Impact of perturbation schemes on the ensemble prediction in a coupled Lorenz model

Qian Zou,[a,b] Quanjia Zhong,[a] Jiangyu Mao,[a] Ruiqiang Ding,[c] Deyu Lu,[a,b] Jianping Li,[d,e] and Xuan Li[f]

[a] *State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029, *China*

[b] *College of Earth Science, University of Chinese Academy of Sciences, Beijing* 100049, *China*

[c] *State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing* 100875, *China*

[d] *Frontiers Science Center for Deep Ocean Multispheres and Earth System (FDOMES)/Key Laboratory of Physical Oceanography/Institute for Advanced Ocean Studies, Ocean University of China, Qingdao* 266100, *China*

[e] *Laboratory for Ocean Dynamics and Climate, Pilot Qingdao National Laboratory for Marine Science and Technology, Qingdao* 266237, *China*

[f] *Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai* 200438, *China*

(*submitted to* ***Advances in Atmospheric Sciences***) *(Second revision)*

May 19, 2022

in press

ABSTRACT

Based on a simple coupled Lorenz model, we investigate how to consider a suitable initial perturbation scheme for ensemble forecasting in a multiscale system involving slow dynamics and fast dynamics. Four initial perturbation approaches are used in the ensemble forecasting experiments: random perturbation (RP), the bred vector (BV), the ensemble transform Kalman filter (ETKF) and the nonlinear local Lyapunov vector (NLLV) methods. Results show that, regardless of the method used, the ensemble averages behave indistinguishably from the control forecasts during the first few time steps. Due to different error growth in different time-scale systems, the ensemble averages perform better than the control forecast after a very short period of lead time in a fast subsystem, but after a relatively long period of time in a slow subsystem. As a result of coupled dynamic processes, whether adding perturbations to fast variables or to slow variables can contribute to an improvement in the forecasting skill for fast variables and slow variables. When it comes to the initial perturbation approaches, the NLLVs show higher forecasting skill than BVs or RPs overall. NLLVs and ETKFs had nearly equivalent prediction skill, and NLLVs won by a narrow margin. In particular, when adding perturbations to slow variables, independent perturbations (NLLVs and ETKFs) perform much better in the ensemble prediction. These results are simply implied in a real coupled air–sea model. For the prediction of oceanic variables, independent perturbations (NLLVs) and adding

in press

perturbations to oceanic variables will be expected to perform better in the ensemble prediction.

Key words: Ensemble prediction; The nonlinear local Lyapunov vector (NLLV); The ensemble transform Kalman filter (ETKF); Coupled air–sea models

https://doi.org/10.1007/s00376-022-1376-z

Article Highlights:

- This study explores ensemble prediction in a multiscale system which involve slow dynamics and fast dynamics by multiple initial perturbation schemes;
- The advantages of an ensemble forecast become apparent after a very short period of time in a fast subsystem, but after a relatively long period of time in a slow subsystem.
- When adding perturbations to slow variables in a multiscale system, independent perturbations (NLLVs and ETKFs) perform much better in the ensemble prediction.

in press

*Corresponding author: Ruiqiang Ding
Email: drq@bnu.edu.cn

in press

1. Introduction

In recent years, air–sea coupled models which describe the interactions between the atmosphere and the ocean have been more extensively applied to simulate weather and climate phenomena (Bender et al. 2007; Larson and Kirtman 2017; Mogensen et al. 2017; Zou et al. 2016). Air–sea coupling plays an important role in the simulation of weather and climate (Dong et al. 2017; Thompson et al. 2018). In the air–sea interface, it involves material and energy exchange, with a lot of complex physical processes (Soloviev et al. 2014). The coupled models can describe these coupled feedback processes better than atmosphere-only models (Perlin et al. 2020). Hence, the simulation of the weather and climate phenomena can be improved by using a coupled air-sea model (Dong et al. 2017; Fu and Wang 2004; Ratnam et al. 2008; Wang et al. 2005).

