Local Diagnostics to Measure the Efficiency of the Ensemble in Representing the Error Space

Istvan Szunyogh[1] **and Elizabeth Satterfield**[2]

[1]*Texas A&M University, College Station, TX, USA*
szunyogh@tamu.edu

[2]*Current affiliation: National Research Council/Naval Research Lab, Monterey, CA, USA*
elizabeth.satterfield.ctr@nrlmry.navy.mil

ABSTRACT

This paper employs local, in physical space, diagnostics to analyze the contribution of model errors to the forecast uncertainty in experiments with a reduced (T62L28) resolution version of the NCEP GFS model. It is shown that in the extratropics the effect of model errors on the dominant patterns of medium range forecast uncertainty is secondary compared to their effects on the magnitude of the errors. It is argued, that this result strongly supports the development of post-processing techniques, which operate on the space of the local ensemble perturbations, and the further development of Stochastic Kinetic Energy Backscatter (SKEB) schemes.

1 Introduction

The main goal of this paper is to demonstrate that the local diagnostic tools we introduced in (3; 4) provide a framework to diagnose the contribution of model errors to the forecast uncertainty. In particular, we demonstrate that in the extratropics, in the medium forecast range and beyond (for forecast times longer than about 72–h), model errors have only modest influence on the dominant error patterns; thus, the most important error patterns can be well predicted by the dynamical instabilities identified under a perfect model hypothesis. Model errors, however, have a strong effect on the magnitude of the forecast errors, making the prediction of the magnitude of the total forecast error and the relative importance of the different error patterns difficult. We argue that because the ensemble provides an efficient representation of the error space, statistical post-processing techniques have great potential to provide useful forecast information about the forecast uncertainty.

The structure of the paper is as follows. Section 2 briefly describes the local diagnostics employed in this paper. Section 3 describes the design of the numerical experiments, while Section 4 presents the key results of the experiments. The paper concludes with a brief discussion of the implications of our results for the representation of the effects of model uncertainties in ensemble forecasting in Section 5.

2 Local diagnostics

The first paper to characterize the model dynamics by the behavior of an ensemble of analyses or forecasts in a local volume around each grid point was (2): the dynamics at model grid point ℓ are characterized by the behavior of the ensemble in a local volume centered at ℓ.

In what follows, we introduce local diagnostics by first defining the *local state vector* $\mathbf{x}_{V_\ell}$, whose components are the grid point variables in a local volume V_ℓ centered at ℓ. The dimension of $\mathbf{x}_{V_\ell}$ is equal to the M_{V_ℓ} number of grid-point variables in V_ℓ. We also assume the availability of a K-member forecast

ensemble and introduce the notation $\mathbf{x}^k_{V_\ell}$, $k = 1,\ldots,K$, for the members of the *ensemble of local state vectors* at location ℓ. Then, the *local ensemble perturbations*, $\mathbf{X}^k_{V_\ell}$, $k = 1,\ldots,K$, can be defined as

$$\mathbf{X}^k_{V_\ell} = \mathbf{x}^k_{V_\ell} - \overline{\mathbf{x}}_{V_\ell}, \tag{1}$$

where

$$\overline{\mathbf{x}}_{V_\ell} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}^k_{V_\ell} \tag{2}$$

is the *local ensemble mean*. (The local ensemble mean at ℓ is the mean state at grid points contained in the local volume centered at ℓ.) With the help of the ensemble of local ensemble perturbations, we can define the *local ensemble based error covariance matrix*,

$$\mathbf{P}_{V_\ell} = \frac{1}{K-1}\sum_{k=1}^{K} \mathbf{X}^k_{V_\ell}\left[\mathbf{X}^k_{V_\ell}\right]^T. \tag{3}$$

