TESTING LINEAR DIAGNOSTICS OF ENSEMBLE PERFORMANCE ON A SIMPLIFIED GLOBAL CIRCULATION MODEL

A Senior Scholars Thesis

by

ETHAN LANE NELSON

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

April 2011

Major: Meteorology

TESTING LINEAR DIAGNOSTICS OF ENSEMBLE PERFORMANCE ON A SIMPLIFIED GLOBAL CIRCULATION MODEL

A Senior Scholars Thesis

by

ETHAN LANE NELSON

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by:

Research Advisor:	Istvan Szunyogh
Director for Undergraduate Research:	Sumana Datta

April 2011

Major: Meteorology

ABSTRACT

Testing Linear Diagnostics of Ensemble Performance on a Simplified Global Circulation Model. (April 2011)

Department of Atmospheric Science
Ethan Lane Nelson

Texas A&M University

Research Advisor: Dr. Istvan Szunyogh
Department of Atmospheric Science

Ensemble weather forecast systems are used to account for the uncertainty in the initial conditions of the atmosphere and the chaotic dynamics of the models. It has been previously found that forecast performance of an ensemble forecast system is inherently flow dependent and that the ensemble predicts potential patterns of forecast errors more reliably than the magnitudes of the errors. A low-resolution global circulation model is implemented to calculate linear diagnostics in the vector space of the ensemble perturbations. It is confirmed that the ensemble provides a good representation of the pattern of forecast uncertainties, but the magnitude of the error projected onto the space of ensemble perturbations is underestimated for the 6 to 120 hour forecast times.

DEDICATION

"For since the creation of the world God's invisible qualities—his eternal power and divine nature—have been clearly seen, being understood from what has been made, so that people are without excuse."

-Romans 1:20, NIV

ACKNOWLEDGMENTS

I would first like to thank Dr. Szunyogh for his guidance, help, and willingness to add me to his team as an undergraduate researcher. I also want to thank Dr. Gyorgyi Gyarmati for her wonderful assistance and advice.

Additionally, I wish to thank Dr. Seung-Jong Baek for the LETKF implementation code, along with the Abdus Salam International Centre for Theoretical Physics Atmospheric General Circulation Model Network for allowing me to use the SPEEDY model for this experiment. I would also like to thank Dr. Elizabeth Satterfield for her help with the diagnostic computational programs.

Lastly, I want to thank my family for their love and support over my lifetime; without their help, I could not be where I am today.

NOMENCLATURE

$\delta\mathbf{x}^{t}$	Ensemble Mean Error
EnKF	Ensemble-based Kalman Filter
GFS	Global Forecast System
K	Ensemble Member Size
LETKF	Local Ensemble Transform Kalman Filter
NCEP	National Centers for Environmental Prediction
NWP	Numerical Weather Prediction
NWS	National Weather Service
$\mathbb{S}_l$.	Linear Space Spanned by the Ensemble Perturbations
SPEEDY	Simplified Parameterizations privitivE-Equation Dynamics Model, a Low Resolution Weather Model
UTC	Coordinated Universal Time
V	A Measure of the Ensemble Spread, or Variance
TV	A Measure of the Difference in the Ensemble Mean and the True Atmosphere; the Skill of the Ensemble
TVS	The Projection of $\delta\mathbf{x}^{t}$ onto the Linear Space $\mathbb{S}_l$.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Weather is an important part of daily life; it affects every human activity from traveling to crop yields to electrical demands. The accurate prediction of weather can save both money and lives. We examine the ability of an ensemble weather forecast system, using linear diagnostics, in predicting the forecast uncertainty.

An overview of numerical weather prediction

Numerical weather prediction is the science of forecasting weather by integrating a mathematical model to simulate the future evolution of the atmosphere over a time period. The skillfulness of these predictions is affected by the accuracy of the observations of the past and present states of the atmosphere; the analysis, which is the model representation of the present atmospheric state based on the past and present observation; and the model representation of the true atmosphere.

