A Method to Improve Prediction of Atmospheric Flow Transitions

P. J. ROEBBER AND A. A. TSONIS

Atmospheric Science Group, Department of Mathematical Sciences, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin

(Manuscript received 3 August 2004, in final form 5 April 2005)

ABSTRACT

Ensemble prediction has become an indispensable tool in weather forecasting. One of the issues in ensemble prediction is that, regardless of the method, the prediction error does not map well to the underlying physics (i.e., error estimates do not project strongly onto physical structures). This paper is driven by the hypothesis that prediction error includes a deterministic component, which can be isolated and then removed, and that removing the error would enable researchers and forecasters to better map the error to the physics and improve prediction of atmospheric transitions. Here, preliminary experimental evidence is provided that supports this hypothesis. This evidence is provided from results obtained from two low-order but highly chaotic systems, one of which incorporates atmospheric flow transitions. Using neural networks to probe the deterministic component of forecast error, it is shown that the error recovery relates to the underlying type of flow and that it can be used to better forecast transitions in the atmospheric flow using ensemble data. A discussion of methods to extend these ideas to more realistic forecast settings is provided.

1. Introduction

Since the midlatitude atmosphere is characterized by variability, persistent flow patterns often lead to anomalous sensible weather (e.g., temperature, precipitation, etc.). Yet, one challenge in short- to medium-range weather prediction is anticipating transitions in the large-scale atmospheric flow. Hence, an improved understanding of the physics and predictability of such flow transitions is desirable.

The fundamental problem of weather forecasting is to identify the range of possible meteorological scenarios that might evolve from a given initial state, and determine whether multiple solutions have high probability (low confidence in an individual solution) or if a single evolution is the most likely (high confidence). This probabilistic view is necessitated by the complexity of atmospheric dynamics (e.g., Lorenz 1963; see a general review in Kalnay 2003). Specifically, limits to deterministic predictability originate from two general sources: model error and initial condition error.

Model error arises because of imperfections in our characterizations of the laws of nature, arising either through parameterization of complex and/or poorly understood physics (such as boundary layer and cloud microphysical processes) or an inability to resolve atmospheric processes smaller than a certain threshold (e.g., atmospheric convection with a 10-km gridpoint model), with resultant upscale error growth.

Initial condition error arises because of the finite spatial availability of observed data (including some variables that are not observed at all), missing data, inaccuracies in the data, and imperfect analysis techniques. All of these errors, even with a perfect model, will grow nonlinearly over time, eventually swamping the forecast signal (Lorenz 1963, 1965, 1969).

The rate of this error growth and hence the lead time at which predictability is lost depends on the stability of the evolving flow (Lorenz 1965), which in addition is affected by factors such as the large-scale flow pattern, season, and geographical domain (Lorenz 1984, 1990). Ensemble forecast systems have been developed as a means to quantify forecast uncertainty, using a variety of methods to simulate analysis and model uncertainties. In the midlatitudes, numerical weather prediction models are often considered to be good enough that to a first approximation, forecast error growth at medium range can be attributed primarily to the process of growing instabilities in a chaotic system that result from

Corresponding author address: Paul J. Roebber, Department of Mathematical Sciences, University of Wisconsin—Milwaukee, 3200 N. Cramer Ave., Milwaukee, WI 53211.
E-mail: roebber@uwm.edu

initial condition uncertainties. This "perfect model" assumption is more prone to failure, however, for short-range forecasts where small-scale features may depend crucially on the physics (e.g., convection). Since the goal of this work is to explore flow transitions, with an inherent time scale beyond that of short-range forecasts, we will not consider model error further in this preliminary work.

A deeper understanding of the physical evolution of error structures can potentially reveal much concerning both the natural dynamics of the atmosphere and model deficiencies (and hence, lead to a better understanding of nature). Currently, there are four primary methods for assessing analysis uncertainty: breeding (used at the U.S. National Centers for Environmental Prediction), singular vectors (used at the European Centre for Medium-Range Weather Forecasts), Kalman filters (currently a topic of active research), and Monte Carlo methods (used at the Canadian Meteorological Centre).

