LINEAR DIAGNOSTICS TO ASSESS THE PERFORMANCE OF AN ENSEMBLE FORECAST SYSTEM

A Dissertation

by

ELIZABETH A. SATTERFIELD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Atmospheric Sciences

LINEAR DIAGNOSTICS TO ASSESS THE PERFORMANCE OF AN ENSEMBLE FORECAST SYSTEM

A Dissertation

by

ELIZABETH A. SATTERFIELD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Istvan Szunyogh
Committee Members,	Gerald R. North
	Ramalingam Saravanan
	Marc G. Genton
Head of Department,	Kenneth Bowman

August 2010

Major Subject: Atmospheric Sciences

ABSTRACT

Linear Diagnostics to Assess the Performance of an Ensemble Forecast System.

(August 2010)

Elizabeth A. Satterfield, B.S., Georgia Institute of Technology;

M.S., University of Maryland

Chair of Advisory Committee: Dr. Istvan Szunyogh

The performance of an ensemble prediction system is inherently flow dependent. This dissertation investigates the flow dependence of the ensemble performance with the help of linear diagnostics applied to the ensemble perturbations in a small local neighborhood of each model grid point location ℓ. A local error covariance matrix $\mathbf{P}_\ell$ is defined for each local region and the diagnostics are applied to the linear space $\mathbb{S}_\ell$ defined by the range of the ensemble based estimate of $\mathbf{P}_\ell$. The particular diagnostics are chosen to help investigate the ability of $\mathbb{S}_\ell$ to efficiently capture the space of true forecast or analysis uncertainties, accurately predict the magnitude of forecast or analysis uncertainties, and to distinguish between the importance of different state space directions. Additionally, we aim to better understand the roots of the underestimation of the magnitude of uncertainty by the ensemble at longer forecast lead times.

Numerical experiments are carried out with an implementation of the Local Ensemble Transform Kalman Filter (LETKF) data assimilation system on a reduced (T62L28) resolution version of the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). Both simulated observations under the perfect model scenario and observations of the real atmosphere are used in these experiments. It is found that (i) paradoxically, the linear space $\mathbb{S}_\ell$ provides an increasingly better estimate of the space of forecast uncertainties as the time evolution of the

ensemble perturbations becomes more nonlinear with increasing forecast time, (ii) $\mathbb{S}_\ell$ provides a more reliable linear representation of the space of forecast uncertainties for cases of more rapid error growth, (iii) the E-dimension is a reliable predictor of the performance of $\mathbb{S}_\ell$ in predicting the space of forecast uncertainties, (iv) the ensemble grossly underestimates the forecast error variance in $\mathbb{S}_\ell$, (v) when realistic observation coverage is used, the ensemble typically overestimates the uncertainty in the leading eigen-directions of $\hat{\mathbf{P}}_\ell$ and underestimates the uncertainty in the trailing directions at analysis time and underestimates the uncertainty in all directions by the 120-hr forecast lead time, and (vi) at analysis time, with a constant covariance inflation factor, the ensemble typically underestimates uncertainty in densely observed regions and overestimates the uncertainty in sparsely observed regions.

DEDICATION

This dissertation is dedicated to my dad,

Charles Russell Satterfield

1946-2009

for his faith and inspiration

and

to my mom, whose support and guidance made this work possible.

ACKNOWLEDGMENTS

A special thanks David Kuhl, Eric Kostelich, and Gyorgyi Gyarmati for their contributions to this work. The research reported in this dissertation was funded by the National Research Foundation (Grant ATM-0935538).

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION*

Lorenz (1963a,b) showed that a nonlinear dynamical system with instabilities has a finite limit of predictability. The loss of predictability is due to small errors in initial conditions, which grow rapidly due to instabilities, and eventually lead to total loss of predictability, an effect known as "chaos". Atmospheric models are chaotic, in other words small errors in initial conditions can lead to large errors in forecasts. Since errors in initial conditions are always present, degradation of forecasts with increasing lead times is an inherent property of atmospheric forecasts, and inevitably results in a complete loss of predictability (Kalnay, 2002). The average limit of atmospheric predictability is generally accepted to be about 10-14 days. However, forecast skill is also dependent on the atmospheric evolution itself, some forecast features can remain predictable for longer than others.

One way to account for the chaotic nature of models is the use of ensemble forecasting techniques. Ensembles of forecasts are started from a perturbed set of initial conditions, representative of the uncertainty at analysis time. The evolved ensemble is meant to provide a representation of uncertainty in space and time. These

---

The journal model is *AMS Monthly Weather Review.*

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Space of Uncertanties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, 138, 962-981,
and
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Magnitude and the Spectrum of Incertainties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, in review,

forecasts started from slight differences in the analysis may be dramatically different or somewhat similar, dependent on the background atmospheric flow. The evolved ensemble can provide guidance about forecast reliability: if the ensemble members evolve very differently, one would be less confident in the forecast than for the case that the ensemble members show good agreement. "Spaghetti plots" used to display operational ensemble products frequently show high predictability in some areas, where the evolved ensemble members remain similar to each other, low predictability in other areas, in which the evolved ensemble members significantly differ (Kalnay, 2002). The mean of the evolved ensemble provides the best forecast in the root-mean-squared error sense at all forecast lead times. By taking the ensemble mean at forecast time, one can effectively filter uncertain features in the ensemble. The ensemble spread (standard deviation of the ensemble) is often assumed to provide a local estimate of the forecast uncertainty. Although this relationship, known as the "spread-skill relationship" has not been firmly established, it is often used to provide guidance for human forecasters (Kalnay, 2002). In general, the ensemble provides a basis for the prediction of the full probability distribution, not only for the first two statistical moments (mean and standard deviation.) Since the early 1990s most major weather prediction centers have implemented ensemble prediction systems to account for the influence of the spatiotemporal changes in predictability on the forecasts [e.g., Kalnay (2002); Palmer and Hagedorn (2006)].

Kuhl et al. (2007), KEA07 hereafter, showed that the spatiotemporal changes in the predictability make the performance of an ensemble prediction system of a finite number of ensemble members inherently flow dependent. Thus, a uniformly good performance of the system over all weather situations cannot be expected. We refer to this flow dependence of the performance of the finite size ensemble as the local predictability of the ensemble performance. The study of KEA07 was based on as-

similating randomly located simulated observations under the perfect model scenario with an implementation of the Local Ensemble Transform Kalman Filter (LETKF) data assimilation system (Hunt et al., 2006; Szunyogh et al., 2008) on a reduced resolution (T62 and 28 vertical levels) version of the model component of the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). Here, we extend the investigation of KEA07, building to a more realistic setting, by first assimilating simulated observations in realistic locations under a perfect model scenario and then assimilating an operationally used set of observations of the real atmosphere.

The main goal of our study is to lay the theoretical foundation of a practical approach to predict the spatiotemporal changes in the performance of an ensemble prediction system. In particular, we define a local volume around each grid point ℓ and define a local state vector that represents the model state in the local volume. Then, we investigate the efficiency of the linear space $\mathbb{S}_\ell$, defined by the range of the ensemble based estimate of the local covariance matrix $\mathbf{P}_\ell$ for the components of the local state vector, in capturing the true forecast uncertainties. Further, we assess the quality of the prediction of the magnitude of the analysis and forecast uncertainties and of the skill of the ensemble in distinguishing between the importance of the different state space directions within $\mathbb{S}_\ell$. The motivation to utilize the local approach here is twofold: first, it provides a natural framework to study the spatial changes in the ensemble performance; second, the experience with ensemble-based Kalman filter data assimilation schemes strongly suggest that an ensemble of practically attainable size can provide a sufficiently accurate estimate of the covariance between different components of the state vectors, at short forecast times, only in localized regions [e.g., Houtekamer and Mitchell (2001); Szunyogh et al. (2005); Anderson (2007); Whitaker and Hamill (2002); Hamill et al. (2001)].

In this study, we use the LETKF algorithm to generate the ensemble initial conditions to investigate the ability of the ensemble to capture the local structure of the forecast errors, and to predict the magnitude and spectrum of uncertainties; however, we expect our results to remain valid for any suitably formed ensemble based Kalman filter scheme. Since we do not attempt to account for the effect of model errors in our formulation of the ensemble, our results should not be used directly to interpret the behavior of an ensemble system that also employs model perturbations, e.g. the ensemble prediction system of the European Center for Medium Range Forecasts (ECMWF, Berner et al. 2009). We believe, however, that our diagnostic approach could be used to validate existing model perturbation schemes.

The structure of this dissertation is as follows. In Chapter II, we introduce the diagnostics we use to assess and explain the performance of the ensemble prediction system at the different locations and times. In Chapter III, we describe the design of the numerical experiments. In Chapter IV, we examine the spatiotemporal evolution of the forecast errors, which is our preferred way to assess the spatiotemporal evolution of predictability, and we analyze the relationship between predictability and our diagnostics. In Chapter V we examine the ability of the ensemble to accurately capture the magnitude and spectrum of forecast uncertainties. In Chapter VI, we summarize our conclusions.

CHAPTER II

DIAGNOSTICS*

We use linear diagnostics applied to the ensemble perturbations in a small local neighborhood of each model grid point to explore the spatio-temporally changing predictive qualities of the ensemble. In particular, we define a local state vector and the associated local covariance matrix to represent the state and the uncertainty in the state estimate at each grid point. In addition, we introduce a set of local diagnostics based on the eigen-solution of the local covariance matrix and a measure of nonlinearity in the evolution of the local state vectors.

A. Local vectors and their covariance

We define a local state vector $\mathbf{x}(\ell)$ with all N state variables of the model representation of the state within a local volume centered at location (grid point) ℓ. For the rest of this dissertation, we will discuss what to do at an arbitrary location ℓ, and so we now drop the argument ℓ from the notation of the local state vectors. The mathematical model we adopt to predict the evolution of uncertainty in a local state estimate (analysis or forecast), $\mathbf{x}^e$, is based on the assumption that the error in the state estimate,

$$\boldsymbol{\xi} = \mathbf{x}^e - \mathbf{x}^t, \tag{2.1}$$

---

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Space of Uncertanties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, 138, 962-981, © Copyright 2010 American Meteorological Society (AMS)
and
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Magnitude and the Spectrum of Incertainties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, in review, © Copyright 2010 American Meteorological Society (AMS)

is a random variable. In Equation (2.1) $\mathbf{x}^t$ is the model representation of the, in practice unknown, true state of the atmosphere. The covariance between the different components of $\boldsymbol{\xi}$ is described by the error covariance matrix $\mathbf{P}_\ell$.

We employ a K-member ensemble of local state estimates, $\mathbf{x}^{e(k)},\ k = 1 \ldots K$, to predict the uncertainty in the knowledge of the local state. The ensemble-based estimate of the covariance matrix $\mathbf{P}_\ell$ is

$$\hat{\mathbf{P}}_\ell = (K-1)^{-1} \sum_{k=1}^{k} \mathbf{x}'^{(k)} \left(\mathbf{x}'^{(k)}\right)^T, \tag{2.2}$$

where the ensemble perturbations $\mathbf{x}'^{(k)}$, k=1...K, are defined by the difference

$$\mathbf{x}'^{(k)} = \mathbf{x}^{e(k)} - \bar{\mathbf{x}},\ k = 1 \ldots K, \tag{2.3}$$

between the ensemble members $\mathbf{x}^{(k)},\ k = 1 \ldots K$, and the ensemble mean,

$$\bar{\mathbf{x}} = K^{-1} \sum_{k=1}^{K} \mathbf{x}^{e(k)}. \tag{2.4}$$

In Equation (2.2), T denotes the matrix transpose. The linear space $\mathbb{S}_\ell$ defined by the range of $\hat{\mathbf{P}}_\ell$ is spanned by the K ensemble perturbations. (We use the subscript ℓ in the notation $\mathbf{P}_\ell$, $\hat{\mathbf{P}}_\ell$, and $\mathbb{S}_\ell$ to emphasize that the linear space we are interested in is defined for the local neighborhood of each grid point.) Based on Equations (2.3) and (2.4) the sum of the ensemble perturbations is zero at all forecast lead times, that is,

$$\sum_{k=1}^{K} \mathbf{x}'^{(k)} = \mathbf{0}. \tag{2.5}$$

Equation (2.5) indicates that the K ensemble perturbations are not linearly independent. Thus, the dimension of the linear space $\mathbb{S}_\ell$ cannot be larger than *K-1.*

To obtain a convenient orthonormal basis in $\mathbb{S}_\ell$ for the definition and computation of our diagnostics, we compute the eigen-solution of $\hat{\mathbf{P}}$. Since $\hat{\mathbf{P}}$ is a non-

negative definite and symmetric N by N matrix, it has N non-negative eigenvalues, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r \ldots \geq \lambda_N \geq 0$, and the N associated eigenvectors, $\mathbf{u}_n$, $n = 1, \ldots, N$, are orthogonal with respect to the Euclidean inner product. That is, when the eigenvectors are chosen to be of unit length with respect to the Euclidean vector norm,

$$\left(\mathbf{u}_i\right)^T \mathbf{u}_j = \delta_{ij}, \tag{2.6}$$

where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. When the number of components of the local state vector is larger than the number of ensemble members ($N > K$), only the first $K - 1$ eigenvalues can be larger than zero. (In what follows, $N > K$ is assumed unless noted otherwise.) In this case, the normalized eigenvectors associated with the first $K - 1$ eigenvalues, $\mathbf{u}_k$, $k = 1, \ldots, K - 1$, define an orthonormal basis in $\mathbb{S}_\ell$. The physical interpretation of the N-vectors $\mathbf{u}_k$, $k = 1, \ldots, K - 1$, is that they represent linearly independent patterns of uncertainty in the ensemble perturbations in the local region at ℓ.

An arbitrary local state vector $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = \bar{\mathbf{x}} + \delta\mathbf{x}, \tag{2.7}$$

where $\delta\mathbf{x}$ is the difference between $\mathbf{x}$ and the ensemble mean $\bar{\mathbf{x}}$. The perturbation vector $\delta\mathbf{x}$ can be further decomposed as,

$$\delta\mathbf{x} = \delta\mathbf{x}^{(\|)} + \delta\mathbf{x}^{(\perp)}, \tag{2.8}$$

where $\delta\mathbf{x}^{(\|)}$ is the component that projects into $\mathbb{S}_\ell$ that is,

$$\delta\mathbf{x}^{(\|)} = \sum_{k=1}^{K-1} \delta x_k^{(\|)} \mathbf{u}_k, \tag{2.9}$$

where the coordinate $\delta x_k^{(\|)},\ k = 1 \ldots K-1$, can be computed by

$$\delta x_k^{(\|)} = \delta \mathbf{x}^T \mathbf{u}_k. \tag{2.10}$$

The vector $\delta \mathbf{x}^{(\perp)}$ is the component of $\delta \mathbf{x}$ that projects into the null space of $\hat{\mathbf{P}}_\ell$. That is, $\delta \mathbf{x}^{(\perp)}$ cannot be represented by the ensemble perturbations. Using this notation, the error $\boldsymbol{\xi}$ in the state estimate $\mathbf{x}^e$ can be decomposed as

$$\boldsymbol{\xi} = \mathbf{x}^e - \mathbf{x}^t = \sum_{k=1}^{K-1} \left[(\delta \mathbf{x}^e - \delta \mathbf{x}^t)^T \mathbf{u}_k \right] \mathbf{u}_k + \left(\delta \mathbf{x}^{e(\perp)} - \delta \mathbf{x}^{t(\perp)} \right) = \delta \boldsymbol{\xi}^{(\|)} + \delta \boldsymbol{\xi}^{(\perp)}, \tag{2.11}$$

where,

$$\delta \boldsymbol{\xi}^{(\|)} = \sum_{k=1}^{K-1} \left[(\delta \mathbf{x}^e - \delta \mathbf{x}^t)^T \mathbf{u}_k \right] \mathbf{u}_k, \quad \delta \boldsymbol{\xi}^{(\perp)} = \delta \mathbf{x}^{e(\perp)} - \delta \mathbf{x}^{t(\perp)}. \tag{2.12}$$

Although the ensemble mean, or the error in the ensemble mean, does not appear directly in the local decomposition of the error (the rhs of Eq 2.11), the ensemble mean provides the reference point for the definition of the basis vectors that span the space $\mathbb{S}_\ell$.

The main focus of our investigation in this dissertation is the linear space $\mathbb{S}_\ell$. We choose diagnostics and design numerical experiments to identify the conditions under which the evolution of the forecast uncertainties can be efficiently described in $\mathbb{S}_\ell$. We emphasize that an ensemble forecast system can, in principle, describe the evolution of the forecast uncertainty even if it is nonlinear. In that case, however, it is not guaranteed that the probability distribution of the uncertainty is Gaussian. In such a case the knowledge of the mean and the covariance matrix may not be sufficient to fully describe the probability distribution. We still restrict our attention to $\mathbb{S}_\ell$, because our objective is to study spatio-temporal changes in the performance of the ensemble. This requires a measure of performance that can be computed for an arbitrary time and location.

B. Diagnostics for the space of the uncertainties

1. Explained variance

Our choice of this measure is the explained variance, which measures the projection of the analysis and forecast errors onto $\mathbb{S}_\ell$. Formally, the explained variance is calculated as

$$EV = \frac{\|\delta\boldsymbol{\xi}^{(\|)}\|}{\|\boldsymbol{\xi}\|} = \frac{\|\delta\boldsymbol{\xi}^{(\|)}\|}{\|\delta\boldsymbol{\xi}^{(\|)} + \delta\boldsymbol{\xi}^{(\perp)}\|}. \tag{2.13}$$

Here $\|\cdot\|$ is the Euclidean vector norm on the space of the local state vectors. (Since $\mathbb{S}_\ell$ is a subspace of the space of the local state vectors, this norm can be used to measure the magnitude of both the error and its projection into $\mathbb{S}_\ell$.)

The larger EV, the more efficient $\mathbb{S}_\ell$ in capturing the uncertain components of the analysis and forecast fields. EV takes its maximum value of one when the entire forecast error projects into $\mathbb{S}_\ell$ ($\delta\boldsymbol{\xi}^{(\|)} = \boldsymbol{\xi}$ and $\delta\boldsymbol{\xi}^{(\perp)} = \mathbf{0}$), and takes its minimum value of zero when the forecast error does not have projection into $\mathbb{S}_\ell$ ($\delta\boldsymbol{\xi}^{(\|)} = \mathbf{0}$ and $\delta\boldsymbol{\xi}^{(\perp)} = \boldsymbol{\xi}$). Our definition of explained variance is similar to the Perturbation versus Error Correlation Analysis (PECA) diagnostic defined in Wei and Toth (2003) when PECA is calculated using an optimally combined perturbation vector. The main difference between the two diagnostics is that we calculate the explained variance using local regions as opposed to the global domain used in Wei and Toth (2003). Finally, we emphasize that the explained variance measures one specific aspect of the performance of the ensemble system and a good performance with respect to the explained variance does not guarantee good performance with respect to another verification score.

2. E-dimension

The ensemble dimension (E-dimension),

$$E = \frac{\left[\sum_{i=1}^{K} \sqrt{\lambda_i}\right]^2}{\sum_{i=1}^{K} \lambda_i}, \tag{2.14}$$

which was introduced by Patil et al. (2001) and discussed in details in Oczkowski et al. (2005), characterizes the local complexity of dynamics. E is a spatio-temporally evolving measure of the steepness of the eigenvalue spectrum, $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_r \ldots \geq \lambda_K$, having smaller values for a steeper spectrum (Szunyogh et al., 2007). For our choice of the perturbations, where the K perturbations are linearly dependent, $\lambda_K = 0$, the largest possible value of E is *K-1*. For a set of linearly independent ensemble perturbations, the maximum value of E is equal to the number of ensemble perturbations, K, which occurs when the uncertainty predicted by the ensemble is evenly distributed between K linear spatial patterns in $\mathbb{S}_\ell$.

3. Linearity of the local dynamics

The period of time in which model dynamics can be approximated as linear is commonly assumed to be 2-3 days (e.g., Palmer et al. 1994). Gilmour et al. (2001), using an objective measure to quantify the importance of nonlinear effects on the dynamics called *relative nonlinearity*, argued that the assumption of linearity may not be valid for longer than 24 hours. To define the relative linearity measure, Gilmour et al. (2001) considered the evolution of twin pairs of ensembles members, which were obtained by adding the same ensemble perturbation to the analysis with both a positive and a negative sign. We introduce a measure of linearity that is motivated by that of Gilmour et al. (2001).

