{x}_1^f)$, where subscript 1 denotes the control and K is the number of ensemble members including the control. The transformation matrix $\mathbf{T}$ is chosen to solve Eq. (1) provided that

$$\mathbf{P}^f = \mathbf{Z}^f(\mathbf{Z}^f)^{\mathrm{T}} \quad \text{and} \tag{7}$$

$$\mathbf{P}^a = \mathbf{Z}^f\mathbf{T}\mathbf{T}^{\mathrm{T}}(\mathbf{Z}^f)^{\mathrm{T}}, \tag{8}$$

where $\mathbf{Z}^f = \mathbf{X}^f/\sqrt{K-1}$. Thus, if the covariance matrix of the raw forecast perturbations $\mathbf{X}^f$ were equal to the true forecast error covariance matrix $\mathbf{P}^f$, and the data assimilation scheme were optimal, then the covariance matrix associated with the transformed perturbations $\mathbf{X}^a$ would be precisely equal to the true analysis error covariance matrix.

Following Bishop et al. (2001), $\mathbf{T}$ is given by

$$\mathbf{T} = \mathbf{C}(\mathbf{\Gamma} + \mathbf{I})^{-1/2}, \tag{9}$$

where columns of the matrix $\mathbf{C}$ contain the eigenvectors of $(\mathbf{Z}^f)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f$, and the nonzero elements of the diagonal matrix $\mathbf{\Gamma}$ contain the corresponding eigenvalues; that is,

$$(\mathbf{Z}^f)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f = \mathbf{C}\mathbf{\Gamma}\mathbf{C}^{\mathrm{T}}. \tag{10}$$

Since $(\mathbf{Z}^f)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f$ is a real symmetric matrix, its eigenvalues are real and its eigenvectors are orthogonal. Note that the solution to (8) is nonunique. If $\mathbf{B}_i$ is a $(K-1)\times(K-1)$ matrix and $\mathbf{B}_i\mathbf{B}_i^{\mathrm{T}} = \mathbf{I}$, then substitution of $\mathbf{T}_i = \mathbf{T}\mathbf{B}_i$ in place of $\mathbf{T}$ in (8) shows that $\mathbf{T}_i$ is also a solution to (8). Differing $\mathbf{B}_i$ correspond to the differing deterministic square root filters discussed in Tippett et al. (2003), of which the ETKF is but a single example. A distinguishing feature of analysis perturbations produced by the ETKF is that they are orthogonal in normalized observation space. To see this, premultiply and postmultiply (10) by $\mathbf{T}^{\mathrm{T}}$ and $\mathbf{T}$, respectively, and substitute $\mathbf{T}$ on the right-hand side by (9). Then we obtain

$$\mathbf{T}^{\mathrm{T}}(\mathbf{Z}^f)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f\mathbf{T} = \mathbf{\Gamma}(\mathbf{\Gamma} + \mathbf{I})^{-1}. \tag{11}$$

The right-hand side of (11) is a diagonal matrix. Thus, from the definition of the ETKF analysis perturbations in (6), the ETKF analysis perturbations are orthogonal under the inner product defined by

$$\{\mathbf{w}; \mathbf{z}\} = \mathbf{w}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{z}, \tag{12}$$

where $\mathbf{w}$ and $\mathbf{z}$ are any two vectors with the length of a state vector. In other words, the analysis perturbations over the observation sites normalized by the square root of the observation error covariance matrix are orthogonal under a Euclidean norm. Note that after normalization each element of the analysis perturbations over the observation sites is dimensionless.

Because the ETKF rotates and rescales perturbations according to the optimal data assimilation equation (1), it is the distribution and quality of observations that controls perturbation amplitude. Furthermore, consistent with the filtering properties of an optimal data assimilation scheme, ensemble variance in directions corresponding to large ensemble variance is reduced by a larger factor than ensemble variance in directions corresponding to small ensemble variance is reduced (see appendix and/or Daley 1991). Consequently, one would expect the aforementioned problems of the breeding methods to be ameliorated by the ETKF method. According to the discussion in the introduction about error variance characteristics of the ETKF and breeding schemes in an idealized system, one would also expect that the ETKF ensemble perturbations to maintain variance in a much wider range of amplifying directions than the breeding ensemble perturbations.

When the number of ensemble perturbations is much smaller than the number of directions to which the forecast error variance projects, (8) significantly underestimates total analysis error variance because it lacks contributions from important parts of the error space. To ameliorate this problem, we multiply the transformed perturbations by an inflation factor so as to ensure that the global 12-h forecast ensemble variance is consistent with the global control forecast error variance at the rawinsonde sites. Denote t_i as the perturbation initialization time. In our experiment, the time interval between t_i and t_{i+1} is 12 h. We hope to find a scalar inflation factor Π_i at time t_i and multiply the transformed perturbation obtained at t_i by Π_i, that is,

$$\mathbf{X}_i^a = \mathbf{X}_i^f\mathbf{T}_i\Pi_i, \tag{13}$$

so that we could expect for the next 12-h forecast, namely, at time t_{i+1},

$$\mathrm{trace}(\langle\tilde{\mathbf{d}}_{i+1}\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\rangle) = \mathrm{trace}(\tilde{\mathbf{H}}\mathbf{P}_{i+1}^e\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I}), \tag{14}$$

where $\langle\cdot\rangle$ represents the expectation operator; $\tilde{\mathbf{H}}$ is the observation operator normalized by the square root of the observation error covariance matrix—that is, $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}$; $\mathbf{P}_{i+1}^e$ is the 12-h ensemble covariance at t_{i+1}; and $\tilde{\mathbf{d}}_{i+1}$ is the innovation vector at t_{i+1} normalized by the square root of the observation error covariance matrix—that is, $\tilde{\mathbf{d}}_{i+1} = \mathbf{R}^{-1/2}(\mathbf{y}_{i+1} - \mathbf{H}\mathbf{x}_{i+1}^f)$, where $\mathbf{y}_{i+1}$ is the observation vector at t_{i+1} and $\mathbf{H}\mathbf{x}_{i+1}^f$ is the 12-h background forecast valid at the time t_{i+1} mapped into observation space by the observation operator $\mathbf{H}$. Equation (14) originates from the relationship among the inno-

vation covariance, the forecast error covariance and the observation error covariance with the assumption that the forecast errors and the observation errors are uncorrelated. Here we adopt normalization by the observation rms error so as to lead dimensionless analysis and reduce the difference in the variance of each innovation element. We also take the trace so that we only consider the consistency of the ensemble variance with the control forecast error variance on a globally averaged basis. At each forecast time, we only have one realization of the innovation vector. Fortunately, for the global observational networks the number of independent scalar innovations within the innovation vector is large (Dee 1995). Consequently, the standard deviation of $\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}$ is small compared to its mean. In this sense, $\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}$ distributes closely around its mean value $\langle\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}\rangle$. Since $\langle\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}\rangle = \mathrm{trace}(\langle\tilde{\mathbf{d}}_{i+1}\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\rangle)$, $\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}$ provides useful estimates of $\mathrm{trace}(\langle\tilde{\mathbf{d}}_{i+1}\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\rangle)$ in Eq. (14). With this assumption, (14) becomes

$$\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1} \approx \mathrm{trace}(\tilde{\mathbf{H}}\mathbf{P}_{i+1}^{e}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I}). \tag{15}$$

Of course, $\tilde{\mathbf{d}}_{i+1}^{\mathrm{T}}\tilde{\mathbf{d}}_{i+1}$ and $\mathrm{trace}(\tilde{\mathbf{H}}\mathbf{P}_{i+1}^{e}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})$ are not available at t_i. To get around this problem, the inflation factor Π_i is obtained in our experiment by assuming that the statistics of the next globally averaged 12-h forecast will be similar to that of the previous 12-h forecast. Specifically, given that the inflation factor at t_{i-1} was Π_{i-1}, the inflation factor at t_i is obtained by first checking if $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i$ is equal to $\mathrm{trace}(\tilde{\mathbf{H}}\mathbf{P}_i^{e}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})$. If not, we need to introduce a parameter α_i so that

$$\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i = \mathrm{trace}(\tilde{\mathbf{H}}\alpha_i\mathbf{P}_i^{e}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I}). \tag{16}$$

Then the inflation factor Π_i is defined as

$$\Pi_i = \Pi_{i-1}\sqrt{\alpha_i}. \tag{17}$$

From (16),

$$\alpha_i = \frac{\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i - N}{\mathrm{trace}(\tilde{\mathbf{H}}\mathbf{P}_i^{e}\tilde{\mathbf{H}}^{\mathrm{T}})} = \frac{\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i - N}{\sum_{i=1}^{K-1}\lambda_i}, \tag{18}$$

where N is the number of observations, and λ_i, $i = 1, \ldots, K-1$ are the diagonal elements of $\mathbf{\Gamma}$ in (10). From Eq. (17), Π_i is a product of these α parameters from the first forecast at time t_1 to that at time t_i; that is,

$$\Pi_i = \sqrt{\alpha_1\alpha_2\cdots\alpha_i}. \tag{19}$$

Although, at the first few initialization cycles, α_i's yielded by (18) vary from order of 100 to order of 0.1 (e.g., for the 16-member ETKF ensemble $\alpha_1 = 520.9$, $\alpha_2 = 1.4$, $\alpha_3 = 0.3$, $\alpha_4 = 0.8$), after just 1 week of 12-h perturbation initialization and forecast cycles, we found that α_i became stuck in a range of values between 0.8 and 1.2 with a mean of 1.0. This fast convergence confirms that the assumptions made in (15)–(17) are valid.

