The Application of Nonlinear Local Lyapunov Vectors to Ensemble Predictions in Lorenz Systems

JIE FENG, RUIQIANG DING, AND DEQIANG LIU

State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

JIANPING LI

State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, and College of Global Change and Earth System Sciences, Beijing Normal University, Beijing, China

(Manuscript received 30 August 2013, in final form 16 April 2014)

ABSTRACT

Nonlinear local Lyapunov vectors (NLLVs) are developed to indicate orthogonal directions in phase space with different perturbation growth rates. In particular, the first few NLLVs are considered to be an appropriate orthogonal basis for the fast-growing subspace. In this paper, the NLLV method is used to generate initial perturbations and implement ensemble forecasts in simple nonlinear models (the Lorenz63 and Lorenz96 models) to explore the validity of the NLLV method.

The performance of the NLLV method is compared comprehensively and systematically with other methods such as the bred vector (BV) and the random perturbation (Monte Carlo) methods. In experiments using the Lorenz63 model, the leading NLLV (LNLLV) captured a more precise direction, and with a faster growth rate, than any individual bred vector. It may be the larger projection on fastest-growing analysis errors that causes the improved performance of the new method. Regarding the Lorenz96 model, two practical measures—namely the spread–skill relationship and the Brier score—were used to assess the reliability and resolution of these ensemble schemes. Overall, the ensemble spread of NLLVs is more consistent with the errors of the ensemble mean, which indicates the better performance of NLLVs in simulating the evolution of analysis errors. In addition, the NLLVs perform significantly better than the BVs in terms of reliability and the random perturbations in resolution.

1. Introduction

The atmosphere is a chaotic system, and even negligible initial errors will give rise to gradual deviation of the forecast state from the true path, eventually resulting in chaos (Lorenz 1963a,b, 1965; Li and Chou 1997). This means that the weather has a predictability limit beyond which forecasts will lose all skill. Based on this, any single forecast is simply an estimate of the future state of the atmosphere within a stochastic framework but provides no information regarding its reliability. Ensemble prediction offers one approach to generate probabilistic forecasts of the future state of the system based on a statistical sampling approach (Leith 1974).

The ensemble method provides an approximate description of the probability density function (PDF) of the forecast state based on a finite number of samples. In a generally perfect model, if a group of initial states sample the uncertainty related to the initial value reasonably well, the states of their individual integration will have the ability to represent the uncertainty of the forecast states. In contrast with a single forecast, ensemble averaging provides a nonlinear filter that reduces error (Toth and Kalnay 1997). However, the ensemble mean simply treats the ensemble forecast as a single forecast, which does not indicate how reliable the result is. The biggest advantage of ensemble prediction is the product of the second moment of the ensemble, or the ensemble spread (Whitaker and Loughe 1998; Zhu et al.

Corresponding author address: Dr. Jianping Li, College of Global Change and Earth System Sciences, Beijing Normal University, 100875 Beijing, China.
E-mail: ljp@bnu.edu.cn

DOI: 10.1175/JAS-D-13-0270.1

© 2014 American Meteorological Society

2002). With this additional information, the quality of forecasts can be significantly enhanced.

Initially, the use of ensemble techniques focused on random samples (Monte Carlo forecasting) as a description of the probability distribution of initial states (Epstein 1969; Leith 1974). However, the atmosphere is an extremely complex system, which has a very high phase-space dimension; that is, the number of random samples must be sufficiently large. Consequently, the cost of performing the requisite number of predictions presents a significant challenge. However, as more and more intensive studies of data assimilation procedures have been completed, analysis errors are no longer considered to be totally unknown. Pure random perturbations may introduce irrelevant directions to a chosen initial state (Houtekamer and Derome 1994). In the early 1990s, ensemble prediction systems were developed and used operationally by weather prediction centers. Different schemes are used to generate the initial perturbations in the control analysis: for example, the bred vector (BV) method (Toth and Kalnay 1993, 1997) at the National Centers for Environmental Prediction (NCEP), the singular vector (SV) method (Molteni and Palmer 1993; Molteni et al. 1996; Buizza 1997) at the European Centre for Medium-Range Weather Forecasts (ECMWF), and the perturbed-observation approach (Houtekamer et al. 1996) at the Canadian Meteorological Centre (CMC). The BV and SV methods have been widely used in predictability and probabilistic prediction of weather and climate (Cai et al. 2003; Yang et al. 2009; Cheng et al. 2010; Buizza 2010). Recently, some methods based on the ensemble transform concept to generate the initial perturbations, such as the ensemble transform Kalman filter (ETKF), ensemble transform (ET), and ensemble transform with rescaling (ETR) methods, have been developed and attempted to improve operational forecasts (Bishop et al. 2001; Wang and Bishop 2003; Wei et al. 2006, 2008; Descamps and Talagrand 2007; Pu and Hacker 2009).

Of several operational ensemble generation schemes, the breeding method is the most time efficient (Wang and Bishop 2003). Toth and Kalnay (1993, 1997) introduced the breeding cycle as a procedure to simulate the behavior of growing modes in the analysis cycle (see Fig. 1). Their rationale was that in a data assimilation cycle, random errors introduced at each analysis stage would evolve into fast-growing directions in the atmospheric flow; that is, the ratio of fast-growing errors to the total error would significantly increase in the control analysis (Toth and Kalnay 1997). Based on these properties of analysis errors, the breeding method was proposed to simulate the analysis cycle and generate fast-growing error modes. In this process, random perturbations are

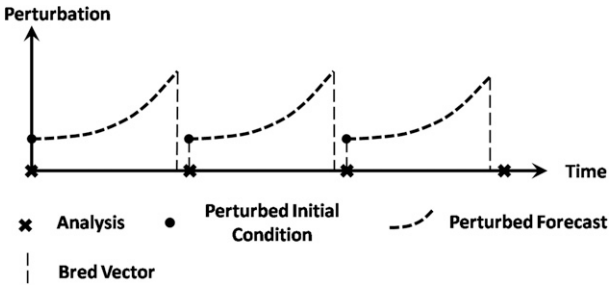

FIG. 1. Schematic representation of the creation of a bred vector for an ensemble forecast (after Toth and Kalnay 1997).

superposed on the initial condition, evolve directly in a nonlinear model of dynamical flow, and are then scaled down to the same size as the initial perturbations at regular intervals. The final perturbations corresponding to each initial state are called BVs, which are added and subtracted from the control analysis to generate the ensemble samples. A more complete description is given in appendix A.

The BVs are a nonlinear extension of the local Lyapunov vectors (LLVs) (Toth and Kalnay 1997; Kalnay 2003). Through the breeding cycles, the BVs all generally turn toward the fastest-growing modes. Toth and Kalnay (1997) state that the BVs have strong similarities in some local regions but are quasi-orthogonal globally. However, the high correlation between the local structures of the BVs may degrade the global orthogonality to some extent. Wang and Bishop (2003) analyzed the orthogonal properties of BVs, concluding that "nearly all of the ensemble forecast variance is contained in a single direction for both the simple breeding and the masked breeding." This result means that BVs may not span the fast-growing subspace efficiently. However, short-term error growth is a local phenomenon. The high correlation of BVs over local regions may limit forecast skill in those regions (Annan 2004, hereafter A04). Several methods have been recently proposed to prevent BVs from collapsing into one dominant vector. For example, A04 attempted to obtain the fastest-growing modes using orthogonal BVs. Primo et al. (2008) introduced a new logarithmic bred vector (LBV) to increase the diversity of the ensemble. Corazza et al. (2003) and Greybush et al. (2013) attempted to keep BVs less correlated by adding a small random perturbation to the perturbed run. These methods improve to some extent the performance of the diversity of the ensemble perturbations. However, they did not effectively eliminate the dependence among the initial perturbations.