The concept of weather prediction over an extended time period was first hypothesized by Cleveland Abbe (Abbe 1901). He realized that the tools needed to estimate the true state of the atmosphere did not yet exist, but knew that physical laws could be used to build a mathematical model to predict the evolution of the atmosphere. Abbe proposed a

---

This thesis follows the style of the *Journal of the Atmospheric Sciences*.

methodology, with many of the governing equations we use in today's models, on how to approach the general topic of weather modeling. Jule Charney, Ragnar Fjörtoft, and John von Neumann created the first numerical prediction of weather in 1950 (Charney, et al. 1950). The first operational model run in real-time took place in Sweden during September 1954 (Kalnay 2007). Nowadays, numerical models of the atmosphere are widely used for operational weather forecasting and research

A numerical weather prediction system consists of three basic components: data assimilation, where weather observations are assimilated through different methods to create an analysis of the atmosphere; the dynamical model composed of the governing equations of the atmosphere to simulate the evolution of the atmospheric state; and a post-processing system, which converts the raw model output into useful forms and variables.

Currently, the Earth has a large global observing network composed of a large variety of instruments. Weather observations inputted into NWP forecast systems are taken by satellites, upper-air balloon radiosondes, automatic sensors at the surface, and other observing platforms. Operational models running on a loop every 6 to 24 hours use present and past observations to predict the future atmospheric state.

NWP forecast systems predict the hydrodynamical and thermodynamical state of the atmosphere. A raw output of the model variables has limited usefulness for weather forecasters. During post-processing, these variables are converted into parameters that can assist the human forecasters to prepare a weather prediction, such as mean wind.

The atmosphere is a chaotic system of infinite dimension; the use of numerical models and noisy observations inevitably introduces error into the prediction, which grows exponentially with forecast time until predictability is completely lost. Numerical weather prediction models are overlaid onto the planet by introducing a gridding system. Depending on the model type and function, the resolution can range anywhere from a hundred of kilometers to hundredths of meters. Models that are used to simulate microphysical processes are required to be of very high resolution to accurately model on that scale. Levels of resolution for different types of models are based on the amount of computational power available and the time allowed to run the model; for example, by the time a global circulation model completed a run at a resolution equal to that of microphysical models, it would be completely useless.

The concept of ensemble modeling

Ensemble forecast systems generate probabilistic estimates of the present and future states of the atmosphere based on the model equations. The model dynamics of the ensemble members is the same as in the single deterministic model forecast systems. However, the difference is that the ensemble of forecasts is started from an ensemble of analyses. The initial conditions are different for each ensemble member and the distribution of the analysis uncertainties they represent is assumed to be normal. Each ensemble member is integrated with the model separately. The mean and the standard deviation of the ensemble members are usually calculated in the post-processing step.

Ensemble predictability

Satterfield and Szunyogh (2010 *a* and *b*), collectively hereafter SS10, demonstrated that the forecast performance of an ensemble is inherently flow dependent. The authors performed experiments to determine the predictability of ensemble uncertainty. The key finding was that the ensemble tested provided a more reliable prediction of the space of uncertainties, or the potential forecast error patterns, than the magnitude of the forecast errors and the relative importance of the different error patterns. Most importantly, they found that the ensemble grossly underestimated the total forecast uncertainty associated with the correctly predicted forecast error patterns.

The previous study used the National Center for Environmental Prediction (NCEP) Global Forecast System (GFS), currently the global operational model run in the United States by the National Weather Service (NWS). Being an operational mode, the NCEP GFS is computationally very expensive. The number of members in an ensemble configuration of this model is strongly limited by the amount of available computational power. SS10 used 40 ensemble members, which is about the same as that currently used in the operational prediction centers.