The way we calculate α_i in (16) can be regarded as an application and extension of the maximum likelihood parameter estimation theory of Dee (1995). This can be understood by noting that because the number of degrees of freedom of $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i$ is large, the mean of its distribution gets close to the peak of its distribution. Thus, α_i is the parameter to make this realization of $\tilde{\mathbf{d}}_i^{\mathrm{T}}\tilde{\mathbf{d}}_i$ the most likely. Finally, we note that, for our experiments, we applied the same method (15)–(19) of selecting inflation factors to both the simple breeding and masked breeding techniques.

3. Numerical experiment design

a. NCAR Community Climate Model

We use version 3 of the Community Climate Model (CCM3) developed at National Center for Atmospheric Research (NCAR). The details of the governing equations, physical parameterizations, and numerical algorithms of CCM3 are presented by Jeffery et al. (1996). We choose the default resolution of T42 and 18 levels.

b. The control analyses

In our experiment, for simplicity we use NCEP–NCAR reanalysis data interpolated to T42 resolution as the control analyses. ''Control forecasts'' are made from these ''control analyses.'' The time period we consider is the boreal summer in the year 2000. We also use the reanalysis data as verifications for ensemble forecasts.

c. Observational network

We will be testing the performance of ensembles with less than 17 members. Because our primary aim is to simply illustrate how the ETKF combines observational and dynamical information, and a 16-member ensemble can only crudely represent the error reducing effect of an observational network with $O(10^5)$ observations, we only employ a simplified observational network for this study. Specifically, the observational network is assumed to consist of measurements of u, v, and T (i.e., wind and temperature) at 200, 500, and 850 hPa at locations corresponding to actual rawinsonde sites (see Fig. 1). We choose model grid points that are nearest to the rawinsonde sites to represent the observation sites. It is further assumed that observations are only taken at 0000 and 1200 UTC. Our ''observations'' are taken to be equal to the values of the reanalysis data at the rawinsonde sites. These pseudo-observations represent an estimate of the mean state of the atmosphere in the grid cells corresponding to the observations. Because the reanalysis data combines true observational data with a dynamically produced first-guess field, the error covariance of the grid-cell state estimates provided by the reanalysis data is unlikely to be the same as the error variance associated with the actual radiosonde observations. To estimate an upper bound on the error variance of the reanalysis data over the observation sites, we first assume that CCM3 12-h forecast errors are uncorrelated with errors of the reanalysis data so that

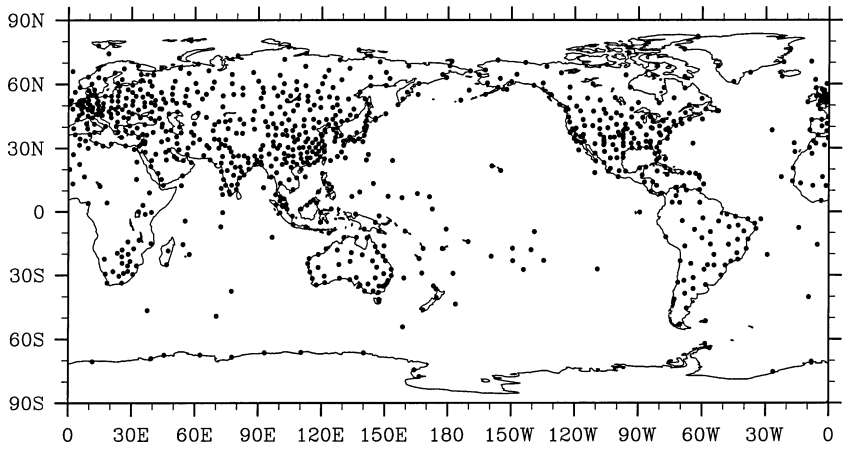

FIG. 1. Network of simulated rawinsonde stations. Black dots denote locations of rawinsonde stations.

$$\langle \mathbf{d}\mathbf{d}^{\mathrm{T}} \rangle = \mathbf{H}\mathbf{P}^{f}\mathbf{H}^{\mathrm{T}} + \mathbf{R}, \tag{20}$$

where $\mathbf{R}$ denotes the error covariance of the reanalysis data in observation space and $\mathbf{d}$ is the innovation vector. Second, we collect 12-h innovations by running a series of 12-h control forecasts for the entire period of boreal summer in 2000. We calculate innovation sample variance for wind and temperature at each observation site by averaging all the squared innovation in summer 2000 at that site. Then we choose the smallest wind and temperature innovation sample variance as the observation error variance for wind and temperature, respectively. The rms wind and temperature observation errors obtained are 2 m s^{-1} and 0.7°C, respectively. Note that these rms errors are smaller than those typically attributed to radiosonde observations in a data assimilation scheme. For simplicity, we also assume there is no error correlation between variables in the reanalysis data. Under this assumption, the observation error covariance matrix $\mathbf{R}$ is diagonal.

d. Construction of initial perturbations

In our experiment, we run 16-member one-sided ensembles for both the ETKF and the breeding methods. We initialize ensemble forecasts at 0000 and 1200 UTC during boreal summer of 2000. Fifteen forecast perturbations at the analysis time are defined as the deviation of the 15 12-h perturbed forecasts from the control forecast. For the ETKF method, from these 15 raw forecast perturbations we use Eqs. (9) and (10) to calculate the 15×15 transformation matrix. An inflation factor is calculated as described in Eqs. (17) and (18). Then the transformation matrix and the inflation factor are postmultiplied to the raw forecast perturbations as in Eq. (13) to obtain 15 initial perturbations. For the simple breeding method, we first calculate a global constant factor as in (3) for each perturbation so that the size of each scaled perturbation is consistent with the size of empirically determined global analysis rms error. We also apply the inflation factor technique as in the ETKF to the simple breeding technique. For the masked breeding method, we choose the square root of the seasonally and vertically averaged initial ensemble wind variance from the 16-member ETKF ensemble (see Fig. 3a) as a mask and smooth it with a spectral filter (Sardeshmukh and Hoskins 1984). The mask e_{ij} used in Eq. (5) for our experiments is shown in Fig. 2. The smoothing effect we choose is equivalent to a triangular truncation of T6. For each raw forecast perturbation the square root of vertically averaged squared wind perturbations is calculated and smoothed with the same spectral filter to obtain y_{ij} in Eq. (5). Then the regional rescaling factor is calculated as in Eq. (5). The inflation factor technique is also applied to the masked breeding method. The initial perturbations of the breeding methods are the raw forecast perturbations multiplied by the rescaling factor and the inflation factor. To initialize the next run, we add the 15 initial perturbations directly to the control analysis.

4. Comparison of initial ensemble variances

Figure 3 shows the square root of the seasonally and vertically averaged initial wind error variance estimated by the ETKF ensemble and the breeding ensembles. For the ETKF ensemble (Fig. 3a), there is significant perturbation amplitude over the Southern Hemisphere but quite small perturbation amplitude over the Eurasian continent. This feature corresponds well to the geographical inhomogeneity of the observation density distribution in Fig. 1. For the simple breeding ensemble (Fig. 3b), initial perturbations in the observation-scarce Southern Hemisphere are much smaller than that of the ETKF. Also, despite the high concentration of rawinsondes over the Eurasian continent, the simple bred-vector initial perturbation amplitude is locally maximized in this region. The masked breeding ensemble (Fig. 3c) is quite similar to the ETKF ensemble as far

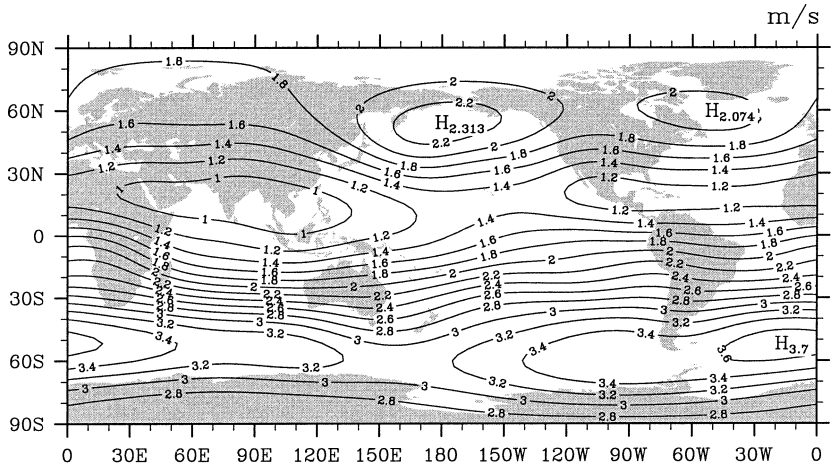

FIG. 2. The breeding mask, defined as smoothed square root of seasonally and vertically averaged initial ensemble wind variance from 16-member ETKF ensemble. Label H indicates local maximum. Contour interval is 0.2 m s^{-1}.

as the estimated average analysis uncertainty is concerned.

The manner in which the ETKF has allowed ensemble spread to be governed by observational density is better seen by plotting maps of vertically and seasonally averaged ratios of ensemble-based analysis rms wind error over ensemble-based 12-h forecast rms wind error. Such maps give a representation of the geographical distribution of the factor that rescales 12-h forecast ensemble spread into initial ensemble spread. Figure 4a displays this ratio for the 16-member ETKF ensembles. The effective rescaling factor for the 16-member ETKF ensemble not only reflects the high concentrations of observations over Europe and North America, but also accounts for the smaller midlatitude concentrations over South Africa, Australia, and South America. The corresponding plot for the simple breeding ensemble (Fig. 4b) shows an approximate constant that does not represent the spatial variation of the observation density distribution. In Fig. 4c the rescaling factor for the masked breeding ensemble is unable to appropriately size the analysis errors differently between the moderately observed continents and the sparsely observed oceans in the Southern Hemisphere midlatitude. This difference is due to the fact that the mask is smoothed, as shown in Fig. 2, and thus relatively small observation densities on Southern Hemisphere midlatitude land masses are neglected.