Early investigators developed the theory of Lyapunov exponents [or the Lyapunov exponent spectrum (LEs)] to quantitatively estimate the expanding or contracting nature of different directions in the phase space of

nonlinear dynamical systems (Benettin 1980; Wolf et al. 1985; Fraedrich 1987; Trevisan and Legnani 1995). The LEs are global properties of the attractor of the dynamical system. Furthermore, local or finite-time Lyapunov exponents have been proposed (Yoden and Nomura 1993; Ziehmann et al. 2000) to measure the short-term growth rate of small initial perturbations. However, these local or finite-time Lyapunov exponents, which are similar to the global Lyapunov exponents, are established based on the assumption that initial perturbations are infinitesimal and therefore can be controlled by linear error growth dynamics. To understand the nonlinear evolution of initial errors, the nonlinear local Lyapunov exponent (NLLE) was proposed to study the predictability of an n-dimensional chaotic system or a single variable within the system (Li et al. 2006; Ding and Li 2007; Rui-Qiang et al. 2008; Ding et al. 2008, 2010; Li and Wang 2008; Li and Ding 2009, 2011, 2013). The NLLE depends on the initial state, initial error, and evolution time. For low-order chaotic systems, the error growth rate mainly depends on the fastest-growing direction. While for high-dimensional chaotic systems, such as the atmosphere, other fast-growing directions also make large contributions to the evolution of initial errors. Therefore, to further reveal the dynamical properties of the error evolution of complex nonlinear systems, the maximal NLLE was expanded in a high-dimensional situation and the NLLE spectrum (NLLEs) was proposed to investigate the nonlinear evolution behaviors of initial perturbations along different directions in phase space (Ding 2007; Li and Wang 2008). This group of orthogonal initial perturbations was defined as the nonlinear local Lyapunov vectors (NLLVs). The first NLLV is referred to as the leading NLLV (LNLLV). NLLVs are a set of orthogonal vectors spanning the fast-growing subspace and would be expected to be a quick and efficient ensemble generation scheme.

Traditional BVs are obtained through natural breeding cycles. However, natural breeding may not be efficient for two reasons. First, the BVs span a relatively small part of the subspace of fast-growing directions owing to their dependence of perturbations. Second, BVs naturally evolved from random perturbations may not be able to provide a stable indication of the fastest-growing direction after a finite breeding period. We attempted to use the comparison and orthogonalization methods to obtain NLLVs to overcome these two problems. The comparison method, which was used by A04, attempts to capture the fastest-growing direction by comparing the growth rate of orthogonal BVs; it is, therefore, more effective than a single breeding cycle. Here, we will use this comparison method for the acquisition of the LNLLV. The additional NLLVs can be obtained successively by using Gram–Schmidt reorthonormalization (GSR) (Wolf et al. 1985; Li and Wang 2008). A detailed description of the GSR and comparison methods is given in section 2. The first few NLLVs are obtained in this way by optimally selecting the fastest-growing modes from the initial perturbations; they are strictly orthogonal and thereby provide an appropriate orthogonal basis for the fast-growing subspace. The NLLV ensemble would be expected to outperform the BV ensemble. A comparison of the performances of the BV and NLLV methods in a perfect system is our primary aim here. Results using the Monte Carlo method are also included for reference.

The remainder of this paper is arranged as follows. Section 2 provides a brief description of the NLLE, NLLEs, and NLLVs, while section 3 introduces the idealized models and the initialization of the analysis states. A description of the BV and NLLV ensemble generation schemes is given in section 4, and the performance of the different ensemble methods are quantitatively compared in section 5. The discussion and our conclusions are presented in section 6.

2. Defining the nonlinear local Lyapunov vector

Consider a general n-dimensional nonlinear dynamical system whose evolution is governed by

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}), \tag{1}$$

where $\mathbf{x} = [x_1(t), x_2(t), \ldots, x_n(t)]^{\mathrm{T}}$ is the state vector at time t, the superscript T is the transpose, and $\mathbf{F}$ represents the dynamics. The evolution of a small error vector $\boldsymbol{\delta} = [\delta_1(t), \delta_2(t), \ldots, \delta_n(t)]^{\mathrm{T}}$, superimposed on a state $\mathbf{x}$, is governed by the nonlinear equations:

$$\frac{d}{dt}\boldsymbol{\delta} = \mathbf{J}(\mathbf{x})\boldsymbol{\delta} + \mathbf{G}(\mathbf{x}, \boldsymbol{\delta}), \tag{2}$$

where $\mathbf{J}(\mathbf{x})\boldsymbol{\delta}$ are the tangent linear terms and $\mathbf{G}(\mathbf{x}, \boldsymbol{\delta})$ are the high-order nonlinear terms of the error $\boldsymbol{\delta}$. Integrating Eq. (2) from $t = t_0$ to $t_0 + \tau$, we obtain

$$\boldsymbol{\delta}(t_0 + \tau) = \boldsymbol{\eta}[\mathbf{x}(t_0), \boldsymbol{\delta}(t_0), \tau]\boldsymbol{\delta}(t_0), \tag{3}$$

where $\boldsymbol{\eta}[\mathbf{x}(t_0), \boldsymbol{\delta}(t_0), \tau]$ is defined as the nonlinear propagator (Ding and Li 2007), which differs from the linear propagator $\mathbf{M}[\mathbf{x}(t_0), \tau]$ obtained from the tangent linear equation of Eq. (2) (Yoden and Nomura 1993; Ziehmann et al. 2000). In a chaotic system, after several regular rescalings, any initial perturbation started a short time

before t_0 will converge to the fastest-growing direction $\boldsymbol{\delta}_1(t_0)$, which is defined as the LNLLV.

Once $\boldsymbol{\delta}_1(t_0)$ has been acquired, the maximal NLLE can be computed from the rate of change of its norm using Eq. (3):

$$\lambda_1[\mathbf{x}(t_0), \boldsymbol{\delta}_1(t_0), \tau] = \frac{1}{\tau} \ln \frac{\|\boldsymbol{\delta}_1(t_0+\tau)\|}{\|\boldsymbol{\delta}_1(t_0)\|}, \tag{4}$$

where $\lambda_1[\mathbf{x}(t_0), \boldsymbol{\delta}(t_0), \tau]$ depends on the initial state $\mathbf{x}(t_0)$ in phase space, the initial error vector $\boldsymbol{\delta}(t_0)$, and the evolution time τ.

If LNLLV along the fastest-growing direction has been obtained, additional NLLVs $[\boldsymbol{\delta}_2(t_0), \boldsymbol{\delta}_3(t_0), \ldots, \boldsymbol{\delta}_m(t_0)]$ along other directions can be successively obtained by the GSR (Li and Wang 2008). In GSR procedure, the direction of the first vector is never affected. The second vector is projected onto the subspace orthogonal to the first vector, and iteratively the nth vector is projected onto the subspace orthogonal to the previous $n-1$ vectors. If $\boldsymbol{\delta}'_1(t_0), \boldsymbol{\delta}'_2(t_0), \ldots, \boldsymbol{\delta}'_m(t_0)$ are m perturbations through breeding procedure, the GSR would provide the following orthogonal set $\boldsymbol{\delta}_1(t_0), \boldsymbol{\delta}_2(t_0), \ldots, \boldsymbol{\delta}_m(t_0)$:

$$\boldsymbol{\delta}_1(t_0) = \boldsymbol{\delta}'_1(t_0),$$

$$\boldsymbol{\delta}_2(t_0) = \boldsymbol{\delta}'_2(t_0) - \frac{(\boldsymbol{\delta}_1(t_0), \boldsymbol{\delta}'_2(t_0))}{(\boldsymbol{\delta}_1(t_0), \boldsymbol{\delta}_1(t_0))} \boldsymbol{\delta}_1(t_0),$$

$$\vdots$$

$$\boldsymbol{\delta}_m(t_0) = \boldsymbol{\delta}'_m(t_0) - \frac{(\boldsymbol{\delta}'_n(t_0), \boldsymbol{\delta}_{n-1}(t_0))}{(\boldsymbol{\delta}_{n-1}(t_0), \boldsymbol{\delta}_{n-1}(t_0))} \boldsymbol{\delta}_{n-1}(t_0) - \cdots - \frac{(\boldsymbol{\delta}'_n(t_0), \boldsymbol{\delta}_1(t_0))}{(\boldsymbol{\delta}_1(t_0), \boldsymbol{\delta}_1(t_0))} \boldsymbol{\delta}_1(t_0), \tag{5}$$

where $(\cdot, \cdot)$ signifies the inner product. In practice, $\boldsymbol{\delta}_1(t_0), \boldsymbol{\delta}_2(t_0), \ldots, \boldsymbol{\delta}_m(t_0)$ are orthogonal to each other, representing the vectors along the directions from the fastest-growing direction to the fastest-shrinking direction. The first few of them will be utilized as ensemble initial perturbations. We may also determine the ith NLLE $\lambda_i(\mathbf{x}(t_0), \boldsymbol{\delta}(t_0), \tau)$ directly from the growth rate of vector $\boldsymbol{\delta}_i(t_0)$:

$$\lambda_i[\mathbf{x}(t_0), \boldsymbol{\delta}_i(t_0), \tau] = \frac{1}{\tau} \ln \frac{\|\boldsymbol{\delta}_i(t_0+\tau)\|}{\|\boldsymbol{\delta}_i(t_0)\|} \quad (i = 2, 3, \ldots, m). \tag{6}$$

The NLLEs depend on $\mathbf{x}(t_0)$, and the sum of them determines the divergence or expansion of a phase-space volume element surrounding the initial states. Therefore, the first few NLLVs corresponding to the largest NLLEs properly span the unstable subspace of initial states. The LNLLV is derived through the breeding method; therefore, it presents some similarity to the BV. However, the largest difference between the NLLV method and the BV method is that besides the LNLLV other orthogonal fast-growing directions are also considered in the generation of ensemble perturbations. Because the first few NLLVs point to independent-growing directions, the perturbation growth of analysis error would be much more likely to be simulated.