The evolution of the kth ensemble member, $\mathbf{x}_g^{e(k)}$ is governed by the equation

$$\mathbf{x}_g^{e(k)} = \mathbf{F}(\bar{\mathbf{x}}_g^a + \delta\mathbf{x}_g^{a(k)}) = \mathbf{F}(\bar{\mathbf{x}}_g^a) + \mathbf{L}\delta\mathbf{x}_g^{a(k)} + \mathbf{N}(\bar{\mathbf{x}}_g^a, \delta\mathbf{x}_g^{a(k)}), \tag{2.15}$$

where $\mathbf{L}$ is a linear operator and $\mathbf{N}$ is a nonlinear function of the analysis perturbation $\delta\mathbf{x}_g^{a(k)}$ and the analysis $\bar{\mathbf{x}}_g^a$. The subscript g indicates that Equation (2.15) is for global state vectors instead of the local state vectors we consider elsewhere in this dissertation. Equation (2.15) is based on the Taylor expansion of the dynamics $\mathbf{F}$ about $\bar{\mathbf{x}}^a$ in the direction of the analysis perturbation $\delta\mathbf{x}^{a(k)}$. Also, when $\mathbf{x}^{e(k)}$ refers to an ensemble member at analysis time, $\mathbf{F}$ and $\mathbf{L}$ are the identity and $\mathbf{N}$ is zero. Since

$$\mathbf{x}_g^e = \mathbf{F}(\bar{\mathbf{x}}^a) \tag{2.16}$$

and

$$\bar{\mathbf{x}}_g^e = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_g^{e(k)}, \tag{2.17}$$

taking the ensemble mean of the two sides of Equation (2.15), we obtain

$$\bar{\mathbf{x}}_g^e = \mathbf{x}_g^e + \frac{1}{K}L\sum_{k=1}^{K}\delta\mathbf{x}_g^{a(k)} + \frac{1}{K}\sum_{k=1}^{K}\mathbf{N}(\mathbf{x}_g^a, \delta\mathbf{x}_g^{a(k)}). \tag{2.18}$$

Applying the global equivalent of Equation (2.5),the second term on the rhs. of Equation (2.18) is zero. Thus, introducing the notation

$$\mathbf{e} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{N}(\mathbf{x}_g^a, \delta\mathbf{x}_g^{a(k)}) \tag{2.19}$$

for the ensemble mean of the nonlinearly evolving component of the forecast ensemble members, Equation (2.18) yields

$$\mathbf{e} = \bar{\mathbf{x}}_g^e - \mathbf{x}_g^e. \tag{2.20}$$

We define the *local relative nonlinearity* measure ρ_ℓ as the ratio between the magni-

tude of $\mathbf{e}$ and the ensemble average of the magnitude of the ensemble perturbations for the local state vectors, that is,

$$\rho_\ell = \frac{\|\bar{\mathbf{x}}^e - \mathbf{x}^e\|}{\frac{1}{K}\sum_{k=1}^{K}\|\delta\mathbf{x}^{e(k)}\|}. \tag{2.21}$$

(Notice that Equation (2.21) is based on local state vectors.) One important difference between our measure and that of Gilmour et al. (2001) is that instead of a pair of perturbations, we consider a K-member ensemble. We make this choice because the K-member analysis ensemble generated by the LETKF algorithm is not composed of pairs, but has the property that the sum of the ensemble perturbations is zero at analysis time. We also note that our definition is based on local state vectors and also applied to global state vectors for comparison. The study of Gilmour et al. (2001) used global state vectors, although their measure, in principal, could be applied locally as well.

C. Diagnostics for the magnitude of the uncertainties

In section B, we focused on investigating the efficiency of the linear space $\mathbb{S}_\ell$ defined by the range of $\hat{\mathbf{P}}_\ell$ in capturing the error in the deterministic prediction, $\mathbf{x}^e$ started from the ensemble mean analysis $\bar{\mathbf{x}}^a$. In this section, our goal is to investigate (i) the accuracy of the ensemble prediction of the expected value of the magnitude of the uncertainty and (ii) the accuracy of the ensemble prediction of the spectrum of uncertainties within $\mathbb{S}_\ell$. To achieve this goal, we apply diagnostics to the

$$\delta\mathbf{x}^t = \mathbf{x}^t - \bar{\mathbf{x}} \tag{2.22}$$

difference between the model representation of the true state and the ensemble mean state estimate, instead of $\boldsymbol{\xi}$. The difference $\delta\mathbf{x}^t$ is often referred to in the literature

as the error in the ensemble mean forecast. This terminology is justified when the ensemble mean is used as a deterministic forecast, which is motivated by the fact that the mean of a perfectly designed ensemble would be the most accurate deterministic forecast in the root-mean-square sense (Leith, 1974). In our study, however, we consider $\bar{\mathbf{x}}$ to be the prediction of the mean of a probability distribution. Since, except for the analysis time, $\delta\mathbf{x}^t$ is expected to be nonzero even if $\bar{\mathbf{x}}$ is a perfect prediction of the mean of the probability distribution, we refer to $\delta\mathbf{x}^t$ as either the difference between the ensemble mean and the model representations of the true state or the local forecast uncertainty.

The motivation to apply diagnostics to $\delta\mathbf{x}^t$, instead of $\boldsymbol{\xi}$, is that a verifiable optimality condition between $||\delta\mathbf{x}^t||$, the magnitude of $\delta\mathbf{x}^t$, and the ensemble variance $V_\ell = trace(\hat{\mathbf{P}}_\ell)$ exists at all forecast times: because V_ℓ is a prediction of the variance $TV_\ell = E[(\delta\mathbf{x}^t)^2]$, where $(\delta\mathbf{x}^t)^2 = (\delta\mathbf{x}^t)^T(\delta\mathbf{x}^t) = ||\delta\mathbf{x}^t||^2$, its expected value, $V = E[V_\ell]$, should satisfy the equation

$$V = TV, \tag{2.23}$$

where $TV = E[TV_\ell]$. (Hereafter, $E[\cdot]$ denotes the expected value.) Verifying the relationship defined by Eq. (2.23), which is often referred to as the spread-skill relationship, is one of the most widely used diagnostics for the validation of an ensemble prediction system. In contrast, all we know about the magnitude $||\boldsymbol{\xi}||$ of $\boldsymbol{\xi}$ is that it should satisfy $E[\boldsymbol{\xi}^2] = V$ ($\boldsymbol{\xi}^2 = \boldsymbol{\xi}^T\boldsymbol{\xi} = ||\boldsymbol{\xi}||^2$) at analysis time and, under an ergodic hypothesis, $E[\boldsymbol{\xi}^2] = 2V$ once the forecast time is so long that predictability is completely lost (Leith, 1974). That is, no verifiable diagnostic relationship exists between $\boldsymbol{\xi}^2$ and V at the intermediate forecast lead times. In addition, the first $K-1$

(nonzero) eigenvalues of $\hat{\mathbf{P}}_\ell$ satisfy the equation

$$V = \sum_{k=1}^{K-1} \lambda_k. \tag{2.24}$$

We introduce the notation $TVS = E[(\delta\mathbf{x}^{t(\|)})^2]$ for the portion of the variance that can be explained by $\mathbb{S}_l$. In the optimal case, $\delta\mathbf{x}^t$ would fully project onto $\mathbb{S}_\ell$, satisfying $\delta\mathbf{x}^t = \delta\mathbf{x}^{t(\|)}$, leading to $TVS = TV$. But, when part of $\delta\mathbf{x}^t$ is not captured by the ensemble, $\mathbf{x}^{t(\perp)} \neq \mathbf{0}$, which leads to $TVS < TV$.

When the ensemble correctly represents the variance of $\delta\mathbf{x}^{t(\|)}$, $V = TVS$. For a given ensemble system, V can be either smaller or larger than TVS. In the former case ($V < TVS$) the ensemble underestimates the manitude of the uncertainty that can be explained by $\mathbb{S}_\ell$, while in the later case the ensemble overestimates the magnitude that can be explained by $\mathbb{S}_\ell$. It may even happen that the ensemble satisfies the optimality condition of Eq. (2.23) for the wrong reason, overestimating the true variance in $\mathbb{S}_\ell$ to compensate for the variance lost by not capturing all true error directions. This situation occurs when the ensemble variance is tuned to satisfy Eq. (2.23) at a given forecast time (e.g., 48-hr), but the ensemble cannot fully capture $\delta\mathbf{x}^t$. Such a situation can be diagnosed by verifying that $V > TVS$. Analyzing the results of our numerical experiment we always make a three-way comparison between TV, TVS, and V.

D. Diagnostics for the spectrum of uncertainties

While the explained variance diagnostic of section B quantifies the efficiency of the space $\mathbb{S}_\ell$ in capturing the space of uncertainty in the state estimate, the comparison of TV, TVS and V quantifies the quality of the ensemble in predicting the magnitude of the uncertainty. These diagnostics, however, do not provide information about the performance of the ensemble in *distinguishing between the relative importance of*

the different error patterns within $\mathbb{S}_\ell$. To introduce a diagnostic that can measure the performance of the ensemble in quantifying the contributions of the different error patterns to the total error within $\mathbb{S}_\ell$, we first recall that the eigenvalue λ_k, is the ensemble-based prediction of the variance of the uncertainty in the k-th eigendirection [1]. We choose the *d-ratio*,

$$\left\{ d_k = \frac{\left(\delta x_k^t\right)^2}{\lambda_k} : \qquad k = 1, \ldots K-1 \right\}, \tag{2.25}$$

which was first introduced in Ott et al. (2002), to measure the accuracy of the prediction of the variance in the k-th eigendirection. In Eq. (2.25), $\delta x_k^t = (\delta \mathbf{x}^t)^T \mathbf{u}_k$ is the k-th coordinate of $\delta \mathbf{x}^{t(\|)}$ in the coordinate system $\{\mathbf{u}_k : k = 1, \ldots, K-1\}$. Since the d-ratio is defined independently for each eigendirection, it is more appropriate to talk about a *spectrum of the d-ratio*. It can be shown that if the ensemble correctly predicts, in a statistical sense, the uncertainty in the k-th direction, the expected value, $E[d_k]$, of d_k at a given time and location is equal to one:

$$E\big[d_k\big] = E\left[\frac{\left(\delta x_k^t\right)^2}{\lambda_k}\right] = E\left[\frac{\left(\delta x_k^t\right)^2}{E\Big[\left(\delta x_k^t\right)^2\Big]}\right] = \frac{1}{E\Big[\left(\delta x_k^t\right)^2\Big]} E\Big[\left(\delta x_k^t\right)^2\Big] = 1, \tag{2.26}$$

In Eq. (2.26) we made use of the fact that λ_k is the ensemble based prediction of $E\Big[\left(\delta x_k^t\right)^2\Big]$, thus, for a correct prediction the two quantities must be equal. Since we have a verifiable optimality condition only for the expected value of d_k, we cannot use d_k to measure the performance of the ensemble at a given time and location. Instead, we collect statistical samples of d_k and obtain estimates of the expected value by computing the sample means. A sample mean smaller than one indicates that the ensemble tends to overestimate the uncertainty in the k-th direction, while

[1] Graphically, the vectors $\sqrt{\lambda_k}\mathbf{u}_k$ are the principal axes of the ellipsoid defined by $(\mathbf{x} - \overline{\mathbf{x}})^T (\mathbf{P})^{-1} (\mathbf{x} - \overline{\mathbf{x}}) = 1$. This ellipsoid represents states of equal probability in $\mathbb{S}_\ell$.

a sample mean larger than one indicates that the ensemble tends to underestimate the uncertainty in the k-th direction.

Finally, we note that the condition $E[d_k] = 1$ is always satisfied when the random variable,

$$d_k^{1/2} = \frac{\delta x_k^t}{\sqrt{\lambda_k}} \tag{2.27}$$

has an expected value equal to zero, $E[d_k^{1/2}] = 0$, and a variance equal to one, $E[d_k^{1/2} - E[d_k^{1/2}]] = 1$, since

$$E[d_k] = E^2[d_k^{1/2}] + E^2[d_k^{1/2} - E[d_k^{1/2}]]. \tag{2.28}$$

A random variable similar to $d_k^{1/2}$ was first introduced for the verification of the ensemble forecast of a scalar variable by Talagrand et al. (1999) and was later named the *reduced centered random variable* (RCRV) by Candille and Talagrand (2005) and Descamps and Talagrand (2007). The difference between $d_k^{1/2}$ and the RCRV is that while RCRV is for a scalar atmospheric state variable, $d_k^{1/2}$ is for a vector, the local state vector, and a spectrum of scalar ratios is obtained by first projecting the centered local state vector on the principal components of the ensemble-based estimate of the background error covariance matrix and then performing the reduction by the ensemble spread only in that particular eigendirection.

CHAPTER III

EXPERIMENT DESIGN*

We carry out numerical experiments both under the perfect model scenario and in a realistic NWP setting. In the perfect model experiments, we generate simulated observations of the hypothetical "true" trajectory of the atmospheric state, where the time series of "true" states, $\mathbf{z}$, is generated by a 60-day model integration of the GFS model at T62L28 resolution starting from an operational NCEP analysis truncated to T62L28 resolution. We first repeat the experiment of KEA07 to verify that the findings of that paper remain valid for the different time period investigated here (January and February of 2004 instead of January and February of 2000).[1] Then, we build to the realistic NWP setting in two steps: first we replace the randomly located simulated observations by simulated observations taken by a realistic observing network, then we replace the simulated observations with observations of the real atmosphere. On the one hand, comparing results from the two perfect model experiments, we can detect features, which are caused by the spatiotemporal inhomogeneities in a realistic observing network. On the other hand, comparing the results from the experiment

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Space of Uncertanties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, 138, 962-981, © Copyright 2010 American Meteorological Society (AMS)
and
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Magnitude and the Spectrum of Incertainties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, in review, © Copyright 2010 American Meteorological Society (AMS)

[1]Another important difference between the experiment design of the two studies is that we use a later version of the LETKF. Most importantly, the LETKF used in this study provides more accurate analyses in the polar regions. We also note that the current implementation of the LETKF defines the local region in terms of distance rather than by model grid points. Since we generate randomly placed simulated observations to cover 10% of the model grid, as in KEA, our experiment considers a greater number of observations for analyses at higher latitudes.

that assimilates observations of the real atmosphere to the results of the experiment that uses identically distributed simulated observations, we can detect features that are due to model errors.

A. Observational data sets

1. Randomly placed simulated observations

The "truth", $\mathbf{z}$, is taken to be an integration of the GFS model starting from the operational NCEP analysis at 0000 UTC 1 January 2004. At each grid point and model level, we generate simulated observations of the two horizontal components of the wind, the temperature, and the surface pressure by perturbing the "true" states with normally distributed, zero mean assumed observational errors with standard deviations of 1 K, 1.1 m/s, and 0.6 hPa for temperature, wind, and surface pressure, respectively. Next, similar to Szunyogh et al. (2005) and KEA07, we randomly choose 2000 soundings, to reflect a 10% observational coverage of the model grid. By choosing observations randomly, we ensure that the simulated observing network has little systematic impact on the geographical distribution of analysis and forecast errors.

a. Simulated observations at realistic locations

In the second set of experiments, we assimilate simulated observations at the locations of routine non-radiance observations of the real atmosphere. These simulated observations are generated by adding random observational noise, created by using the standard deviation of the estimated observational error provided with each observation by NCEP, to the "true" grid point values of the surface pressure, the temperature, and the two horizontal components of the wind vector. The location and type of observations is obtained from a database that includes all nonradiance

observations operationally assimilated at NCEP between 000UTC 1 January 2004 and 000UTC 15 February 2004, with the exception of satellite radiances, but including satellite derived winds. We also exclude all surface observations, except for the surface pressure and the scatterometer wind measurements over oceans.

b. Observations of the real atmosphere

Finally, the observations of the real atmosphere, which are used to obtain the type and location for the simulated observations at realistic locations, are assimilated.

B. Selection of the LETKF parameters

For each observational data set, an analysis is obtained at the native model resolution every 6 hours. Diagnostics are computed at a reduced $2.5^\circ \times 2.5^\circ$ grid resolution. We assimilate observations between 1 January 2004 0000 UTC and 15 February 2004 0000 UTC. In these experiments, we apply a multiplicative covariance inflation at each analysis step to increase the magnitude of the estimated analysis uncertainty to compensate for the loss of ensemble variance due to sampling errors, the effects of nonlinearities and model errors. In essence, the covariance inflation factor controls the magnitude of the analysis ensemble perturbations. In our code, the covariance inflation factor, $\rho = \rho(\sigma, \varphi)$, is a function of the model vertical coordinate σ and the geographical latitude φ. That is, ρ is constant in the zonal direction at a given model level and latitude. We tuned the covariance inflation factor independently for the three experiments, trying to ensure that the ensemble of analysis perturbations satisfies the condition $V \approx TVS$ in each experiment.

The parameters of the LETKF used in this experiment are the following:

- The ensemble has $K = 40$ members.

- Observations are considered for assimilation in a 800 km horizontal radius of the grid point, where the state is estimated.

- Observations have equal weight within a 500 km radius of the given grid point, beyond which the weight of the observations tapers linearly to zero at 800 km.

- Observations are considered in a vertical patch radius centered at the grid point. This layer has depth 0.35 scale height between model levels 1 to 15 and gradually increases to 2 at the top of the model atmosphere.

- Surface pressure is assimilated at the first model level and temperature, and zonal and meridional winds are assimilated at all 28 model levels.

- For simulated randomly distributed observations, we use a 25% covariance inflation at all vertical levels in all geographic regions. For the simulated observations taken at realistic locations, the covariance inflation is 2.5% at all vertical levels in the SH extratropics and 10% in the NH extratropics. In the Tropics, the covariance inflation varies from 2.5% to 7.5%. For the conventional observations of the real atmosphere, the covariance inflation tapers from 50% at the surface to 40% at the top of the model atmosphere in the SH extratropics and from 100% to 60% in the NH extratropics, and changes smoothly in the tropics (between 25°S and 25°N) from the values of the SH extratropics to the values of the NH extratropics.

C. Initialization

In the two sets of experiments which assimilate observations in realistic locations, high-frequency oscillations (typically associated with gravity waves) are filtered from all background ensemble members with a digital filter scheme (Huang and Lynch,

1993), which is part of the NCEP GFS model and can be turned on or off by choice. (Unlike in the original formulation of the digital filter algorithm, where a filtered analysis is produced, the NCEP filter provides only a filtered background field.) We use the filter with a 3 hr cutoff frequency. We find that turning the digital filter on in these two sets of experiments leads to a major improvement of the analyses.

In the experiments with randomly placed observations, turning the digital filter on degrades the analysis in the Tropics (Figure 1). More precisely, the surface pressure errors with the digital filter turned on (top panel of Figure 1) have a clear wavenumber two pattern in the tropics. A more careful examination of the structure of the error fields reveals that the digital filter wipes out the *semidiurnal tidal wave* [2]. We illustrate this effect of the filter by showing the spectrum of Fourier amplitudes of the time series of surface pressure at the location 0°N, 160°W for the nature run and the analyses prepared with and without the use of the digital filter (Figure 2). The 12-hr frequency oscillation characteristic of the semidiurnal tidal wave is not present in the run that uses the digital filter, even though this oscillation is the dominant signal in the nature run and in the analysis cycles that do not use the digital filter. We note that the digital filter initialization also has a negative effect on the analysis of the semidiurnal tidal wave in the two experiments that assimilate observations at realistic locations. In those experiments, however, the problem does not get exposed because the beneficial effect of the filter from removing spurious gravity waves, that are not present in the experiment based on uniformly distributed observations, outweighs the degradation from wiping out the semidiurnal tidal wave.

The semidiurnal tidal wave is primarily caused by the absorption of solar ra-

[2]The possibility that the error field in the top panel of Figure 1 may be associated with the semidiurnal tidal wave signal was first pointed out to us by Nedjeljka Zagar of the University of Ljubljana.

diation by ozone in the stratosphere and the atmosphere. The response to this stratospheric excitation propagates downward in the form of an inertia-gravity wave (Chapman and Lindzen, 1970). Our conjecture is that the digital filter affects this inertia-gravity wave. We also suspect that applying a digital filter initialization to the analysis increment instead of filtering the 6-hr background forecast, which is the general practice for variational data assimilation schemes, would eliminate the negative effect of the filter on the semi-diurnal tidal wave as suggested by Sankey et al. (2007) [3].

D. Forecasts

We prepare the deterministic forecasts daily, started from the mean analysis at 0000UTC and 1200UTC, and output every 12 hours. These model integrations provide the state estimate $\mathbf{x}^{a,f}$. At analysis time and at short forecast lead times (while the time evolution of the ensemble perturbations stays linear), this state estimate provides our best deterministic estimate of the state. At longer lead times $\mathbf{x}^{a,f}$ simply represents a forecast for which the analysis was drawn from a probability distribution that is consistent with our estimate of the analysis uncertainty.