However, the simulation of weather and climate phenomena using coupled air-sea models involves many uncertainties, including initial condition uncertainty (Lorenz 1969, 1982) and model uncertainty (Leutbecher and Palmer 2008). The Ensemble prediction technology has been developed to deal with these uncertainties (Demeritt et al. 2007; Ehrendorfer 1997; Leith 1974). It generates ensemble members by adding perturbations to the analysis state (Magnusson et al. 2008). The ensemble mean of ensemble members can reduce the errors compared to a single forecast, and we can quantitatively estimate the probability density of a forecast state with a finite number of ensemble members (Feng et al. 2014; Froude et al. 2007; Leutbecher and

in press

Palmer 2008).

Here, we mainly focus on the ensemble prediction in relation to initial condition uncertainty. The key to constructing initial perturbations is to generate several initial states which can represent real initial uncertainty (Zhang and Krishnamurti 1999). Many ensemble initial perturbation methods have been developed in succession, such as the Monte Carlo method (also called the random perturbation (RP) method (Leith 1974)), the bred vector (BV) method (Toth and Kalnay 1993, 1997), the singular vector (SV) method (Palmer 1992), the ensemble transform Kalman filter (ETKF) method (Wang and Bishop 2003), the ensemble transform with rescaling (ETR) method (Wei et al. 2008; Wei et al. 2006), the conditional nonlinear optimal perturbations (CNOPs) method (Mu and Jiang 2008) and the nonlinear local Lyapunov vector (NLLV) method (Ding et al. 2017; Feng et al. 2014; Feng et al. 2016; Feng et al. 2018).

A surge of studies have focused on ensemble prediction in atmosphere-only or ocean-only models, but it has not been explored extensively in air–sea coupled models. Ensemble prediction in coupled models seems more complex because of the different time scales between the ocean and the atmosphere (Liu et al. 2013). An initial error can also evolve on different time scales (Vannitsem 2017). In addition, the feedback process between the coupled components makes the system highly sensitive to errors (Zhang et al. 2005). Hence, important issues in ensemble forecasting in coupled models which contain feedback processes at different time

in press

scales remain to be explored.

Therefore, in this paper, we determine how to add appropriate ensemble initial perturbations to a multiscale system, based on multiple initial perturbation methods. The system is called the coupled Lorenz model, with a slow subsystem coupled with a fast subsystem (Boffetta et al. 1998; Ding and Li 2012). The fast subsystem fluctuates approximately 10 times faster than the slow subsystem, which is close to the relative time-scale between the atmosphere and the ocean (Wang et al. 2002). Therefore, we can assume the coupled Lorenz model as a toy coupled air-sea model.

The remainder of this paper is organized as follows. Section 2 introduces the coupled Lorenz model and the algorithms to obtain the BVs, ETKFs and NLLVs. Section 3 presents properties of RPs, BVs, ETKFs and NLLVs in the multiscale system. Section 4 is a summary and discussion of our major findings.

2. Model and methodology

2.1. Coupled Lorenz model

The model used in this study is the coupled Lorenz model. It couples two simple Lorenz63 model (Lorenz 1963), with different time scales. The first characterizes the slow dynamics and the second characterizes the fast dynamics (Boffetta et al. 1998; Ding and Li 2012). It is governed by the equations

in press

$$
\begin{cases}
\dfrac{dX_1^{(s)}}{dt} = \sigma\left(X_2^{(s)} - X_1^{(s)}\right), \\
\dfrac{dX_2^{(s)}}{dt} = \left(-X_1^{(s)}X_3^{(s)} + r_s X_1^{(s)} - X_2^{(s)}\right) - \varepsilon_s X_1^{(f)} X_2^{(f)}, \\
\dfrac{dX_3^{(s)}}{dt} = X_1^{(s)}X_2^{(s)} - bX_3^{(s)}, \\
\dfrac{dX_1^{(f)}}{dt} = c\sigma\left(X_2^{(f)} - X_1^{(f)}\right), \\
\dfrac{dX_2^{(f)}}{dt} = c\left(-X_1^{(f)}X_3^{(f)} + r_f X_1^{(f)} - X_2^{(f)}\right) + \varepsilon_f X_1^{(f)} X_2^{(s)}, \\
\dfrac{dX_3^{(f)}}{dt} = c\left(X_1^{(f)}X_2^{(f)} - bX_3^{(f)}\right),
\end{cases} \tag{1}
$$