Although the technical details of each of these methods are distinct (see Kalnay 2003 for an overview), they each rely on a single concept: initial analysis uncertainties are simulated by adding small perturbations (within the limits of the uncertainty of the analysis) to the unperturbed (control) analysis. It is important to recognize that these techniques are intended to simulate analysis uncertainty, not necessarily to reveal intrinsic dynamics. Tribbia and Baumhefner (2004) argue that these uncertainties seed the baroclinically active region of the spectrum and organize over time on the synoptic scales, extracting energy from the larger-scale flow rather than the smaller scales.

In the breeding method, analysis errors are considered to be filtered forecast errors. As such, much of the analysis error may represent phase error, a typical problem in data-sparse regions. The singular vector method attempts to directly calculate forecast error, under the assumption that the components of the error that grow most rapidly will dominate. Since singular vectors are mathematical constructs rather than physical structures, the actual error may project more or less strongly onto these growing patterns. Not surprisingly, experience with this method suggests that these patterns do not "map" directly onto locations that are consistent with synoptic experience as to the origins of forecast uncertainties (e.g., the position of an upper-level jet streak). The Kalman filter method uses estimation theory to more accurately express the breeding method's fundamental hypothesis that analysis errors are filtered forecast errors. As such, this method likely preferentially identifies phase errors, which are consistent with the underlying attractor. The Monte Carlo method perturbs all available observations simultaneously with random numbers of realistic amplitude (with respect to error statistics). Hence, no direct dynamics can be inferred from this method.

It is well known that, given a nonlinear system, the evolution of small fluctuations is given by the linearized equations

$$\dot{\mathbf{x}}' = J\mathbf{x}', \tag{1}$$

where J is the Jacobian. These fluctuations can be thought of as an unbiased measurement error, e_{nb}, and/or a biased error arising through the data assimilation process, e_{b}, which contribute to the forecast error. Since the Jacobian implies deterministic operations, some recoverable dynamics are included in the error. This conclusion also holds for the more general, nonlinear problem.

In this paper, we propose a method that capitalizes on *recovering* part of the forecast error provided by any of the above ensemble approaches and *relating that error recovery to the underlying physics of the flow*. We demonstrate successful application of this method to two well-known, low-dimensional chaotic systems (including a simple climate model with identifiable flow patterns). The use of simple models to gain insights into complex problems has a rich history in meteorological research [recent examples are the investigations of the dynamics of model error by Nicolis (2003, 2004)]. The low-order modeling approach is undertaken with the expectation that it may reveal general principles, the details of which will need to be determined by resorting to studies using more complex models. This topic is addressed in the concluding discussion.

2. Examples

a. *The logistic equation*

First, we consider a very simple mathematical system, which however can be extremely sensitive to initial conditions. It is the logistic equation

$$X_{n+1} = \mu X_n(1 - X_n), \tag{2}$$

where μ is a parameter. In this very simple discrete system, future values are obtained by successive iterations. The system exhibits a variety of dynamics for different values of μ. For $\mu = 4$ the system is strongly chaotic with a corresponding Lyapunov exponent equal to ln (2). This means that if the initial condition is specified with an error of less than 1%, the forecast is no better than "climate" after approximately six iterations.