Forecast system

Our hypothesis is that some result of SS10 was strongly affected by the number of ensemble members used in the experiment. To test this hypothesis, we replace the state-of-the-art operational global circulation model of NCEP used in SS10 with a simplified global circulation model called SPEEDY (e.g. Molteni 2003 and Kucharski, et al. 2006). The computational cost of running the SPEEDY model is significantly lower than that of the NCEP GFS, but it is still sufficiently realistic to expect that our findings will be applicable to a state-of-the-art ensemble forecast system. SS10 laid the foundation for techniques in predicting and analyzing the ensemble's skill in the prediction of forecast uncertainties; we use these mathematical tools to evaluate the ensemble model runs in our experiment.

The SPEEDY model

The Simplified Parameterizations privitivE-Equation Dynamics (SPEEDY) model is developed and maintained by a group of researchers through the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy. It is a low resolution model with some simplifying assumptions, compared to a state-of-the-art operational NWP model. The model used in SS10 had a horizontal resolution of about 200 kilometers with 28 vertical levels, while SPEEDY model used for this experiment has a resolution of about 400 kilometers with 8 vertical levels.

Local ensemble transform Kalman filter

To obtain an initial condition on the model grid based on the randomly distributed observations, the data are assimilated into a short-term (six-hour in our case) model forecast called the background. In the assimilation process, observations are weighted based on their distance from the grid point and their assumed accuracy. Different data assimilation schemes are used for a variety of applications; one method is the Local Ensemble Transform Kalman Filter (LETKF), which was introduced by Hunt et al. 2007. The LETKF is a relatively new, computationally efficient data assimilation scheme.

Generally, a Kalman Filter is a mathematical technique to estimate the state of a time evolving system by minimizing a cost function that measures the distance of the state estimate from the observations and the background (short term forecast from the analysis at the previous time) (Kalman 1960). A Kalman Filter provides, in addition to the state estimates, an estimate of the uncertainty in the state estimate. To be precise, it provides an estimate of the covariance matrix of the normal distribution that describes the analysis uncertainty. This matrix is called the analysis error covariance matrix. An Ensemble-based Kalman Filter (EnKF) generates an analysis ensemble, which is consistent with the analysis error covariance matrix.

CHAPTER II

METHODS

Linear diagnostics

For this experiment, we test a set of local diagnostics used to assess the performance of the ensemble described by SS10. We follow the derivation of the ensemble variance. The ensemble mean is given as

$$\bar{\mathbf{x}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}^{(\mathrm{k})} \tag{1}$$

where K is the number of ensemble members $\{\ \mathbf{x}^{(\mathrm{k})} : k = 1 \ldots K\ \}$, and the ensemble perturbations are defined as the difference

$$\left\{\mathbf{x}'^{(\mathrm{k})} = \mathbf{x}^{(\mathrm{k})} - \bar{\mathbf{x}} \quad \middle| \quad k = 1 \ldots K\right\}. \tag{2}$$

The evolution of the "true" state of the atmosphere $\mathbf{x}^{\mathrm{t}}$ is simulated by an integration of the NCEP GFS model started from an operational analysis valid at 00 UTC 1 January 2004 to 12 UTC 29 February 2004; the model has a resolution of 144 x 73 x 28 grid points (notated T62L28) but is truncated to the same resolution of the SPEEDY model. Additionally, the difference between the "true" state and the ensemble mean is

$$\delta\mathbf{x}^{\mathrm{t}} = \mathbf{x}^{\mathrm{t}} - \bar{\mathbf{x}}. \tag{3}$$

This quantity is often referred to as the ensemble mean forecast error. We also define $\mathbb{S}_l$ as the linear space spanned by the ensemble perturbations.