where the superscripts (s) and (f) denote the slow dynamics and the fast dynamics, respectively. The physical parameters of the above equation are displayed in Table 1. The relative time scale c is a constant set to 10, indicating that the fast dynamics fluctuate approximately 10 times faster than the slow dynamics. It is near the relative temporal scale of between ocean and the atmosphere, which is about 9 (Wang et al. 2002). The variation in the fast variables changes much faster than the variation in the slow variables (Fig. 1). The uncoupled slow and fast Lorenz models (coupling coefficients $\varepsilon_s = 0, \varepsilon_f = 0$) exhibit chaotic dynamics, with their Lyapunov exponents greater than zero. Setting $\varepsilon_s = 10^{-2}, \varepsilon_f = 10$, the maximal Lyapunov exponent in the coupled Lorenz model has a value of 11.5, close to the value from uncoupled fast Lorenz models (Boffetta et al. 1998) indicating that it is the error growth of the fast system that determines the maximal Lyapunov exponent in the coupled Lorenz model.

The associated attractor of the coupled system seems interesting from the physical parameters given in Table 1. The two-dimensional projections of the attractor are shown in Fig. 2. The fast dynamics appear to show a typical Lorenz model (Fig.

in press

2d–f), whereas the slow dynamics seems much more chaotic, losing the "butterfly" appearance of the original Lorenz63 model (Fig. 2a–c).

2.2 Initial perturbation schemes

We use four methods to generate initial perturbations: RP, BV, ETKF and NLLV. A brief description of the BV, NLLV and ETKF methods follows.

2.2.1 Computation of the BVs

The BV method is based on the rationale that any initial random errors in the basic flow would evolve into the fastest growing directions (leading Lyapunov vectors) in the phase space (Feng et al. 2014; Toth and Kalnay 1993, 1997). The generation of BVs is as described follows. At first, a group of small initial random perturbations are added to the analysis state. After a period of integration (a breeding cycle), the differences between the control and perturbed forecasts are rescaled to the size of the initial perturbations and the rescaled difference fields will be added to the next analysis. After repeating the process for several breeding cycles, the perturbation evolves into a fast-growing perturbation, and the BVs are generated. Following mathematical language is to describe the repeated process:

$$\boldsymbol{x}_p\left(t_i\right)=\boldsymbol{x}_c\left(t_i\right)+\varepsilon_0\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|}, \tag{2}$$

where the $\boldsymbol{x}_c$ and $\boldsymbol{x}_p$ represent the control trajectory and perturbation trajectory, respectively. The term $\varepsilon_0\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|}$ represent the scaling, where ε_0 is a scaling factor and $\boldsymbol{p}$ is the difference between control forecast and perturbated forecast.

in press

2.2.2 Computation of the NLLVs

NLLVs are a nonlinear extension of the Lyapunov vectors (LVs) similar to BVs (Feng et al. 2014; Hou et al. 2018). Compared to BVs, different NLLVs are independent, and represent the fastest direction of error growth in different subspaces of the phase space. The generation of NLLVs is introduced below (Feng et al. 2014; Feng et al. 2016). As shown in Fig. 3, the leading NLLV (NLLV1), which is the fastest growing direction, can be obtained via a breeding process similar to the creation of a BV. In each breeding cycle, the rest of the NLLVs can be obtained via a Gram–Schmidt reorthonormalization (GSR) process (Feng et al. 2014; Wolf et al. 1985). The evolved perturbations (grey dashed lines) are orthogonalized with the leading NLLV (NLLVn are orthogonalized associated with NLLV1, NLLV2, NLLV3, …, NLLVn-1). The orthogonalized perturbations are then scaled back to the initial size and enter the next breeding process. After multiple breeding cycles, the NLLVs are produced. In this paper, the breeding cycle for generating BVs and NLLVs is 0.05 time units (tus) and was repeated for 20 times.