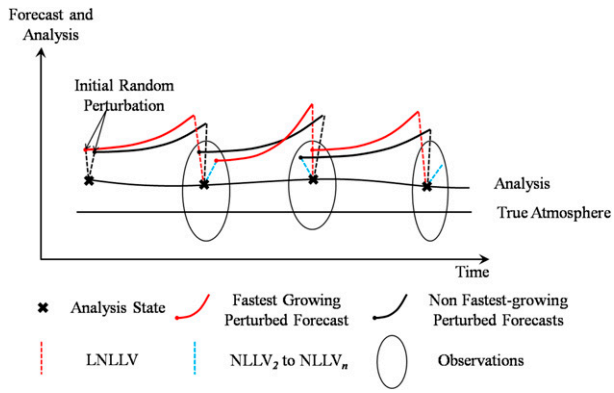

FIG. 2. Schematic representation of the analysis cycle and the creation of NLLVs. In each analysis cycle, the analysis states are updated by combining the background forecasts (black dashed curve) and the observations (open circle). NLLVs are attained to capture the fastest-growing components of the analysis errors. A group of initial random perturbations are added to the initial analysis state and integrated to the end of a breeding cycle similar to the background forecast in analysis cycles. The red solid curve represents the fastest-growing perturbed forecast during each cycle and the black solid curves represent other perturbed forecasts from the second to the nth fastest. The resulting perturbations between the analysis and the perturbed forecasts are then rearranged according to their growth rates during the last cycle. The direction of the fastest-growing mode (LNLLV, red dashed line) is kept, and the other directions (black dashed lines) are orthogonalized successively to acquire NLLV$_2$ to NLLV$_n$ (blue dashed line). Then these perturbations are rescaled to the same size as the initial perturbations. This process is repeated for several cycles to obtain effective NLLVs.

To obtain the fastest-growing LNLLV more effectively than a single BV, we adapted the comparison method of A04 before orthogonalizing the perturbations (as illustrated in Fig. 2). This method includes two steps: a comparison of the growth rates of perturbations in each breeding cycle and a reordering of the perturbations by their growth rate from the largest to the smallest. The NLLVs bear some similarities to the orthogonal BVs as proposed by A04 in the use of the comparison method. However, some differences still exist between them. First, the orthogonal BVs in A04 are mainly used in the local region to increase local diversity, while the NLLVs are globally orthogonal and

represent a nonlinear extension of the LLVs based on the theory of Lyapunov stability. Second, the generated ensemble perturbations in A04 are much less correlated than the BVs but are not strictly orthogonal. A04 reported that the mean correlation between any two ensemble perturbations is 0.17 (i.e., not zero) at the end of the integration time in the Lorenz96 model, implying that these perturbations are not strictly orthogonal. However, the NLLVs are strictly orthogonal (i.e., zero correlation) and so form an appropriate orthogonal basis for the fast-growing subspace.

3. Experimental setup

To determine the differences among these ensemble methods as clearly as possible, we made our comparison based on two ideal models, namely, the Lorenz63 (Lorenz 1963a) and Lorenz96 models (Lorenz 1996). The former can be regarded as a one-point model, while the latter can been seen as resembling a low-order real NWP model. We have taken special account of the generation of the control analyses to ensure that they have no partiality for any initialization methods. Anderson (1997) simply used random perturbations as analysis errors to compare dynamically constrained (i.e., BV and SV) and unconstrained perturbation methods, but these kinds of random perturbations in dynamical systems were recognized to be the nongrowing type on average, which would reduce initially (Toth and Kalnay 1993). Therefore, growing perturbations (i.e., BV and SV) may have no advantage in simulating such analysis errors. Houtekamer and Derome (1994) even used any of the BVs as an analysis error to investigate the improvement of the breeding method to control the forecast. Our experiments found that the higher correlation between the BVs and the analysis error give some unfair advantage to the breeding method over other ensemble methods (not shown). To perform a thorough comparison of the different initialization methods, following Descamps and Talagrand (2007), the control analyses in the BV, NLLV, and random perturbation methods were obtained using the ensemble Kalman filter (EnKF) data assimilation method (Evensen 2003, 2004). The detailed description of this assimilation method is shown in appendix B.

The BV and NLLV cycles are implemented in relation to the analysis cycle. In each analysis cycle, an analysis state is updated by combining the observations and forecast states from a numerical model, as shown in Fig. 2. The forecast error that evolves in time through the atmospheric flow will naturally project into the growing directions. Therefore, growing components will develop within the analysis cycle and dominate the analysis errors. The breeding and NLLV cycles are implemented to simulate the development of growing errors in the analysis cycle and to obtain fast-growing modes associated with the evolving state of the atmosphere. This allows the BVs and NLLVs to sample efficiently the probability of analysis error.

The first model is the Lorenz63 model (Lorenz 1963a):

$$\begin{cases} \dot{x} = -\sigma x + \sigma y \\ \dot{y} = rx - y - xz \\ \dot{z} = xy - bz \end{cases}, \tag{7}$$

where $\sigma = 10$, $r = 28$, and $b = 8/3$, and for which the well-known butterfly attractor exists. We selected a fourth-order Runge–Kutta time integration scheme with a time step equal to 0.01 time units. The linear Lyapunov exponents of this simple model were calculated to be 0.906, 0, and −14.572, under the given parameters (Wolf et al. 1985). This means that there is only one fast-growing dimension on the attractor. Consequently, we performed a two-member experiment of ensemble prediction ($N = 2$, where N denotes the number of the ensemble member in a forecast example) for this model.

The Lorenz96 model is a 40-variable ($m = 40$, where m denotes the dimension of the state vector of the ensemble forecast model) model that has been used by various authors as a low-order proxy for atmospheric prediction and assimilation studies (e.g., Lorenz and Emanuel 1998; Anderson 2001). We denote the state variables $X_i (i = 1, 2, \ldots, 40)$, which are governed by the equations

$$dX_i/dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F. \tag{8}$$

In the case $F = 8$, the model solves Eq. (8) also using a fourth-order Runge–Kutta scheme with a time step of 0.05 time units. This model can be seen as the time evolution of an arbitrary one-dimensional quantity on a constant latitude circle in which a time increment of 0.2 nondimensional units corresponds to about 1 day in terms of the error growth rate (Lorenz 1996).

Here, all experiments are perfect model experiments in which the output of a long-period integration will be regarded as references of the forecasts once the system has reached its attractor. The observations to be assimilated are obtained by adding a random noise to the true state. We assume that there are observations on every grid. If a vector $\mathbf{x}$ denotes the true state vector, and $\mathbf{y}$ is the observation vector, then

$$\mathbf{y} = H\mathbf{x} + \boldsymbol{\varepsilon}, \tag{9}$$

where H is the observation operator that defines the projection of prognostic variables $\mathbf{x}$ to the space of observational data $\mathbf{y}$. The added noise vector $\boldsymbol{\varepsilon}$ has a Gaussian distribution of $\mathbb{N}(0, 1)$ for both the Lorenz63 and Lorenz96 models, where $\mathbb{N}$ denotes the sample

TABLE 1. Ensemble size, number of observation stations, assimilation interval, and total assimilation length of the EnKF assimilation for the Lorenz63 and Lorenz96 models.

Model	Ensemble size	Observation network	Assimilation interval (time units)	Assimilation length (time units)
Lorenz63	20	3	0.1	0.4
Lorenz96	500	40	0.05	10

space of the Gaussian distribution, 0 is the expectation, and 1 is the standard deviation.