We also note that, in the Northern Hemisphere, the masked breeding rescaling factor has an eastward phase shift relative to the ETKF rescaling factor. This could be due to the initial imbalance or rank deficiency of the masked bred perturbations. In any case, the ETKF tendency to place minima in rescaling factors on the eastern boundaries of midlatitude data-sparse regions, that is, west coasts of continents, is consistent with the fact that optimal data assimilation schemes attenuate error variance most in regions where the forecast error variance is large relative to observation error variance (see appendix). The masked breeding ensembles positioning of rescaling minima east of the eastern boundaries of midlatitude data-sparse regions is inconsistent with the way one would expect an optimal data assimilation scheme to reduce error variance.

Another striking difference is located in the Tropics, where one observes local minima in east Africa, west Indonesia and west South America in the masked breeding while no such patterns appear in the corresponding regions of the ETKF. These local minima suggest faster equatorial disturbance growth in the masked breeding ensemble than in the ETKF ensemble. Such equatorial growth is typically attributed to interactions of Kelvin waves, Rossby waves and convection along East African and Andean mountain ranges (McPhaden and Gill 1987; P. Roundy 2002, personal communication; Kleeman 1989). Could it be that masking has increased equatorial wave sources within the perturbation subspace, thus increasing the prominence of such growth within the masked bred mode ensemble subspace? We speculate that it has. The masking procedure does not obey any known balance equations. Whenever masking is applied, it is likely that some sort of balance inherent to the system (e.g., midlatitude quasigeostrophic balance) will be violated. Hamill et al. (2000) also discusses that the use of the regional rescaling process introduces noises, which results in spuriously large ensemble spread. In contrast, since each ETKF analysis perturbation represents a linear combination of balanced forecast perturbations [Eq. (6)], the ETKF analysis perturbations are guaranteed to be in balance provided that 12-h forecast perturbation amplitude is small enough for the tangent linear approximation to be valid.

To further study the ability of the ETKF to allow the ensemble spread to be governed by the geographical

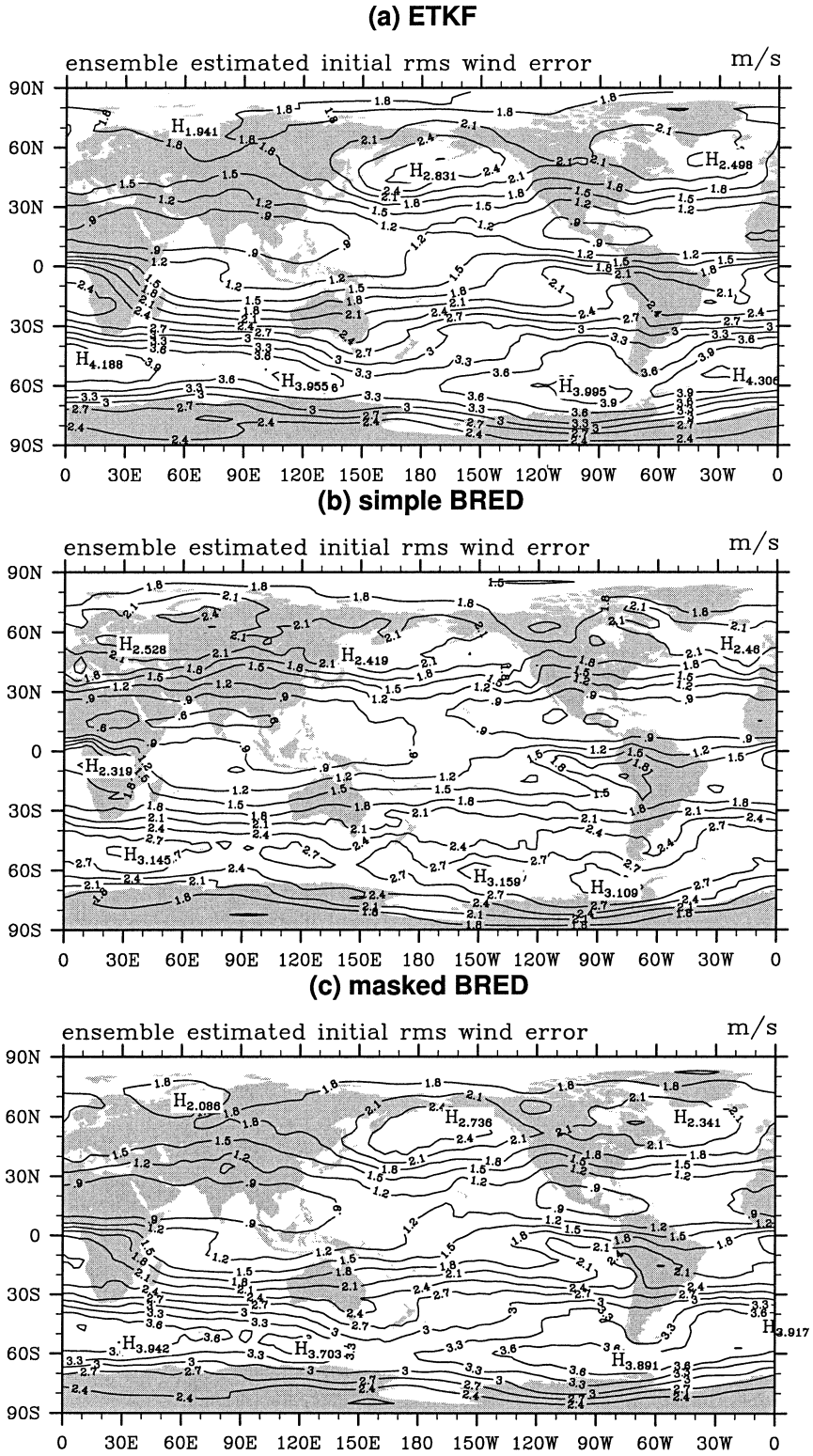

FIG. 3. Square root of seasonally (boreal summer in 2000) and vertically averaged ensemble wind variance of initial ensemble perturbations for (a) 16-member ETKF ensemble, (b) 16-member simple breeding ensemble, and (c) 16-member masked breeding ensemble. Label H indicates local maximum. Contour interval is 0.3 m s^{-1}.

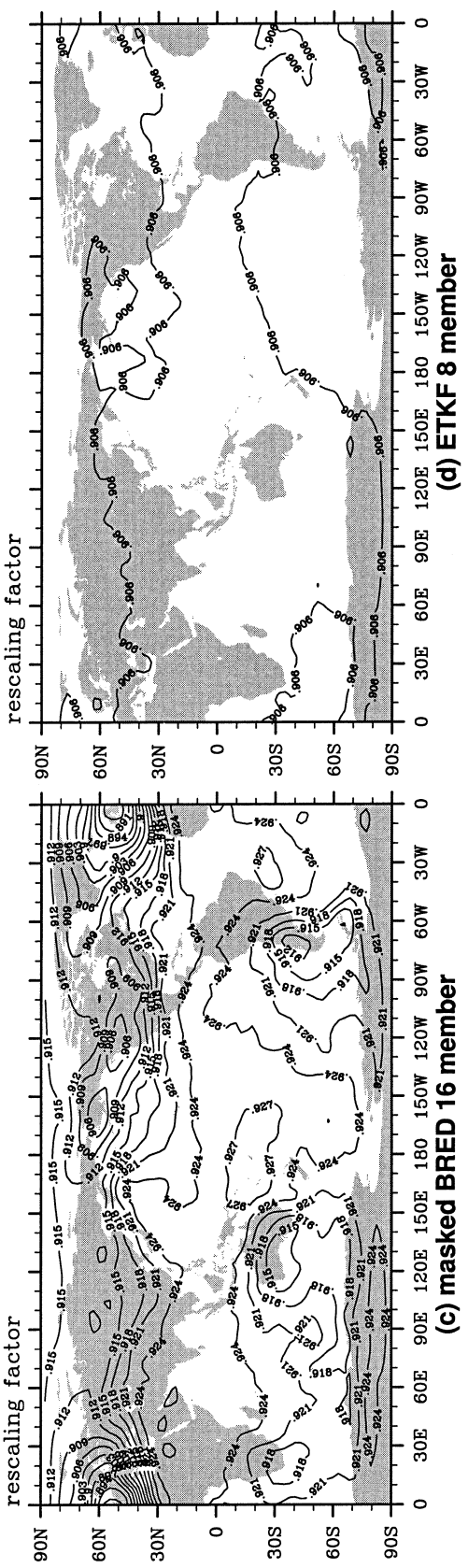

FIG. 4. Seasonally and vertically averaged ratio of square root of initial ensemble wind variance over square root of 12-h ensemble forecast wind variance for (a) 16-member ETKF ensemble, (b) 16-member simple breeding ensemble, (c) 16-member masked breeding ensemble, and (d) 8-member ETKF ensemble. Nine-point local smoothing is applied. Contour interval is 0.003 for (a), (b), and (d); and 0.005 for (c).

inhomogeniety of the observation distribution, we also ran an 8-member ETKF ensemble to see how the ensemble size of the ETKF scheme affected the rescaling factor. Figure 4d shows that over both hemispheres, the rescaling factor of the ETKF 8-member ensemble is not as detailed as that of the ETKF 16-member ensemble. Also, the ETKF 8-member ensemble does not show the land–sea contrast in the rescaling effect over Southern Hemisphere midlatitudes as well as the ETKF 16-member ensemble does. Essentially, this is because problem of spurious long-distance correlations is worse for the ETKF 8-member ensemble than it is for the 16-member ensemble.