The local state vector, local perturbations, and local error covariance matrix can be defined at any forecast lead time, including the analysis time (forecast time zero). We treat the space of ensemble perturbations, defined by the range of P_{V_ℓ}, as a linear (vector) space: we assume that a linear combination of the local ensemble perturbations at an arbitrary location ℓ is also a local perturbation. We introduce the notation S_{V_ℓ} for this vector space. The assumption that S_{V_ℓ} is a linear space implies that the addition of a linear combination of the local ensemble perturbatins to the local mean $\overline{\mathbf{x}}_{V_\ell}$ results in a plausible local atmospheric state. Further, a local state can be written,

$$\mathbf{x}_{V_\ell} = \overline{\mathbf{x}}_{V_\ell} + \mathbf{X}_{V_\ell}\mathbf{w}_{V_\ell}, \tag{4}$$

where, $\mathbf{X}_{V_\ell}$ is a $M_{V_\ell} \times K$ matrix, whose kth column is $\mathbf{X}^k_{V_\ell}$ and $\mathbf{w}_{V_\ell}$ is a *vector of linear weights*. Introducing the notation $S^w_{V_\ell}$ for the space of the K-dimensional weight vectors,$\mathbf{w}_{V_\ell}$, the *matrix of local ensemble perturbations*, $\mathbf{X}_{V_\ell}$, represents a map from $S^w_{V_\ell}$ to $S^K_{V_\ell}$.

Since, typically, $M_{V_\ell} \gg K$, not all local state vectors can be (fully) represented in S_{V_ℓ}. The most straightforward approach to determine the projections of local state vectors is through computing the SVD-expansion of $\mathbf{X}_{V_\ell}$,

$$\mathbf{X}_{V_\ell} = \mathbf{U}_{V_\ell}\Sigma_{V_\ell}\mathbf{V}^T_{V_\ell}. \tag{5}$$

In Eq. (5), the columns of the $M_{V_\ell} \times r$ matrix $\mathbf{U}_{V_\ell}$ are the first r left singular vectors of $\mathbf{X}_{V_\ell}$, where r is the number of non-zero singular values of $\mathbf{X}_{V_\ell}$; the columns of the $K \times r$ matrix $\mathbf{V}_{V_\ell}$ are the first r right singular vectors of $\mathbf{X}_{V_\ell}$, while the diagonal entries of the $r \times r$ diagonal matrix Σ_{V_ℓ} are the non-zero singular values of $\mathbf{X}_{V_\ell}$. [The left singular vectors of $\mathbf{X}_{V_\ell}$ are also the orthogonal eigenvectors of the covariance matrix $\mathbf{P}_{V_\ell}$, the square of the singular values of $\mathbf{X}_{V_\ell}$ are the eigenvalues of $\mathbf{P}_{V_\ell}$, and r is the rank of $\mathbf{P}_{V_\ell}$.] The $\mathbf{u}^{(1)}_{V_\ell}\ldots,\mathbf{u}^{(r)}_{V_\ell}$ left singular vectors of $\mathbf{X}_{V_\ell}$ span the range of $\mathbf{X}_{V_\ell}$; thus, they define an orthonormal basis in $S^K_{V_\ell}$. Since, according to Eq. (1), the sum of the ensemble perturbations is $\mathbf{0}$, the columns of $\mathbf{X}_{V_\ell}$ cannot be all linearly independent, which implies that $r \leq K-1$. Our experience, accumulated by computing the singular value spectrum for different ensembles, suggests that while some of the trailing singular values are typically small, they are sufficiently larger than round-off to be considered non-zero. This motivates us to assume in what follows that $r = K-1$.

We apply diagnostics to the

$$\delta\mathbf{x}^t_{V_\ell} = \mathbf{x}^t_{V_\ell} - \bar{\mathbf{x}}_{V_\ell} \tag{6}$$

difference between the $\mathbf{x}^t_{V_\ell}$ local model representation of the true state and the ensemble mean, $\bar{\mathbf{x}}_{V_\ell}$. The difference $\delta\mathbf{x}^t_{V_\ell}$ is usually interpreted as the error in the ensemble mean forecast. This terminology is justified in forecast applications that use the ensemble mean as a deterministic forecast. [For a perfect ensemble, the deterministic forecast that has the smallest root-mean-square error is the ensemble mean