To make the experimental setup clearer, two tables listing the experiment setup of EnKF data assimilation and the breeding procedures are included (Tables 1 and 2, respectively). The statistical performances of the forecasts are compared, for each initialization scheme, by averaging over a large number of experiments. The number of experiments was $M = 10\,000$ with different initial analysis states for the Lorenz63 model and $M = 500$ for the Lorenz96 model. To ensure that the different experiments were conducted independently, they had different initial states and different observations. For the Lorenz63 model observations were assimilated every 0.1 time units and the performing time was 0.4 time units in each case, while for the Lorenz96 model the assimilation cycle was 0.05 time units (6 h) and was repeated over 10 time units (50 days). Under the parameters as set here, the amplitudes of the initial analysis errors were within the bounds 5%–10% of the natural variability (about 14 for the Lorenz63 model and 2.8 for the Lorenz96 model; in the L_2 norm). The mean root-mean-square error (RMSE) of global analysis is about 0.98 for the Lorenz63 model and about 0.17 for the Lorenz96 model.

Once the analysis states are generated, the ensemble perturbations produced by the random perturbation, BV and NLLV schemes will be added and subtracted from them. The BVs and NLLVs are generated through breeding cycles with regular rescaling. For the Lorenz63 model, the length of a breeding cycle of the NLLV and BV methods is 0.1 time units and the process is repeated over 0.4 time units. For the Lorenz96 model, the length of a breeding cycle of both methods is 0.2 time units (1 day) and the breeding time is 2 time units (10 days). The orthogonalization process for the NLLVs is performed during each breeding cycle. As for the random perturbation method, the random noise is from a Gaussian distribution of $\mathbb{N}(0, 1)$. The ensemble perturbations of the three methods have the same size as that of the analysis errors for both models. The ensemble members were integrated for 4 time units for the Lorenz63 model to implement ensemble forecasts and 2 time units (10 days) for the Lorenz96 model. The initial states to start assimilation in each sample were drawn at an interval of 0.05 time units (0.2 time units) for the Lorenz63 model (the Lorenz96 model) after a first spinup of 1000 time units.

4. Nonlinear model results

a. *Perturbation growth rate*

Before providing a detailed assessment of the quality of the ensemble predictions, the error growth of the different initial schemes will be investigated. For the Lorenz63 model, the growth rates of the vectors from the three schemes were compared by calculating the ratio of the final to initial error (in the L_2 norm) at the integrated interval (i.e., the time of a breeding cycle 0.1 time units after the initial state). The parameters used to generate the initial perturbations were those referred to above. However, we fixed the rescaling size to 1.0, which means that the final size of the error vector is simply the growth rate. Three random vectors are used to generate the LNLLV. All results presented here were averaged over 10 000 samples. Similar to the results from A04, the standard breeding method results in an average ratio of 1.26, while the LNLLV has an average ratio of 1.44. As for the random vector, it initially points to a nongrowing direction on average and thus has an average ratio of only 0.96.

The process for the Lorenz96 model was similar to that for the Lorenz63 model. The rescaling factor was 0.17 (the average RMSE of analysis error). After a breeding period of 2 time units (10 days), the first EOF explains 52% of the variance of the bred vectors, indicating that these modes have a certain degree of similarity in

TABLE 2. Number of experiments, number of ensemble members, rescaling interval, and total breeding length of the BV and the NLLV method for the Lorenz63 and Lorenz96 models.

Model	Number of experiments	Ensemble member	Rescaling interval (time units)	Breeding length (time units)
Lorenz63	10 000	2	0.1	0.4
Lorenz96	500	6	0.2	2

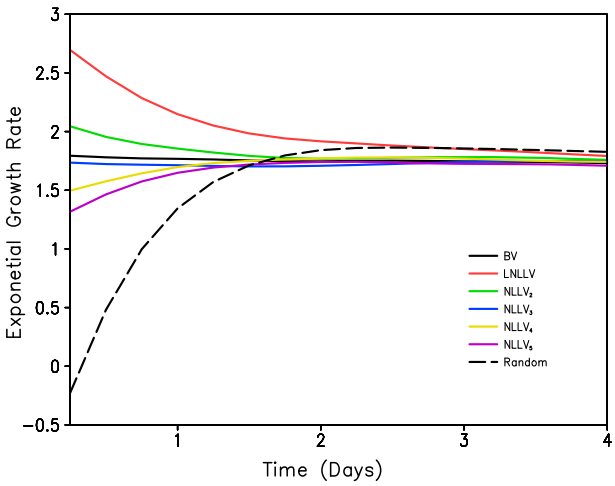

FIG. 3. Average growth rate in the form of Lyapunov exponent as a function of forecast lead time for the bred vector (black solid line), the random vector (black dashed line), and the NLLV (colored solid lines; red, green, blue, yellow, and purple from LNLLV to $NLLV_5$).

patterns that represent the unstable modes determined by the background flow (Corazza et al. 2003). A group of 10 vectors were used for the generation of the NLLVs and BVs. The comparison of growth rates from the different schemes is shown in Fig. 3. Here, the Lyapunov exponent was used to measure the perturbation growth.

If the size of the perturbation at time t is expressed by $\|\Delta x(t)\|$, then the Lyapunov exponential growth can be calculated as

$$\lambda = \frac{1}{t - t_0} \ln \frac{\|\Delta x(t)\|}{\|\Delta x(t_0)\|}, \tag{10}$$

where t_0 denotes the initial time. Only the results of the first five NLLVs (LNLLV to $NLLV_5$) are shown here. Each performance is the average of 500 samples. The LNLLV appears to remain the fastest during the first 4 days. The NLLVs represent different directions ordered by their growth rates. Among the first five NLLVs, the growth rates of the first two both over take those of the BVs during the first 2 days, while the growth rates of the last three NLLVs are slightly smaller than that of the BV. These results indicate that the NLLVs span the fast-growing subspace well. After that period, with the growth of the perturbations, the growth rates of all perturbations slow and become very close. The random perturbations are initially characterized by negative growth rates; the growth becomes enhanced (superexponential) (Trevisan and Legnani 1995) after about 1 day. This may be related to the fast growth of the small-scale errors. After small-scale error saturation had occurred in a short time, the growth rate gradually slowed down and mainly depended on the error growth in the large scales (Lacarra and Talagrand 1988; Molteni and Palmer 1993). We chose three pairs of members ($N = 6$) to implement ensemble forecasts for the Lorenz96 model. The use of more ensemble members only limitedly improves the prediction skill for both the BV and NLLV methods.

b. *The Lorenz63 model*

The RMSE and pattern anomaly correlation (PAC) offer good measures of overall forecast performances (Buizza et al. 2005). Figures 4a and 4b shows the mean error (in the L_2 norm) and the mean correlation (cosine of the angle of forecast state vector and true state vector), respectively, as a function of forecast lead time. The corresponding measurements of control forecasts are also shown for comparison. The ensemble averaging significantly improves the performance of the forecasts as a whole, although different ensemble methods, because of their distinct initialization schemes, have different forecast skills. On account of the analysis cycles that produce the initial states, growing errors have gradually accounted for a significant proportion of the analysis errors. Random perturbations, as nongrowing errors, simulate the probabilistic distribution of analysis errors in the suboptimal directions and thus have the worst ensemble performance. The LNLLV has a larger projection onto the fast-growing component of the analysis errors, and so behaves better than the traditional BV, especially for medium and long-range forecasts.

c. *The Lorenz96 model*

The Lorenz96 model is a relatively complex model that resembles a real NWP model and can be used to assess different aspects of forecast performance. As, in the real world, uncertainties in the initial states are inevitable, we assumed that the control analysis is the most real initial state we could obtain. Here, all results were derived from the comparison with true references.