At each grid point of the SPEEDY model configuration, a state vector is defined that describes the model representation of the state variables in a cube centered around the location extending to the top layer of the model. Variables used to define the state vector for this experiment are temperature (K), surface pressure (Pa), meridional wind velocity ($m \cdot s^{-1}$), and zonal wind velocity ($m \cdot s^{-1}$). A prediction in the ensemble uncertainty of a local state is an estimate of the local error covariance matrix $\widehat{\mathbf{P}}_l$

$$\widehat{\mathbf{P}}_l = \frac{1}{K-1}\sum_{k=1}^{K} \mathbf{x}'^{(k)}\left(\mathbf{x}'^{(k)}\right)^T, \tag{4}$$

where T is the matrix transpose.

Ensemble variance—the spread—can be calculated in two ways: either by

$$V_l = trace\left(\widehat{\mathbf{P}}_l\right), \tag{5}$$

or by an eigendecomposition of $\widehat{\mathbf{P}}_l$ leading to

$$V_l = \sum_{k=1}^{K-1} \lambda_k, \tag{6}$$

where λ_k is the k-th nonnegative eigenvalue in descending order. Additionally, the total forecast error variance, is defined as

$$TV_l = E[(\delta \boldsymbol{x}^t)^2]. \tag{7}$$

The forecast error variance, commonly referred to as the skill of the forecast, can be decomposed into two components: one that projects onto the space of ensemble perturbations $\mathbb{S}_l$ and another that projects onto the null space of the locally estimated covariance matrix $\widehat{\mathbf{P}}_l$. The component projected onto the null space of $\widehat{\mathbf{P}}_l$ is not fully discussed in this paper; simply, it is the linear space orthogonal to the space of ensemble perturbations. Eigendecomposition of the local estimated covariance matrix yields *K* eigenvalues of which the first *K* - 1 are normalized by

$$\boldsymbol{u}_k = \lambda_{\boldsymbol{k}} \left[\sum\nolimits_{i=1}^{K-1} \lambda_i^2\right]^{-1/2}. \tag{8}$$

The component of $\delta\boldsymbol{x}^t$ projected onto the linear space of perturbations is computed by

$$\delta\mathbf{x}^{t(\parallel)} = \sum_{k=1}^{K-1} [(\delta\mathbf{x}^t)^T \mathbf{u}_k]\, \mathbf{u}_k. \tag{9}$$

Diagnostically, this projection is defined as

$$TVS = E\left[\left(\delta\mathbf{x}^{t(\parallel)}\right)^2\right]. \tag{10}$$

TVS = *V* is optimal, in which the ensemble perturbations capture the magnitude of uncertainty explained by $\mathbb{S}_l$. Moreover, *TVS* = *TV* occurs when the forecast error space is predicted by the ensemble perturbation space $\mathbb{S}_l$, or when $\delta\mathbf{x}^t$ projects completely onto $\mathbb{S}_l$. An overestimation or underestimation in the magnitude of uncertainty explained by the space $\mathbb{S}_l$ occurs when *V* > TVS or *V* < *TVS,* respectively. Thus, the ultimately desired case of a fully predicted error in both pattern and magnitude is indicated by *V* = *TVS*.

State variables like wind velocities and surface pressure cannot simply be added in calculating total errors because of inconsistent units; as such, the diagnostics are calculated by transforming the state variables into a term with which the Euclidian norm squared has unit energy ($J \cdot kg^{-1}$). Transformation weights in the rescaling are those derived by Oczkowski, et al. 2005 as

$$\|\cdot\|_{E_l}^2 = \sum_l \left(U_l^T D_l U_l + V_l^T D_l V_l + \frac{C_P}{T_r} T_l^T D_l T_l \right) + R_d T_r \ln\,(P_s), \tag{11}$$

where U_l, V_l, and T_l are the perturbations from the truth or ensemble mean of the zonal and meridional wind velocities and temperatures at level l; P_l is the pressure at level l; P_s is the surface pressure; $D_l = \sqrt{(P_{l+1} - P_l)/P_s}$; C_p = 1004 $J \cdot kg^{-1} \cdot K^{-1}$; T_r the reference temperature, chosen as 273 K; and R_d = 287 $J \cdot mol^{-1} \cdot K^{-1}$.