(1)]. In our study, however, we strictly view $\bar{\mathbf{x}}_{V_\ell}$ as the prediction of the mean of the probability distribution of the forecast state. Since, except for the analysis time, $\delta\mathbf{x}^t_{V_\ell}$ is expected to be nonzero even if $\bar{\mathbf{x}}_{V_\ell}$ is a perfect prediction of the mean of the probability distribution, we simply refer to $\delta\mathbf{x}^t$ as the difference between the ensemble mean and the model representations of the true state. We measure the magnitude of $\delta\mathbf{x}^t_{V_\ell}$ by

$$TV_{V_\ell} = \left(\delta\mathbf{x}^t_{V_\ell}\right)^2 = \left(\delta\mathbf{x}^t_{V_\ell}\right)^T \delta\mathbf{x}^t_{V_\ell}. \tag{7}$$

The $\delta\mathbf{x}^t_{V_\ell}$ difference can be decomposed as

$$\delta\mathbf{x}^t_{V_\ell} = \delta\mathbf{x}^{t(\|)}_{V_\ell} + \mathbf{x}^{t(\perp)}_{V_\ell}. \tag{8}$$

Here, $\delta\mathbf{x}^{t(\|)}_{V_\ell}$ is the component of $\delta\mathbf{x}^t_{V_\ell}$ that projects onto $S^K_{V_\ell}$ and $\mathbf{x}^{t(\perp)}$ is the part of $\delta\mathbf{x}^t$ that is orthogonal to $S^K_{V_\ell}$. The left singular vectors provide a convenient basis to compute $\delta\mathbf{x}^{t(\|)}$ by

$$\delta\mathbf{x}^{t(\|)}_{V_\ell} = \sum_{k=1}^{K-1} \left[(\delta\mathbf{x}^t_{V_\ell})^T \mathbf{u}^{(k)}_{V_\ell}\right] \mathbf{u}^{(k)}_{V_\ell}. \tag{9}$$

The magnitude of $\delta\mathbf{x}^{t(\|)}_{V_\ell}$ is

$$TVS_{V_\ell} = \left(\delta\mathbf{x}^{t(\|)}_{V_\ell}\right)^2 = \left(\delta\mathbf{x}^{t(\|)}_{V_\ell}\right)^T \delta\mathbf{x}^{t(\|)}_{V_\ell}. \tag{10}$$

TVS_{V_ℓ} measures the extent to which the linear space spanned by the ensemble perturbations can represent the possible states of the atmosphere: $TVS_{V_\ell} \leq TV_{V_\ell}$ and when $TVS_{V_\ell} = TV_{V_\ell}$, the vector space $S^K_{V_\ell}$ correctly captures the space of forecast uncertainties. (The degree to which $TVS_{V_\ell} = TV_{V_\ell}$ holds, is not necessarily a measure of the quality of the ensemble, because it is not guaranteed that a state computed by Eq. (4) from the the nonlinearly evolving ensemble perturbations is a possible state of the atmosphere.) We compare TV_{V_ℓ} and TVS_{V_ℓ} to

$$VS_{V_\ell} = \frac{1}{K}\sum_{k=1}^{K}\left(\mathbf{X}^k_{V_\ell}\right)^T \mathbf{X}^k_{V_\ell} = \mathrm{trace}\left(\mathbf{P}_{V_\ell}\right) = \frac{1}{K-1}\sum_{k=1}^{K-1}\left(\sigma^{(k)}\right)^2, \tag{11}$$

which is a measure of the ensemble spread in the local volume V_ℓ. In Eq. (11), $\sigma^{(1)}, \ldots, \sigma^{(r)}$ are the first $K-1$ singular values of $\mathbf{X}^k_{V_\ell}$.