The superiority of the 16-member results over the 8-member results indicates that the sensitivity of the ETKF ensemble rescaling factors to variations in observational density would be further improved by moving to a 32-member ensemble. As the ensemble size increased, one would have to include satellite observations in the observation operator in order to avoid the ensemble variance becoming unrealistically large over the ocean basins. Spurious long-distance covariances associated with small ensemble size are primarily responsible for the lack of oceanic ensemble spread in the 8- and 16-member results.

5. Maintenance of variance along orthogonal basis vectors

As mentioned in the introduction, in a system with fixed and perfect linear dynamics and fixed observation distribution and error statistics, provided that the initial ensemble perturbations span the vector subspace of unstable normal modes of the linear dynamics operator, the ETKF ensemble would eventually maintain error variance in all amplifying normal modes, whereas the simple breeding technique would eventually only maintain error variance in the direction corresponding to the most rapidly amplifying normal mode. To see whether such profoundly different error variance maintenance characteristics would be present with an imperfect time-varying dynamics propagator and with the number of ensemble members less than that of the growing normal modes of the linear dynamics operator, we examine the mean eigenvalue spectra of the ensemble covariance matrices (see Fig. 5). For each ensemble generation technique, heights of the 15 bars correspond to 15 seasonally averaged diagonal elements of $\mathbf{\Gamma}$ in Eq. (10) for 12-h forecasts. The spectrum of the ETKF eigenvalues is much flatter than that from both breeding methods. There is nearly zero variance in the last six and three uncorrelated and orthogonal directions for the simple breeding and the masked breeding, respectively. In other words, while there are large amounts of ensemble forecast variance present in all 15 uncorrelated orthogonal directions of the ETKF ensemble, nearly all of the ensemble forecast variance is contained in a single direction for both the simple breeding and the masked breeding.

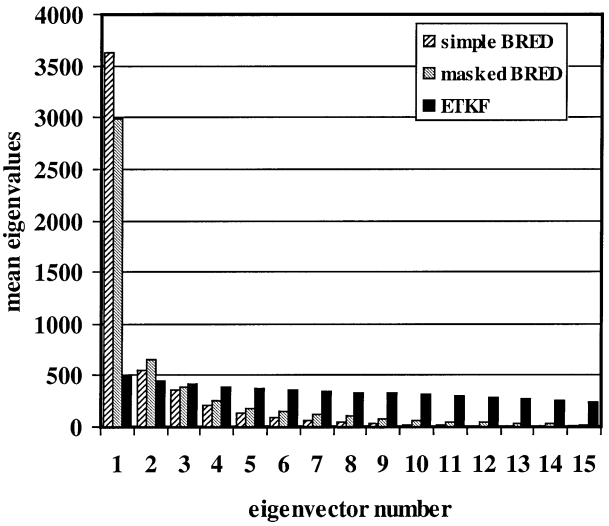

FIG. 5. Seasonal mean spectra of eigenvalues of ensemble-based 12-h forecast covariance matrices normalized by observation error covariance in observation sites for 16-member ETKF, simple breeding and masked breeding ensembles.

The ETKF analysis perturbations are obtained by solving the error variance update equation for an optimal data assimilation scheme, that is, Eq. (1). From Eqs. (6) and (9), forecast perturbations are first rotated and then rescaled. The fact that the ETKF eigenvalue spectrum is relatively flatter is due to the filtering effect of Eq. (1) (see appendix). From Eq. (7), the ETKF estimated forecast error covariance matrix at the observation sites normalized by the observation error covariance is

$$\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^{\mathrm{T}} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{Z}^f(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{Z}^f)^{\mathrm{T}}, \tag{21}$$

where $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}$. Define $\mathbf{E} = \tilde{\mathbf{H}}\mathbf{Z}^f\mathbf{C}\mathbf{\Gamma}^{-1/2}$, then (21) becomes

$$\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^{\mathrm{T}} = \mathbf{E}\mathbf{\Gamma}\mathbf{E}^{\mathrm{T}}. \tag{22}$$

Because $\mathbf{E}^{\mathrm{T}}\mathbf{E} = \mathbf{I}$ (but $\mathbf{E}\mathbf{E}^{\mathrm{T}} \neq \mathbf{I}$), (22) is an approximate eigenstructure expression for the estimated $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^{\mathrm{T}}$. From Eqs. (6) and (9), and the definition of $\mathbf{E}$, the ETKF analysis perturbation in the observation space normalized by the square root of $\mathbf{R}$ before applying the inflation factor is

$$\tilde{\mathbf{H}}\mathbf{Z}^a = \mathbf{E}\mathbf{\Gamma}^{1/2}(\mathbf{\Gamma} + \mathbf{I})^{-1/2}. \tag{23}$$

Equation (23) shows that the ETKF analysis perturbations point in the same directions as the eigenvectors of the estimated forecast error covariance matrix $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^{\mathrm{T}}$. From (23) the ETKF estimated analysis error covariance matrix at the observation sites is

$$\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^{\mathrm{T}} = \mathbf{E}\mathbf{\Gamma}(\mathbf{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^{\mathrm{T}}, \tag{24}$$

which is consistent with (A5) in the appendix. Comparing (22) and (24), the error variance in the direction of the ith eigenvector is reduced by multiplying by a

factor $1/(\lambda_i + 1)$ where λ_i is the ith eigenvalue of the estimated forecast error covariance matrix $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$. Thus, in normalized observation space, the eigenvalue spectrum of the analysis error covariance matrix is flatter (''whiter'') than the corresponding forecast error covariance matrix eigenvalue spectrum. The above analysis shows that rescaling by differing factors in differing directions is required to solve (1). Since the breeding scheme rescales by a similar factor in all directions, the filtering effect of the ETKF does not exist in the breeding ensemble. Figure 5 illustrates these facts.

The lack of error variance in the trailing eigenvectors of the breeding ensembles indicates that there would be little point in having more than one or two ensemble members in the breeding ensembles. For the ETKF, since each direction contributes comparable amounts to the ensemble forecast variance, larger size ensemble would provide improved estimates of error variance. This hypothesis is supported by the superiority of the 16-member ETKF over the 8-member ETKF in the sensitivity of the rescaling factor to the observation distribution shown in Figs. 4a and 4d.

Equation (24) can also help us understand differences between the inflation factors used in the 8- and 16-member ETKF ensembles. For the 16-member ETKF ensemble, the mean value of the inflation factor on the initial variance, that is, Π_i^2 [see Eqs. (8) and (13)], after α_i converged, was found to be equal to 280 while for the 8-member ETKF ensemble it was 592 (note that the mean inflation factor for individual perturbations, that is, Π_i, is 16.6 for the 16-member ETKF and 24.2 for the 8-member ETKF). The fact that the inflation factor on initial variance for the 16-member ensemble is about half (280/592) of the inflation factor required for the 8-member ensemble indicates that the primary reason for Eq. (8) producing too small analysis error variance estimates is that the rank of the ensemble-based covariances is much smaller than the rank of the true forecast error covariance matrix. To understand this result in terms of (24) note that at the perturbation initialization time, the inflation factor applied 12 h earlier makes the sum of the diagonal elements of $\mathbf{\Gamma}$ in Eq. (22) approximately the same for both the 8-member and 16-member ETKF ensembles. Typically, $K - 1$ of the elements of $\mathbf{\Gamma}$ are much larger than 1 and hence for the 8-member ensemble the trace of $\mathbf{\Gamma}(\mathbf{\Gamma} + \mathbf{I})^{-1}$ is about 7 whereas for the 16-member ensemble the trace is about 15. To make the sum of the diagonal elements of $\mathbf{\Gamma}$ consistent with the control forecast error variance in the next 12-h forecast, the 8-member ETKF initial perturbations consequently need to be inflated by a factor about 2 times as big as that for the 16-member ETKF ensemble. In this way one sees that the inflation factor will diminish as the number of ensemble members increases.