The LETKF contains a multiplicative covariance inflation factor for analyses to handle some inherent errors, preventing an uncontrolled divergence of the Kalman filter. Covariance inflation is used to increase analysis uncertainty as a result of the model and sampling errors. Our implementation contains a consistent covariance inflation factor ρ for every point at each vertical level in the model. This is applied at all analysis times to each ensemble member by

$$\mathbf{x}'^{(k)} = \rho \cdot \mathbf{x}'^{(k)} \tag{12}$$

Methods of experimentation

We use the SPEEDY model with 96 x 48 x 7 grid point resolution to perform all the experiments, with an implementation of the LETKF for analysis and ensemble creation. The number of ensemble members used in the experiment is 40, which was found to be a sufficient number by Kuhl, et. al. (2007) and was also used in SS10. A

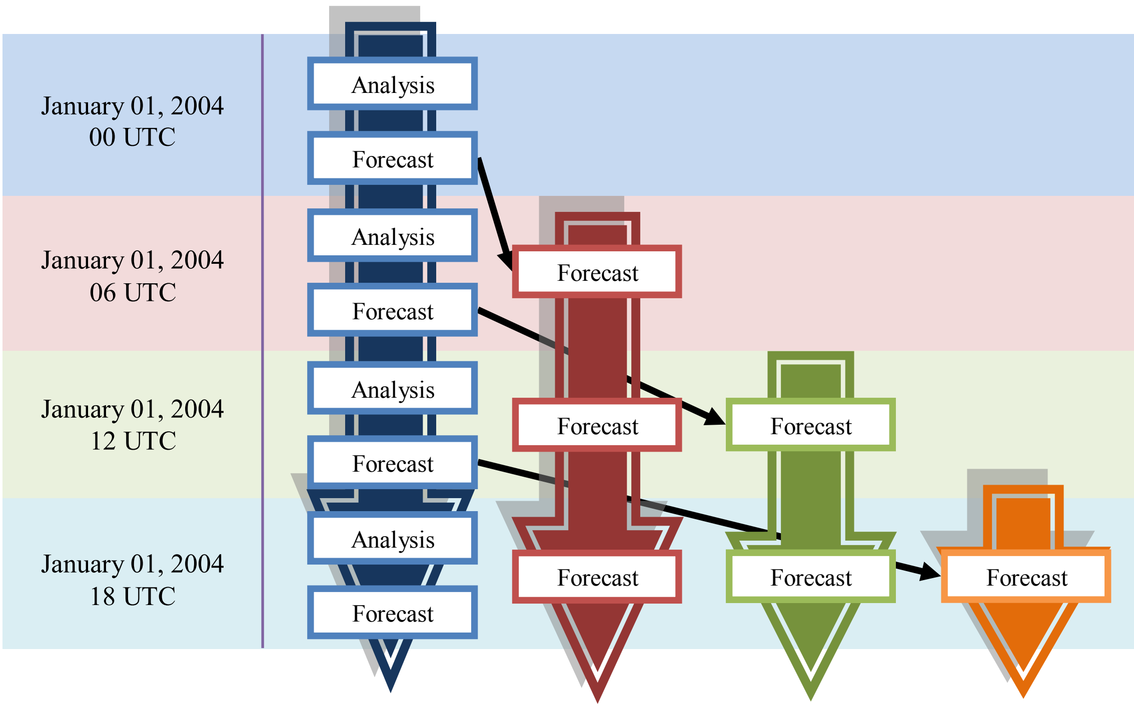

FIG 1. Depiction of the experiment forecast cycle. Five day forecast cycles are started from an analysis every six hours which serves as the initial background.

deterministic forecast is generated every 6 hours from 00 UTC 01 January 2004 to 12 UTC 29 February 2004. Ensemble analyses are created at every 6 hour interval with the deterministic forecast serving as the state estimate. Each analysis is used as the background for a 120-hour forward integration of the SPEEDY model. Figure 1 depicts our method of experimentation for the first four forecast cycles.