We introduce the notations TV, TVS and VS, respectively, for the expected values of TV_{V_ℓ}, TVS_{V_ℓ} and VS_{V_ℓ}. For an ensemble that correctly represented the probability distribution of the state, or at least the second moment of the probability distribution, $VS = TV$ would hold, because from a probabilistic point of view, the $\delta\mathbf{x}^t_{V_\ell}$ perturbation is a random draw from the probability distribution that the ensemble perturbations are supposed to estimate. In addition, when the ensemble correctly represents the variance of $\delta\mathbf{x}^{t(\|)}_{V_\ell}$, $VS = TVS$. For a given ensemble system, VS can be either smaller or larger than TVS. In the former case ($VS < TVS$), the ensemble underestimates the magnitude of the uncertainty that can be explained by a linear combination of the ensemble perturbations, while in the later case the ensemble overestimates the magnitude of the uncertainty that can be explained by linearly combining the ensemble perturbations. The combination of $VS = TV$ and $VS > TVS$ indicates that a significant part of $\delta\mathbf{x}^t_{V_\ell}$ does not project onto $S^K_{V_\ell}$, which can be a result of either the specific nature of the nonlinearities in the dynamics or imperfections of the ensemble (e.g., not having a sufficient number of ensemble members or a design flaw in the generation of the initial conditions). For the ensemble we discuss in this paper, $VS < TVS$.

3 Experiment design

Our experiment design is identical to that of (3; 4). All experiments are carried out with a reduced resolution (T62L28) 2004 version of the model component of the operational NCEP GFS. Observations

are assimilated with the implementation of the LETKF data assimilation scheme on the NCEP GFS, which was described in detail in (5). We generate a $K = 40$ member analysis ensemble, which is then used to generate the background ensemble of the next analysis cycle and to provide the initial conditions of the 360-h ensemble forecasts used for the computation of TV, TVS and VS at the different forecast times. All numerical experiments are carried out for the time period 1 January 2004 0000 UTC and 29 February 1800 UTC. The diagnostics are computed based on the forecasts started at 0000 and 1200 UTC between 11 January and 15 February on a $2.5^\circ \times 2.5^\circ$ resolution grid.

To analyze the effects of model errors on the diagnostics, two experiments are carried out under the perfect model hypothesis, so the effects of model errors on the third (realistic) experiment can be assessed. The true state for the perfect model experiments was generated by a continuous integration of the model started from the operational analysis of NCEP at January 1 2004, 0000 UTC. The difference between the two perfect model experiments is that one of them assimilates vertical soundings of the model atmosphere at 2000 fixed, randomly selected locations, while the other assimilates observations whose types and locations are identical to those in the realistic experiment. In the third experiment, all non-radiance observations, which were assimilated at NCEP in real time, are assimilated with the LETKF. To compute the diagnostics for this experiment , high (T254L64) resolution operational NCEP analyses truncated to 2.5° x 2.5° resolution are used as proxy for the "true" state.

In our implementation of the LETKF, we apply a multiplicative covariance inflation at each analysis step to increase the magnitude of the estimated analysis uncertainty. Thus, 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. We tuned the covariance inflation factor independently for the three experiments, so that $V \approx TVS$ in each experiment. We note that this tuning condition is different from the one used in the common practice of ensemble forecasting, where the magnitude of the analysis ensemble perturbations is tuned to satisfy $VS \approx TV$. Our motivation to use $V \approx TVS$ instead of $V \approx TV$ is that in our setting the magnitude of the variance inflation affects, in addition to the analysis perturbations, the mean analysis: our experience is that increasing VS beyond TVS, leads to an increase of the analysis error (see Figure 1 of (4)). This increase results in an increase of TV at analysis time and at short lead times. Thus, increasing the variance inflation not only has the undesirable effect of increasing the analysis error, but is also fails to improve the relationship between VS and TV.

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. To properly weight the different model variables in the computation of the variances, 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 variance has dimension of energy. The use of this scaling to compute scalar products of perturbations of the state vector in a primitive equation model was first suggested by (6), and has been routinely used in the generation of ensemble perturbations at ECMWF. The expected value in the computation of TV, TVS and VS is estimated by the sample mean over all locations (grid points), ℓ, in the NH extratropics ($\varphi \geq 30^\circ N$) and over all forecasts started between 0000 UTC 11 January and 1200 UTC 15 February.