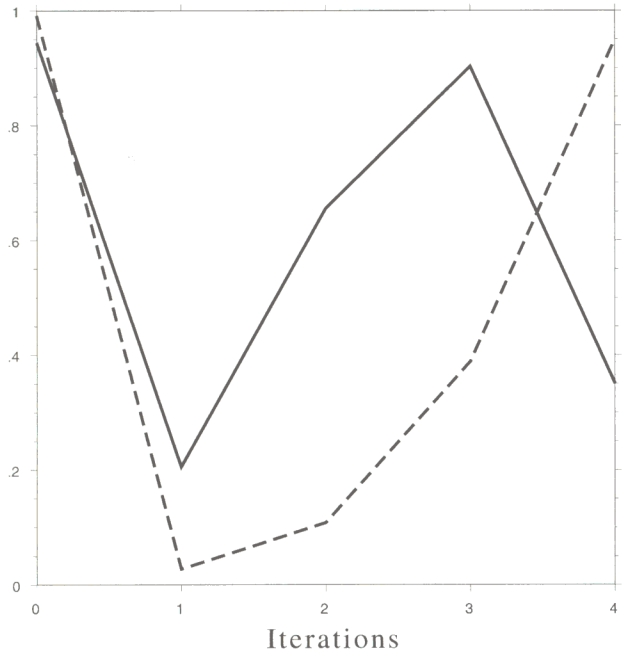

FIG. 1. Logistic equation (solid line) and associated "perturbed" initial state (dashed), through four iterations.

TABLE 1. Correlation coefficient between forecasts and "truth" for the logistic and Lorenz (1984) models. The Lorenz (1984) model verifications are stratified according to the low- and high-amplitude flow patterns and pattern transitions (low to high and high to low amplitude).

Model	Deterministic forecast	Neural network
Logistic equation—Two iteration simulations	0.80	0.90
Lorenz (1984)—Two time units (low amplitude)	0.77	0.87
Lorenz (1984)—four time units (low amplitude)	0.35	0.67
Lorenz (1984)—two time units (high amplitude)	0.95	0.97
Lorenz (1984)—four time units (high amplitude)	0.78	0.84
Lorenz (1984)—two time units (transitions)	0.83	0.94
Lorenz (1984)—four time units (transitions)	0.61	0.83

For the experiments with the logistic model, the operational data assimilation cycle was mimicked: an analysis was constructed, consisting of a weighted average of observations (incorporating an unbiased error $e_{nb,o}$) and a first guess (obtained from a previous short-term forecast from the model). This analysis was used as the basis for the next forecast, and the short-term forecast from that run was then cycled back into the succeeding analysis. For all our tests, the set of forecasts were extended from two to four iterations into the future, a time scale for which substantial error growth occurs in the logistic equation with these error characteristics (Fig. 1).

Error recovery was investigated using neural networks. Neural networks are nonlinear tools that are often used to model nonlinear deterministic dynamics when the actual equations describing the dynamics are not available. In its simplest form the architecture of a neural network involves a set of inputs $p(i)$ and an output Q. Using a training sample the inputs are assigned optimum weights, $w(i)$, and then the inner product, $P = \Sigma p(i)w(i)$, is estimated. The final output Q is then obtained by passing P through a nonlinear function $f(x)$. Such a network is called a two-layer network since it involves a layer for the inputs and a layer for the output. Neural networks can become more complex by adding hidden layers or by including more than one output. In general, with the help of known outputs (from the training sample) the network, initially set to a random state, modifies its structure (changes the weights) in such a way as to improve its performance. If the network architecture is rich enough, this procedure leads the network to a state where inputs are successfully mapped into outputs for all chosen training (input–output) pairs. Essentially, the network describes a nonlinear model fit of the error. Further information on neural networks can be found in Tsonis (1992), Roebber et al. (2003), Marzban (2003), and references therein.

A generalized feed-forward network with one hidden layer (and 70 nodes) was used for the error recovery experiments. Nine inputs were used: the initial analysis, the observations (including error) at the initial time, and the lagged forecasts at two, three, and four time steps validating at the verification time. The cost function used in the training was the mean square error.

Correlation coefficients between "forecasts" from the logistic equation and the verification (using the unperturbed initial conditions) show that, at two time steps, substantial predictability exists ($r = 0.80$; Table 1). Despite this high predictability, the forecasts can be improved still further (~17% relative to the two iteration forecast, based on the increase in accounted variance) through the use of the trained neural network. This exceeds by 8% the predictability obtained with a linear regression whose inputs are the set of lagged forecasts with regression coefficients optimized using the training dataset (not shown). Hence, in this highly chaotic system, it is still possible to recover a significant fraction (~48%, based on the amount of accounted variance relative to the maximum possible improve-