In addition to the state estimate, the LETKF also generates an ensemble of analyses to estimate the uncertainty in the state estimate. These analyses serve as initial conditions for the ensemble of forecasts. Ensemble forecasts are obtained once daily, started from the ensemble of analyses, at 0000UTC and output every 12 hours. Both the deterministic forecast and the ensemble forecasts are carried out to a five day lead time. Unlike the experiments which use realistically placed observations, forecasts for the experiment that assimilates observations in random locations are

[3] We are currently in the process of developing such an initialization algorithm for the LETKF scheme.

run without the use of the digital filter. We note that turning the digital filter off in this experiment slightly increases the forecast error up to 12-hr lead times, after which the filter has no effect on the forecast errors.

Forecast error statistics are computed by comparing the deterministic forecasts, $\mathbf{x}^{a,f}$, to the "true" states, $\mathbf{z}$. Forecasts started from analyses generated by assimilating conventional observations of the real atmosphere are verified using the high (T254L64) resolution operational NCEP analyses truncated to 2.5° x 2.5° resolution as proxy for the "true" state. (The local state vectors are defined on the 2.5° x 2.5° grid and not on the nominally higher resolution native computational grid of the model.) These operational analyses were obtained by NCEP assimilating a large number of satellite radiance observations in addition to the conventional observations used in our experiments. Forecast error statistics are generated for the 36-day period, 0000 UTC 11 January 2004 - 0000 UTC 15 February 2004.

E. Selection of the parameters of the diagnostics

We define the local state vector by all temperature, wind, and surface pressure grid point variables in a cube that is defined by 5x5 horizontal grid points and the entire column of the model atmosphere. Computing projections in the vector space $\mathbb{S}_\ell$ requires the definition of a scalar product on $\mathbb{S}_\ell$. In this dissertation, we follow the approach of Oczkowski et al. (2005) and Kuhl et al. (2007): we use the Euclidean scalar product, but before we compute it, we transform the ensemble perturbations to ensure that all vector components have the same physical dimension. In particular, we choose the transformation weights so, that the square of the Euclidean norm, computed by taking the scalar product of a transformed ensemble perturbation by itself, has dimension of energy. The use of this transformation to compute scalar

products of the perturbations of the state vector of a primitive equation model was first suggested by Talagrand (1981). Based on our previous experience (Oczkowski et al., 2005, KEA07), for a specific choice of the local volume, the spatiotemporal variability in the E-dimension and the explained variance for the first 5 forecast days are associated with synoptic scale atmospheric processes. The choice of the local state vector also determines the minimum number of ensemble members required to provide a sufficiently accurate estimate of the local error covariance matrix $\mathbf{P}_\ell$. The problem of achieving a proper balance between the size of the local region and the number of ensemble members was investigated in detail in the context of ensemble-based Kalman filtering by Szunyogh et al. (2005). Since we expect the dynamically active number of degrees of freedom to be the largest at analysis time, we chose the number of ensemble members to be $K = 40$, a choice that was found to be sufficient for an efficient data assimilation with the LETKF using similar size local regions by Szunyogh et al. (2005). Since the dimension of the local state vector is $N = 1975$, our choices for the local state vector and the ensemble size satisfy the condition K<N, which was assumed in Chapter II.

To illustrate the only modest sensitivity of our diagnostics to the selection of the local region size, which is a desirable property considering that the choice of the local region is somewhat arbitrary, we show snapshots of the E-dimension and explained variance for two different size local regions at the 120-hr forecast time (Figures 3 and 4). The results indicate, in agreement with Oczkowski et al. (2005), that while the values of E-dimension somewhat increase, and the values of explained variance somewhat decrease, with doubling the size of the local region in both horizontal directions, the general location of the low-dimensional, and high explained variance, regions are mostly unaffected. The figures also show that smaller region size, which is used in the rest of the paper, provides a sharper resolution of the diagnostics.

CHAPTER IV

NUMERICAL EXPERIMENTS FOR THE SPACE OF UNCERTAINTIES*

A. Forecast errors

First, to illustrate the general spatial distribution of the errors in the $\mathbf{x}_g^e$ state estimate we examine the absolute error in the analyses and forecasts of the meridional wind component at 500 hPa. We choose the meridional wind instead of the more commonly used geopotential height, because this way we can use the same quantity to characterize the errors in the Tropics and the extratropics. Plots of the absolute error are obtained by computing the time average of $\|\boldsymbol{\xi}\|$ at each location (grid point). Figure 5 shows the time mean absolute error at analysis time and at the 72-hr forecast lead time for all three experiments. The results obtained by assimilating simulated observations in randomly placed locations show that the largest analysis errors are in the Tropics and the smallest analysis errors are in mid-latitude storm track regions, in agreement with Szunyogh et al. (2005). Forecast errors become dominant in the storm track regions within 48-72 hours. In comparison, when simulated observations are placed in realistic locations, the results show that the distribution of the magnitude of the analysis errors is strongly modulated by the observation density: the lowest errors are over continents in the Northern Hemisphere and the highest errors are over Antarctica and in the oceanic region between Cape Horn and the Antarctic Peninsula.

We see strong similarities in the spatial distribution of the errors at analysis

---

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Space of Uncertanties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, 138, 962-981, © Copyright 2010 American Meteorological Society (AMS)

time and for short term forecasts in both experiments that assimilate observations in realistic locations. This similarity indicates that observation density plays a more dominant role than model error in determining the large scale spatial variation of the analysis and the short term forecast errors. Nevertheless, the results obtained by assimilating observations of the real atmosphere show that the magnitude of the forecast error is almost double the forecast error found in the experiments which used simulated observations. In all three experiments, we find rapid growth of forecast errors in the mid-latitude storm track regions, which become the dominant region of forecast error by the 72-hr time. The dominance of the synoptic scale error structures at and beyond the 72-hr forecast time is also well illustrate by Figures 6 which show the evolution of the spectral distribution of the forecast errors: initially the magnitude of the errors is similar in the 1-5, 6-10, and 21-42 wavenumber ranges and slightly higher in the 11-20 wave number range, but at longer lead times the errors in the 6-10 wave number range become dominant (the dominance of the 6-10 wavenumber range occurs more rapidly for the experiments that assimilate realistically placed observations). Errors in the wave number range 21-42 saturate much faster than the errors at the larger scale. The increasing wavelength of the dominant error structures contributes to the general tendency of a decrease of the E-dimension. Also, the increasing dominance of the synoptic scale error structures leads to the development of local minima of the E-dimension in the regions of the main synoptic scale features.

B. E-dimension and explained variance

Szunyogh et al. (2005) showed that for lower values of E-dimension, the ensemble more certainly captured the structure of the background error. KEA07 extended the E-dimension diagnostic to study predictability of the performance of ensemble forecasts

and found that, in the extratropics, atmospheric instabilities[1] that led to fast local error growth also led to low E-dimension and, therefore, to increased certainty that a greater portion of the forecast error was efficiently captured by the ensemble.

We investigate the relationship between the E-dimension, explained variance, and forecast errors with the help of joint probability distribution functions (JPDFs). The JPDF shown in Figure 7 is obtained by calculating the number of occurrences in each bin defined by $\Delta E \times \Delta EV$, where ΔE denotes the bin increment for E-dimension and ΔEV denotes the bin increment for the explained variance. The number of occurrences is then normalized by $\Delta E \times \Delta EV \times n$, where n is the total sample size, which is equal to the total number of grid points in a geographic region multiplied by the total number of verification times. This normalization ensures that the integral of the plotted values over all bins is equal to one. At analysis time, we find lower values of E-dimension corresponding to higher values of explained variance for the experiments which use realistically placed observations (two lower left panels of Figure 7) than for the experiment that uses randomly placed simulated observations. As forecast lead time increases, lower values of E-dimension have a greater probability of corresponding to high value of explained variance. In good agreement with KEA07, we find that at the 120-hr lead time (right panels of Figure 7), *the lower the E-dimension, the greater the probability that explained variance is high. We find this relationship independent of experiment and geographic region.* We also note that the evolution of the E-dimension and explained variance is similar to that in KEA07 for all three experiments. Our local results are in agreement with the findings of Wei and

[1]Here we use the term "instability" in the mathematical sense, that is, it refers to the divergence of nearby model trajectories in state space. These instabilities are not always directly related to atmospheric instabilities characterized by the generation of transient kinetic energy (e.g., baroclinic instability, barotropic instability and convection).

Toth (2003) for the global ensemble perturbations and errors, which showed higher PECA values with increasing lead time due to the collapse of the phase space of ensemble perturbations and forecast errors into a smaller dimensional subspace.

A unique feature of the results for the experiments which use real observations (bottom two panels of Figure 7) is that the values of explained variance never reach their theoretical upper limit of one. This behavior is most likely due to the effects of the model errors, since in the two experiments that use simulated observations the largest values of the explained variance are near one. We cannot determine, however, based on the results of our experiments, whether this reduction in the maximum of the explained variance occurs because some of the forecast errors are orthogonal to the model attractor, thus an ensemble of model forecasts cannot capture them, or because our approach to generate the ensemble perturbations does not enable the members of the forecast ensemble to explore that part of the model attractor that includes the true system state.

The results shown in Figure 7 suggest that the E-dimension may be a good linear predictor of the lower bound of the explained variance at a given time and location. To further explore this idea, we break up the 36 day data set into two sets of 18 days, and we try to find a quantitative relationship between the E-dimension and the lower bound of the explained variance based on the first 18 days (training period), which we can then use to predict the lower bound of the explained variance based on the E-dimension for the second 18-day period. To obtain a prognostic equation, we first order the values of E-dimension for the training period and divide them into 100 bins, each containing an equal number of data points. Each bin provides a pair of data: a value of E defined by its mean for the bin and the minimum of the explained variance in the bin, $min_E EV$. Based on this paired data set the correlation between E and $min_E EV$ is 0.9224, 0.6248, and 0.7353 at the 120-hr lead time for experiments

that assimilate simulated observations in random locations, simulated observations in realistic locations, and observations of the real atmosphere respectively, which are statistically significant at the 99.99% level by a t-test. This supports our hypothesis that E may be a good linear predictor of $min_E EV$. Thus, we compute the linear regression coefficients between the paired data over all of the bins, that is, we obtain a and b, such that

$$min_E EV = aE + b. \tag{4.1}$$

The results are shown in Figure 8. The correlation values (0.906, 0.8206, and 0.669) between the predicted and the actual minimum of the explained variance indicate a linear relationship, which is statistically significant at the 99.99% confidence level by a t-test. Nevertheless, when we use the related linear regression coefficients to predict $min_E EV$ based on E, the predicted values are clearly overestimated in some cases. This problem is most noticeable, for the experiment that assimilates observations of the real atmosphere at analysis time. A more careful inspection of Figure 7 suggests that this overestimation of the minimum explained variance by the linear model may be due to a few statistical outliers of $min_E EV$ in the training data sets. To dampen the effects of the outliers, we repeat our calculations using the 5th percentile value, $EV_{(5)}$, of the explained variance instead of $min_E EV$. That is, we obtain a and b such that,

$$EV_{(5)} = aE + b,\ P(EV \leq EV_{(5)}|E) = 0.05, \tag{4.2}$$

where $P(\cdot|E)$ is the conditional probability given E. Since we have a hundred data points in each bin, the 5th percentile of the explained variance is defined by the 5th member of the ascending ordered set of values of explained variance. The results are summarized in Figure 9, which shows that the linear expression based on the E-dimension provides a better prediction of the 5th percentile than the minimum of

EV. The prediction of the 5th percentile of explained variance is especially accurate at longer forecast lead times, e.g. at the 120-hr lead time the correlation between the predicted and the true value of the explained variance is greater than 0.94 and the root-mean-square error of the prediction is less than 0.02 in all three experiments.

To further illustrate the predictive value of the linear relationship we found, we plot a snapshot of the actual values of explained variance and the predicted 5th percentile, at the 120 hr forecast time for the experiment which assimilates conventional observations of the real atmosphere (Figure 10). The upper panel of Figure 10 shows an area of local low-dimensionality that develops in the northeast Pacific in an atmospheric region dominated by a strong ridge. The location of the low-dimensional area suggests that there is an uncertainty in the forecast of the ridge and that this uncertainty is dominated by a few state space directions. The second panel of Figure 10 shows that, as expected based on Figure 7, the ensemble is efficient in capturing the space of uncertainties in the low-dimensional region ($EV \approx 0.93$) . When we use the linear expression of Equation 4.2, we obtain a value of 0.85 for the 5th percentile (which is slightly lower that the actual value of 0.93, as expected) which correctly predicts a good performance of the ensemble in capturing the uncertain forecast patterns in the ridge region. We emphasize that a good prediction of the uncertain forecast patterns is not a guarantee of a good prediction of the magnitude or the probability distribution of the forecast uncertainty.

C. Explained variance and forecast error

Figure 11 shows the JPDF for the explained variance and the state estimation error in the NH extratropics. While there is no obvious relationship between the magnitude of the analysis errors and the explained variance (left panels), similar to the results

of KEA07, *the ensemble captures the patterns associated with larger forecast errors more efficiently.* In addition, both the minimum and the maximum of explained variance increase with forecast time in all three experiments, which indicates that $\mathbb{S}_\ell$ provides an increasingly better representation of the space of forecast uncertainty with increasing forecast time (results are shown only for the 120-hr forecast time).

Figure 12 shows the mean E-dimension for the bins in the JPDF for the analysis and forecast error and the explained variance. Interestingly, the distribution of E-dimension with explained variance at analysis time is more similar for the two experiments which assimilate realistically distributed observations. For these two experiments, we find locations where the explained variance is high and the E-dimension is low, but the analysis error is relatively large. These are locations where the ensemble efficiently captures the space of uncertainties, but there are no observations available to take advantage of this information. Such locations do not exist for the experiment that assimilates randomly placed observations, as in that experiment the observational coverage is sufficiently dense at all locations to effectively remove the background errors at locations of high explained variance (low E-dimension).

At the 120-hr forecast time, the findings of KEA07 extend to the more realistic settings: the instabilities that lead to large forecast errors also lead to low E-dimension, and therefore, to higher explained variance. However, because low E-dimension can occur even if the forecast error is small, the E-dimension cannot be used to predict the magnitude of the forecast error. That is, we may not be able to predict that the forecast error will be large at a given time and location, but we know that *if the error will be large, the ensemble will provide a reliable representation of the potential forecast error patterns.*

D. Local linearity

As pointed out before, analyzing Figure 11, within the five day forecast range we investigated in this dissertation, the minimum and maximum of the explained variance increase with increasing forecast time. In other words, the linear space $\mathbb{S}_\ell$ provides an increasingly better representation of the space of the forecast errors for longer forecast lead times. This is a counterintuitive result, because the ensemble perturbations are expected to evolve linearly only for the shortest forecast lead time, when their magnitude is still small. In this case they can be considered a representation of the tangent space to the nonlinear system trajectory. The evolution of the ensemble perturbations is expected to become more nonlinear as their magnitude is growing with increasing forecast time. In the remainder of this section, we investigate the roots of this seemingly paradoxical result.

A potential resolution of the paradox is that while the evolution of the ensemble perturbations becomes less linear with increasing forecast time with respect to a global measure of nonlinearity, there may be local regions where linearity becomes stronger with increasing forecast time. To explore this possibility, we first investigate the local variability of the strength of nonlinearity. We do this by comparing the mean and the standard deviation of the local values of the relative nonlinearity for the different forecast lead times (Figure 13). While the globally averaged values of relative nonlinearity computed with and without localization are similar, the standard deviations of the values computed using localization show a strong local variability in the degree of linearity. This variability is especially striking for the experiment which assimilates randomly placed observations: values of the relative nonlinearity as low as its mean at the 24-hr lead time are within one standard deviation at 120 hours. Conversely, values as high as the mean at 84 hours are within one standard deviation

at the 24-hr forecast lead time. While the local variability of the relative nonlinearity is not negligible, the results do not support a scenario in which nonlinearity could locally decrease at most locations. To confirm that high explained variance is not due to strong local linearity, we also prepared scatter plots for the explained variance and the relative nonlinearity (Figure 14). It is obvious, based on the scatter plot, that a weak nonlinearity is not a necessary condition for high explained variance. Our results are also in line with those who found that a tangent linear model can, qualitatively, well predict the fastest growing error patterns for much longer times than it could be expected based on the nonlinearity index (e.g., Reynolds and Rosmond 2003).

Our finding that while model errors double the magnitude of the forecast errors, they do not decrease much the efficiency of the ensemble based linear space in representing the space of forecast errors, makes us speculate that a large part of the model errors may simply act as an additional stochastic forcing.

CHAPTER V

NUMERICAL EXPERIMENTS FOR THE MAGNITUDE AND SPECTRUM OF UNCERTAINTIES*

A. Prediction of the magnitude of forecast error

Figure 15 shows the evolution of TV, TVS and V. In this figure the expected value is estimated by taking the spatial average over all grid points in the NH extratropics (30° N - 90° N) and the temporal average over all forecasts started between 0000 UTC 11 January 2004 and 0000 UTC 15 February 2004.

First we compare the total forecast error variance, TV, and the forecast error variance in $\mathbb{S}_\ell$, TVS. In the two perfect model experiments, TVS is only slightly smaller than TV, which indicates that $\mathbb{S}_\ell$ provides an efficient representation of the error space. As expected, in the third experiment, where part of TV is due to model errors, the difference between TV and TVS is larger than in the two perfect model experiments.

Interestingly, the difference between TVS and V at longer lead times is much larger than the difference between TV and TVS. In other words, *although the linear space $\mathbb{S}_\ell$ spanned by the ensemble perturbations provides a good representation of the space of forecast uncertainties, the ensemble severely underestimates the total variance in $\mathbb{S}_\ell$*. Even though, this underestimation is more serious in the experiment where model errors have an effect on the total error variance TV, the underestimation in the two perfect model experiments is also significant. Thus, the underestimation of

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Magnitude and the Spectrum of Incertainties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, in review,

the forecast error variance cannot be fully explained by the lack of accounting for the effect of model errors in our ensemble. Also, the underestimation cannot be explained by an initial underestimation of the variance of the error in $\mathbb{S}_\ell$, as initially the ensemble variance V is tuned to be about equal to TVS. (Further inflating the analysis ensemble members would lead to accounting for part of $\delta\mathbf{x}^{(\perp)}$ by over inflating $\delta\mathbf{x}^{(\|)}$.)

When we fit TV to an exponential curve, $TV = x_o \exp^{at}$, we find the fastest growth rate for the experiment that assimilates simulated observations in random locations, where the doubling time ($T_d = a^{-1} ln2$) is $T_d = 25.5774$ hours ($T_d = 19.7478$ hours for TVS and $T_d = 23.1049$ hours for V). The growth rate is the slowest for the experiment that uses observations of the real atmosphere. For this experiment, $T_d = 33.1649$ hours ($T_d = 25.7676$ hours for TVS and $T_d = 42.7869$ hours for V). For the experiment which assimilates realistically placed simulated observations, the doubling time is $T_d = 26.9707$ hours ($T_d = 22.3596$ hours for TVS and $T_d = 26.9707$ hours for V). Interestingly, the initial growth rate of TVS and V is faster than exponential growth: up until the 36-hr lead time, the time evolution of the growth rate of TVS and V can be best approximated by a second-order polynomial.

Figure 16 shows the zonal power spectrum of the meridional wind at 500 hPa averaged over all latitudes in the NH extratropics and over time. Left panels show results which were obtained by computing the power spectra for each ensemble perturbation, then taking the ensemble mean of the spectra. Right panels show the spectra for the $\delta\mathbf{x}^t$ difference between the model representation of the truth and the ensemble mean. The shape of the spectra and the time evolution of the spectra is very similar all panels with the exception of the ensemble spectra for the real observations (lower left panel). Unlike in the other panels, where a dominant range of wavenumber 6-10 emerges from a relatively flat initial spectrum, at analysis time, for the real ob-

servations the spectrum of the ensemble perturbations has a well pronounced peak at wavenumber 2. A more detailed investigation of this spectrum revealed that this peak is associated with large ensemble variance in a region over the arctic north of Russia. While the qualitative behavior of the spectra of $\delta\mathbf{x}^t$ and the ensemble perturbations is similar, as can be expected based on Figure 15, the amplitudes are larger for $\delta\mathbf{x}^t$ than for the ensemble perturbations. By computing the growth rates for the wave number k=8, we find that the differences between the predicted and true uncertainties are due to differences in the magnitude of the two quantities at analysis time, rather than to a faster growth rate of the true uncertainty. This result suggests that the analysis ensemble systematically underestimates the analysis uncertainty at the synoptic (baroclinically most active) scale. Because we tuned the analysis ensemble to satisfy $V \approx TVS$ at analysis time, there are two potential explanations for the underestimation of the uncertainty at the synoptic scales. First, because the computations of V and TVS involve spatial averaging over large regions, we may underestimate the uncertainty in regions where it is associated with synoptic scale features due to an overestimation of the uncertainty in regions where it is dominantly related to some different scales. Second, the underestimation of the uncertainty at the synoptic scale could be due to the overestimation of the importance of local error patterns associated with different scales in $\mathbb{S}_\ell$. The results we later show in this dissertation for the spectrum of uncertainties clearly suggest the first scenario. For the case of the real observations, the overestimation of the importance of wavenumber 2 features at higher latitudes also point to the first scenario.