4 Results

Our results are summarized in Fig. 1. This figure extends the results of (3; 4) from a maximum forecast time of 120-h to a maximum forecast time of 360-h. The shape of the curve that shows the time evolution of TV is very similar in the three experiments, except that the initial value of TV becomes larger with the addition of each new realistic feature, leading to a faster saturation of TV in the more realistic experiments. The increase in the initial magnitude of TV is particularly large between the perfect model

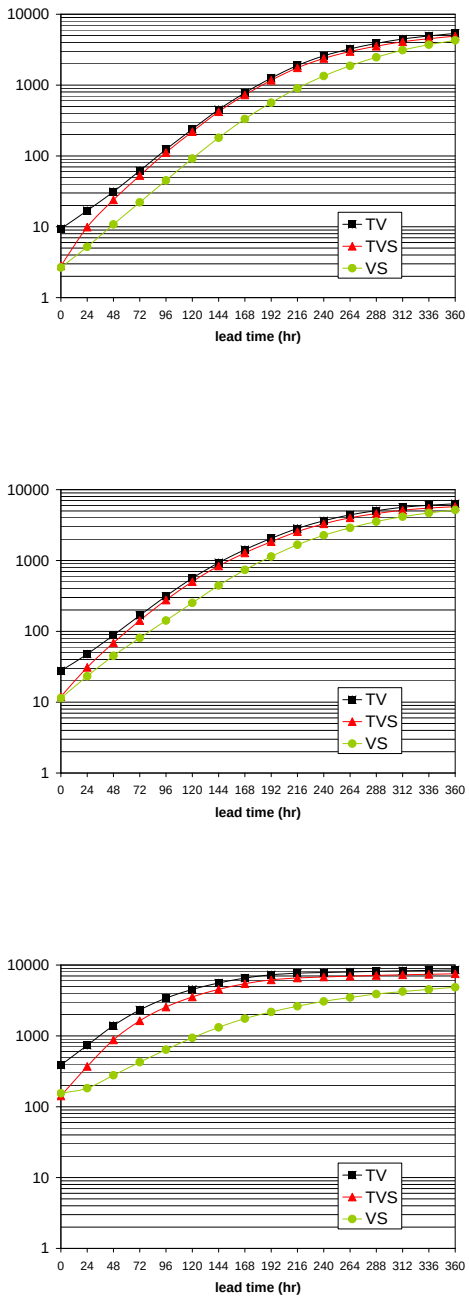

Figure 1: The time evolution of TV, TVS and VS for he perfect model experiment with randomly distributed observations (top); the perfect model experiment with realistically distributed observations (middle); and the realistic experiment with observations of the real atmosphere.

experiment with realistic observation distribution and the realistic experiment. This suggests that, as expected, the contribution of model errors to the analysis uncertainty is larger than the contribution of the nonuniform spatial distribution of the observations.

Perhaps the most interesting feature in Fig. 1 is the behavior of the curves that shows the time evolution of TVS: while TVS is initially smaller than TV, the two quantities become very similar after about 72-120-h in all three experiments. In a statistical sense, the linear space $S^K_{V_\ell}$ provides, after an initial transient period, an almost perfect representation of the space of forecast uncertainties, where the difference between the ensemble mean and the true system may fall. In (4), we showed that by the end of the initial (72-120-h long) transient period, synoptic scale patterns dominate both the difference between the truth and the ensemble mean and and the ensemble perturbations [see Figure 3 of (4)]. This result suggests that the dominant uncertainties are associated with synoptic scale features and $S^K_{V_\ell}$ provides an efficient representation of these patterns of uncertainties. The behavior we find is in good agreement with that of (7), who found that forecast uncertainties in spectral space were propagating toward the synoptic scale with time, where the growth of their magnitude became exponential.