Observations of virtual temperature and both components of the wind are generated and assimilated at every grid point at the seven pressure levels (100, 200, 300, 500, 700, 850, and 925 hPa). Additionally, surface pressure observations are simulated at each surface grid point. These observations are perturbed from the generated values with zero mean and standard deviation 1 unit for each state variable to form a normal distribution of ensemble member perturbations. Experiments using an observation at every grid point are the equivalent of having a sounding of the vertical profile of the atmosphere for over 4600 points inputted to the model, which is much greater than the number used in operational weather models.

Diagnostics are computed for both the analyses and the deterministic forecasts for all forecasts started between 00 UTC 11 January 2004 and 12 UTC 15 February 2004. Local state vectors and diagnostics are computed in 5 x 5 grid point cubes extending upward through all seven levels.

CHAPTER III

RESULTS

We tune the covariance inflation factor to satisfy the condition $V \approx TV$ at analysis times. We find that this condition is satisfied when the inflation factor is 130%. In our calculations, V, TV, and TVS are averaged over the northern extratropics (30°N - 90ºN) for each time. Forecasts with common lead times are then temporally averaged.

Figure 2 shows the evolution of V, TV, and TVS. For forecasts with lead times of 6 to 24 hours, TVS accelerates towards TV and away from V, signifying that the space $\mathbb{S}_l$

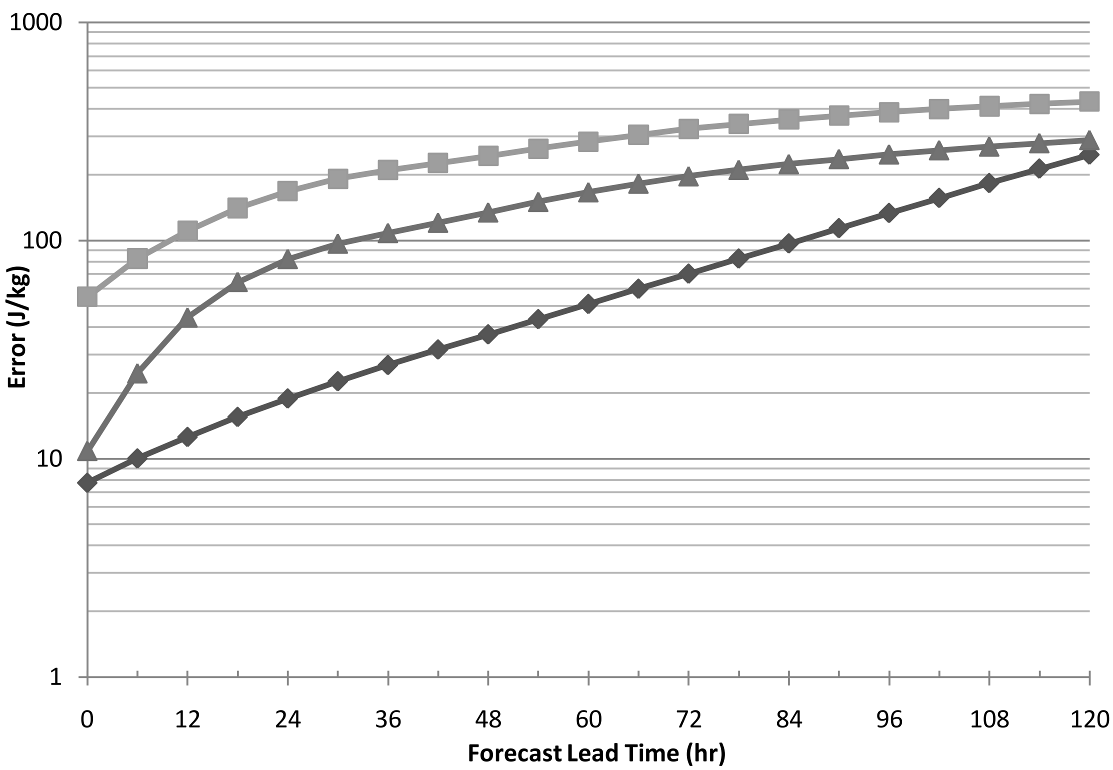

FIG 2. Graph showing the evolution of V (diamonds), TV (squares), and TVS (triangles). The diagnostics are calculated for all cycles initiated between 00 UTC 11 January 2004 and 12 UTC 15 February 2004.