1) OVERALL MEASURES OF ENSEMBLE PERFORMANCE

Similarly to the Lorenz63 model, the mean RMSE and PAC were also used here to assess overall performance (Fig. 5). Besides, the ensemble spread is added in Fig. 5a as a comparison with the RMSE of ensemble mean. The spread is defined as the standard deviation of the perturbed ensemble members. The three ensemble schemes evidently outperform the single control forecast in ensemble skill. The gain in predictability from running an ensemble is about 12–24 h on forecast days 7–10. For the first few days (about 4 days), ensemble averages are indistinguishable from the control forecasts. The initial

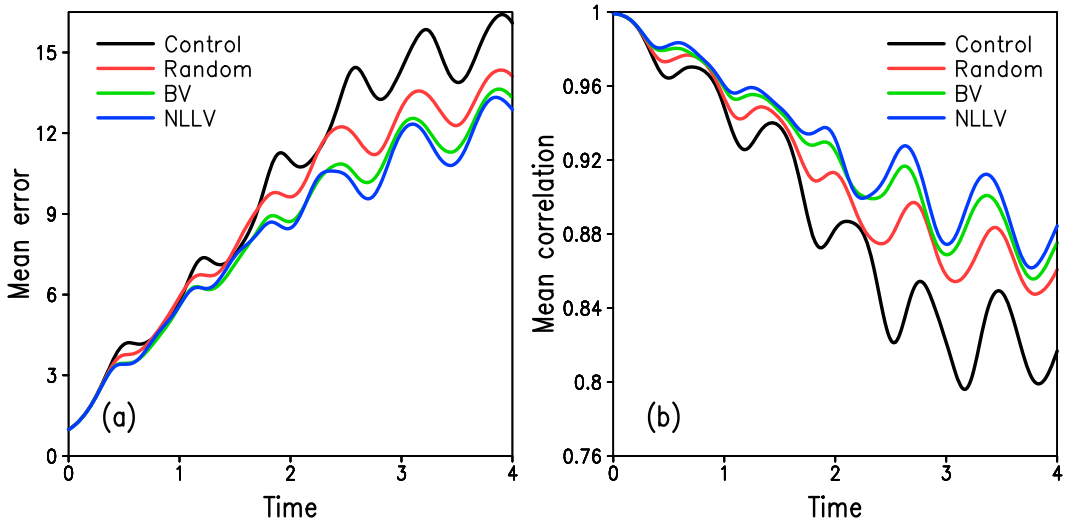

FIG. 4. (a) Mean error and (b) mean correlation of 10 000 samples as a function of lead time from the Lorenz63 model for the control run (black line), random perturbation method (red line), BV method (green line), and NLLV method (blue line).

interval can be considered to be a linear stage, during which positive and negative twin perturbations cancel each other out almost completely. However, when the errors gradually increase and enter the nonlinear stage, ensemble averaging plays a much more important role in nonlinear filtering to reduce error growth.

The ensemble skill of the BVs did not show an evident improvement over that of the random vectors. The reason for this may be that during the initial linear stage, the random vectors have already pointed to growing directions through the dynamical flow during the forecast stage, which also spans the subspace of growing errors. However, random vectors have a lower correlation with the analysis errors than bred vectors, and so perform slightly less well than bred vectors, especially in the PAC score. Compared with these two schemes, the ensemble mean of NLLVs is more skillful for 5–10-day forecasts. This can be explained by the fact that the NLLVs sample analysis errors in more mutually orthogonal directions than do the other two schemes. The

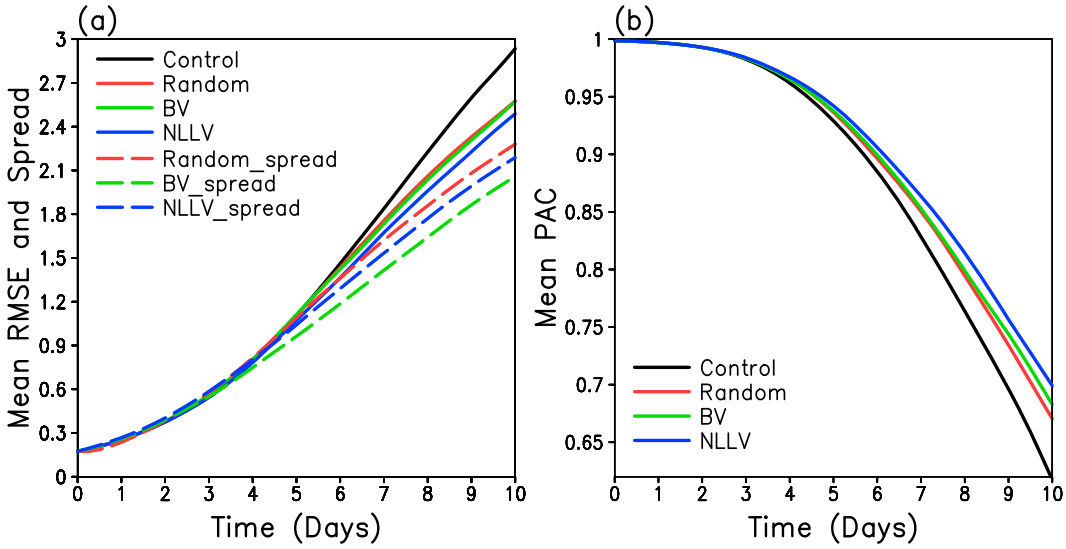

FIG. 5. (a) Mean RMSE (solid lines) and ensemble spread (dashed lines), and (b) mean PAC (pattern anomaly correlation) of 500 samples as a function of lead time from the Lorenz96 model for the control run (black), random perturbation method (red), BV method (green), and NLLV method (blue).

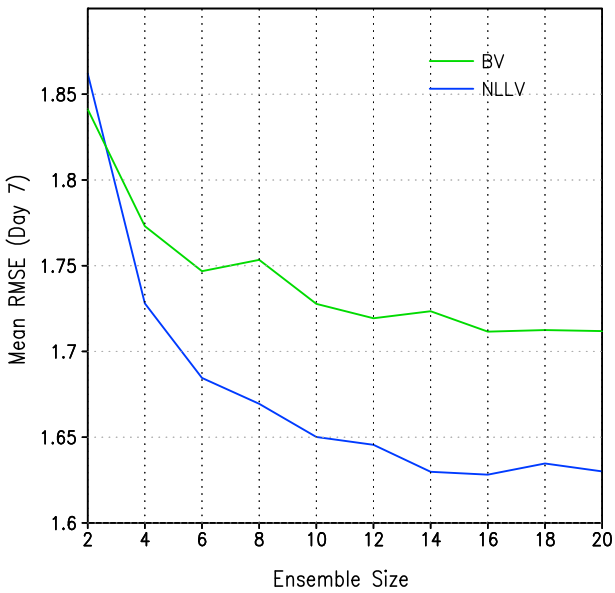

FIG. 6. Mean RMSE of 500 samples at day 7 as a function of the number of BV (green line) and NLLV (blue line) ensemble members from the Lorenz96 model.

probability distribution of the analysis errors are more likely to be well simulated by the orthogonal basis the NLLVs form.

The spread is a measure of ensemble reliability. For a statistically reliable ensemble system, ensemble spread should be close to the error of the ensemble mean (Buizza et al. 2005). Figure 5 shows that the growth of the RMSE exceeds that of the spread for all three methods. The use of the comparison method in generating the ensemble perturbations allows the NLLV method to exhibit the largest perturbation growth during the first 2 days; after this, the random method exhibits the largest spread overall, but the spatial distribution of its spread has the lowest correlation with that of the ensemble forecast error (Fig. 6). The ensemble spread of the NLLVs remained larger than that of the BVs from days 1 to 10. This was attributed to the NLLVs sampling the analysis errors in directions much less correlated than used in the breeding schemes.

To distinguish the respective effects of the comparison and the orthogonalization processes on ensemble forecasting, three experiments were implemented and their results were listed in Table 3. By comparing experiments 1 and 2, the only use of the orthogonalization procedure could reduce about 3%–4% of the RMSE of ensemble mean relative to the BV method. The further use of the comparison method as in experiment 3 could have another nearly 1% improvement for the ensemble-mean skill. These results indicate that the global orthgonality is more critical than the comparison method for improving

TABLE 3. Mean RMSE of ensemble mean for the Lorenz96 model at days 3, 6, and 9 for different experiments. Each experiment was implemented with different initial perturbations, including three bred vectors (expt 1), three orthogonal NLLVs (expt 2), and three orthogonal NLLVs selected from ten perturbations using the comparison method (expt 3).

	RMSE of ensemble mean		
Initial perturbations	Day 3	Day 6	Day 9
Expt 1	0.572	1.435	2.318
Expt 2	0.552	1.380	2.251
Expt 3	0.545	1.364	2.228

ensemble forecasting in the Lorenz96 model. But the two methods may have different relative performance based on different models and different assimilation systems. This needs a further investigation in future research.

We also compared the forecast error for the BV and NLLV methods at day 7 with the same number of ensemble members (Fig. 6). It can be seen that the ensemble-mean skill for the BVs and NLLVs are very close when two ensemble members are used. However, with the increase of the number of ensemble members, the forecast error of the NLLV and BV schemes are both decreased, and the performance of NLLVs is more significant than the BVs. It indicates that the addition of the rest NLLVs have larger contributions to the improvement of prediction skill than the BVs. It can be attributed to the orthogonality of the NLLVs that better spanning the fast-growing subspace though the rest of them point to mild-growing modes.

2) THE VALIDATION SCORES

Reliability and resolution are considered to be two important attributes of ensemble predictions (Murphy 1973). Here, two scores are used to assess the performances of the ensemble forecasts in statistical reliability and resolution. One is the spread–skill correlation and the other is the classical Brier score.