It is important to note, that even at analysis time, where $S^K_{V_\ell}$ does a relatively poor job of capturing $\delta\mathbf{x}^t_{V_\ell}$ compared to the longer forecast lead times, it provides a sufficiently accurate representation of the error space for an efficient data assimilation. (In essence, the LETKF data assimilation scheme seeks for an analysis increment that best fits the projection of the vector of innovations on $S^K_{V_\ell}$.)

The fact, that the presence of model errors does not have an important effect on the efficiency of $S^K_{V_\ell}$ at capturing the dominant patterns of forecast uncertainties, indicates that model errors have little effect on these patterns. A further examination of Fig. 1 reveals, however, that model errors have an important effect on the magnitude of the patterns of forecast uncertainty: while the ensemble spread (VS) underestimates TVS in all three experiments, the problem is most severe for the case, in which errors are present. This behavior suggests that model errors make an important contribution to the growth of TVS; or in other words, to the amplification of the errors in $S^K_{V_\ell}$.

5 Discussion

Our results suggest that the dominant patterns of medium range forecast uncertainties in the extratropics can be efficiently captured by a small ($K < 100$) ensemble generated without explicitly accounting for the model uncertainty, even if model errors are present. The results also show, however, that model errors contribute to the underestimation of the forecast uncertainty in the ensemble space. Since the ensemble underestimates the magnitude of the forecast uncertainty in the ensemble space even in the perfect model scenario, we conjecture that the underestimation of the magnitude of the uncertainty is, in part, due to the lack of a representation of the effects of the initial condition uncertainties that fall outside of the ensemble space. While the magnitude of the uncertainty in the individual directions is small, their combined effect on the later evolution of the dominant patterns of uncertainty is important. In addition, we conjecture that model errors play a similar role to the initial condition uncertainties that fall outside of the ensemble space, with the important addition that model error related uncertainties are generated continuously at all forecast times.

On the practical side, our results show that while the dominant patterns of forecast uncertainty in the extratropics can be efficiently predicted without accounting for the model uncertainties, predicting the magnitude of the patterns of uncertainties remains a major challenge, which cannot be addressed without the development of techniques for the representation of the effects of model errors on the magnitude of the dominant forecast error patters. Our results suggest two potential approaches to improve the prediction of the magnitude of the forecast uncertainties. First, since the ensemble provides a good representation of the local space of uncertainties, linear post-processing techniques operating on $S^K_{V_\ell}$

may have great potential to improve the prediction of the magnitude of the uncertainties. Second, the development of stochastic parameterization schemes that can correctly represent the bulk effect of the physics-related uncertainties in the many state space directions on the dominant directions of uncertainties would be highly desirable. Refining the existing Stochastic Kinetic Energy Backscatter(SKEB) schemes would be a logical first step in that direction.

Finally, we note that extending our results to the tropics is not trivial, as the dynamics of uncertainties in the tropics is dominated by convectively coupled waves instead of the synoptic scale baroclinic waves that dominate the evolution of forecast uncertainties in the extratropics.

Acknowledgements

The research reported in this paper was funded by the National Research Foundation (Grant ATM-0935538). A portion of this research was performed while ES held a National Research Council Research Associateship Award at the Naval Research Laboratory.

References

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

[2] Patil, D., B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. *Phys. Rev. Lett.,* **86,** 5878–5881.

[3] Satterfield, E. and I. Szunyogh, 2010: Predictability of the performance of an ensemble forecast system: Predictability of the space of uncertainties. *Mon. Wea. Rev.,* **138,** 962–981.

[4] Satterfield, E., and I. Szunyogh, 2011: Assessing the performance of an ensemble forecast system in predicting the magnitude and the spectrum of analysis and forecast uncertainties. *Mon. Wea. Rev.,* **139,** 1207–1223.

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

[6] Talagrand, O., 1981: A study of the dynamics of four-dimensional data assimilation. *Tellus,* **33A,** 43-60.

[7] Tribbia, J. J. and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Mon. Wea. Rev.,* **126,** 3292-3302.