6. Comparison of ensemble perturbation growth characteristics

a. Analysis error covariance singular vectors

Ensemble forecasts should be able to reliably identify forecasts where the chance of large forecast errors is larger than usual. Because rapid amplification of analysis error can lead to large forecast errors, it is desirable for an ensemble to contain perturbations representative of likely analysis errors that can grow quickly. Ehrendorfer and Tribbia (1997) and Houtekamer (1995) pointed out that if one could obtain an accurate estimate of the inverse analysis error covariance matrix, one could find the initial vectors that evolve into the leading eigenvectors of the forecast error covariance matrix under any norm of interest. These vectors are consistent with the analysis error covariance statistics and are called the analysis error covariance singular vectors (AECSVs; Ehrendorfer and Tribbia 1997; Hamill et al. 2003, hereafter HSW). Barkmeijer et al. (1998, 1999) refers to these vectors as Hessian singular vectors (HSVs) since the inverse of the analysis error covariance is estimated by the Hessian of the 3DVAR cost function. For an ensemble that provides norm independent estimates of $\mathbf{P}^f$, for example, the ETKF, it is trivial to perform eigenvector decompositions of its estimated $\mathbf{P}^f$ to find the linear combination of initial ensemble perturbations that evolve into the leading eigenvectors of its estimated $\mathbf{P}^f$ under any norm of interest. This is done by first finding a linear combination of the forecast ensemble members that is equal to the leading eigenvectors of the estimated $\mathbf{P}^f$ under the norm of interest. Under the assumption of linear dynamics, the same linear combinations of the initial ensemble perturbation members will give the initial structure that evolve to the leading eigenvectors of the estimated $\mathbf{P}^f$. Thus, to the extent that one accepts the ensemble estimated analysis error covariance, the ensemble perturbations also provides AECSVs. Different from Barkmeijer et al. (1998, 1999), the AECSVs provided by the ensemble is flow dependent (see similar calculation and discussion in HSW). Because, as indicated by sections 4 and 5 (also indicated by the mean and forecast error variance estimations in sections 7 and 8), the ETKF estimates the analysis error covariance more accurately than the breeding schemes, the ETKF provides more accurate AECSVs than the breeding schemes.

b. Total energy norm singular vectors

Palmer et al. (1998) argued that total energy norm singular vectors (TESVs) provide a reasonable approximation to the forecast error covariances on the grounds that analysis errors appeared to be spectrally white in the total energy norm. They also argue that because the amplification rate of the dominant singular vectors is 3 to 4 times larger than that of the breeding vectors, the dominant singular vectors will explain more forecast

errors than the dominant breeding vectors provided the projection of the analysis errors onto the dominant singular vectors is not less than 3 to 4 times the projection of the analysis errors onto the breeding vectors. In this section, we show that this problem proposed by Palmer et al. (1998) that the bred vector grows too slow can be ameliorated by the ETKF, which can be regarded as a modified version of the breeding scheme.

As in all singular-vector-related techniques (Molteni et al. 1996; Ehrendorfer and Tribbia 1997; Houtekamer 1995; Barkmeijer et al. 1998, 1999), to find the fastest growing direction under a particular norm, we assume that for a short term forecast the dynamics operator is approximately linear; that is,

$$\mathbf{X}^f = \mathbf{M}\mathbf{X}^a, \tag{25}$$

where $\mathbf{M}$ is the linear dynamics propagator that maps analysis perturbations $\mathbf{X}^a$ into forecast perturbations $\mathbf{X}^f$. We employ an approximate energy norm defined as

$$\frac{1}{2}u'^2 + \frac{1}{2}v'^2 + \frac{c_p}{T_r}T'^2, \tag{26}$$

where u', v', and T' are wind and temperature perturbations; c_p is the specific heat at constant pressure; and T_r is the reference temperature (Palmer et al. 1998). To find the fastest growth in terms of this approximate energy norm, one needs to select the direction in the initial ensemble perturbation subspace so that the total energy amplification in this direction is maximized, that is,

$$\max \frac{\mathbf{b}^{\mathrm{T}}(\mathbf{X}^f)^{\mathrm{T}}\mathbf{S}\mathbf{X}^f\mathbf{b}}{\mathbf{b}^{\mathrm{T}}(\mathbf{X}^a)^{\mathrm{T}}\mathbf{S}\mathbf{X}^a\mathbf{b}}, \tag{27}$$

where $\mathbf{b}$ is a vector combining initial ensemble perturbations to find out the fastest growing direction and $\mathbf{S}$ is a matrix used to define approximate perturbation energy norm. Solving for $\mathbf{b}$ in Eq. (27) is the same as solving for $\mathbf{b}$ in function

$$F = \mathbf{b}^{\mathrm{T}}(\mathbf{X}^f)^{\mathrm{T}}\mathbf{S}\mathbf{X}^f\mathbf{b} - \lambda[\mathbf{b}^{\mathrm{T}}(\mathbf{X}^a)^{\mathrm{T}}\mathbf{S}\mathbf{X}^a\mathbf{b} - 1] \tag{28}$$

subject to the conditions that $\partial F/\partial\mathbf{b} = \mathbf{0}$ and $\partial F/\partial\lambda = 0$, which lead to an eigenvalue problem for λ. The largest value of λ is the maximum growth in terms of total energy norm. It is easy to show that the leading eigenvalue of $\boldsymbol{\beta}^{-1/2}\mathbf{D}^{\mathrm{T}}(\mathbf{X}^f)^{\mathrm{T}}\mathbf{S}\mathbf{X}^f\mathbf{D}\boldsymbol{\beta}^{-1/2}$ is the solution, where $\boldsymbol{\beta}$ and $\mathbf{D}$ are the eigenvalue and eigenvector matrices of $(\mathbf{X}^a)^{\mathrm{T}}\mathbf{S}\mathbf{X}^a$. Similar calculation is shown in Bishop and Toth (1999).

Figure 6 shows the maximal energy growth for 1- and 2-day forecast lead times. The ETKF ensemble has larger maximal growth at all lead times. The reason for this may be related to the fact that in a time-invariant basic state, perturbations that maximize energy norm growth require a complete set of normal modes of $\mathbf{M}$ for their construction (Farrell 1988, 1989). In a time-invariant basic state, the breeding ensemble perturbation subspace would only contain the most rapidly growing normal mode, whereas the ETKF ensemble perturbation subspace would contain a wider range of normal modes. Thus, in a time-invariant basic state, supernormal mode growth rates could be found within the ETKF ensemble perturbation subspace whereas finite time growth rates would be bounded by the normal mode growth rate in a breeding ensemble perturbation subspace. Figures 5 and 6 suggest that a similar lack of independent perturbation structures within the ensemble subspace of bred perturbations may constrain maximal growth within the bred-vector ensemble subspace.

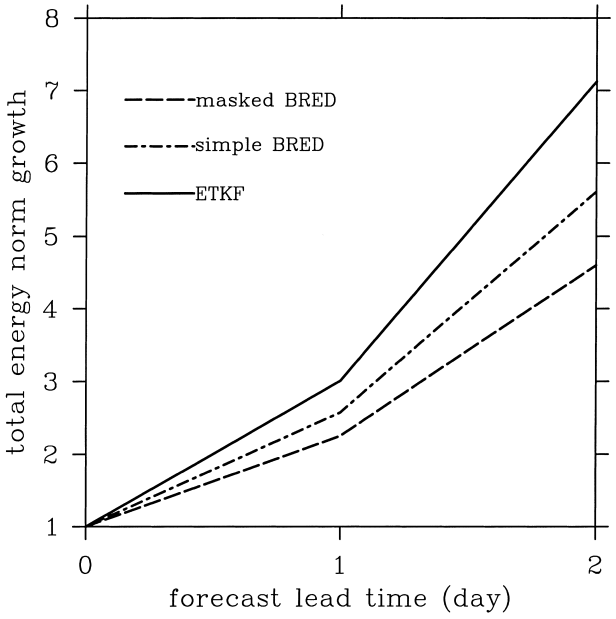

FIG. 6. Fastest growth in ensemble perturbation subspace in total energy norm from 1- to 2-day forecast lead time.

Note that the masked bred-vector perturbations have even smaller maximal growth than the simple bred-vector perturbations. This suggests that while masking increases the number of directions in which variance is maintained in a bred-vector ensemble (see Fig. 5), these additional directions of variance are incapable of increasing mean maximal energy growth. We speculate that the masking is introducing unbalanced perturbations into the ensemble that generally attenuate rapidly with time. Note that the above discussion is based on the belief in the assumption of linear dynamics. How well the results from this linear dynamics approximation corresponds to the actual nonlinear dynamics needs to be further evaluated.

7. Root-mean-square error of the ensemble mean

In our experiment, the perturbed initial conditions are not centered on the control analysis as the usual positive/negative pair perturbations. So, for short-term forecasts the ensemble mean is different from the control fore-

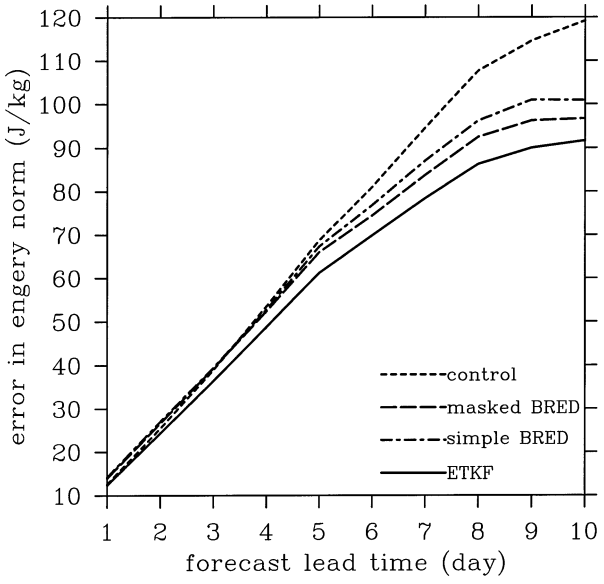

FIG. 7. Globally averaged ensemble mean forecast error (200, 500, and 850 hPa) in terms of the approximate energy norm as a function of forecast lead time. The corresponding measurement of control forecast error is also shown as a comparison.

cast.[4] The ensemble means are validated against the NCEP–NCAR reanalysis data at 1-, 2-, 3-, . . . , 10-day forecast lead time at the observation sites. In Fig. 7 we plot the 200-, 500-, and 850-hPa globally averaged ensemble mean forecast error in terms of the approximate energy norm defined in section 6 [Eq. (26)]. The corresponding measurements of control forecast errors are also shown for comparison. It turns out that the ETKF ensemble mean is more accurate than both the simple breeding and the masked breeding for 1-day through 10-day forecasts. The ensemble mean of the masked breeding is only slightly more skillful than that of the simple breeding. Compared to the control, the ETKF ensemble mean is more accurate than the control for 1-day through 10-day forecasts while the breeding ensemble means are only more accurate than the control after 4-day forecast. The more accurate ETKF ensemble mean indicates that the ETKF initial perturbations sample the analysis errors better than both breeding schemes. This can be explained by the fact that the ETKF samples analysis errors in much more orthogonal and uncorrelated directions than do the breeding schemes (Fig. 5). Calculations on the correlations (not shown here) of the forecast errors from ensemble members show that the forecast errors of the ETKF ensemble members are less correlated than those of both masked and simple breeding ensemble members.