The first measure of statistical reliability focuses on spread and skill correlations. The average error of the ensemble-mean forecasts from the three schemes is used as a reference field to ensure fair comparisons; that is, the forecast skill is defined as the average distance of the respective ensemble-mean forecasts from the true state. Regions with high forecast skill (small average error) regularly have a relatively small spread and vice versa (Buizza et al. 2005). The spread–error correlation over the whole field (40 grids) was computed as a measure of transient consistency. The day-to-day variations of this measure for the BV, random perturbations, and NLLV methods are shown in Fig. 7. As a comparison, the results of the linear LLVs are also given. Similar to the

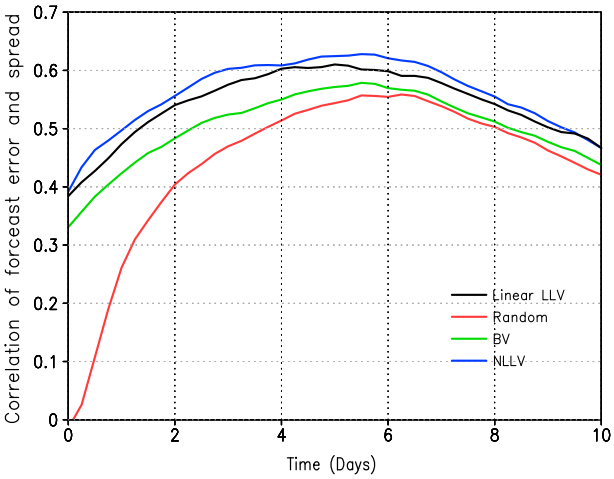

FIG. 7. The correlation coefficients between the ensemble error and the ensemble spread as a function of time for the four schemes, linear LLVs (black line), random perturbations (red line), BVs (green line), and NLLVs (blue line). All results are the average of 500 samples.

NLLVs, they are also orthogonal perturbations, but they are obtained using an integration of the tangent linear model (TLM) of Lorenz96.

The forecast error field at time 0 can be considered to be the analysis error amplitude. As the ensemble perturbations are designed to represent analysis errors, the ensemble spread should have a similar pattern to the analysis distance field. As is indicated in Fig. 7, the NLLV scheme has the highest correlation coefficient at the initial time, which means that this method performed best at simulating the initial analysis errors. Furthermore, this method maintains its advantage through the whole forecast period. Because the effects of nonlinear terms were introduced in the NLLVs, the spread of the NLLV ensemble better reflected the nonlinear error growth of the ensemble mean than did the LLVs from 1 to 10 days. It is worth noting that the initial coefficient of the Monte Carlo method approximates to zero, which implies that random perturbations are almost incapable of capturing the mode of the analysis errors. The BV method performs better than the Monte Carlo method. The initial perturbations of these four schemes evolve to become increasingly similar to the forecast errors under the driving effects of background flow. The correlations between the spread and skill of these four schemes spontaneously peaks at about day 5, which may be due to reaching the end of the linear stage and the increasingly strong nonlinear effects. At long forecast ranges, the correlation between spread and skill must decrease to zero (Whitaker and Loughe 1998).

The other score used was the classical Brier score, which measures both the reliability and resolution of the ensemble predictions. This score assesses the performance of the probability forecast of the occurrence of a binary event Φ. In the experiments with the Lorenz96 model, the evaluation was made on each grid $X_i (i = 1, \ldots, 40)$. The Brier score is computed for the event $2.3 \leq X_i \leq 5.9$ (the climatological mean to the distance of one standard deviation), which can be assumed to be normal for warm events. It means that if the value of X_i falls between 2.3 and 5.9, the observed event occurs and vice versa. The final score was calculated as the average of 500 samples. Conventionally, the basic Brier score (BS) can be decomposed into its reliability (BS_{rel}) and resolution (BS_{res}) components (Murphy 1973; Stephenson et al. 2008). Further detail regarding this score is given in appendix B. In brief, BS and BS_{rel} were negatively oriented, which means that smaller values indicate a better performance, while for BS_{res} the opposite was the case. The results are presented in Fig. 8.

As is shown in Fig. 8a, for the basic Brier score, the NLLVs performed better than either the BVs or random vectors, and the advantage of BVs over random vectors was very small. Figure 8b presents the evolution of the two components of BS as a function of time. Overall, the best performance was achieved by the NLLV ensemble, and the figure also indicates that the superiority of the NLLVs over the BVs comes mainly from their reliability. It may be explained that the stronger correlations and smaller ensemble diversity of BVs result in synergic deviation of ensemble forecast states from the observed state. So the BV ensemble may not describe the probability of real events efficiently. On the other hand, the superiority of the NLLVs over random vectors comes mainly from their resolution. This phenomenon may arise because random members sample the true state from a largely unconstrained distribution, which, over longer periods, approximately provides the mean probability of the event occurring. Overall, the relative performance can be ranked as NLLV > BV > Random.

5. Discussion and conclusions

The NWP models have extremely high degrees of freedom, which leads to huge difficulties in sampling all of the directions in the phase spaces. Therefore, one can argue that it is appropriate to sample the directions in a subspace that is relatively more important. One proposal is to employ a breeding cycle similar to the analysis cycle to generate perturbations largely projecting onto the growing components of the analysis errors (Toth and Kalnay 1997). This method dramatically improved the level of the ensemble prediction using finite members and has become a useful operational forecast tool. However, independent breeding cycles may cause a

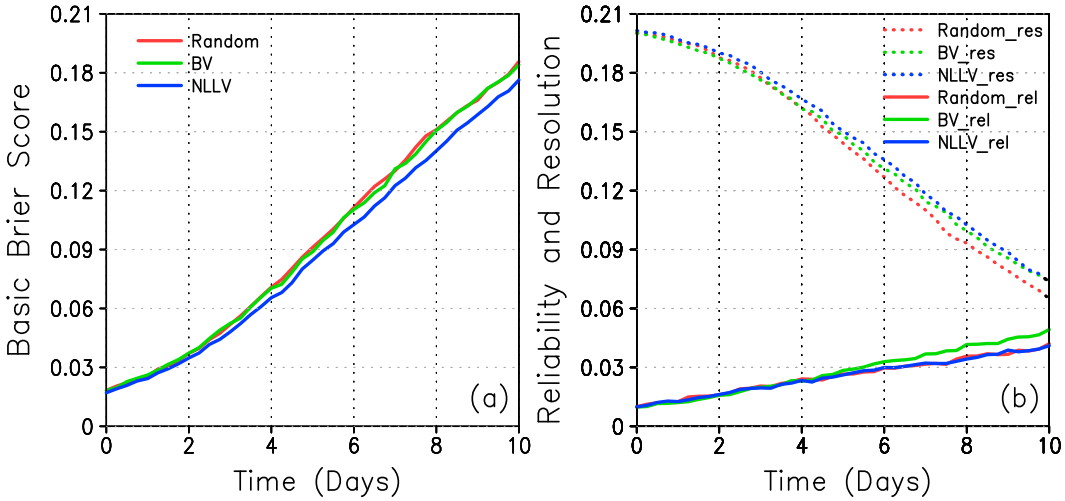

FIG. 8. From the Lorenz96 model: (a) average basic Brier score of 500 samples as a function of lead time for the random perturbation method (red line), BV method (green line), and NLLV method (blue line); (b) resolution (dotted) and reliability (solid) contributions to the Brier score. Values are computed for the event $2.3 \leq X_i \leq 5.9$ ($i = 1, \ldots, 40$).

certain degree of global similarity of bred vectors, especially in local regions, where the stronger correlations are in fact repeated sampling, which degrades the effectiveness of the ensemble. In the present paper, the NLLV ensemble generation technique was proposed to sample the probability distribution of analysis errors using a set of independent and orthogonal directions. Generally, compared to the BVs, the globally orthogonality of the NLLVs can be helpful to reduce the dependence among perturbations and increase the ensemble spread. The comparison method could also be used to optimally select the fastest-growing perturbations. Therefore, the NLLVs performed better than the BVs in simulate the analysis errors in the fastest-growing subspace.

The search for the fastest-growing subspace is improved by using more initial perturbations in the NLLV scheme than in the BV scheme. This increases the computing time (to about 4 times that required by the BV scheme for the Lorenz96 model). However, the breeding cycles of the NLLVs are independent of each other before orthogonalization and thus can be implemented in parallel to save computational time. Therefore, the increased computation cost of the NLLVs in comparison with the BVs is mainly attributed to the comparison and orthogonalization procedures. Our experiment using the Lorenz96 model has a computational expense for generating the NLLV initial perturbations that is only about 8% greater than that of generating the BV initial perturbations.