B. Prediction of the spatiotemporal changes in the magnitude of forecast error

The usual candidate for a predictor of $||\delta \mathbf{x}^t||$ at location ℓ is the ensemble spread, $V_\ell^{1/2}$, at the same location. These two quantities are known to have a positive correlation, which is typically low at analysis time and asymptotes to a level of about 0.5 by about 72-hr lead time (e.g. Barker 1991, Houtekamer 1993, and Whitaker and Loughe 1998). KEA07 found that the correlation for our system was in good agreement with those earlier results. Since a correlation of 0.5 for a sample size of N=129,600 suggests the existence of a linear relationship between the $V_\ell^{1/2}$ and $||\delta \mathbf{x}^t||$, a prediction of $||\delta \mathbf{x}^t||$ based on the $V_\ell^{1/2}$ with a linear regression may seem to be a natural choice. Computing the correlation for our experiments, we find that it is largest for the case of simulated observations at random locations (0.58), slightly lower for the case of realistically placed simulated observations (0.52), and much lower for the realistic case (0.26). This relatively large drop in the correlation for the realistic case would itself provide an argument against using a linear regression to predict $||\delta \mathbf{x}^t||$. An even more problematic feature of the relationship between $||\delta \mathbf{x}^t||$ and $V_\ell^{1/2}$, which is illustrated by Figure 17, is that for larger values of $V_\ell^{1/2}$, $||\delta \mathbf{x}^t||$ varies over a much wider range of values. To better understand the problematic aspects of this result, we recall from linear statistics, that univariate regression predicts the value of a random variable y based on a predictor x by the $E[y|x]$ conditional expectation of y given x (e.g., p.264 in Rao 1973). Thus, a prediction of $||\delta \mathbf{x}^t||$ with a prediction of its conditional expectation does not reflect the large potential magnitude of the forecast error for a large value of the spread. This motivates us to investigate the relationship between $V_\ell^{1/2}$ and the upper bound of $||\delta \mathbf{x}^t||$, instead of the expectation of $||\delta \mathbf{x}^t||$ given $V_\ell^{1/2}$.

We start our investigation by breaking up the 36 day data set of 120-hr forecasts into two sets of 18 days, and search for a quantitative linear relationship between V_ℓ

and the upper bound of $||\delta \mathbf{x}^t||^2$ given V_ℓ based on the first 18 days (training period). We then use the functional relationship found for this training period to predict the upper bound of $||\delta \mathbf{x}^t||^2$ based on V_ℓ for the second 18-day period. In order to dampen the effects of outliers, we use the 95th percentile of the bin values of $||\delta \mathbf{x}^t||^2$ instead of the bin maximum. To obtain a qualitative prognostic relationship, we first order the values of V_ℓ for the training period and divide them into 100 bins, each containing an equal number of data points. Each bin provides a pair of data: a value of V_ℓ defined by its mean for the bin and the 95th percentile of $||\delta \mathbf{x}^t||^2$.

Based on the training data set, the correlation between the bin mean of V_ℓ and the bin 95th percentile of $||\delta \mathbf{x}^t||^2$ is 0.9889, 0.8176, and 0.9206 at the 120-hr lead time for experiments that assimilate simulated random locations, simulated observations in realistic locations, and observations of the real atmosphere respectively, which are statistically significant at the 99.99% level by a t-test. This suggests that V_ℓ may be a good linear predictor of $||\delta \mathbf{x}^t||^2$, even in the realistic case. Thus, we use the linear regression coefficients obtained for the training data set to predict the 95th percentile of $||\delta \mathbf{x}^t||^2$ for the second 18 days. The results are summarized in Figure 18. The correlation values (0.9518, 0.8685, and 0.8568) between the predicted and the actual 95th percentile values of $||\delta \mathbf{x}^t||^2$ indicate a linear relationship, which is statistically significant at the 99.99% confidence level by a t-test. Encouraged by the strong linear predictive relationship we find at 120-hr, we turn our attention to the shorter lead times. At analysis time, we find lower correlation values between the bin mean of V_ℓ and the bin 95th percentile of $||\delta \mathbf{x}^t||^2$ (0.2472, 0.7076, and 0.2182) as well as lower correlation values between the predicted and the actual 95th percentile values of $||\delta \mathbf{x}^t||^2$ (0.2902, 0.6573, and 0.4161) than at 120-hr lead time. As can be expected, the correlation values increase with forecats lead time, by the 48-hr lead time both perfect model experiments show correlation values greater than 0.75 for

both the training period and the prediction. For the case of realistic observations, the correlation values remain relatively low untill around the 96-hr lead time, where we find correlation values of 0.7259 for the training period and 0.6382 for the prediction of 95th percentile values of $||\delta \mathbf{x}^t||^2$. We recall that 96-hr is about the lead time at which the peak in the power series of the ensemble perturbations moves into the k=8 range, in better agreement with $||\delta \mathbf{x}^t||$ (Figure 16).

C. Spectrum of the d-ratio

We now turn our attention to investigating the efficiency of the ensemble in distinguishing between the importance of the eigendirections (error patterns in physical space) in $\mathbb{S}_\ell$. We first compute the spectrum of d-ratio d_k using the same definition of the local volume as in our calculations of E-dimension and explained variance. The results are summarized in Figure 19.

First we discuss the results for analysis time (left panels of Figure 19). We find that for the experiment which assimilates simulated observations in random locations, the ensemble only slightly underestimates the error in the directions it captures (upper left panel). When realistically placed simulated observations are used, the ensemble tends to underestimate uncertainty in all captured directions, with the exception of the few leading directions (middle left panel). In the experiment with observations of the real atmosphere, the uncertainty is underestimated in all directions captured by the ensemble. The similarity between the shape of the spectra in the two experiments that assimilate observations at realistic locations and the flat spectrum in the third experiments, where observations are nearly uniformly distributed, suggests that the larger underestimation of error variance in the trailing directions is due to the uneven distribution of observations.

As forecast time increases the underestimation of the error by the ensemble becomes increasingly more severe in all directions and in all experiments. We show results for 120-hr forecast time (right panels of Figure 19). In the experiment with randomly distributed observations, the spectrum remains flat (top right panel), while in the experiment with realistically distributed simulated observations the slope of the spectrum does not increase. In the realistic case, in contrast, the spectrum becomes much steeper indicating an increasingly more severe underestimation toward the trailing directions. Comparing the spectra from the two experiments with realistically distributed observations, we conclude that model errors lead to a more severe underestimation of the forecast uncertainty in the trailing directions.

In order to obtain d-ratio figures whose meteorological (physical) meaning is easier to interpret, we now change the definition of the local volume: we investigate a single variable at a single level using 5 by 5 horizontal grid points. In these calculations N=25 ($N < K$), hence, the upper bound for the E-dimension in $\mathbb{S}_\ell$ is 25. The variable and levels we choose for this analysis are the surface pressure, the temperature at 850 hPa, the two horizontal wind components at 500 hPa, and the geopotential height at 250 hPa. Figure 20 shows the time mean of this ratio in the leading direction, d_1, for the temperature at 850 hPa. For the experiment which assimilates randomly placed observations, initially d_1 is highest near the poles. The main regions of enhanced baroclinicity over Japan and off the coast of Newfoundland also show underestimation. The latter result suggests that when the distribution of the observations is nearly uniform, the use of a zonally constant covariance inflation factor in the analysis leads to an underestimation of the uncertainty in the dynamically more active regions. In contrast, for the two experiments which assimilate realistically placed observations d_1 tends to reflect the local observation density: the uncertainty is underestimated in regions of high observation density, such as Europe,

Japan, and the United States, and underestimated in regions of lower observational density, such as the Southern Hemisphere and the oceanic regions. This result is an indication that our zonally constant covariance inflation factor cannot be tuned to be optimal everywhere when there are zonal changes in observation density. Thus we conjecture that implementing a spatially varying adaptive covariance inflation technique, such as described in Anderson (2009) or a localized version of Li et al. (2009), may lead to an improvement of the analyses and the short term ensemble forecasts. The time averaged spectrum of the d-ratio for a particular grid point (Figure 21) at analysis time shows that for the two experiments which use simulated observations, the ensemble, on average, overestimates the uncertainty in all directions, except for the trailing directions, where uncertainty is underestimated. For the experiment that uses observations of the real atmosphere, the ensemble underestimates uncertainty in most directions, more severely for the trailing directions. For the same grid point at the 120-hr lead time the two experiments that use simulated observations show overestimation in all directions. For observations of the real atmosphere, the ensemble overestimates the uncertainty in the leading directions and underestimates uncertainty in the trailing directions.

D. Relationship between E-dimension and d-ratio

Since in Chapter IV we found the E-dimension a good predictor of the performance of $\mathbb{S}_\ell$ in capturing the forecast error, we have been hoping to find a similar relationship between the E-dimension and the quality of the prediction of the magnitude and the spectrum of uncertainties by the ensemble. While all of our attempts at finding a qualitative relationship between the E-dimension and the performance of the ensemble in predicting the magnitude of the uncertainty have failed, we have found interesting

differences between the spectra of d-ratios for different values of the E-dimension.

To explore the relationship between the E-dimension and the spectrum of d-ratios, we output values of E-dimension and the corresponding values for the spectrum of the d-ratio. The data pairs are then ordered by E-dimension values and divided equally into 100 bins. We find that at analysis time the spectrum is better behaved for lower values of the E-dimension. For instance, while for the bin with the lowest value of E-dimension (Figure 22 upper left panel) the spectrum is relatively flat and the values are near one, for the bin with the highest values of E-dimension the underestimation of the uncertainty by the ensemble is more severe. Interestingly, at 120-hr lead time the spectra are better behaved for the higher values of E-dimension. For instance, for the same two examples compared at analysis time, the underestimation of the forecast, with the exception of a few leading directions, is more severe for the regions of low E-dimension. These results indicate that while the spectrum of the d-ratio benefits from better representation of the space of uncertainties in the low E-dimensional regions at analysis time, having a more diverse distribution of the uncertainty at a longer forecast lead time improves the representation of the forecast uncertainty. The practical implications of this result, when combined with the results of Chapter IV, is that while a forecaster should trust the 96-120 hr ensemble predictions of the possible error patterns in the lower dimensional regions more, he or she should also keep in mind that the seemingly unimportant patterns of uncertainty play a more important role in reality than suggested by the raw ensemble forecast.

E. E-dimension, d-ratio, and forecast error

To further investigate the problem of underestimating the uncertainty with the an ensemble that effectively captures the space of uncertainties, we plot the eigenvalue

spectrum (normalized by the leading eigenvalue) and the percentage of TVS,

$$TVS(k) = \frac{\sum_{i=1}^{k} \delta^2 \mathbf{x}^{t(||)}}{\sum_{i=1}^{K} \delta^2 \mathbf{x}^{t(||)}}, \tag{5.1}$$

captured by the first k eigendirections. Figure 23, which is obtained with the same bin averaging techniques as Figure 22, shows the results for the minimum, median, and maximum bin values of the E-dimension. At analysis time, these three values of the E-dimension are 29.3679, 34.1271, and 36.3424 for the simulated observations in random locations, 14.4292, 25.9455, and 35.4192 for the simulated observations in realistic locations, and 21.8493, 31.88, and 36.9536 for observations of the real atmosphere. At the 120-hr lead time, the minimum, median, and maximum bin values of E-dimension are 6.34464, 14.0873, and 24.9761 for simulated observations in random locations, 8.22311, 17.0315, and 26.0225 for simulated observations in realistic locations, and 10.3602, 17.7928, and 25.7089 for observations of the real atmosphere. We find that, low values of E show a quicker saturation of the percentage of TVS compared to the eigenvalue spectrum. For example, for the experiment which assimilates simulated observations in random locations, at the 120-hr lead time, at the point where the eigenvalue spectrum approaches zero, only approximately 90% of TVS has been captured by the ensemble. For the experiment that assimilates observations of the real atmosphere, at the 120-hr lead time, even for the lowest bin value of E-dimension, we find that all directions captured by the ensmeble are necessary to capture 100% of TVS, but the eigenvalue spectrum saturates around n=15 (in agreement with Figure 22, which shows that the ensemble underestimation increases sharply for trailing directions). These results support the use of linear post-processing techniques to increase the ensemble spread in the trailing directions at longer forecast lead times.

CHAPTER VI

CONCLUSIONS*

In this dissertation, we studied the spatiotemporally changing nature of predictability by coupling a reduced resolution version of the model component of the NCEP GFS with the LETKF data assimilation scheme. Our focus was on exploring the predictive value of the linear space, $\mathbb{S}_\ell$, spanned by the ensemble perturbations in capturing the forecast uncertainties for the first five forecast days. We employed a hierarchy of increasingly more realistic experiment designs in order to be able to investigate the effects of the temporal and spatial distribution of the observations and the effects of model errors. While we found that the distribution of the observations modulates the distribution of the analysis and forecast errors and model errors lead to a doubling of the average and maximum magnitude of the errors, the following main findings apply to all experiment set ups, including the one that assimilated observations of the real atmosphere:

- We have found that $\mathbb{S}_\ell$ provides an increasingly better representation of the space of uncertainties with increasing forecast time.

- We have shown that the improving performance of $\mathbb{S}_\ell$ with increasing forecast time is not due local linear error growth, but rather to nonlinearly evolving forecast errors that have a growing projection on the linear space $\mathbb{S}_\ell$.

---

*Portions of this chapter have been reprinted from:
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Space of Uncertanties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, 138, 962-981, © Copyright 2010 American Meteorological Society (AMS)
and
"Predictability of the Performance of an Ensemble Forecast System: Predictability of the Magnitude and the Spectrum of Incertainties" by E.A. Satterfield and I. Szunyogh, 2010, *Mon. Wea. Rev.*, in review, © Copyright 2010 American Meteorological Society (AMS)

- We have shown that the E-dimension is strongly anti-correlated with the error variance explained by the ensemble, thus, it is a reliable linear predictor of the performance of $\mathbb{S}_\ell$ in capturing the most important forecast error patterns. What makes this finding especially valuable is that E-dimension is always low when the forecast error is large, that is, we can have high-confidence in the skill of the ensemble in capturing the error patterns associated with the large error in the deterministic forecast and we can use a simple linear regression to provide a quantitative prediction of this skill at a given time and location.

- The results suggest that predicting the magnitude of the forecast uncertainty and the relative importance of the different patterns of uncertainty is, in general, a more difficult task than predicting the space of uncertainty (the collection of the potential patterns of uncertainty).

- While the ensemble, which is tuned to provide near to optimal performance at analysis time, provides a good representation of the space of forecast uncertainty, it severely underestimates not only the total magnitude of the uncertainty, but also the magnitude of the uncertainty that projects onto the space spanned by the ensemble perturbations.

- The ensemble tends to more severely underestimate the forecast uncertainty in the directions (for the patterns of uncertainty) that are present in the ensemble with a small amplitude. This problem is more pronounced at locations where the E-dimension is low (where a very few patterns dominate the predicted uncertainty), which are, interestingly, also the locations where the ensemble provides the best representation of the space of uncertainties.

Some of our results point to shortcomings of our particular system. Most importantly, using a zonally invariant multiplicative covariance inflation factor leads to an underestimation of the local magnitude of the analysis uncertainty in regions of dense observations and an overestimation of the local magnitude of the analysis uncertainty in regions of sparse observations. Also, the system underestimates the analysis uncertainty at the synoptic scales, which leads to an underestimation of the forecast uncertainty at the same scale. These shortcomings of the system can be, most likely, eliminated by implementing a more sophisticated covariance inflation scheme on the LETKF. Some of our other results, on the other hand, have potentially broader implications and, we hope to investigate the generality of these findings in the near future utilizing the THORPEX Interactive Grand Global Ensemble (TIGGE) data base. If we found that the results of this dissertation are general, that would support the approach followed by many synopticians in interpreting the raw ensemble forecasts, which is based on paying more attention to the predicted patterns of uncertainties than to the uncalibrated raw quantitative ensemble-based measures of the uncertainty. Such findings would provide an additional argument for the use of post-processing techniques to enhance the ensemble based forecasts. The fact that a linear space $\mathbb{S}_\ell$, provides a good representation of the uncertainty in the medium forecast range confirms that linear statistical techniques, such as those based on reforecasts (e.g., Hamill et al. 2004; Hamill and Whitaker 2006; Hamill et al. 2008; Hagedorn et al. 2008), have a great potential to improve ensemble-based forecasts.

Finally, we believe that our reasoning behind using a simple linear regression to predict the upper bound of the forecast uncertainty based on the spread is sufficiently general to be valid for any ensemble forecast system. This relationship, along with the strong predictive linear relationship between the E-dimensions and the performance of the ensemble in capturing the patterns of uncertainty, could be implemented in

operations after a proper tuning of the linear regression coefficients.

REFERENCES

Anderson, J. (2009). Spatially-varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, 72–83.

Anderson, J. L. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, **230**(1-2), 99 – 111.

Barker, T. (1991). The relationship between spread and forecast error in extended-range forecasts. *J. Climate*, **4**, 733–742.

Berner, J., Shutts, G., Leutbecher, M., and Palmer, T. (2009). A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.

Candille, G. and Talagrand, O. (2005). Evaluation of probabilistic prediction systems for a scalar variable. *Quart. J. Roy. Meteor.*, **131**(609), 2131–2150.

Chapman, S. and Lindzen, R., 1970: *Atmospheric Tides*. D. Reidel Press, Dordrecht, Holland, 200 pp.

Descamps, L. and Talagrand, O. (2007). On some aspects of the definition of initial conditions for ensemble prediction. *Mon. Wea. Rev.*, **135**, 3260–3272.

Gilmour, I., Smith, L. A., and Buizza, R. (2001). Linear regime duration: Is 24 hours a long time in synoptic weather forecasting? *Journal of Atmospheric Sciences*, **58**, 3525–3539.

Hagedorn, R., Hamill, T., and Whitaker, J. (2008). Probabilistic forecast calibration using ecmwf and gfs ensemble forecasts. part i: two-meter temperatures. *Mon. Wea. Rev.*, **136**, 2608–2619.

Hamill, T., Hagedorn, R., and Whitaker, J. (2008). Probabilistic forecast calibration using ecmwf and gfs ensemble forecasts. part ii: precipitation. *Mon. Wea. Rev.*, **136**, 2620–2632.

Hamill, T. and Whitaker, J. (2006). Probabilistic quantitative precipation forecasts based on reforecast analogs: theory and application. *Mon. Wea. Rev.*, **134**, 3209–3229.

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

Hamill, T., Whitaker, J., and Wei, X. (2004). Ensemble reforecasting: improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132**, 1434–1447.

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

Houtekamer, P. L. (1993). Global and local skill forecasts. *Mon. Wea. Rev.*, **121**(6), 1834–1846.

Huang, X.-Y. and Lynch, P. (1993). Diabatic digital filtering initialization: Application to the hirlam model. *Mon. Wea. Rev.*, **121**, 589–603.

Hunt, B., Kostelich, E., and Szunyogh, I. (2006). Efficient data assimilation for spatio-

temporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**(1-2), 112–126.

Kalnay, E., 2002: *Atmospheric Modelling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, UK, 341 pp.

Kuhl, D., Szunyogh, I., Kostelich, E., Gyarmati, G., Patil, D., Oczkowski, M., Hunt, B., Kalnay, E., Ott, E., and Yorke, J. (2007). Assessing predictability with a local ensemble Kalman filter. *J. Atmos. Sci.*, **64**, 1116–1140.

Leith, C. E. (1974). Theoretical skill of monte carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.

Li, H., Kalnay, E., and Miyoshi, T. (2009). Simultaneous estimation of covariance inflation and observation errors within ensemble Kalman filter. *Quart. J. Roy. Meteor.*, **135**, 523–533.

Lorenz, E. (1963a). Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.

Lorenz, E. (1963b). The predictability of hydrodynamic flow. *Trans. NY Acad. Sci., Series II*, **25**, 409–432.

Oczkowski, M., Szunyogh, I., and Patil, D. J. (2005). Mechanisms for the development of locally low dimensional atmospheric dynamics. *J. Atmos. Sci.*, **62**, 1135–1156.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A. (2002). A local ensemble Kalman filter for atmospheric data assimilation. *arXiv.org:physics/0203058*.