8. Comparison of ensemble predictions of innovation variance

a. Concept of ensemble spread precision

Studies by Toth et al. (2001) and Zhu et al. (2002) suggest that evaluating the ability of an ensemble to predict case-dependent forecast uncertainty is a critical criterion to evaluate an ensemble forecast system. Just like the true state can be regarded as a random variable around the forecast, the true forecast error variance can be regarded as a random variable around the ensemble variance. An accurate prediction of forecast error variance is one in which the true forecast error variance distributes closely to the ensemble variance, that is, the variance of the forecast error variance around the ensemble variance is small. We refer to the ability of an ensemble to get forecast error variance right on every day at every grid point variable as ''the precision'' of the ensemble variance.

As we shall discuss in future work, information about the degree of ensemble variance precision can be used to increase the accuracy of error probability density functions derived from ensemble variances. Here, we introduce new tests of ensemble variance precision. Tests such as rank histograms (Hamill 2001), the relative operating characteristics (ROC) curve (Mason 1982; Richardson 2000), and the Brier Skill Score (Brier 1950; Atger 1999) will not be considered here because they require a PDF to be derived from the ensemble forecast and, in our view, the derivation of such PDFs is nontrivial. [Note that the simple method of constructing PDFs by assuming that each ensemble member represents a random draw from the distribution of true forecasts is inappropriate for ETKF perturbations. As shown in Eqs. (23) and (24), each ETKF analysis perturbation divided by $\sqrt{K-1}$ has an amplitude equal to one standard deviation of the error distribution that occurs in its direction. Random perturbations would not be governed by such a constraint. For similar reasons, breeding perturbations should not be considered to be random either.] Future work will be devoted to the derivation of such PDFs. The spread–skill correlation is not used because, according to Whitaker and Loughe (1998), even for a perfect ensemble the magnitude of the correlation between spread and skill need not be large.

b. Test 1: Resolved range of innovation variance

Consider a scatterplot of points for which the ordinate and abscissa of each point is respectively given by the squared 500-hPa U wind innovation and 500-hPa U wind ensemble variance at 1-day forecast lead time at one midlatitude observation location.[5] Points collected correspond to all midlatitude radiosonde stations and all

[4] Comparison on paired and unpaired forecasts for a given computational expense is carried out in a forthcoming paper.

[5] Examples of similar scatterplots are given in Majumdar et al. (2002).

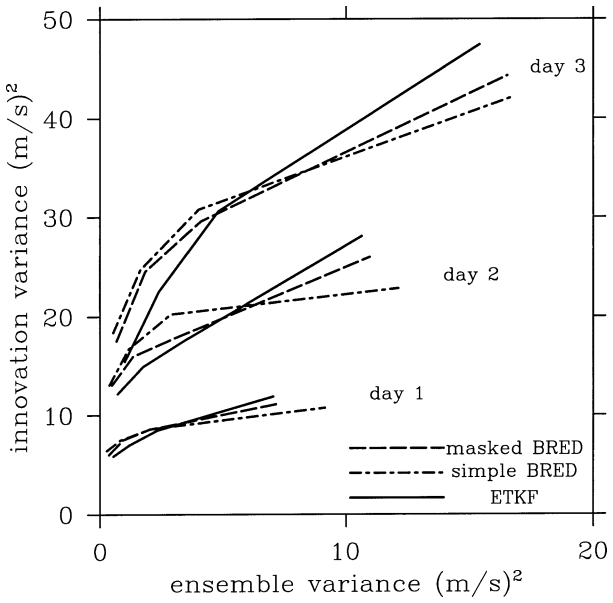

FIG. 8. This figure is plotted by first drawing a scatterplot (not shown) of squared 500-hPa *U* wind innovation vs 500-hPa *U* wind ensemble variance at a particular forecast lead time for each midlatitude observation location for all forecasts during boreal summer of 2000, dividing the points into four equally populated bins, arranged in order of increasing ensemble variance, and then averaging the squared innovation and ensemble variance in each bin. What is shown is averaged squared innovation vs the averaged ensemble variance from 1- to 3-day forecasts.

1-day 500-hPa *U* forecasts throughout the Northern Hemisphere summer of 2000. To begin, we divide the points into four equally populated bins, arranged in order of increasing ensemble variance. Then we average the squared innovation and ensemble variance in each bin, respectively. Connecting the points corresponding to these average values then yields a curve describing the relationship between the sample innovation variance and the ensemble variance. Figure 8 provides examples of such curves.

To interpret the curves in Fig. 8, note that if forecast errors were uncorrelated with observation errors then innovation variance would be equal to the sum of forecast error variance and observation error variance. Thus, if observation error variance and model errors were also uncorrelated with forecast error variance[6] and if the ensemble variance were equal to the forecast error variance, this curve would be a straight line with a slope of 45°. However, when the true forecast error variance is a random variable distributed about the given ensemble variance then the slope would be less than 45°.

To see that inaccurate predictions of forecast error variance tend to lead to relatively small slopes, note that if the ensemble variance were uncorrelated with innovation variance then the innovation variances would be the same for all values of the ensemble variance. In such a situation, relatively small ensemble variances would underpredict innovation variance, whereas relatively large ensemble variances would overpredict innovation variance. Thus, we expect inaccurate predictions of forecast error variance corresponding to relatively small (large) ensemble spread to be negatively (positively) biased. As the accuracy of ensemble-based predictions of forecast error variance improves, this spread-dependent bias will diminish, and the range of innovation variances distinguishable from ensemble variance will increase. Thus, by observing the range of the sample innovation variance in such plots, the degree of ensemble–spread precision is indicated.

Lines in Fig. 8 demonstrate the relationship between the sample innovation variance and the ensemble variance for 500-hPa *U* for 1–3-day forecasts (curves for 4–10 days are not shown), where we consider four equally populated bins with about 2800 points in each bin. The ETKF ensemble variance can resolve a wider range of innovation variance than both breeding methods.[7] This result is further confirmed when we increase the number of bins to 64 (~175 points in each bin). The dashed curves in Fig. 9 are the relationship between the sample innovation variance and the ensemble variance for 500-hPa *U* for 2-day forecasts when the number of bins increases to 64. Revealed by the dashed curves, the range of the sample innovation variance from the ETKF ensemble is about 40% and 100% greater than those from the masked breeding and the simple breeding ensembles, respectively. When the ensemble variance is larger than 15 $(\text{m s}^{-1})^2$, both the simple breeding and masked breeding completely lose their ability to distinguish differing forecast error variances. Even for a 9-day forecast lead time, the 64-bin ETKF dashed curve (see dashed curves in Fig. 10) allows innovation variance to be distinguished over a 17% larger range than either of the corresponding breeding curves.

c. Test 2: Sensitivity to bin size

If the variance of the true forecast error variance about the ensemble variance is large, then as the sample size in each bin is reduced the relationship between the sample innovation variance and the ensemble variance becomes noisier. Thus, the rate at which this relationship becomes noisy as the sample size in each bin is reduced is another crude estimate of the accuracy of forecast error variance prediction method. Figure 9 shows such a sensitivity test for the 500-hPa *U* at 2-day forecast lead time. The solid line corresponds to 4 bins of about 2800 realizations in each bin while the dashed line corresponds to 64 bins of about 175 realizations in each

[6] Note that the variance of the error of representation (often included in observation error variance estimates) and model error are both likely to be correlated with forecast error variance.

[7] Note that the variations of the slopes for each curve are expected to be because the forecast error variance could correlate with the model errors and/or the error of representations in observations.