in press

2.2.3 Computation of the ETKFs

The ETKF method is initially introduced by Bishop et al. (2001). The method is derived from ensemble-based data assimilation theory, which is associated with the Kalman filtering (Wang and Bishop, 2003; Wei et al., 2006; Wu et al., 2015). Similar to the ensemble Kalman filter (EnKF), ETKF apply Kalman filtering to generate a sample analysis ensemble. However, the ETKF use the forecast error covariance matrix only to estimate the analysis error covariance through a transformation matrix, not updating the mean state (Wang and Bishop, 2003; Zhou et al., 2019). The equation for the ETKF algorithm is as follows:

$$\boldsymbol{X}_{\mathrm{a}} = \boldsymbol{X}_{\mathrm{f}}\boldsymbol{T}, \tag{3}$$

where $\boldsymbol{X}_{\mathrm{a}}$ and $\boldsymbol{X}_{\mathrm{f}}$ is denoted as the analysis perturbation and forecast perturbation matrix and $\boldsymbol{T}$ is a transformation matrix. The detail computation process follows Hunt et al. (2005). Localization is not used here. A multiplicative covariance inflation factor (with a value of 1.3) is applied. The observation was produced by adding a random perturbation (following standard Gaussian distribution) to true state. Moreover, we use an ensemble size of 20, assimilated every 0.05 tus and the performing time is over 1 tus.

Studies have shown that the ETKF can be used for generate ensemble perturbations and have a better preformation on sampling the analysis uncertainties than most ensemble generation schemes (Wei et al., 2006; Feng et al., 2016). One of greatest qualities for ETKFs is that they are orthogonal in observation space (Wang and Bishop, 2003; Wei et al., 2006; Feng et al., 2016).

2.3 Experimental design

To make the performance of the evolution of the initial perturbations in a multiscale system as clear as possible, we undertook several ensemble forecasting experiments in the coupled Lorenz model, based on RP, BV, ETKF and NLLV

in press

methods. The model is integrated by Fourth Order Runge-Kutta method with a time step of 0.005 tus in all experiments. The procedure for the ensemble forecasting experiments is shown in Fig. 4. The first 10000 steps involve a spin up of the coupled Lorenz model. After the spin up, we use a 200-step ensemble Kalman filter (EnKF) data assimilation scheme (Evensen 2003, 2004) to create the initial analysis state. The parameter set of EnKF assimilation procedure is same to the ETKF scheme. The assimilation cycle is 0.05 tus, which is perfect to project to 6 hours window in real world. Hence, 1 tus is assumed to be equal to 5 days in real world in this paper. At the same time of assimilation process, the BV and NLLV perturbations are calculated based on the assimilated data as a basic flow. Then the ensemble perturbations created by the RP, BV, ETKF and NLLV methods are added to the analysis state in pairs (both positive and negative perturbations are added). The Ensemble perturbation vectors are scaled to 1×10^{-2}. The integration from the analysis state is the control forecast. And the perturbed forecasts are ensemble members. Increasing the number of ensemble members, the prediction level of ensemble forecast which is drove by BVs, NLLVs and ETKFs showed an improvement (not shown). Thus, the ensemble size is 6 pairs in this paper (with positive and negative perturbations superimposed in pairs). We run 10000 samples of the ensemble forecast (repeating the assimilation/breeding processes and forecasting processes). The initial value of each sample has one step interval. The initial states of 10000 samples include a representative range of coupled model states.

in press

2.4 Verification method

To evaluate the reliability of the ensemble predictions, a classical Brier score is applied to assess the relative skill of the BV compared with that of NLLV and ETKF. For any event ϕ, the Brier score (Brier 1950) is computed as:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2, \tag{4}$$

where N is the number of samples, f_i denotes the probability of the *i*-th sample for event ϕ prediction, and o_i denotes the probability of the *i*-th sample actually occurring for event ϕ (which can take on values of only 0 or 1).