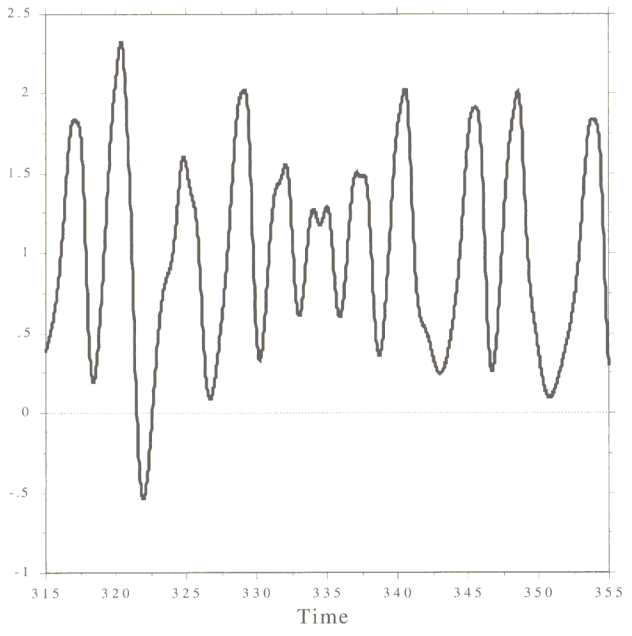

FIG. 2. Lorenz climate model integration for a 40 time unit period. The value of X is shown. High-amplitude oscillations are apparent from time = 315–330, followed by low-amplitude oscillations through time = 340, and back again to high.

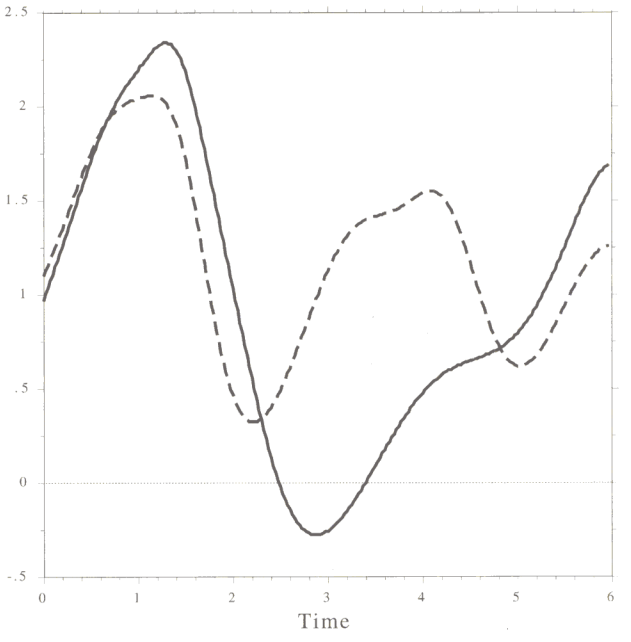

FIG. 3. Lorenz climate model integrations for a six time unit period, starting from an initial state and a second integration from a slightly perturbed initial state. The value of X, the meridional temperature gradient from the model, is shown.

ment over the two iteration deterministic forecasts) of the total forecast error using the nonlinear technique.

b. *Lorenz (1984) climate model*

While the above is an example of a general chaotic system on which one could test our hypothesis, it has little relationship to atmospheric phenomena. As such, it leaves unanswered the question as to whether this approach can be applied to large-scale atmospheric flows. For this reason we consider another example, which represents a low-order general circulation model (Lorenz 1984, 1990). It is also highly chaotic and is described by the following equations:

$$dX/dt = -Y^2 - Z^2 - aX + aF,$$

$$dY/dt = XY - bXZ - Y + G,$$

$$dZ/dt = bXY + XZ - Z. \qquad (3)$$

These equations can be derived from the mean and perturbation quasigeostrophic potential vorticity equations, using a truncated Fourier series expansion of the geostrophic streamfunction (Holton 1992). The independent variable t represents time, while X, Y, and Z represent the meridional temperature gradient (or equivalently from thermal wind balance, the strength of the zonal flow), and the amplitudes of the cosine and sine phases of a chain of superposed large-scale eddies, respectively. The term F represents the meridional gradient of diabatic heating, and the value to which X would be driven in the absence of coupling between the westerly current and the eddies. The term G is the asymmetric thermal forcing, representing the longitudinal heating contrast between land and sea, and it is the value to which Y would be driven in the absence of coupling. The coupling is represented by the terms XY, XZ, and $-Y^2 - Z^2$, and results in amplification of the eddies at the expense of the westerly current.