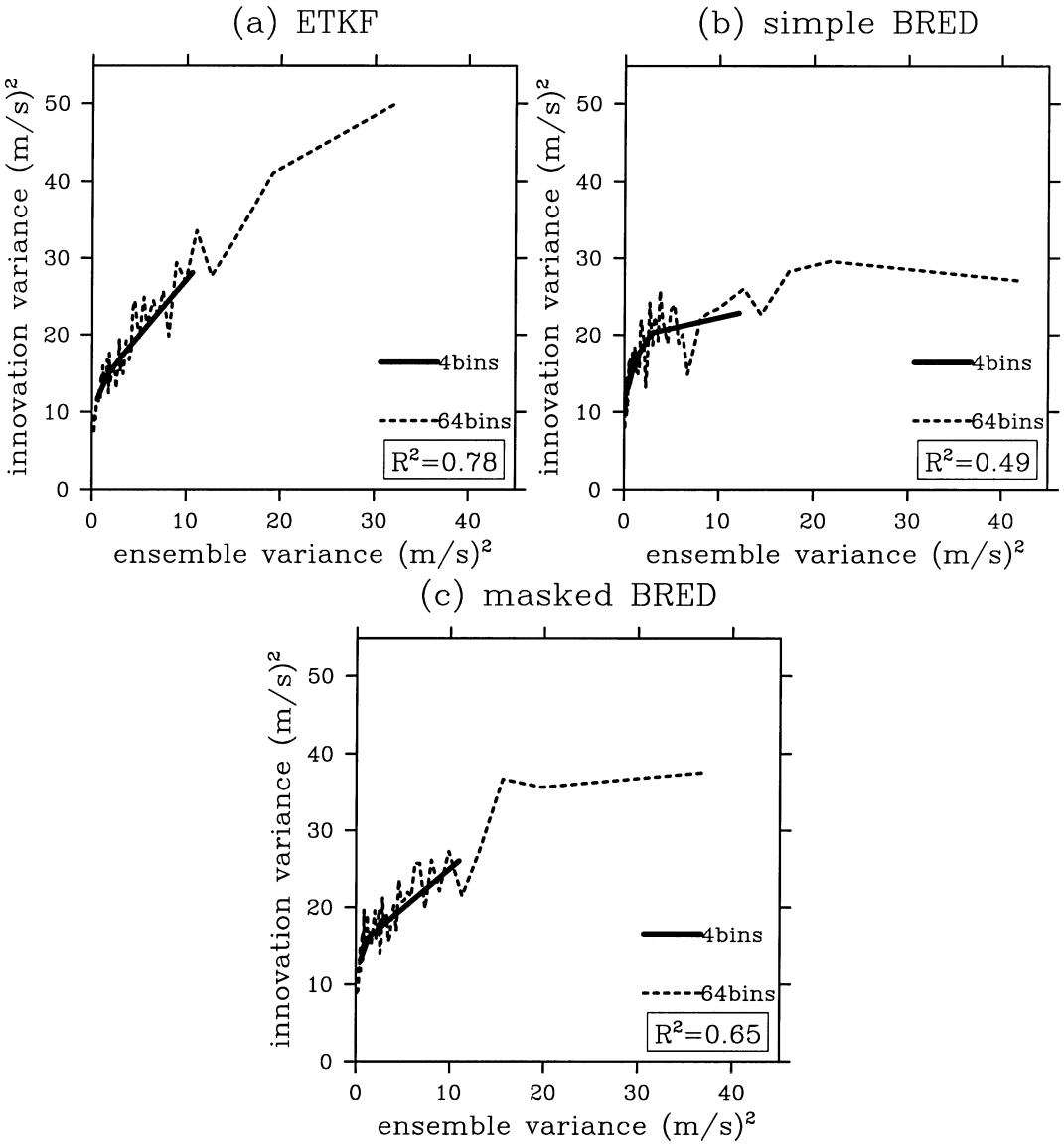

FIG. 9. Sensitivity test on relationship in Fig. 8 to bin size for 2-day forecast lead time for (a) 16-member ETKF ensemble, (b) 16-member simple breeding ensemble, and (c) 16-member masked breeding ensemble. The *R*-square values are shown on right bottom.

bin. The *R*-square values for the dashed curve relative to the solid curve are 0.78, 0.65, and 0.49 for the ETKF, masked breeding, and simple breeding ensembles, respectively. Thus, as the sample size in each bin is decreased from 2800 to 175, the relationship between the predicted and the realized innovation variance becomes the noisiest for the simple breeding ensemble. As the forecast lead time goes up to 9 or 10 days, the *R*-square values for the three ensemble methods become similar (see Fig. 10 for 9-day forecast). As mentioned above, even in this case, the 64-bin ETKF curve for the 9-day forecast still allows innovation variance to be distinguished over a larger range than either of the corresponding breeding curves.

d. Test 3: Kurtosis

Consider a distribution of innovations corresponding to a single value of ensemble variance. Assume that the innovations are normally distributed with variance equal to the sum of the forecast error variance and observation error variance. An accurate ensemble variance would be equal to the forecast error variance for all of the innovation realizations and the distribution of innovations corresponding to the ensemble variance value would be normally distributed. However, if the ensemble-based forecast error prediction was inaccurate, then the innovation realizations would correspond to a distribution of forecast error variances. In particular, it

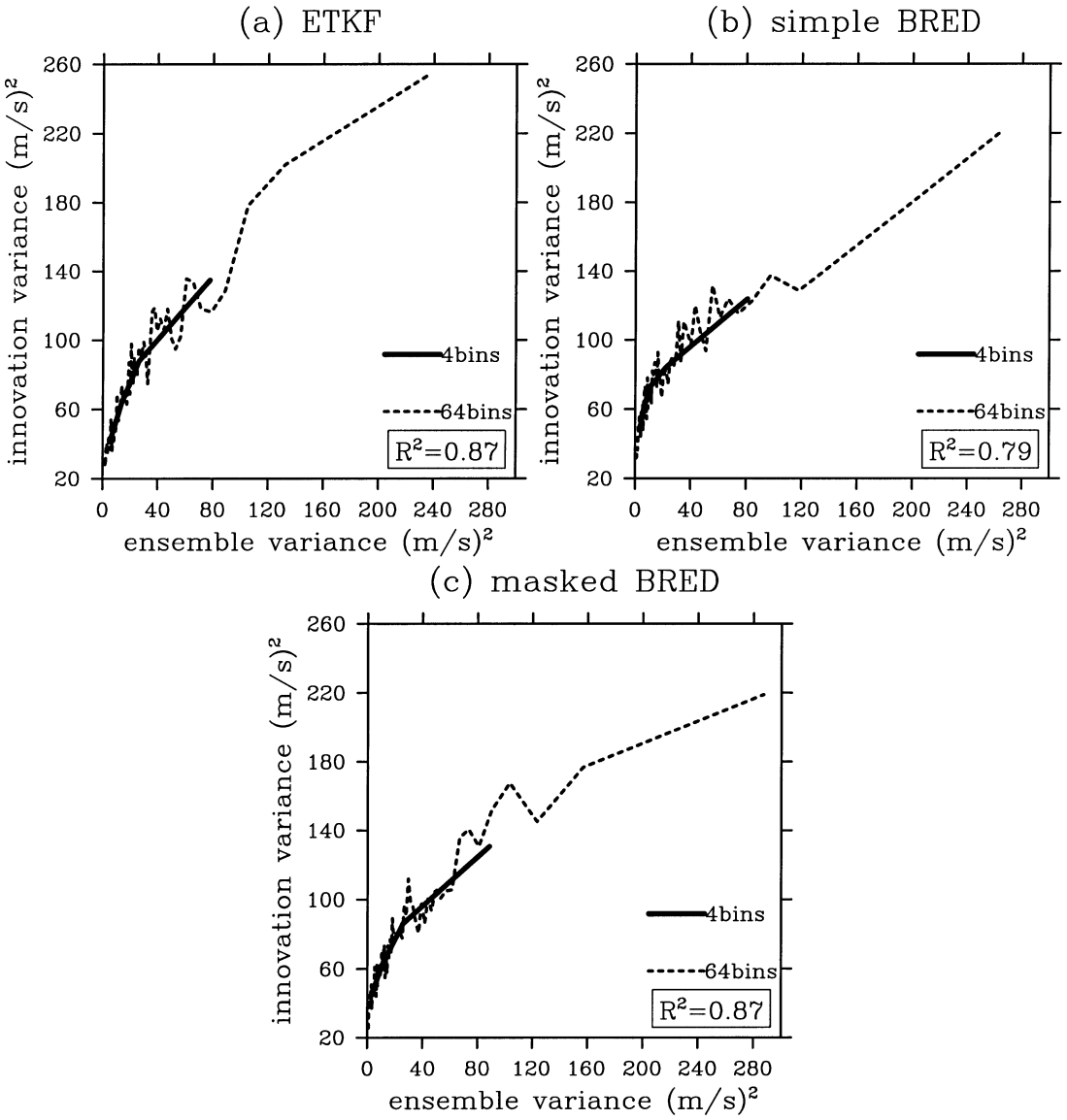

FIG. 10. As in Fig. 9, except for 9-day forecast lead time.

would be likely that innovations corresponding to innovation variances substantially larger than the mean of the innovation variances would occur. Consequently, the likelihood of finding innovations with extremely large magnitude within the distribution would increase. A simple measure of the distribution that is sensitive to such extreme values is the kurtosis. The kurtosis is the expected value of the fourth power of realizations normalized by their standard deviation (Goedicke 1953). The kurtosis tends to be larger for distributions with relatively heavy tails (i.e., many extreme values) than it is for distributions with thin tails. For a normally distributed random variable, the kurtosis is equal to 3. For finite sample sizes, the true mean and kurtosis are unknown. Nevertheless, one can approximate them with the sample mean and the sample kurtosis. We first calculate the innovation kurtosis in each of the four bins (about 2800 innovations in each bin) and then average the four kurtosis values. Figure 11 shows the 500-hPa U innovation kurtosis averaged over four bins as a function of forecast lead time. The ETKF kurtosis is smaller than that of both the breeding methods throughout 1- to 10-day forecasts.

We conclude this section by noting that the results from all three of our ensemble precision measures indicate that the ETKF ensemble variance is more accurate than the ensemble variances produced by the breeding ensembles.