provides a good representation of the error space, while the ensemble further underestimates the variance in $\mathbb{S}_l$. For forecasts past 24 hour lead times, *TVS* mimics behavior of *TV* and indicates that the ensemble represents the general pattern of forecast uncertainty. *TV* and *TVS* increase linearly through the whole forecast cycle, while *V* increases exponentially ($V = 9.07 \cdot e^{0.028 \cdot t}$ provides the best fit). The difference between *V* and *TVS* greatly decreases past the 84 hour lead time; this means that the ensemble better predicts the total variance in $\mathbb{S}_l$ and the magnitude of the forecast errors for longer lead times.

Figure 3 shows the differences between *TV* and *TVS* for the time period of calculated diagnostics and Figure 4 the differences between *TVS* and *V*. For earlier lead times, the differences between all the diagnostics are increasing; however, *TVS* slows its rate of growth after 84 hours, decreasing the difference between *TVS* and *V*. Largest rates of

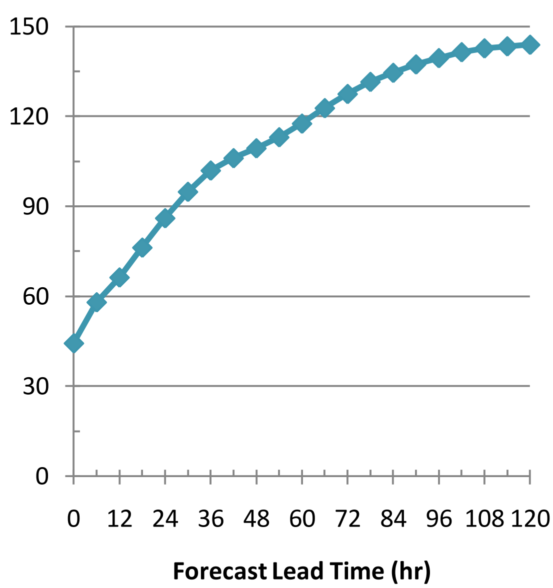

FIG 3. Differences between *TV* and *TVS*.

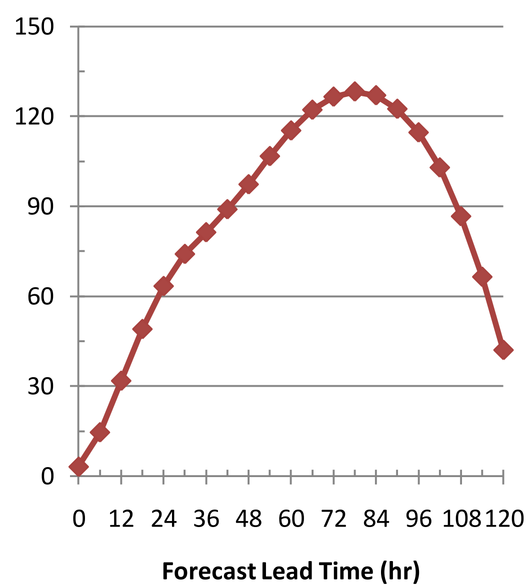

FIG 4. Differences between *TVS* and *V*.

change in Figure 4 occur between 6 to 30 and 102 to 120 hour lead times. It is of particular interest to note that the difference in *TV* and *TVS* begins to maintain a constant value past the 108 hour forecast time.