Palmer, T. and Hagedorn, R., editors, 2006: *Predictability of Weather and Climate*. Cambridge University Press, Cambridge, UK, 718 pp.

Palmer, T. N., Buizza, R., Molteni, F., Chen, Y.-Q., and Corti, S. (1994). Singular vectors and the predictability of weather and climate. *Royal Society of London Philosophical Transactions Series A*, **348A**, 459–475.

Patil, D. J., Hunt, B. R., Kalnay, E., Yorke, J. A., and Ott, E. (2001). Local low dimensionality of atmospheric dynamics. *Phys. Rev. Let.*, **86**, 5878–5881.

Rao, C., 1973: *Linear Statistical Inference and its Applications*. Wiley, New York, 2 edition, 625 pp.

Reynolds, C. A. and Rosmond, T. E. (2003). Nonlinear growth of singular-vector-based perturbations. *Quart. J. Roy. Meteor.*, **129**(594), 3059–3078.

Sankey, D., Ren, S., Polavarapu, S., Rochon, Y. J., Nezlin, Y., and Beagley, S. (2007). Impact of data assimilation filtering methods on the mesosphere. *J. Geophys. Res.*, **112**, D24104.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Kalnay, E., Hunt, B. R., Ott, E., Satterfield, E. A., and Yorke, J. A. (2008). A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, **60A**, 113–130.

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

Szunyogh, I., Satterfield, E. A., Aravequia, J. A., Fertig, E. J., Gyarmati, G., Kalnay, E., Hunt, B. R., Kostelich, E. J., D. D. Kuhl, E. O., and Yorke, J. A., 2007: The local ensemble transform Kalman filter and its implementation on the NCEP

global model at the University of Maryland. In *Proc. Flow Dependent Aspects of Data Assimilation*, Reading, UK, ECMWF, 47–63.

Talagrand, O. (1981). A study of the dynamics of four-dimensional data assimilation. *Tellus*, **33**, 43–60.

Talagrand, O., Vautard, R., and Strauss, B., 1999: Evaluation of probabilistic prediction systems. In *Proc. Workshop on Predictability*, Reading, UK, ECMWF, 1–25.

Wei, M. and Toth, Z. (2003). A new measure of ensemble performance: Perturbation versus error correlation analysis (peca). *Mon. Wea. Rev.*, **131**, 1549–1565.

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

Whitaker, J. and Loughe, A. (1998). The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.*, **126**(12), 3292–3302.

APPENDIX A

FIGURES

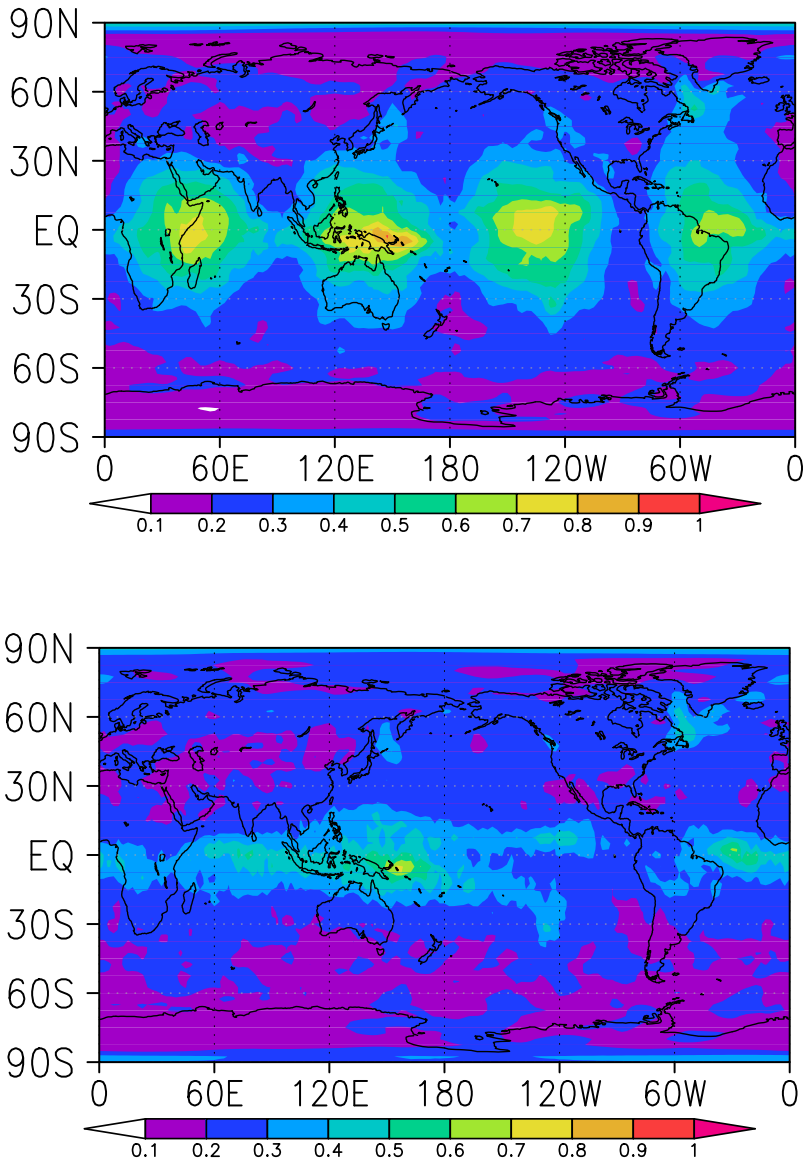

Fig. 1. The time mean absolute error of the surface pressure analysis assimilating simulated observations at random locations with digital filter initialization (top) and without digital filter initialization (bottom). The results shown in this figure use a 10% covariance inflation at all vertical levels and in all geographic regions. The average is taken over all analyses between 01 January 2004 0000UTC and 29 February 2004 1800UTC

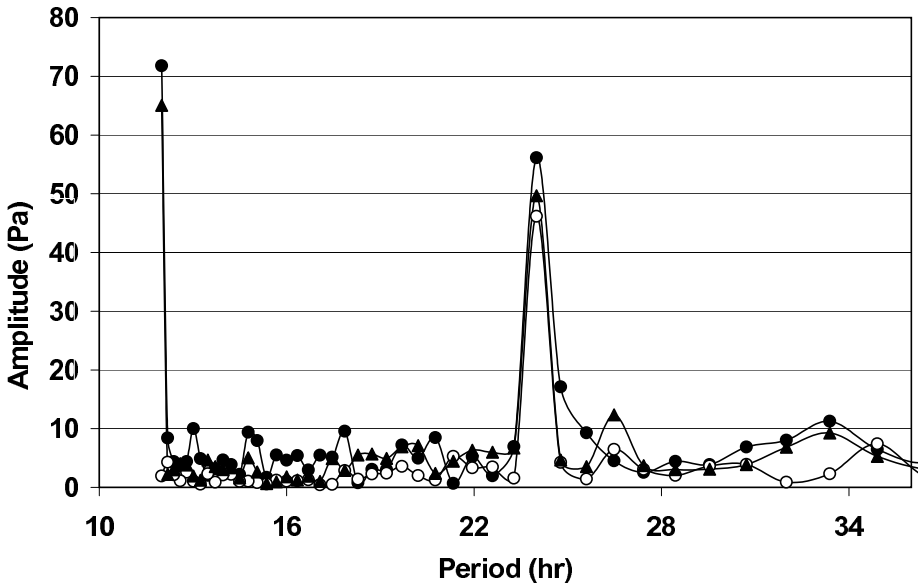

Fig. 2. Fourier analysis of the surface pressure analyses at the location 0°S, 160°W between 14 January 2004 0600UTC and 15 February 2004 0000UTC assimilating simulated observations at random locations without digital filter initialization (closed circles), with digital filter initialization (open circles). The spectrum for the nature run is also shown (triangles). The results shown in this figure are obtained with a 10% covariance inflation at all vertical levels and in all geographic regions.

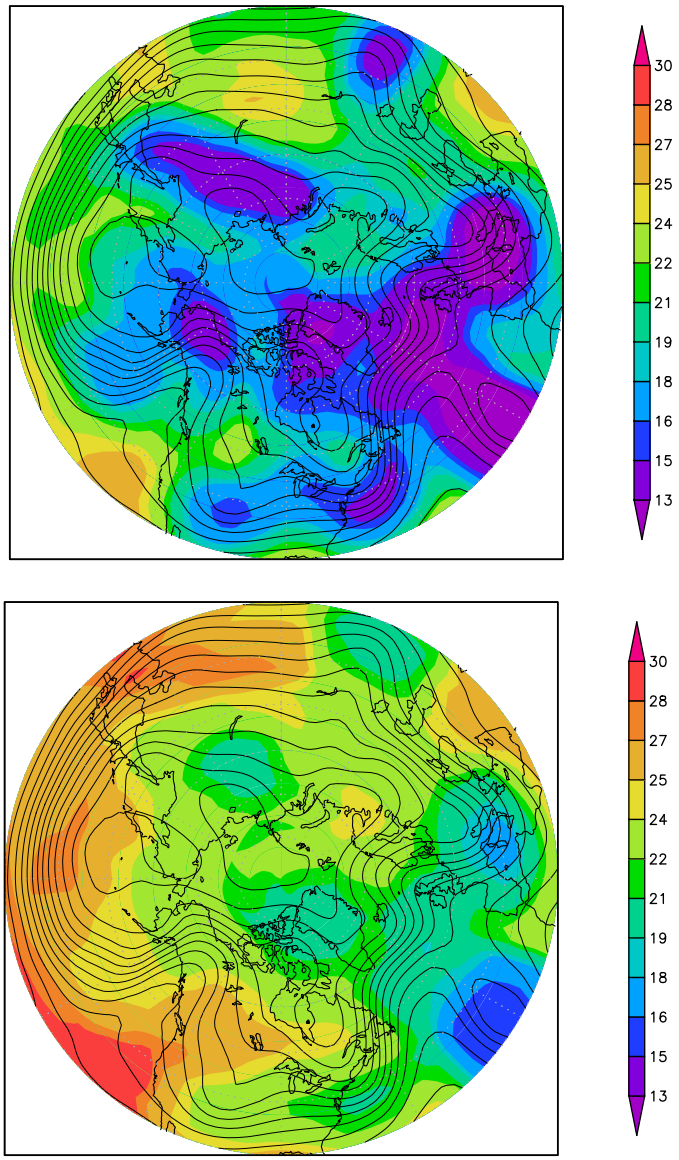

Fig. 3. E-dimension (shades) and geopotential height control (contours) at the 250-hPa level shown for the experiment that assimilates conventional observations for a local region size of 5x5 (top panels) and 10x10 (bottom panels). Results are shown for the 5-day forecast started on 4 Feb 2004.

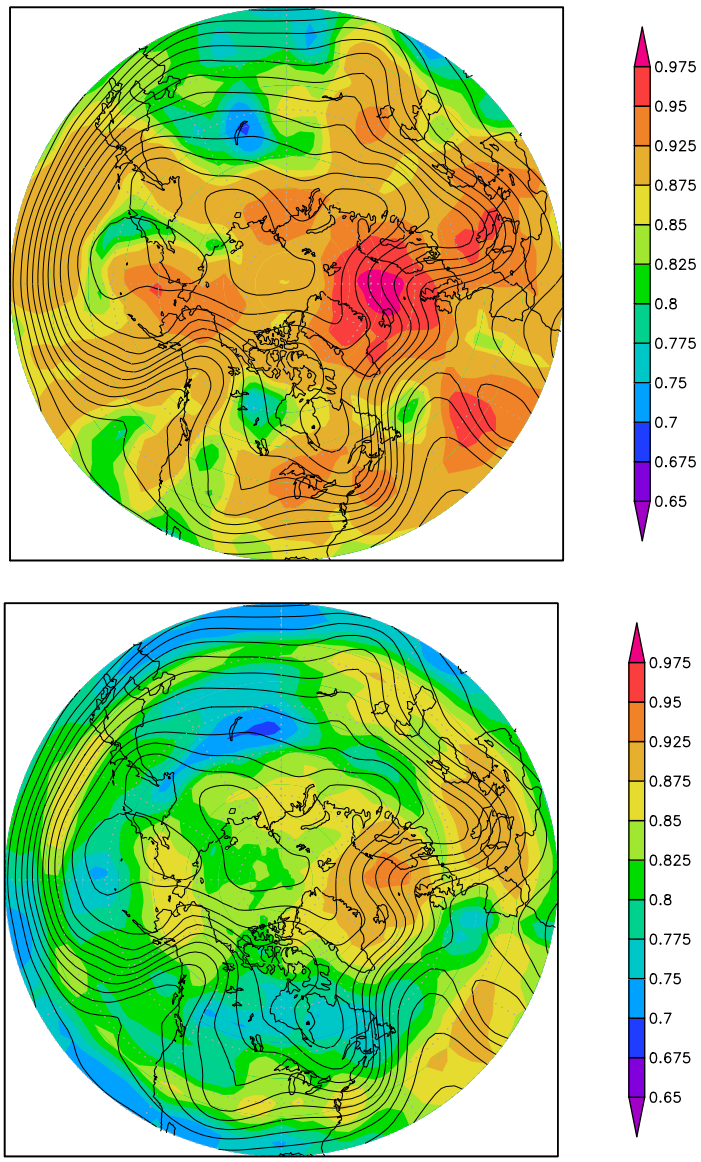

Fig. 4. Explained variance (shades) and geopotential height control (contours) at the 250-hPa level shown for the experiment that assimilates conventional observations for a local region size of 5x5 (top panels) and 10x10 (bottom panels). Results are shown for the 5-day forecast started on 4 Feb 2004.

.

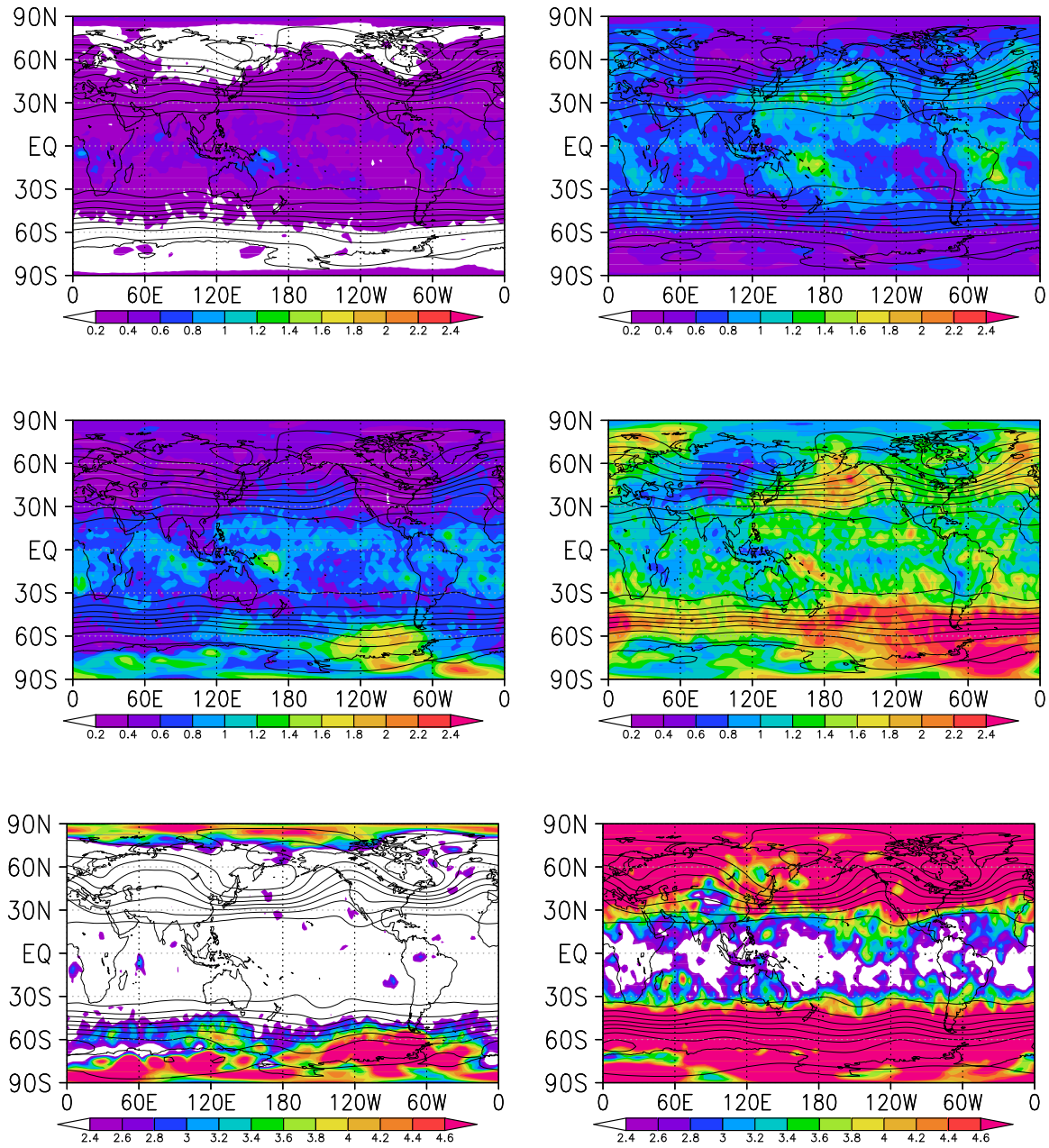

Fig. 5. Time-mean absolute analysis/forecast error of the meridional wind component (shades) and the geopotential height of the verifying analysis/forecast(contours) at the 500 hPa pressure level. Results are shown for the analysis (left) and the 72-hr forecast (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). The average is taken over all forecasts started between 11 January 2004 0000UTC and 15 February 2004 0000UTC. Note the different scale for the forecast errors in the bottom panels.

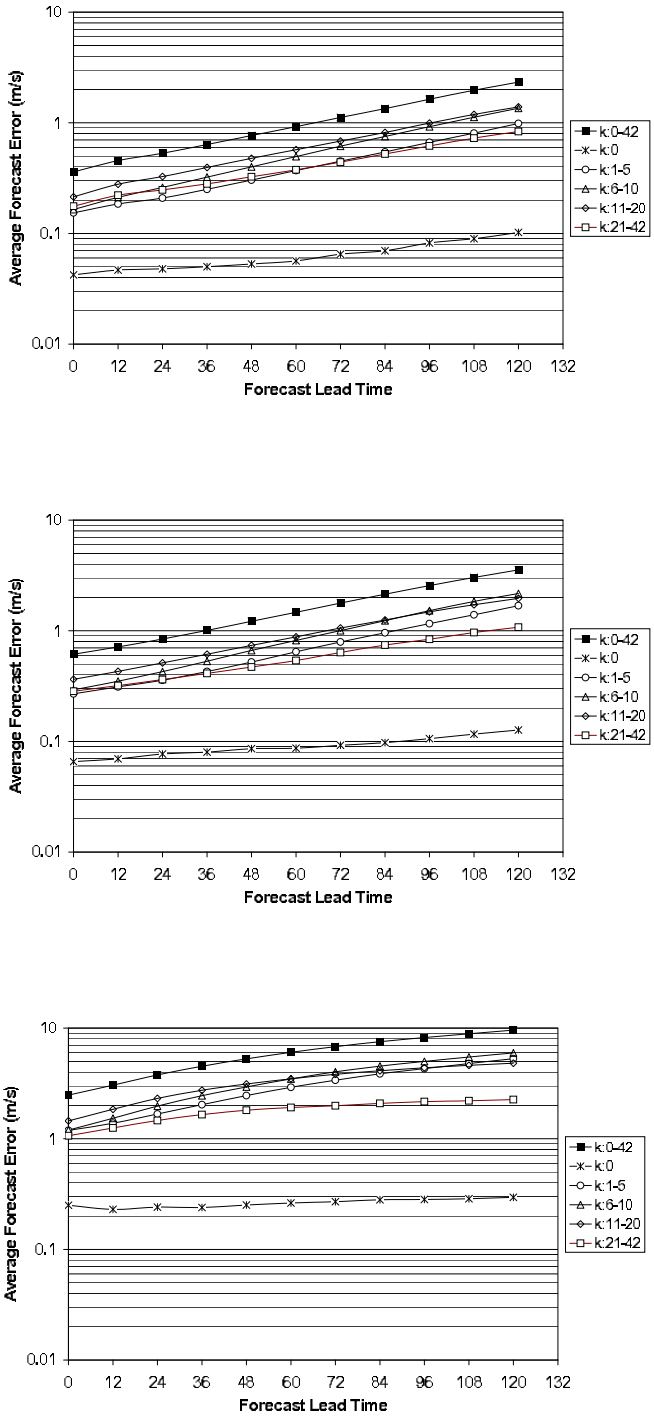

Fig. 6. Dependence of the time mean forecast error on the forecast lead time for the meridional wind component at 500 hPa level in the NH extratropics. The evolution of the forecast error is shown for different ranges of the zonal wavenumber k

.