9. Computational expense

The main computations required for the ETKF scheme are the $(K - 1) \times (K - 1)$ inner products in

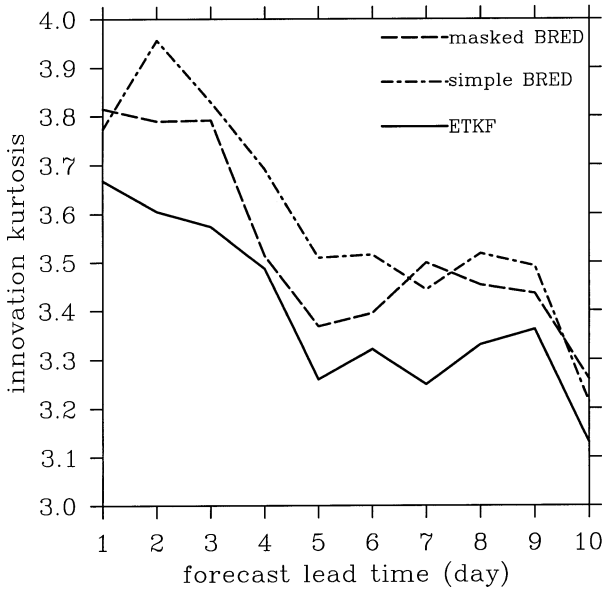

FIG. 11. Innovation kurtosis for 500-hPa U averaged on four bins defined in Fig. 8 for 1- to 10-day forecast lead time.

observation space required to form the elements of $(\mathbf{Z}^f)^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f$ [Eq. (10)] together with its eigenvector decomposition. Because K is the number of ensemble members, the number of operations required for these computations is much less than that required for the generation of the singular vector perturbations (Molteni et al. 1996) and the perturbed observation perturbations (Houtekamer et al. 1996). In our experiment, the computational expense for generating the ETKF initial perturbations is only 6% and 4% more expensive than that of generating the simple breeding and masked breeding initial perturbations, respectively.

10. Summary

The ETKF ensemble generation technique is similar in spirit to the breeding technique in that it views analysis perturbations as filtered forecast perturbations. Instead of multiplying each forecast perturbation a global or regional rescaling factor, the ETKF ensemble generation scheme transforms forecast perturbations into analysis perturbations via linear combinations of forecast perturbations that solve the error variance update equation for the optimal data assimilation within the ensemble perturbation subspace. Consequently, the ETKF analysis perturbation is able to reflect the density and accuracy of observations. Although the breeding modulated by a mask seems able to resolve the spatial inhomogenieties of the observation distribution, the mask is only two-dimensional and time invariant. Besides, it can produce initial perturbation imbalance that excites unrealistic wave growth. Provided 12-h forecast perturbations are balanced and of linear amplitude then the ETKF analysis perturbations are balanced. In the ETKF scheme, directions corresponding to large forecast error variance in observation space are attenuated more than directions corresponding to small forecast error variance. Thus, while the error variance of a 16-member breeding ensemble is concentrated in a single direction, the ETKF 16-member ensemble error variance was spread with comparable amounts among 15 independent, orthogonal, and uncorrelated directions. The application of a mask on the breeding only slightly ameliorates the rank deficiency problem of the simple breeding method.

The fastest growth under the energy norm within the ensemble perturbation subspace of the ETKF scheme is larger than the breeding schemes. The application of the mask actually degrades the fastest growth rate of the simple breeding scheme. The ensemble mean of the ETKF ensemble is more accurate than both breeding methods. It is argued that imperfections in the correspondence between ensemble variance and forecast error variance would be indicated by (a) a relatively small range over which sample pseudoinnovation variance could be predicted from ensemble variance, (b) a high sensitivity of the relationship between ensemble variance and sample pseudoinnovation variance to the number of realizations in each bin, and (c) distributions within each bin of pseudoinnovations whose mean fourth powers (kurtosis) are relatively high (see section 8b for the definition of a bin). Investigation along these lines indicates that the ETKF estimates of forecast error variance are considerably more accurate than those of the breeding techniques. In other words, the ensemble spread of the ETKF ensemble is better able to distinguish the case-to-case forecast uncertainty than the breeding ensembles.

Besides the superior ensemble forecast skills of the ETKF scheme over the breeding schemes, the computational expense of the ETKF is also quite small ($<6\%$ more than the breeding ensemble). Thus, the ETKF ensemble generation scheme would be straightforward to employ operationally. When applied to operational forecast centers, more ensemble members and more realistic observation networks need to be included.

Acknowledgments. The authors gratefully acknowledge financial support from Pennsylvania, Power and Light, ONR Grant N00014-00-1-0106 and ONR Project Element 0601153N, Project Number BE-033-0345. It is doubtful that our integrations could have been performed without the generous and expert computational assistance of Robert Hart, Vijay Agarwala, and Jan Dutton. Frank Peacock also made important contributions at the early stages of this work. Thanks also goes to Tom Hamill and one anonymous reviewer for their patient and thoughtful reviews.

APPENDIX

Filtering Effect of Optimal Data Assimilation

For an optimal data assimilation scheme, analysis error covariance matrix $\mathbf{P}^a$ is related to the forecast error covariance matrix $\mathbf{P}^f$ and observation error covariance matrix $\mathbf{R}$ by

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f, \quad \text{(A1)}$$

where $\mathbf{H}$ is the linear observation operator that maps model variables to observed variables. Define $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}$. Then premultiply and postmultiply (A1) by $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{H}}^T$, respectively. Then (A1) becomes

$$\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T = \tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T - \tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I})^{-1}\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T, \quad \text{(A2)}$$

where $\mathbf{I}$ is the identity matrix; $\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T/\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$ is the analysis/forecast error covariance matrix in the observation subspace and normalized by the observation error covariance. To see the filtering effect, we express (A2) in terms of the eigenvectors of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$. Suppose each column of $\mathbf{E}$ is an eigenvector of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$, and $\mathbf{\Gamma}$ is a diagonal matrix containing corresponding eigenvalues of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$. Then

$$\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T = \mathbf{E}\mathbf{\Gamma}\mathbf{E}^T, \quad \text{and} \quad \text{(A3)}$$

$$(\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T + \mathbf{I})^{-1} = \mathbf{E}(\mathbf{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T. \quad \text{(A4)}$$

Using (A3) and (A4), (A2) becomes

$$\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T = \mathbf{E}[\mathbf{\Gamma} - \mathbf{\Gamma}^2(\mathbf{\Gamma} + \mathbf{I})^{-1}]\mathbf{E}^T = \mathbf{E}\mathbf{\Gamma}(\mathbf{\Gamma} + \mathbf{I})^{-1}\mathbf{E}^T. \quad \text{(A5)}$$

Comparing Eqs. (A3) and (A5), error variance on the ith eigenvector is reduced by multiplying a factor $1/(\lambda_i + 1)$, where λ_i is the ith eigenvalue. Thus, error variance in the direction of the leading eigenvector is attenuated the most. Consequently, the eigenvalue spectrum of $\tilde{\mathbf{H}}\mathbf{P}^a\tilde{\mathbf{H}}^T$ is whiter than that of $\tilde{\mathbf{H}}\mathbf{P}^f\tilde{\mathbf{H}}^T$.

REFERENCES

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

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Wea. Rev.*, **127,** 1941–1953.

Barkmeijer, J., M. van Gijzen, and F. Bouttier, 1998: Singular vectors and estimates of the analysis error covariance metric. *Quart. J. Roy. Meteor. Soc.*, **124,** 1695–1713.

——, R. Buizza, and T. N. Palmer, 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125,** 2333–2351.

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.

——, C. A. Reynolds, and M. K. Tippett, 2003: Optimization of the fixed global observing network in a simple model. *J. Atmos. Sci.*, in press.

Brier, G. W., 1950: Verification of forecasts espressed in terms of probabilities. *Mon. Wea. Rev.*, **78,** 1–3.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52,** 1434–1456.

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

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

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

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

Farrell, B. F., 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45,** 163–172.

——, 1989: Optimal exictation of baroclinic waves. *J. Atmos. Sci.*, **46,** 1193–1206.

Goedicke, V. A., 1953: *Introduction to the Theory of Statistics.* Academic Press, 286 pp.

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

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

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

——, 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.

——, C. Snyder, and J. S. Whitaker, 2003: Ensemble forecasts and the properties of flow-dependent analysis-error covariance singular vectors. *Mon. Wea. Rev.*, in press.

Houtekamer, P. L., 1995: The construction of optimal perturbations. *Mon. Wea. Rev.*, **123,** 2888–2898.

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

——, 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.

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

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

Jeffery, T. K., J. H. Hack, B. B. Gordon, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, 1996: Description of the NCAR community climate model (CCM3). NCAR Tech. Note NCAR/TN-420+STR, 152 pp.

Kalman, R., 1960: A new approach to linear filtering and predicted problems. *J. Basic Eng.*, **82,** 35–45.

——, and R. Bucy, 1961: New results in linear filtering and prediction theory. *J. Basic Eng.*, **83,** 95–105.

Kleeman, R., 1989: A modeling study of the effect of the Andes on the summertime circulation of tropical South America. *J. Atmos. Sci.*, **46,** 3344–3362.

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

Majumdar, S. J., C. H. Bishop, and B. J. Etherton, 2002: Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Mon. Wea. Rev.*, **130,** 1356–1369.

Mason, I. B., 1982: A model for the assessment of weather forecasts. *Aust. Meteor. Mag.*, **30,** 291–303.

McPhaden, M. J., and A. E. Gill, 1987: Topographic scattering of equatorial waves. *J. Phys. Oceanogr.*, **17,** 82–96.

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

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

Richardson, D. S., 2000: Skill and economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **126,** 649–669.

Sardeshmukh, P., and B. J. Hoskins, 1984: Spatial smoothing on the sphere. *Mon. Wea. Rev.,* **112,** 2524–2529.

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

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.

——, Y. Zhu, and T. Marchok, 2001: The use of ensembles to identify forecasts with small and large uncertainty. *Wea. Forecasting,* **16,** 436–477.

Whitaker, J. S., and A. F. Loughe, 1998: The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.,* **126,** 3292–3302.

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

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.,* **83,** 73–83.