With $F = 8$ and $G = 1$ (representing perpetual winter conditions; see Lorenz 1984, 1990; Roebber 1995), the above system exhibits two very distinct flow patterns (Fig. 2). The first, which consists of low-amplitude fluctuations represents a steady, zonal jet. The second, consisting of high-amplitude fluctuations, corresponds to alternating strong and weak zonal jets. Another regime is that of a blocking event, but its frequency of occurrence for these parameters is very small and highly transient. For our analysis, this pattern will not be considered.

Predictability in this system is strongly dependent on initial conditions (Fig. 3). In this example, a forecast starting with a small error in the initial state results in an error that is quite large by three time units (representing 120 time steps), the result of a forecast transition from a high-amplitude to a low-amplitude flow that

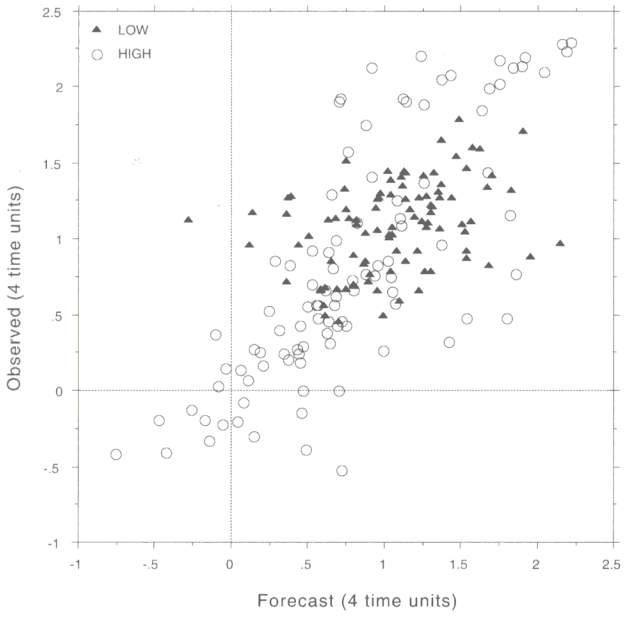

FIG. 4. Scatterplot of Lorenz model forecasts (four time units from initial state) and "observed" values of X (meridional temperature gradient) for the two flow patterns.

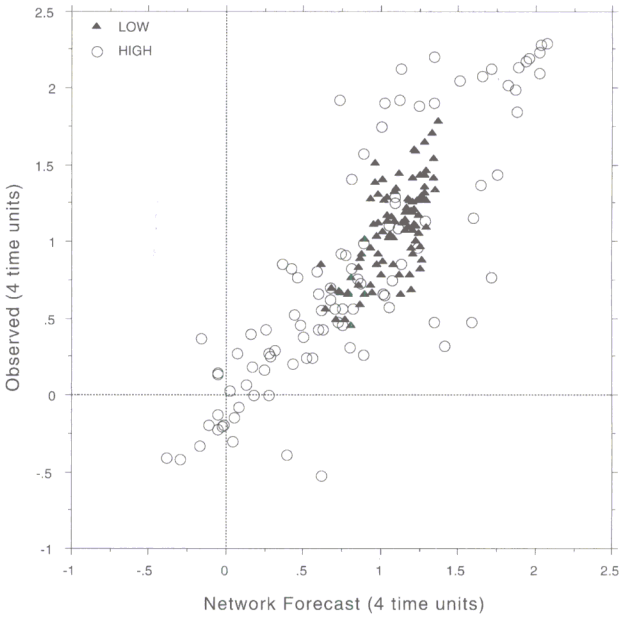

FIG. 5. Scatterplot of neural network corrected forecasts (four time units from initial state) and observed values of X (meridional temperature gradient) for the two flow patterns. Note the substantial reduction of variance in low-amplitude cases relative to Fig. 4.

did not verify. In other instances, of course, pattern transitions will be correctly forecast and the errors will be less.