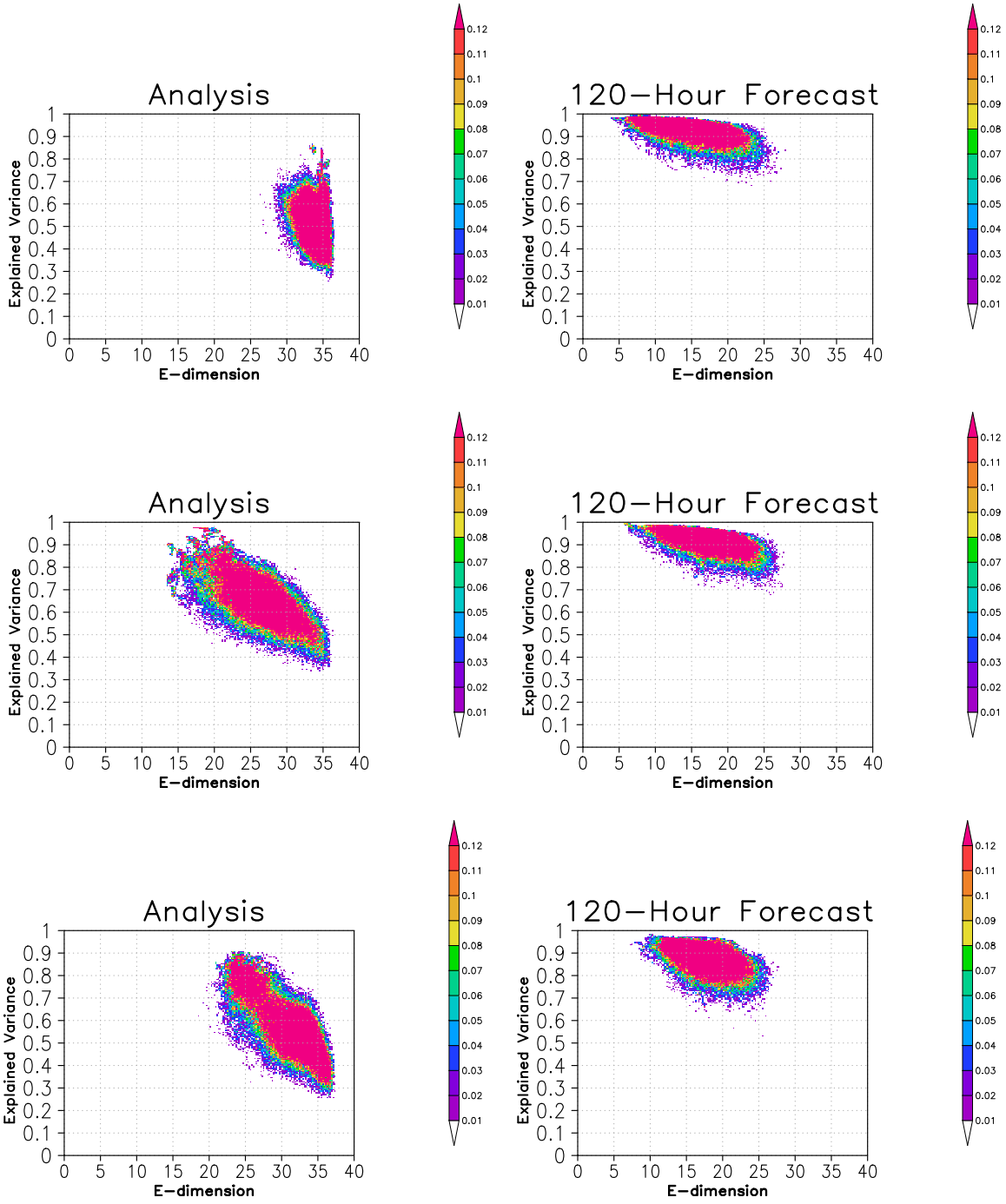

Fig. 7. Joint probability distribution of the E-dimension and the explained variance in the NH extratropics. The bin increments are 0.005 for the explained variance and 0.2 for the E-dimension. Shown are the distributions for the analysis (left) and the 120-hr forecast lead time (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel).

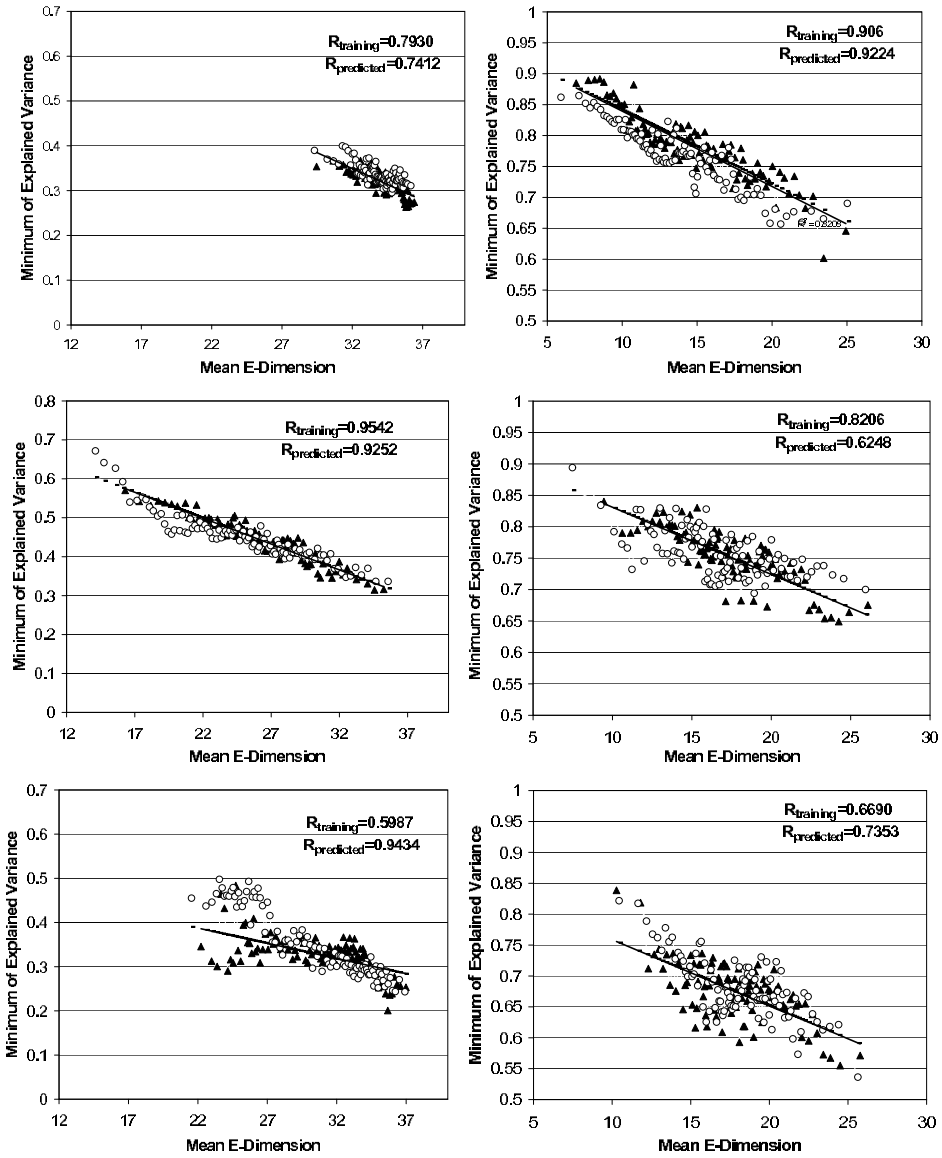

Fig. 8. Mean E-dimension and the minimum of explained variance of data divided equally into 100 bins for the NH extratropics for the first 18 days (triangles). The linear regression curve fitted to these data is shown by a solid straight line. If the prediction of the explained variance by the linear statistical model was perfect, the actual values for the second 18 days (open circles) would fall on this line. Shown are the distributions for the analysis (left) and the 120-hr forecast lead time (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). The legends show the correlation between E and $min_E EV$ in the training data set ($R_{training}$) and between E and the predicted value of $min_E EV$.

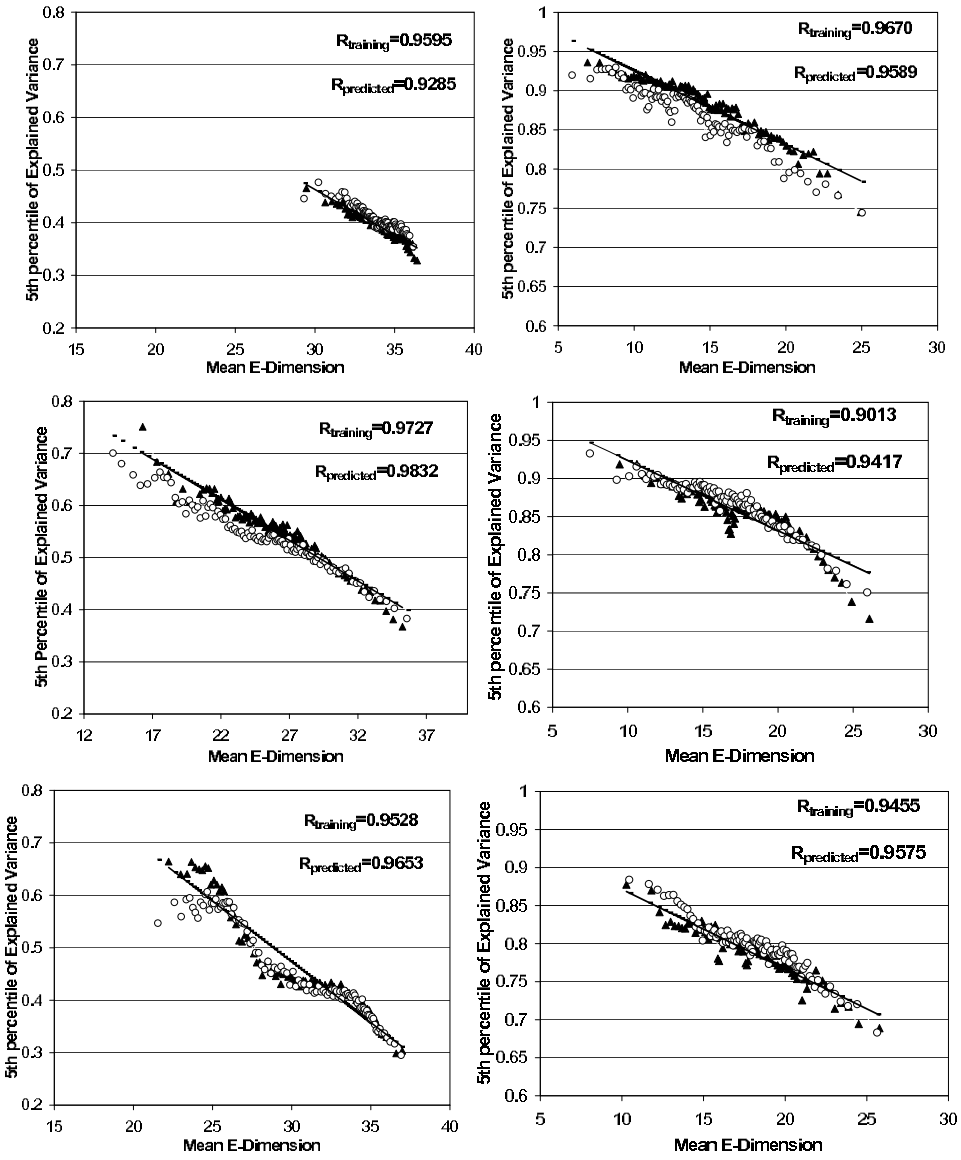

Fig. 9. Mean E-dimension and the 5th percentile of explained variance of data divided equally into 100 bins for the NH extratropics for the first 18 days (triangles). The linear regression curve fitted to these data is shown by a solid straight line. If the prediction of the explained variance by the linear statistical model was perfect, the actual values for the second 18 days (open circles) would fall on this line. Shown are the distributions for the analysis (left) and the 120-hr forecast lead time (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). The legends show the correlation between E and $min_E EV$ in the training data set ($R_{training}$) and between E and the predicted value of $min_E EV$.

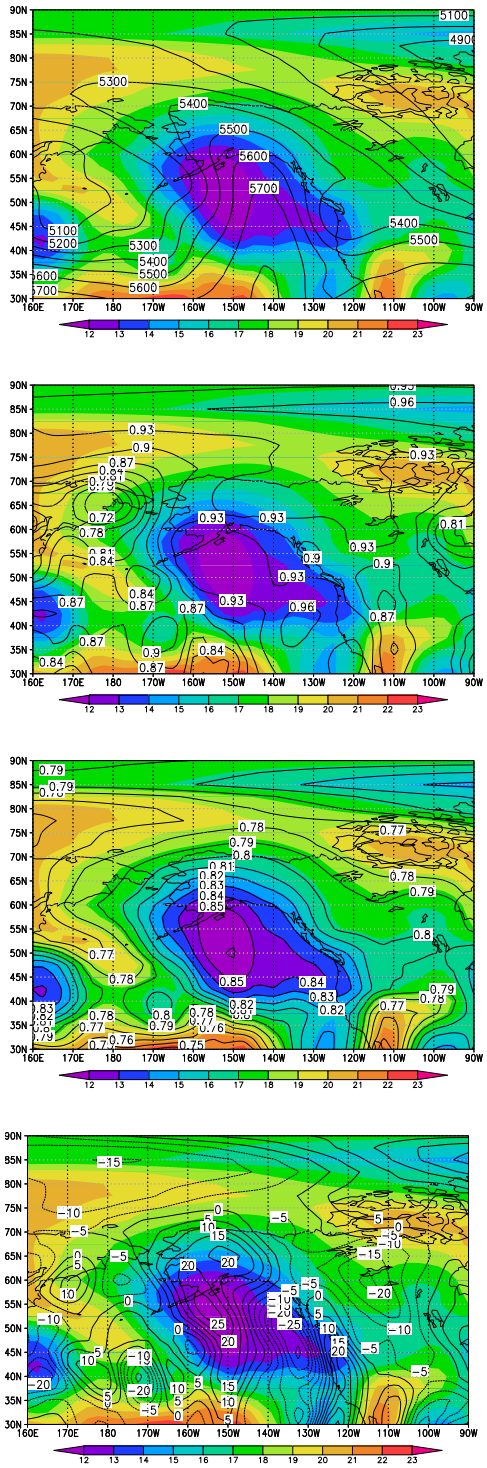

Fig. 10. Shown by contours are the 500 hPa geopotential heights (a) the actual values of explained variance (b)the predicted values of explained variance (c) and the error in the 500 hPa meridional wind (d) for the 120-hr forecast started on 20 January 2004 0000 UTC for the experiment that assimilates observations of the real atmosphere. Values of the E-dimension are shown by color shades in all four panels.

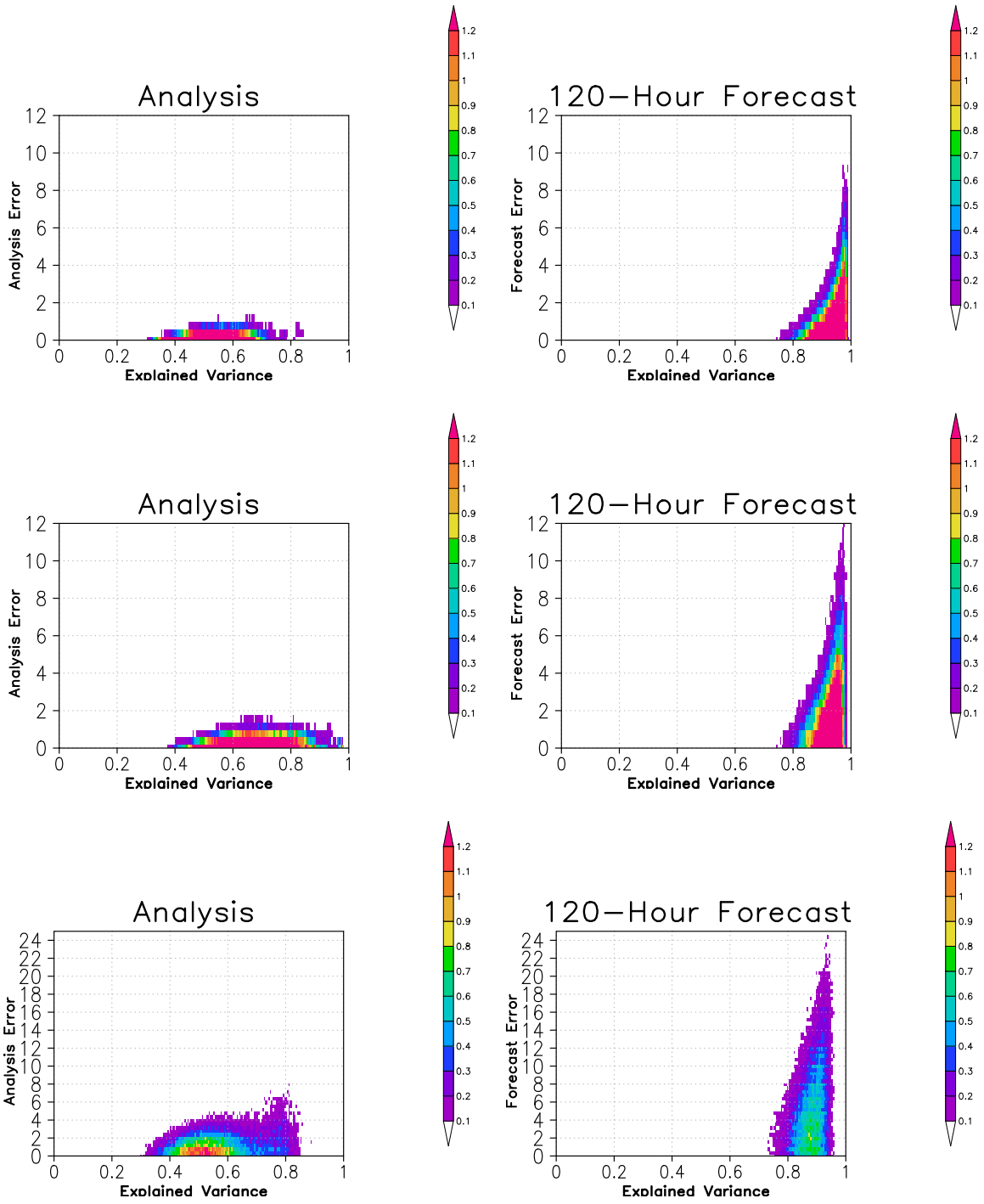

Fig. 11. Joint probability distribution of the analysis/forecast errors and the explained variance. The bin increments are 0.005 for the explained variance and 0.4 for the forecast error. Results are shown for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel) for analysis (left) and the 120-hr forecast (right). Note the different scale for the forecast errors in the bottom panel.