For the low-order models discussed here, the NLLV method has shown a relatively better performance than those of the BV and the Monte Carlo methods. The Monte Carlo method generally needs many more samples because of its defects in the recognition of the specification of initial errors, which leads to a large computational cost. On the other hand, random sampling degrades its resolution for actual events (Fig. 8b). The significant advantage of BVs is the ease of implementation. However, in our experiments with the Lorenz96 model, the BVs, which have stronger global similarity and smaller ensemble spread, may not describe the probability of real events efficiently and thus have lower reliability.

Our experiments provide a preliminary demonstration of the GSR method as a quick and effective method for separating independent ensemble modes in the Lorenz 3- and 40-variable models. A high-dimensional numerical model will greatly increase the dimension of the ensemble perturbations and also bring certain difficulties in utilizing the GSR method. However, the dimension of the fast-growing subspace is much smaller than the model dimension. Orthogonal perturbations can capture this subspace to the greatest extent, and orthogonalization is only needed to generate the first few fastest-growing NLLVs, thus greatly saving computation time. The present computing power is sufficient to satisfy the requirements of the orthogonalization in high-order models. Therefore, the NLLV scheme could be expected to be quick and effective for generating ensemble perturbations in a high-dimension numerical model. Nevertheless, some other techniques mentioned in the introduction have been proposed to increase the diversity of traditional BVs. These schemes may be able

to improve the prediction skill of the traditional BV method. It is worthwhile investigating the performance of NLLVs and comparing with these methods in more complex models.

Acknowledgments. We wish to thank Dr. Zhina Jiang for many useful comments and assistance with this work. This research was jointly supported by the NSFC projects (41375110, 41175069) and the 973 projects of China (2012CB955200 and 2010CB950400).

APPENDIX A

Ensemble Generation Scheme of Bred Vector

Initially, Toth and Kalnay (1993) used a global rescaling factor to generate each ensemble perturbation when using the BV method. In general, the value of the factor is the same size as the empirically determined global analysis RMSE (Wang and Bishop 2003). However, global rescaling is not optimal for reflecting the characteristics of the observations over local regions because it appears that regions where observations are limited usually have larger analysis errors, and larger regional perturbation amplitudes are therefore expected, and vice versa. Based on this relationship, Toth and Kalnay (1997) introduced the smooth regional rescaling factor.

Here, only simple breeding is considered using a global rescaling factor since all the variables in the Lorenz96 model are equivalent and the observations have the same level for each grid point. The procedure can be easily expressed by the equation

$$\boldsymbol{\delta}^a = \boldsymbol{\delta}^f \cdot \frac{c}{d}, \tag{A1}$$

where $\boldsymbol{\delta}^f$ and $\boldsymbol{\delta}^a$ are the perturbation vectors of forecast and analysis, respectively, and d represents the amplitude of the $\boldsymbol{\delta}^f$, while c is the value of the rescaling factor. The detailed processes used to generate BVs are schematically described in Fig. 1 (Toth and Kalnay 1993, 1997). The initial analysis state (in this paper, the true state was used) was perturbed by adding a random vector of the same size as the analysis error. The model was then integrated from the corresponding perturbed states. At the end of each cycle, the perturbation was rescaled to the initial size and again superposed on the new reference to integrate. This breeding cycle is repeated to the initial state for forecasting. In this paper, the time of the breeding cycle of the Lorenz63 and Lorenz96 models was 0.1 and 0.2 time units (1 day), respectively.

APPENDIX B

The Ensemble Kalman Filter

The generation of initial analysis states in these experiments was based on the algorithm described by Evensen (2003, 2004). For a given observation time, the ensemble of forecast states are obtained as a reference. The ensemble matrix is then defined as

$$\mathbf{X}^f = (\mathbf{x}_1^f, \mathbf{x}_2^f, \ldots, \mathbf{x}_N^f). \tag{B1}$$

Here, if $\overline{\mathbf{x}^f}$ is denoted as the mean of the ensemble, then an ensemble perturbation matrix can be expressed as

$$\mathbf{X}'^f = (\mathbf{x}_1^f - \overline{\mathbf{x}^f}, \mathbf{x}_2^f - \overline{\mathbf{x}^f}, \ldots, \mathbf{x}_N^f - \overline{\mathbf{x}^f}). \tag{B2}$$

The variance covariance matrix is

$$\mathbf{P}^f = \frac{1}{N-1}\mathbf{X}'^f\mathbf{X}'^{f\mathrm{T}}. \tag{B3}$$

The $\mathbf{P}^f$ is used to update the Kalman gain $\mathbf{K}$, which is written as

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}. \tag{B4}$$

Here, $\mathbf{K}$ is actually a weighting according to the ratio of the forecast and observational error covariance; $\mathbf{R}$ is the observational error covariance matrix. The observations are assimilated to produce a new analysis of the state:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i^f). \tag{B5}$$

The $\mathbf{y}_i (i = 1, 2, \ldots, N)$ is a set of perturbed observations that corresponds to each previous forecast of the ensemble. They are defined as

$$\mathbf{y}_i = \mathbf{y} + \boldsymbol{\varepsilon}_i. \tag{B6}$$

The $\boldsymbol{\varepsilon}_i$ values are independent realizations of the noise $\mathbb{N}(0, 1)$ and are equivalent to $\boldsymbol{\varepsilon}$ in section 3. Here, the average of all analysis states of the ensemble (i.e., $\overline{\mathbf{x}^a}$) is regarded as the initial analysis state when performing the forecasts. To avoid the problem of undersampling, a 7% variance inflation factor was applied to $\mathbf{X}'^f$ for both of the Lorenz models. The ensemble size of the EnKF assimilation for the Lorenz63 and Lorenz96 models is much larger than the model dimension that has been efficient to estimate the covariance matrix of the background error. Therefore, the localization method for the EnKF system of both models was not used.

APPENDIX C

Brier Score Decomposition

The algorithm of the Brier score used here is based on Stephenson et al. (2008). Assume that the unit probability interval [0, 1] is divided to m bins. The total forecasts whose number is n are placed in corresponding bins by their forecast probability. Denote the n_k probability forecasts that have fallen in the kth bin, then $n = \sum_{k=1}^{m} n_k$. In the kth bin, the probability of forecasts and observations read f_{kj} and o_{kj} (0 or 1), $j = 1, 2, \ldots, n_k$, respectively. The basic Brier Score can be written

$$\mathrm{BS} = \frac{1}{n}\sum_{k=1}^{m}\sum_{j=1}^{n_k}(f_{kj} - o_{kj})^2. \tag{C1}$$

The Brier score can be decomposed into its reliability, resolution, and the uncertainty components:

$$\mathrm{BS} = \mathrm{BS}_{\mathrm{rel}} - \mathrm{BS}_{\mathrm{res}} + \mathrm{BS}_{\mathrm{unc}}, \tag{C2}$$

$$\mathrm{BS}_{\mathrm{rel}} = \frac{1}{n}\sum_{k=1}^{m} n_k(\overline{f_k} - \overline{o_k})^2, \tag{C3}$$

$$\mathrm{BS}_{\mathrm{res}} = \frac{1}{n}\sum_{k=1}^{m} n_k(\overline{o_k} - \overline{o})^2, \tag{C4}$$

where $\overline{f_k}$ is the average forecast probability and $\overline{o_k}$ is the relative frequency that the observed event occurred in the kth bin; $\overline{o}$ is the mean probability for the event to occur. They are calculated using the equations

$$\overline{f_k} = \frac{1}{n_k}\sum_{j=1}^{n_k} f_{kj}, \tag{C5}$$

$$\overline{o_k} = \frac{1}{n_k}\sum_{j=1}^{n_k} o_{kj}, \tag{C6}$$

$$\overline{o} = \frac{1}{n}\sum_{k=1}^{m} n_k\overline{o_k}. \tag{C7}$$

It is BS, $\mathrm{BS}_{\mathrm{rel}}$, and $\mathrm{BS}_{\mathrm{res}}$ that are used in the present paper. The first two are negatively oriented and the third is positively oriented.

REFERENCES

Anderson, J. L., 1997: The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. *Mon. Wea. Rev.,* **125,** 2969–2983, doi:10.1175/1520-0493(1997)125<2969:TIODCO>2.0.CO;2.

——, 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* **129,** 2884–2902, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Annan, J. D., 2004: On the orthogonality of bred vectors. *Mon. Wea. Rev.,* **132,** 843–849, doi:10.1175/1520-0493(2004)132<0843:OTOOBV>2.0.CO;2.