Experiments with the Lorenz model were designed as follows. The model was run without interruption for an extended period (1164 time units) and the data recorded (base run). The model was then rerun 194 times out to six time units ($6 \times 194 = 1194$), with the initial conditions obtained from the corresponding values of the base run with a small error superimposed (mean absolute error for the initial conditions of the resulting 194 runs is 0.08 in X, Y, and Z). Of these 194 runs, 98 (96) were high-amplitude (low amplitude) flows.

Table 1 shows that for the low-amplitude pattern at two time units, predictability is relatively high ($r = 0.77$), but that it is still possible to recover a substantial portion of the error using neural networks sorted according to the flow (note that the network design is the same as that used for the logistic equation, except with seven inputs representing the initial analysis and observed values for X, Y, and Z at lead times of either two or four units, plus the forecast value for X at the verification time). In contrast, predictability is much higher for the high-amplitude pattern and little additional predictability can be achieved. By four time units, these differences are accentuated (Fig. 4), with limited correspondence between the forecast and observed states for the low-amplitude pattern. There is, however, substantial error recovery using a neural network with the initial analysis, observations (including error) at the initial time, and the forecast as inputs (Table 1). Figure 5 is similar to Fig. 4, but shows the corrected forecasts using the neural network. Comparison between Figs. 4 and 5 clearly shows that the error variance of the high-amplitude pattern remains approximately the same following nonlinear correction, but that the error variance of the low-amplitude cases decreases significantly.

Although this is a simple model, similar success in correcting forecasts with neural networks have been obtained in realistic forecast settings (e.g., Hall et al. 1999; Roebber et al. 2003; Marzban 2003). This finding indicates, as suggested in the introduction, that the error recovery is flow dependent, a result that is consistent with previous research that documents flow-dependent predictability in the observed atmosphere (e.g., Nutter et al. 1998). This suggests that it should be possible to use this *differential error recovery* to map the error to the underlying physical characteristics of the zonal jets. Since the varying flow is dictated by the physics of the system, this would indicate that error recovery can be used to understand the physics. Here, the low-amplitude state in the model occurs for relatively large available potential energy and small kinetic energy, such that a flow transition also reflects a substantial transition in the energy budget. Importantly, this result also holds for pattern transitions [in which

the initial flow is a low (high)-amplitude flow that transitions to a high (low)-amplitude flow by forecast verification time], such that an additional 9%–12% of the variance can be accounted for at two–four time units in these cases compared to those in which a particular pattern is maintained throughout the forecast interval (Table 1).

These results suggest that differential error recovery can be used within an ensemble forecast context to forecast flow transitions. To test this idea, the Lorenz model was used to generate ensemble forecasts for a set of events in which both patterns are verified. The reduction of the variance in the model ensemble forecasts (as opposed to the error, which is unknown in the forecast mode) using the previously trained neural network (with a median ensemble size of eight members) is used to classify the forecast as representing either the low- or high-amplitude pattern (Table 2). The classification is performed using a simple probabilistic model assuming that the error recovery is normally distributed (a hypothesis that cannot be rejected at the 1% level of significance, based upon the Kolomogorov–Smirnov test of normality). Standard measurements of diagnostic skill (probability of detection, false alarm rate, bias, and true skill statistic) show that this technique is highly successful in the context of this two-pattern low-order model. (Note that the verification statistics for the ensemble mean forecast are considerably lower for this sample, with a true skill statistic of only 0.32.)

These experiments indicate that the theoretical possibility of partial error recovery can be realized even in highly nonlinear systems, and that it may be possible to apply techniques of this kind to the output of ensemble systems to better predict and understand atmospheric flow transitions.