3. Results

Before evaluating the quality of the ensemble predictions, the errors from the control forecast are going to be investigated. We assume that the model is perfect, and the true state is a long run of the model for each sample. As shown in Fig. 5, there exists a large difference between the control forecast and the true value. The evolution of the control and true value show rapid fluctuating changes over the whole system (Fig. 5a). When separating the coupled Lorenz system into a fast subsystem and a slow subsystem, similar characteristics are found in the fast subsystem compared to the whole system (Fig. 5b). However, these two time series in the slow subsystem show slow fluctuating changes, and they show a significant distinction until 4 tus (Fig. 5c). Given that there exists a relatively large difference between the control forecast and the true state, we use the Lyapunov exponential form error growth rate to

in press

measure the variation of forecast error for a control run. It is found that the initial error for analysis shows positive growth over time. The forecast error for a control run mainly comes from the fast subsystem. The variation in forecast error for the control run is different in the slow and fast subsystem, error growing much faster in fast subsystem than in slow subsystem (Fig. 5d). The equation for the Lyapunov exponential form error growth rate is as follows:

$$\lambda = \frac{1}{t - t_0} \ln \frac{\| \mathrm{V} x(t) \|}{\| \mathrm{V} x(t_0) \|}, \tag{4}$$

where t_0 is the initial time, $\| \mathrm{V} x(t) \|$ denotes the error size in the L_2 norm at time t.

Studies have proved that the ensemble forecast improves the quality of the control forecast (Ndione et al. 2020; Toth and Kalnay 1997). Running an ensemble of forecasts from adding perturbations to initial conditions, the ensemble mean can improve the prediction by filtering out unpredictable components, and the spread among the forecasts can provide a probability prediction (Toth and Kalnay 1993). In order to explore appropriate ensemble initial perturbations configuration in a multiscale system, many ensemble forecast experiments are conducted in this part, with multiple perturbation methods (RP, BV, ETKF and NLLV). The root-mean-square error (RMSE) for the ensemble mean and the ensemble spread are used to measure the forecast skill from the experiments. For a “perfect ensemble”, the ensemble spread will be close to the RMSE of the ensemble mean for all forecast times (Buckingham et al. 2010; Magnusson et al. 2008; Palmer et al. 2006).

in press

Additionally, considering different error growth in the fast subsystem and slow subsystem, we shall discuss them separately. In Fig. 6, the mean RMSE (solid lines) and ensemble spread (dashed lines) are plotted for the control (black), RP (red), BV (blue), ETKF (purple) and NLLV (green) after adding perturbations to all variables. The RMSE is oscillating at short lead time. It is possible that the RMSE oscillation at short lead- time relate our temporal scale, which is similar to diurnal cycling. In the first 0.5 tus, the RMSE for the NLLV, ETKF, BV, and RP ensembles are similar to that of the control run. This is mainly because positive and negative perturbations superimposed on the control run cancel each other out at the initial time (errors grow linearly at the initial time (Ding and Li 2007)). Soon after, regardless of the perturbation method, the ensemble forecast can effectively reduce forecast errors from the control run in general. In the RMSE for the ensemble mean, the results from NLLVs are the lowest, followed by ETKFs, BVs, RPs, and the control forecast. Among them, NLLVs and ETKFs have nearly the same forecast ability. These two methods have obviously better predictive skill than BVs in the two main periods: 0.5–2 tus and 4.5–8 tus (smaller RMSE for ensemble mean and bigger ensemble spread) (Fig. 6a). During the period 0.5–2 tus, the better predictive skill of NLLVs and ETKFs over the whole system is reflected mainly in the reduction in forecast errors in fast subsystem (Fig. 6b). And it is reflected in the reduction in forecast errors in slow subsystem during the period 4.5–8 tus (Fig. 6c).

Now we wonder whether to add perturbations to different variables of this

in press

system can achieve improvements from BVs to ETKF and NLLVs. Good ensemble perturbations should reflect the initial uncertainty of analysis (Toth and Kalnay 1993). Different perturbation methods have different ability in capturing the initial uncertainties. Owing to different error growth for initial perturbations in fast and slow subsystem (Fig. 5d), the prediction skill for different perturbation methods may differ when adding different timescale perturbations. Here, three error-addition schemes are used in this study: adding perturbations to both fast and slow variables, adding perturbations only to fast variables, and adding perturbations only to slow variables. It is shown that whether adding perturbations to fast variables or to slow variables contributes to an improvement in the forecasting skill for fast variables due to the feedback process between the coupled components (Fig. 7). When adding perturbations only to fast variables, the ensemble skills of all perturbation methods are improved in the prediction of fast variables after 0.4 tus (Fig. 7b). However, when adding perturbations only to slow variables, only NLLVs and ETKFs can improve the prediction skill of fast variables during the period 0.4–0.8 tus (Fig. 7c). In other words, only better independent perturbations superimposed on the slow subsystem can improve the forecasting skill of the fast subsystem.