Configuring the energy rescaling to omit certain state variables at different times (e.g. only upper level winds, surface pressure and temperatures at all levels, low level winds and high level temperatures, etc.) showed trends similar to what was found when including all state variables at all levels, though with lower values because of a smaller total error. As a result, it is demonstrated that these diagnostics are not strongly dominated by one state variable but are rather an accurate portrayal of the general of all state variables.

A large obstacle encountered over the course of this project was the compatibility of the SPEEDY model (a simplified model used mostly in a testing environment) with diagnostics applied to state-of-the-art weather models. It took a considerable amount of time to code and debug the programs; inevitably, this is one of the consequences of computer-based programming. Additionally, some of the simplifications used in the SPEEDY model could be influencing a higher overall forecast error.

CHAPTER IV

CONCLUSIONS

We have shown that the linear diagnostics of SS10 are applicable to the simplified global circulation model used in this study. These diagnostics indicated a good representation of the pattern of uncertainties by the ensemble forecasts; however, for the lead time period used in this study (6 hours up to 120 hours), the magnitude of the uncertainties projected onto the linear space $\mathbb{S}_l$ was underestimated. One cause for this could be the large area over which the diagnostics are spatially averaged. A closer examination in the areas of error could lead to places with consistently underestimated forecast uncertainties, which often appear as a result of incorrect resolution of orography. In these cases, a bias correction factor would be implemented in these areas.

Analysis schemes have the ability to tune forecast uncertainties; for this study, a multiplicative covariance inflation factor was used to analyses so that the initial forecasts would provide an accurate estimate of the magnitude of forecast uncertainty projected onto the linear space $\mathbb{S}_l$. But, the estimated magnitude of this uncertainty is largely underestimated for longer lead times from 6 to 114 hours, yet it decreases with time. The trend present in our experiments suggests a possibility that the ensemble may estimate the magnitude of the forecast space better for forecasts further out than 120 hours.

We propose that further investigation of the linear diagnostics' behaviors at these longer forecast lead times is required to determine whether the diagnostics remain a valid predictor of ensemble performance at these times. If validity is confirmed, a post process could be implemented to enhance ensemble forecasts leading to an improved interpretation of the forecasts.

REFERENCES

Abbe, C., 1901: The physical basis of long-range weather forecasts. *Mon. Wea. Rev.*, **29**, 551-561.

Charney, J., R. Fjörtoft, and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237-254.

Hunt, B., E. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatio-temporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112-126.

Kalman, R. E, 1960: A new approach to linear filtering and prediction problems. *Transaction of the ASME--Journal of Basic Engineering*, **82,** 35-45.

Kalnay, E., 2007: *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge, UK: Cambridge University Press.

Kucharski, F., F. Molteni, and A. Bracco, 2006: Decadal interactions between the western tropical pacific and the north atlantic oscillation. *Cli. Dyn.*, **26**, 79-91.

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

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. *Cli. Dyn.*, **20**, 175-191.

Oczkowski, M., I. Szunyogh, and D. Patil, 2005: Mechanisms for the development of locally low-dimensional atmospheric dynamics. *Jour. Atmo. Sci.*, **62**, 1135-1156.

Satterfield, E., and I. Szunyogh, 2010b: Assessing the performance of an ensemble forecast system in predicting the magnitude and the spectrum of analysis and forecast uncertainties. *Mon. Wea. Rev.*, in press.

——, and ——, 2010a: Predictability of the performance of an ensemble forecast system: predictability of the space of uncertainties. *Mon. Wea. Rev.*, **138**, 962-981.

CONTACT INFORMATION

Name: Ethan Lane Nelson

Professional Address: c/o Dr. Istvan Szunyogh
Department of Atmospheric Science
MS 3150
Texas A&M University
College Station, TX 77843

E-mail Address: ethan.nelson@tamu.edu

Education: B.S., Meteorology, Texas A&M University, May 2012
Minor, Mathematics
Undergraduate Research Scholar