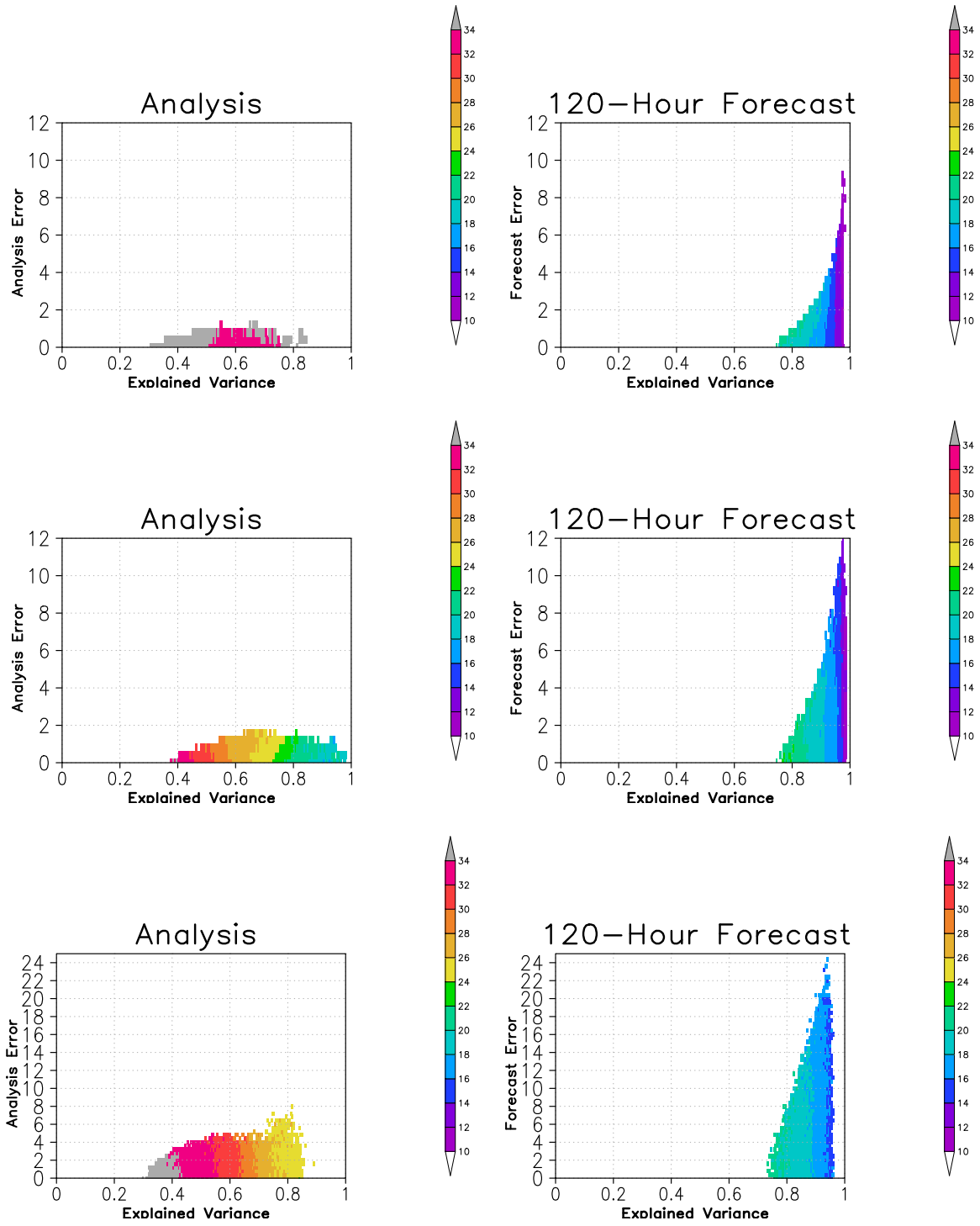

Fig. 12. Color shades indicate the mean E-dimension for each nonempty bin in Figure 11. Shown are the distributions for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). Note the different scale for the forecast errors in the bottom panel.

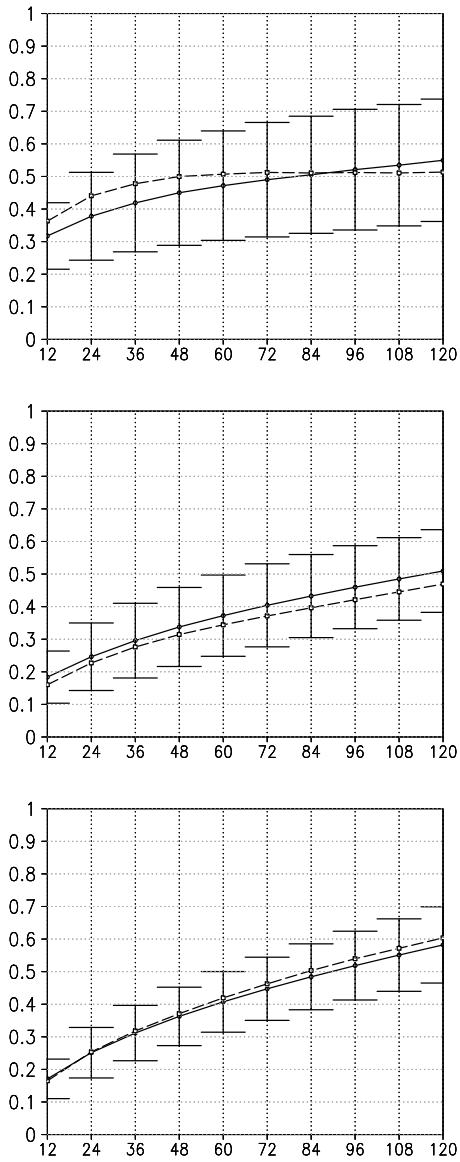

Fig. 13. Time-mean of the globally averaged relative nonlinearity at each forecast lead time. Results are shown for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). The solid line shows the results obtained when the relative nonlinearity is first calculated for local regions consisting of 5x5 model grid points and then the results are averaged over all local regions. The associated standard deviation is also shown. The dashed line shows the nonlinearity index computed for the global state vector.

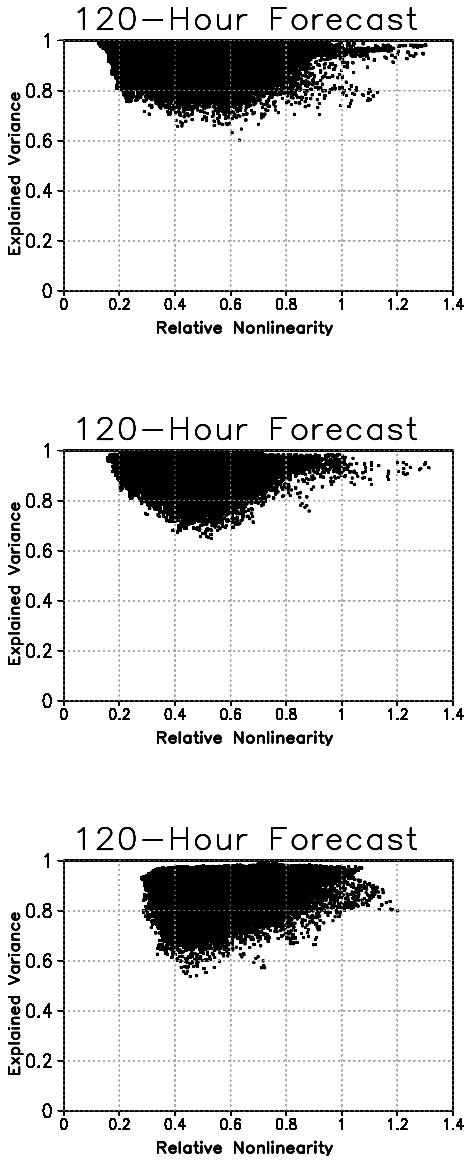

Fig. 14. Scatter plot of the NH explained variance and the relative nonlinearity. Results are shown for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel).

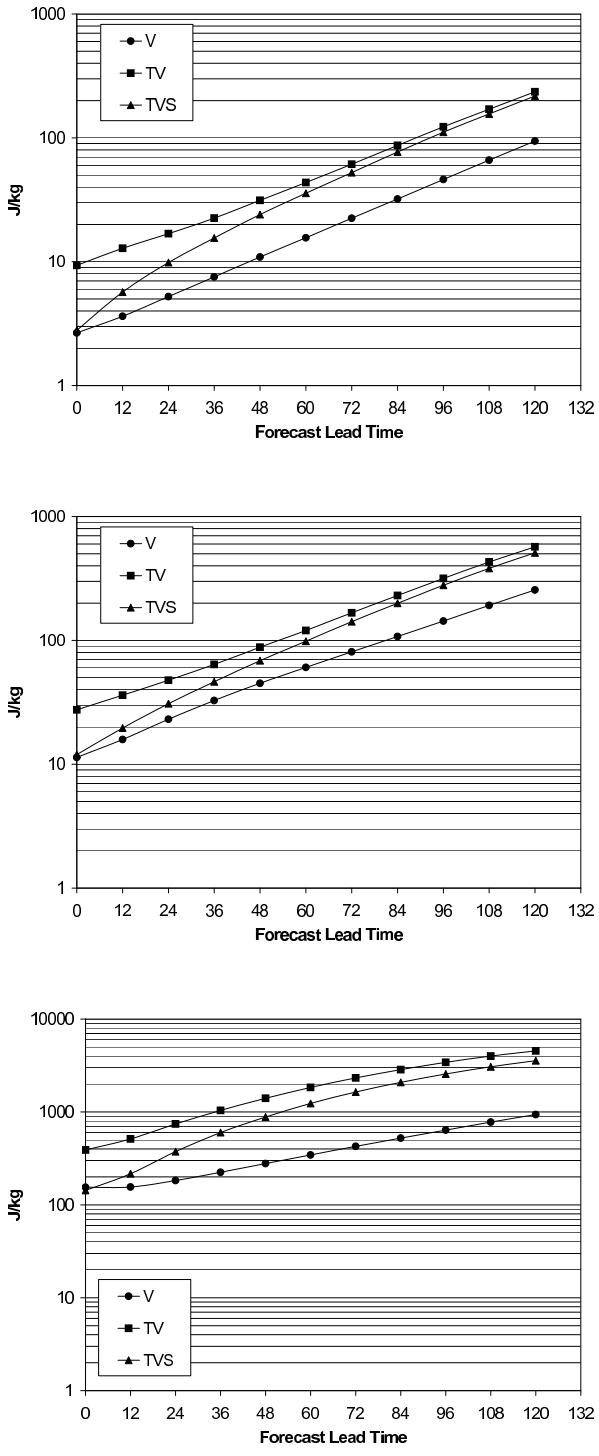

Fig. 15. The time evolution of TV (squares), TVS (triangles), and V (circles) for the NH extratropics. Results are shown for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and observations of the real atmosphere (bottom panel). Note the different scale in the bottom panel.

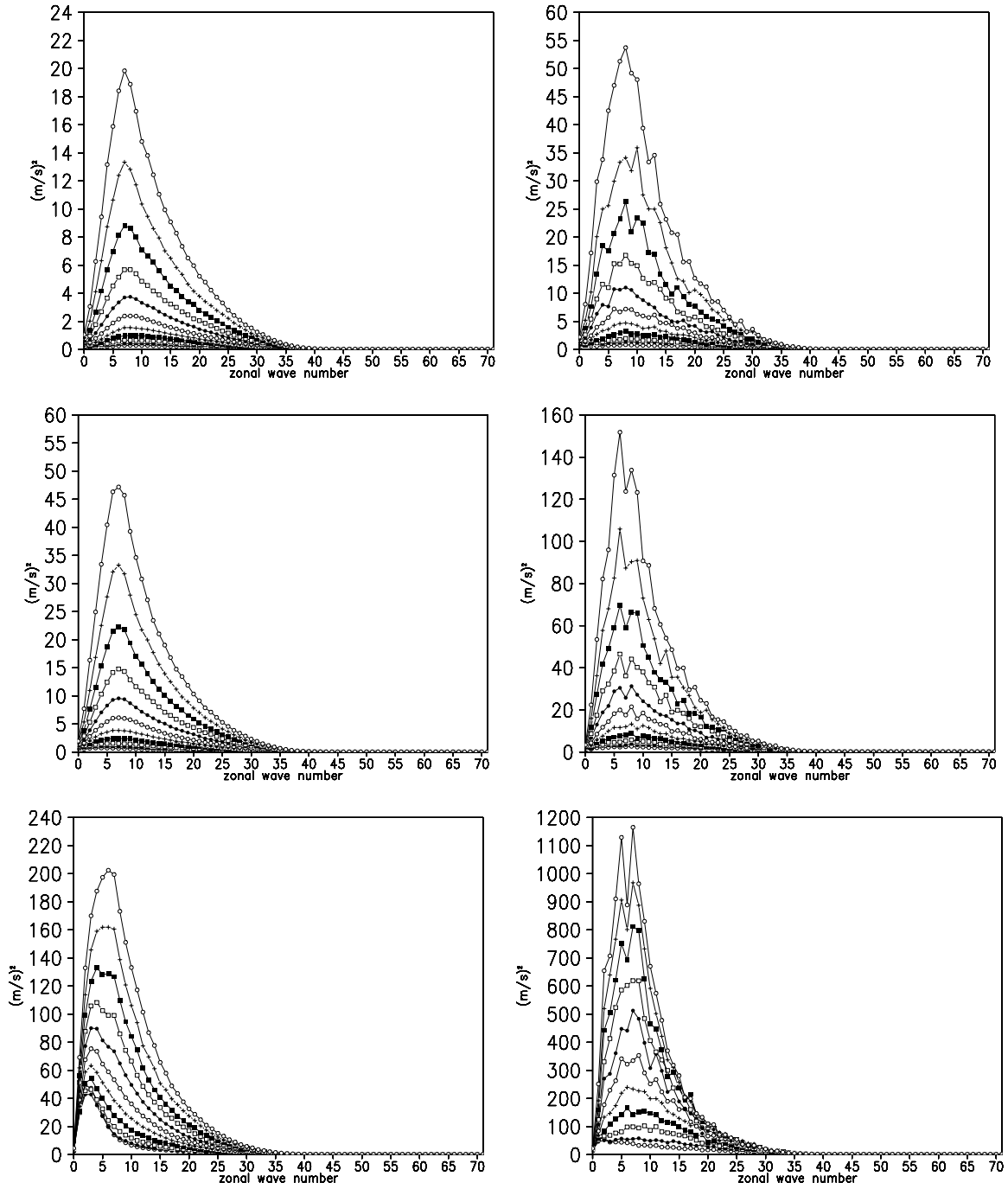

Fig. 16. The zonal power spectrum of the meridional component of the wind averaged over all latitudes in the NH extratropics and over time. Results are shown for 00-hr through 120-hr forecast lead times at 12-hr increments, averaged over the ensemble perturbations (left) and for $\delta\mathbf{x}^t$ (right) for the experiments which assimilate simulated observations in random locations (top panels), simulated observations in realistic locations (middle panels), and observations of the real atmosphere (bottom panels).

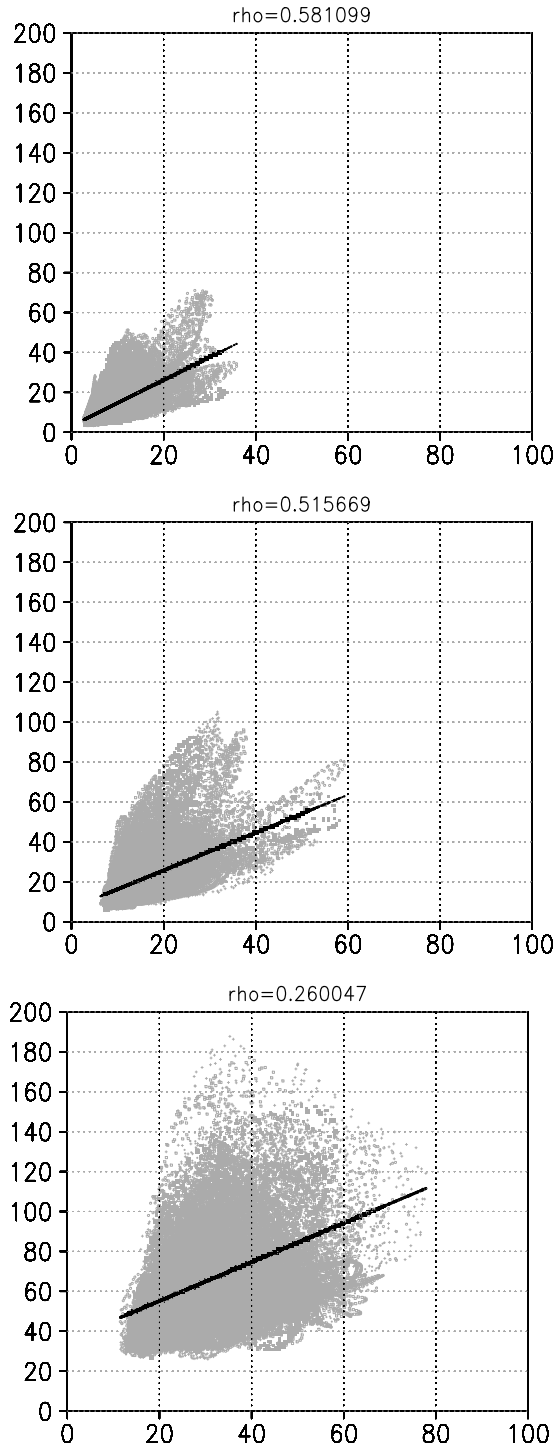

Fig. 17. Linear regression for ensemble skill based on spread. Shown are the NH results for the actual values (gray dots) and predicted values (black line) at the 5-day lead time for experiments that assimilate simulated observations in random locations (top panel), simulated observations in realistic locations (middle panel) and observations of the real atmosphere (bottom panel).

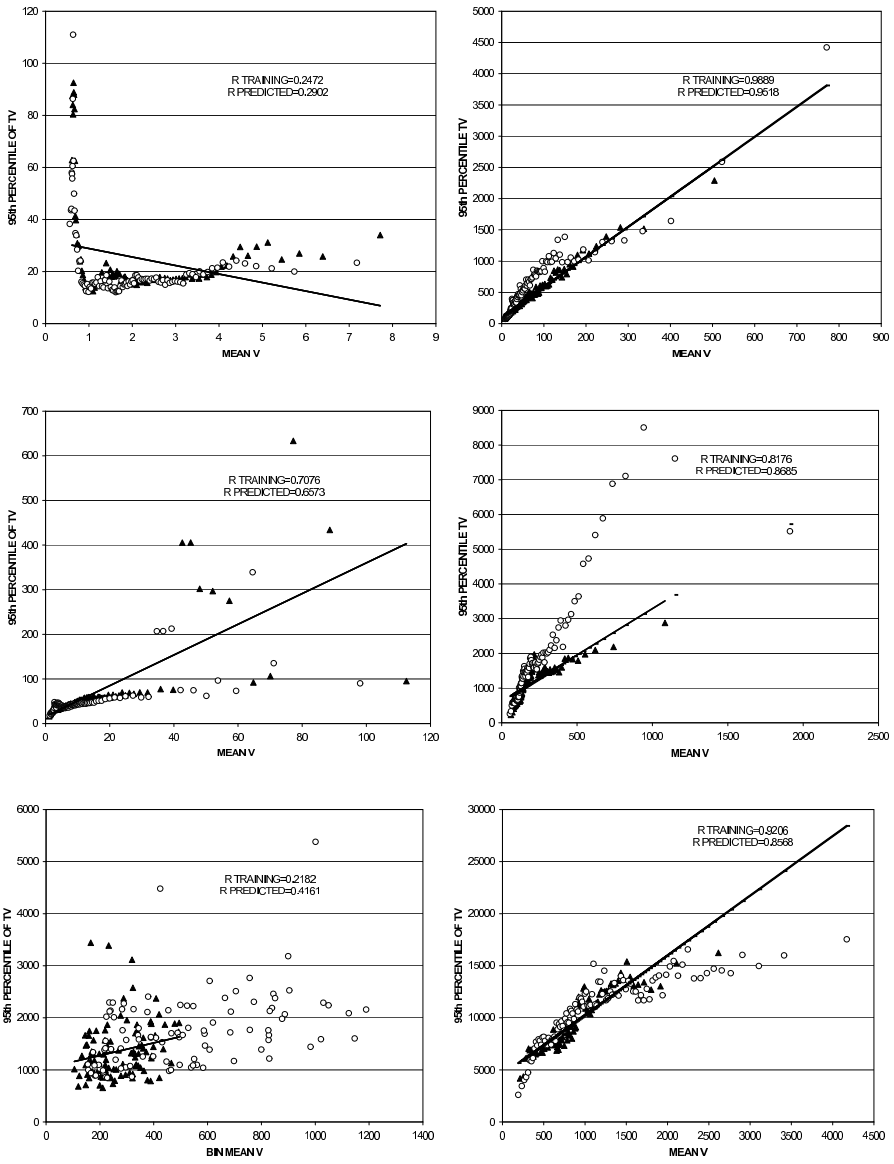

Fig. 18. Mean V and the 95th percentile of TV of data divided equally into 100 bins for the NH extratropics for the first 18 days (triangles). The linear regression curve fitted to these data is shown by a solid straight line. If the prediction of the 95th percentile of TV by the linear statistical model was perfect, the actual values for the second 18 days (open circles) would fall on this line. Shown are the distributions for the analysis (left) and the 120-hr forecast lead time (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and observations of the real atmosphere (bottom panel). The legends show the correlation between V and 95th percentile of TV in the training data set ($R_{training}$) and between V and the predicted value of the 95th percentile of TV.