TABLE 2. Error recovery and flow pattern diagnosis in the Lorenz model. The reduction in the standard deviation of the ensemble forecasts using the neural network for the two patterns at two time ranges and the contingency table of pattern diagnoses and associated verification statistics are shown. The probability of detection (low) = 0.77 and (high) = 0.90. The false alarm rate (low) = 0.11 and (high) = 0.21. The bias (low) = 0.86 and (high) = 1.14. The true skill statistic = 0.67.

Reduction in std dev of X (meridional thermal gradient).

	Final = high amplitude	Final = low amplitude
Two time unit ensembles	0.0607	0.1089
Four time unit ensembles	0.0675	0.2608

Ensemble pattern diagnosis—Contingency table (No. of cases)

		Diagnosed final pattern	
Pattern at final time		Low amplitude	High amplitude
	Low amplitude	73	22
	High amplitude	9	83

3. Discussion

We have investigated an approach in which a nonlinear method is used to recover the deterministic component of the forecast error. Using two low-order nonlinear models we show that not only this is possible, but that this error recovery reduces the forecast error enough to allow us to map the remaining error to the underlying physics and to improve the prediction of transitions.

Although the method is straightforward, some additional aspects must be considered before it can be applied to realistic ensemble forecast systems. First, the simple ensemble system studied in this preliminary analysis effectively represents a single point; hence, this local error correction method must be generalized to account for the effects of error propagation in real systems. Work that might establish the suitability of a more general method is under way, using another simple (but nonlocal) chaotic model (Lorenz 1995).

Second, this study has been conducted under the constraints of the perfect model assumption. As noted in the introduction, model error imposes an important limit on forecast accuracy. When this work is extended to real ensemble systems, the effects of model error on predictability will also need to be addressed.

Last, in order to represent every part of the multidimensional input space and to protect against memorization, neural networks require large data samples for training and cross validation. In the context of a hemispheric or global model, where the dimensionality is necessarily large, it will likely be necessary to take advantage of atmospheric teleconnections to constrain the problem. The results of these refinements will be reported in a future paper.

REFERENCES

Hall, T., H. E. Brooks, and C. A. Doswell III, 1999: Precipitation forecasting using a neural network. *Wea. Forecasting,* **14,** 338–345.

Holton, J. R., 1992: *An Introduction to Dynamic Meteorology*. 3d ed. Academic Press, 511 pp.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Lorenz, E. N., 1963: Deterministic non-periodic flow. *J. Atmos. Sci.,* **20,** 130–141.

——, 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus,* **17,** 321–333.

——, 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289–307.

——, 1984: Irregularity: A fundamental property of the atmosphere. *Tellus,* **36A,** 98–110.

——, 1990: Can chaos and intransitivity lead to interannual variability? *Tellus,* **42A,** 378–389.

——, 1995: Predictability—A problem partly solved. *Proc. ECMWF Seminar on Predictability,* Vol. I, Reading, United Kingdom, ECMWF, 1–18.

Marzban, C., 2003: Neural networks for postprocessing model output: ARPS. *Mon. Wea. Rev.,* **131,** 1103–1111.

Nicolis, C., 2003: Dynamics of model error: Some generic features. *J. Atmos. Sci.,* **60,** 2208–2218.

——, 2004: Dynamics of model error: The role of unresolved scales revisited. *J. Atmos. Sci.,* **61,** 1740–1753.

Nutter, P. A., S. L. Mullen, and D. P. Baumhefner, 1998: The impact of initial condition uncertainty on numerical simulations of blocking. *Mon. Wea. Rev.,* **126,** 2482–2502.

Roebber, P. J., 1995: Climate variability in a low-order coupled atmosphere-ocean model. *Tellus,* **47A,** 473–494.

——, S. L. Bruening, D. M. Schultz, and J. V. Cortinas Jr., 2003: Improving snowfall forecasting by diagnosing snow density. *Wea. Forecasting,* **18,** 264–287.

Tribbia, J. J., and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Mon. Wea. Rev.,* **132,** 703–713.

Tsonis, A. A., 1992: *Chaos: From Theory to Applications*. Plenum, 274 pp.