Benettin, G., L. Galgani, A. Giorgilli, and J. M. Strelcyn, 1980: Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica,* **15,** 9–20, doi:10.1007/BF02128236.

Bishop, C., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Buizza, R., 1997: Potential forecast skill of ensemble prediction, and spread and skill distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.,* **125,** 99–119, doi:10.1175/1520-0493(1997)125<0099:PFSOEP>2.0.CO;2.

——, 2010: Horizontal resolution impact on short and long range forecast error. *Quart. J. Roy. Meteor. Soc.,* **136,** 1020–1035, doi:10.1002/qj.613.

——, P. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.,* **133,** 1076–1097, doi:10.1175/MWR2905.1.

Cai, M., E. Kalnay, and Z. Toth, 2003: Bred vectors of the Zebiak–Cane model and their potential application to ENSO predictions. *J. Climate,* **16,** 40–56, doi:10.1175/1520-0442(2003)016<0040:BVOTZC>2.0.CO;2.

Cheng, Y., Y. Tang, P. Jackson, D. Chen, and Z. Deng, 2010: Ensemble construction and verification of the probabilistic ENSO prediction in the LDEO5 Model. *J. Climate,* **23,** 5476–5479, doi:10.1175/2010JCLI3453.1.

Corazza, M., and Coauthors, 2003: Use of the breeding technique to estimate the structure of the analysis "errors of the day." *Nonlinear Processes Geophys.,* **10,** 233–243, doi:10.5194/npg-10-233-2003.

Descamps, L., and O. Talagrand, 2007: On some aspects of the definition of initial conditions for ensemble prediction. *Mon. Wea. Rev.,* **135,** 3260–3272, doi:10.1175/MWR3452.1.

Ding, R. Q., 2007: *Nonlinear Error Dynamics and Predictability Study* (in Chinese). Chinese Academy of Sciences, 192 pp.

——, and J. P. Li, 2007: Nonlinear finite-time Lyapunov exponent and predictability. *Phys. Lett.,* **364A,** 396–400, doi:10.1016/j.physleta.2006.11.094.

——, ——, and K. Ha, 2008: Trends and interdecadal changes of weather predictability during 1950s–1990s. *J. Geophys. Res.,* **113,** D24112, doi:10.1029/2008JD010404.

——, ——, and K. Seo, 2010: Predictability of the Madden–Julian oscillation estimated using observational data. *Mon. Wea. Rev.,* **138,** 1004–1013, doi:10.1175/2009MWR3082.1.

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus,* **21,** 739–759, doi:10.1111/j.2153-3490.1969.tb00483.x.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–367, doi:10.1007/s10236-003-0036-9.

——, 2004: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.,* **54,** 539–560, doi:10.1007/s10236-004-0099-2.

Fraedrich, K., 1987: Estimating weather and climate predictability on attractors. *J. Atmos. Sci.,* **44,** 722–728, doi:10.1175/1520-0469(1987)044<0722:EWACPO>2.0.CO;2.

Greybush, S. J., E. Kanlnay, M. J. Hoffman, and R. J. Wilson, 2013: Identifying Martian atmospheric instabilities and their physical origins using bred vectors. *Quart. J. Roy. Meteor. Soc.,* **139,** 639–653, doi:10.1002/qj.1990.

Houtekamer, P. L., and J. Derome, 1994: Prediction experiments with two-member ensembles. *Mon. Wea. Rev.,* **122,** 2179–2191, doi:10.1175/1520-0493(1994)122<2179:PEWTME>2.0.CO;2.

——, L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124,** 1225–1242, doi:10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2.

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

Lacarra, J.-F., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus,* **40A,** 81–95, doi:10.1111/j.1600-0870.1988.tb00408.x.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.,* **102,** 409–418, doi:10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2.

Li, J. P., and J. F. Chou, 1997: Existence of atmosphere attractor. *Sci. China,* **40D,** 215–224.

——, and S. H. Wang, 2008: Some mathematical and numerical issues in geophysical fluid dynamics and climate dynamics. *Commun. Comput. Phys.,* **3,** 759–793.

——, and R. Q. Ding, 2009: Studies of predictability of single variable from multi-dimensional chaotic dynamical system (in Chinese with English abstract). *Chin. J. Atmos. Sci.,* **33,** 551–556.

——, and ——, 2011: Temporal–spatial distribution of atmospheric predictability limit by local dynamical analogues. *Mon. Wea. Rev.,* **139,** 3265–3283, doi:10.1175/MWR-D-10-05020.1.

——, and ——, 2013: Temporal–spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans. *Int. J. Climatol.,* **33,** 1936–1947, doi:10.1002/joc.3562.

——, ——, and B. H. Chen, 2006: Review and prospect on the predictability study of the atmosphere (in Chinese). *Review and Prospects of the Developments of Atmosphere Sciences in Early 21st Century,* China Meteorology Press, 96–104.

Lorenz, E. N., 1963a: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

——, 1963b: The predictability of hydrodynamic flow. *Trans. New York Acad. Sci. Ser. II,* **25,** 409–433.

——, 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus,* **17,** 321–333, doi:10.1111/j.2153-3490.1965.tb01424.x.

——, 1996: Predictability: A problem partly solved. *Proc. Seminar on Predictability,* Vol. I, Reading, United Kingdom, ECMWF, 1–18.

——, and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399–414, doi:10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2.

Molteni, F., and T. N. Palmer, 1993: Predictability and finite time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298, doi:10.1002/qj.49711951004.

——, R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The new ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119, doi:10.1002/qj.49712252905.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.,* **12,** 595–600, doi:10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

Primo, C., M. A. Rodríguez, and J. M. Gutiérrez, 2008: Logarithmic bred vectors. A new ensemble method with adjustable spread and calibration time. *J. Geophys. Res.,* **113,** D05116, doi:10.1029/2007JD008998.

Pu, Z. X., and J. Hacker, 2009: Ensemble-based Kalman filters in strongly nonlinear dynamics. *Adv. Atmos. Sci.,* **26,** 373–380, doi:10.1007/s00376-009-0373-9.

Rui-Qiang, D., L. Jian-Ping, and H. Kyung-Ja, 2008: Nonlinear local Lyapunov exponent and quantification of local predictability. *Chin. Phys. Lett.,* **25,** 1919–1922, doi:10.1088/0256-307X/25/5/109.

Stephenson, D. B., C. A. S. Coelho, and I. T. Jolliffe, 2008: Two extra components in the Brier score decomposition. *Wea. Forecasting,* **23,** 752–757, doi:10.1175/2007WAF2006116.1.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74,** 2317–2330, doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

——, and ——, 1997: Ensemble Forecasting at NCEP: The breeding method. *Mon. Wea. Rev.,* **125,** 3297–3318, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Trevisan, A., and R. Legnani, 1995: Transient error growth and local predictability: A study in the Lorenz system. *Tellus,* **47A,** 103–117, doi:10.1034/j.1600-0870.1995.00006.x.

Wang, X., and C. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.,* **60,** 1140–1158, doi:10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2.

Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. H. Bishop, and X. Wang, 2006: Ensemble transform Kalman filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus,* **58A,** 28–44, doi:10.1111/j.1600-0870.2006.00159.x.

——, ——, ——, and ——, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus,* **60A,** 62–79, doi:10.1111/j.1600-0870.2007.00273.x.

Whitaker, J. S., and A. F. Loughe, 1998: The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.,* **126,** 3292–3302, doi:10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2.

Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano, 1985: Determining Lyapunov exponents from a time series. *Physica D,* **16,** 285–317, doi:10.1016/0167-2789(85)90011-9.

Yang, S.-C., C. Keppenne, M. Rienecker, and E. Kalnay, 2009: Applications of coupled bred vectors to seasonal-to-interannual forecasting and ocean data assimilation. *J. Climate,* **22,** 2850–2870, doi:10.1175/2008JCLI2427.1.

Yoden, S., and M. Nomura, 1993: Finite-time Lyapunov stability analysis and its application to atmospheric predictability. *J. Atmos. Sci.,* **50,** 1531–1543, doi:10.1175/1520-0469(1993)050<1531:FTLSAA>2.0.CO;2

Zhu, Y. J., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.,* **83,** 73–83, doi:10.1175/1520-0477(2002)083<0073:TEVOEB>2.3.CO;2.

Ziehmann, C., L. A. Smith, and J. Kurths, 2000: Localized Lyapunov exponents and the prediction of predictability. *Phys. Lett.,* **271A,** 237–251, doi:10.1016/S0375-9601(00)00336-4.