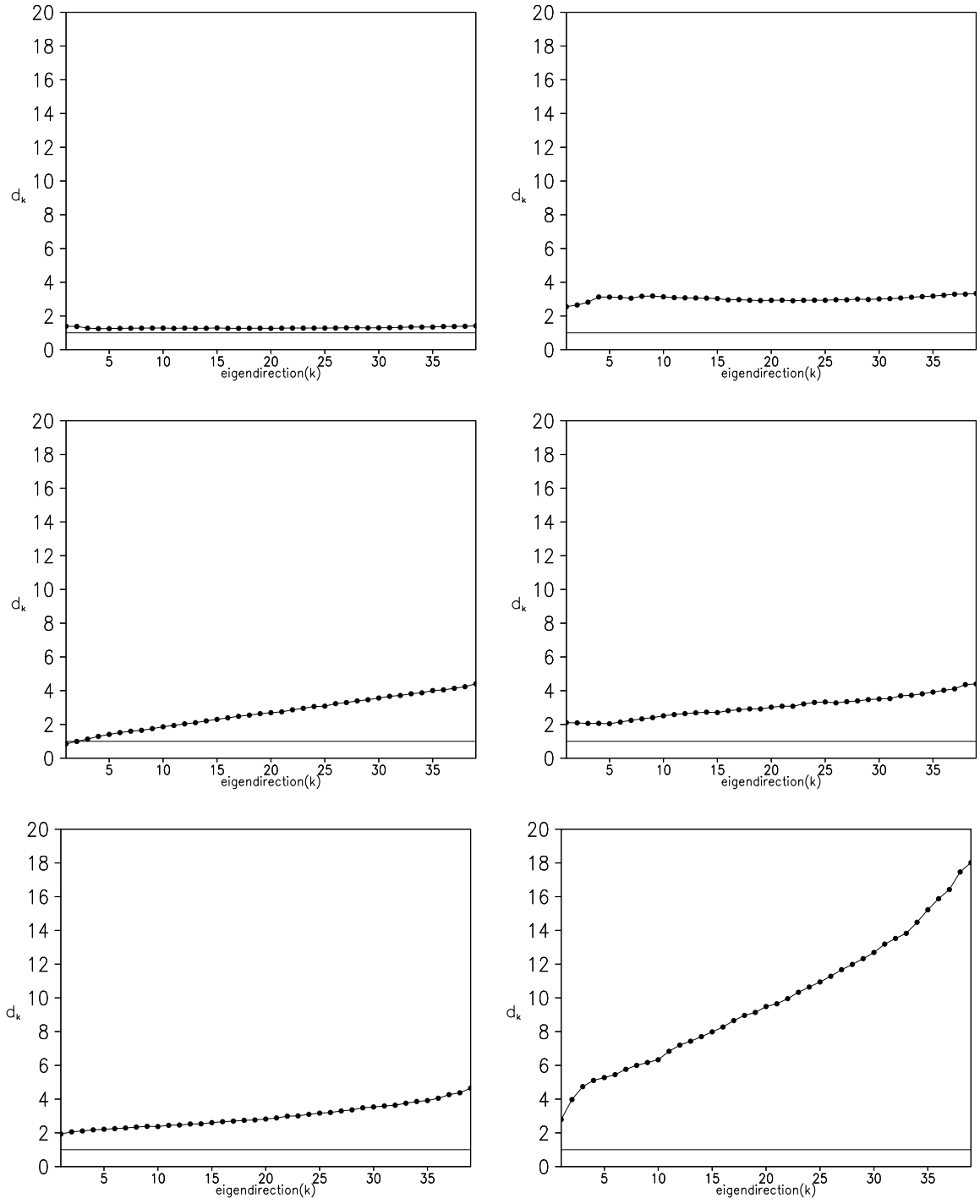

Fig. 19. The time mean of the Northern Hemisphere average spectrum of the ratio d_k, calculated for all assimilated variables in the local regions with energy rescaling. Results are shown for analysis time (left) and the 5-day lead time (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and observations of the real atmosphere (bottom panel). The average is taken over all forecasts started between 11 January 2004 0000UTC and 15 February 2004 0000UTC.

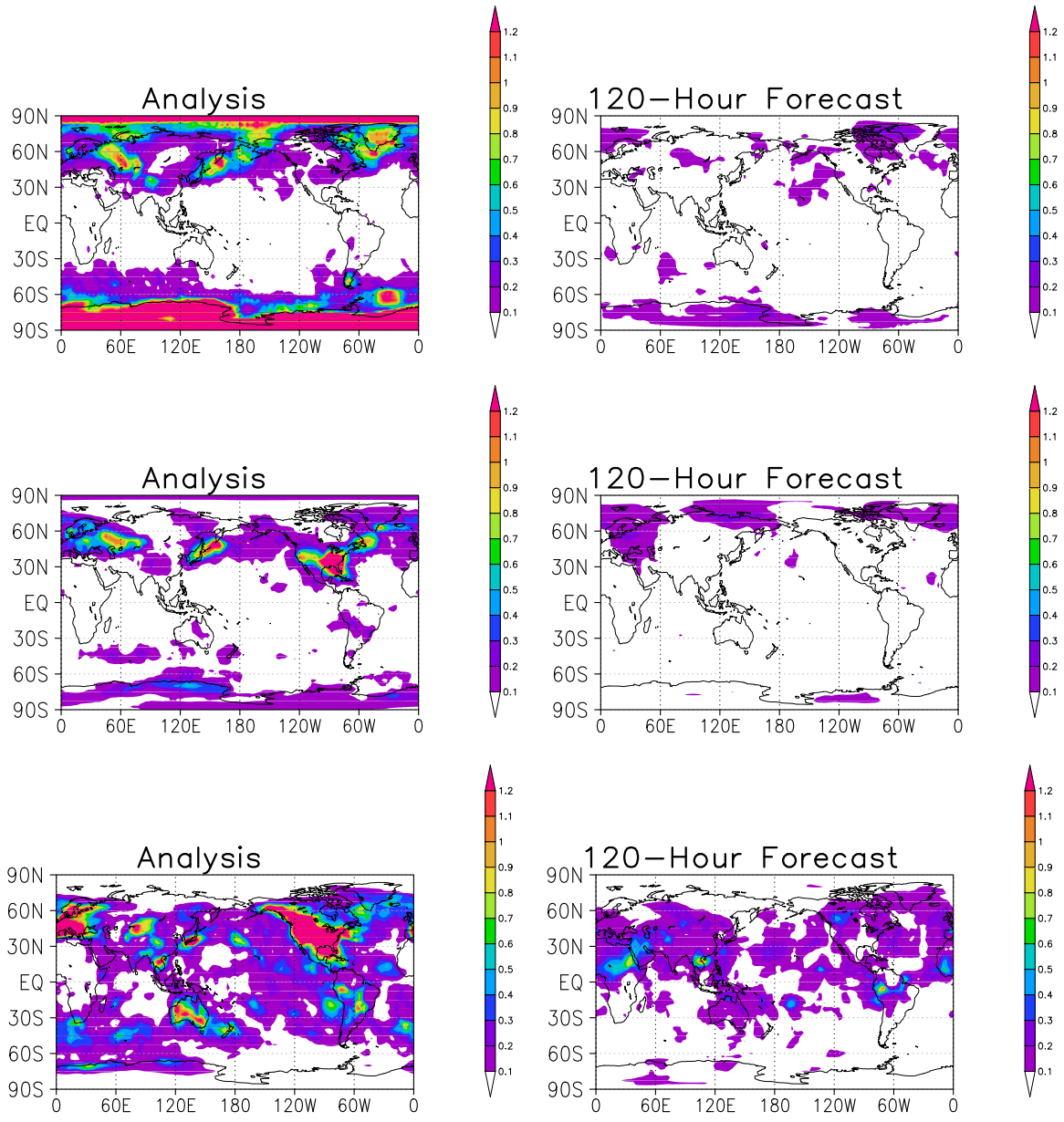

Fig. 20. The time average of the ratio d_k in the leading direction for the temperature at 850 hPa. Results are shown for analysis time (left) and the 5-day forecast (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and observations of the real atmosphere (bottom panel). The average is taken over all forecasts started between 11 January 2004 0000UTC and 15 February 2004 0000UTC.

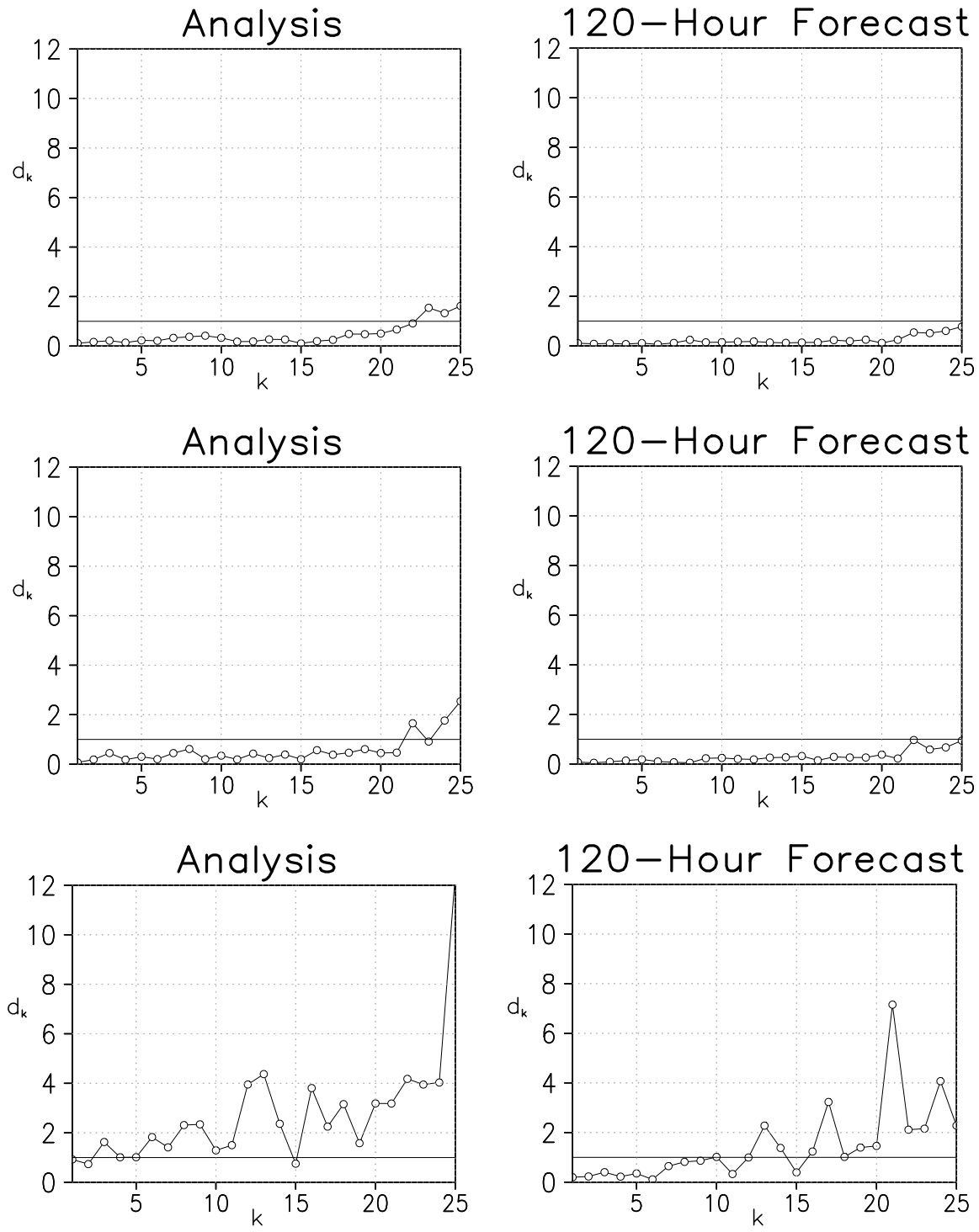

Fig. 21. The spectrum of the ratio d_k at the point 60°N 120°W for the temperature at 850 hPa. Results are shown for the analysis time (left) and the 5-day forecast (right) for experiments that assimilate randomly distributed simulated observations (top panel), simulated observations at the locations of conventional observations (middle panel), and conventional observations of the real atmosphere (bottom panel). The average is taken over all forecasts started between 11 January 2004 0000UTC and 15 February 2004 0000UTC.

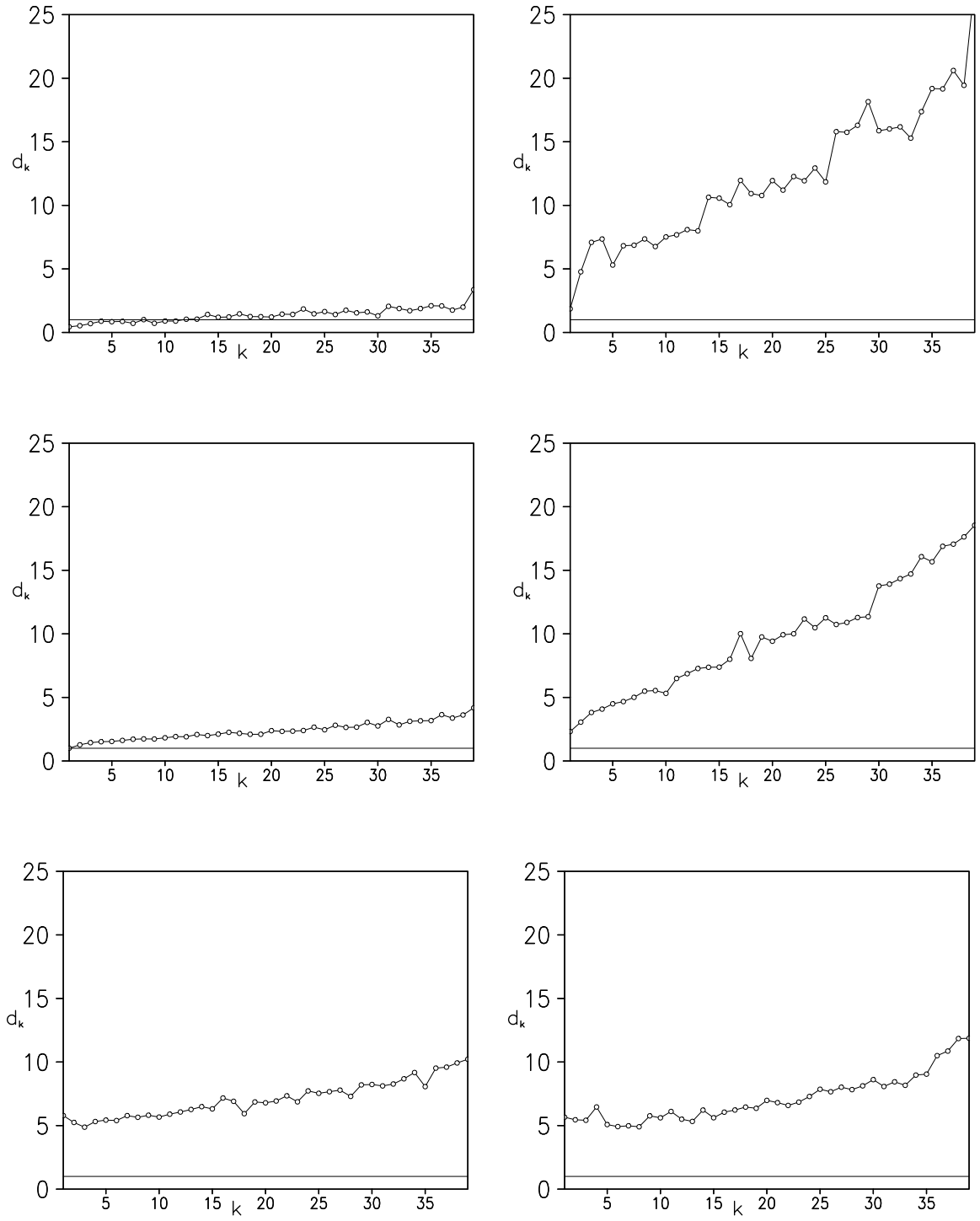

Fig. 22. The time mean of the Northern Hemisphere average spectrum of the ratio d_k, calculated for all assimilated variables in local regions with energy rescaling. Results are shown for observations of the real atmosphere for the minimum bin average of E-dimension (top panels), median bin average of E-dimension (middle panels), and maximum bin average of E-dimension (bottom panels) for analysis time (left panels) and the 120-hour forecast lead time (right panels).

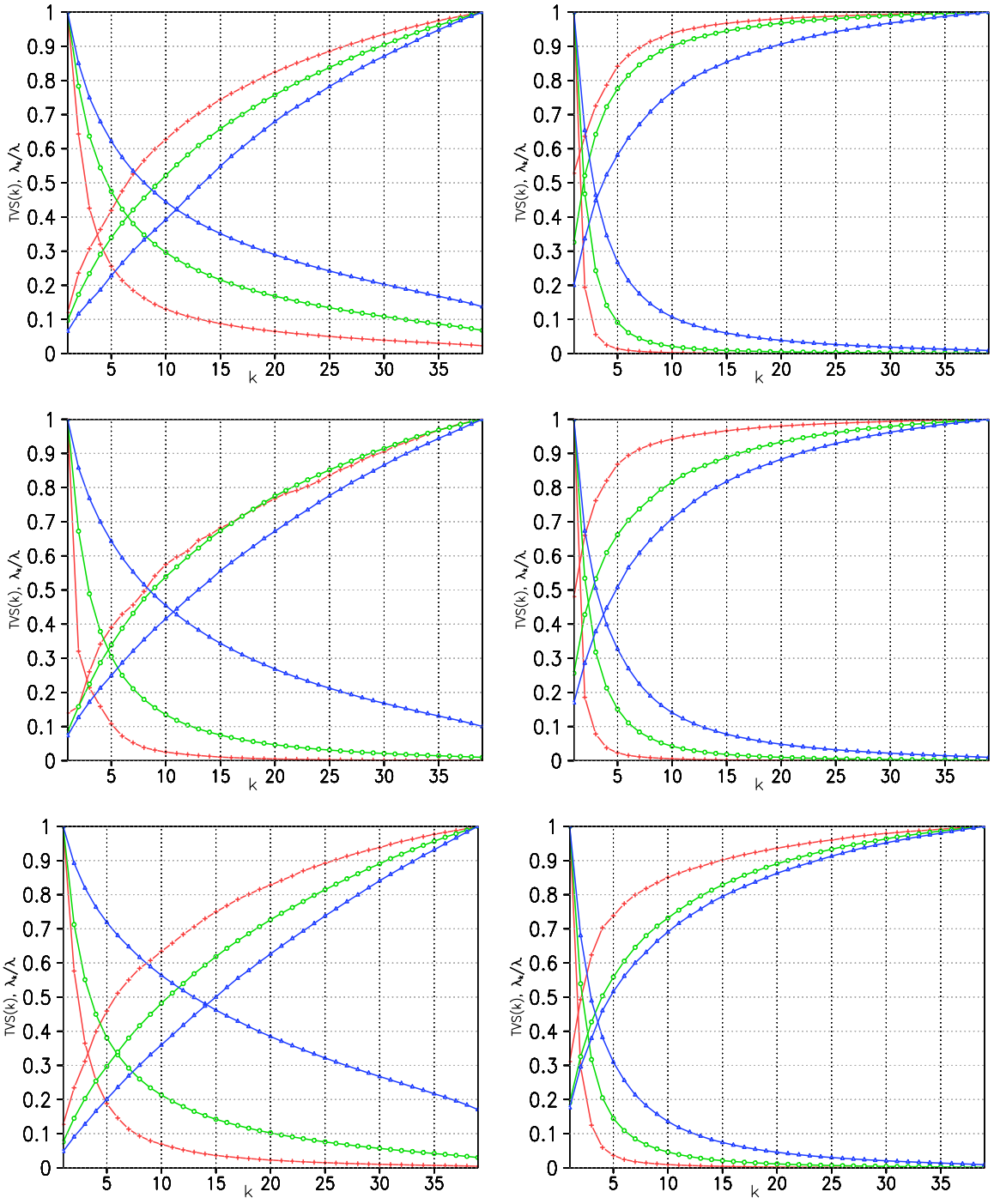

Fig. 23. The eigenvalue spectrum (normalized by the leading eigenvalue) and the percentage of TV_S for low (red plus signs), median (green open circles), and high (blue triangles) values of E for the Northern Hemisphere. Shown are the results at analysis time (left) and at 120-hour forecast lead time (right) for the experiments which assimilate simulated observations in random locations (top panels), simulated observations in realistic locations (middle panels), and observation of the real atmosphere (bottom panels).

VITA

Elizabeth A. Satterfield received her Bachelor of Science in applied mathematics from the Georgia Institute of Technology in 2002. She entered the Atmospheric and Oceanic Science Department of the University of Maryland in August 2004 and received her Master of Science degree in December 2008. In January 2009 she transfered to Texas A&M, where she received her PhD in May 2010. Her research interests include ensemble predictability and data assimilation. She recently accepted a postdoctoral research position at the Naval Research Laboratory in Monterey, CA.

Ms. Satterfield may be reached at the Department of Atmospheric Sciences, MS 3150, Texas A&M University, College Station, TX, 77843-3150. Her email address is easatterfield@tamu.edu