The ensemble forecast of slow variables behaves differently with fast variables. The advantages of the ensemble forecast over the control forecast become apparent up to 4 tus (Fig. 8). When adding perturbations only to fast variables, the forecasting skills of BVs, ETKFs and NLLVs are equivalent (Fig. 8b). However, when adding

in press

perturbations only to slow variables, large differences are shown between the BVs with NLLVs and ETKFs, indicating that more independent perturbations have better prediction skill for the slow subsystem (Fig. 8c). Because of the feedback process between the coupled components, perturbations to both fast variables and slow variables contribute to an improvement in the forecasting skill for slow variables.

In general, the ensemble forecast shows a different performance in different time-scale systems. After a while the ensemble forecast starts to show better prediction skill than the control run. The advantages of the ensemble forecast become apparent after a very short period of time in a fast subsystem, but after a relatively long period of time in the slow subsystem. The reason for this difference is associated with the different error growths of different time-scale systems. In fast dynamics, errors from the analysis state grow quickly, whereas they will grow relatively slowly in slow dynamics. Besides, adding perturbations to both fast variables and slow variables contributes to an improvement in the forecasting skill for fast variables and slow variables, indicating that uncertainty in both fast and slow variables plays a role in the prediction of fast variables and slow variables. When adding perturbations to slow variables, it seems that independent perturbations (NLLVs and ETKFs) perform much better than the other types (BVs or RPs) in the prediction of both fast variables and slow variables. This is probably mainly because highly independent perturbations can better capture initial uncertainty information. Additionally, NLLVs seems to win ETKFs by a narrow margin (Fig. 6-8). To further confirm this, we conduct an

in press

independent samples t-test with RMSE of 10000 samples for the ETKF method and NLLV method. The RMSE data is from the experiments same as in Fig. 6a. The mean RMSE from NLLV is less than that from ETKF (with a difference of -0.1429), exceeding the 90% confidence level (with a probability value of 0.0812; not shown). Therefore, of the two independent perturbations, NLLV is better than ETKFs.

Other evidence also shows that the independent perturbations (NLLVs) show better forecasting quality than BVs. Figure 9 provides the distribution of RMSE and ensemble spread from 10000 samples for NLLV and BV predictions. At the beginning of the ensemble forecast, the forecast errors for both NLLVs and BVs are concentrated mainly around the diagonal, indicating that the forecasting skill of NLLVs is roughly equal to that of BVs (Fig. 9a). The number of samples with a prediction error by NLLVs less than that by BVs increase over time, reaching 58% of total samples at 6 tus (Fig. 9b). The number of samples with an ensemble spread by NLLVs is greater than that by BVs at any time (Fig. 9d–f). It is concluded that compared to BVs, NLLVs tend to have smaller RMSE for the ensemble mean and bigger ensemble spread, indicating a better ensemble prediction performance.

The Brier score (BS) is commonly used in evaluating the quality of probabilistic forecasts generated by ensembles (Stephenson et al. 2008). We choose the event ϕ_1 (for $X_3^{(f)}$ is the climatological mean to the distance of one standard deviation) (Fig. 10a) and event ϕ_2 (for $X_3^{(s)}$ is the climatological mean to the distance of one standard deviation) (Fig. 10b) to calculate the basic BS from the average of 10000 samples.

in press

The smaller the value of BS, the better the forecasting skill of the ensemble forecast. As shown in Fig. 10, the NLLVs are more skillful than the BVs or RPs and their performance are similar to ETKFs.

The other verification method used was the Talagrand diagram (also called a rank histograms), which can characterize the reliability of an ensemble forecast (Candille and Talagrand 2010; Talagrand et al. 1997). For a reliable ensemble forecasting system, the observation must fall with equal probability into any of the $N+1$ intervals divided by the N ensemble forecast values (Talagrand et al. 1997). Considering an ensemble forecasting system with N members, the predicted value of $X_3^{(s)}$ can be defined as $P_{i,j}$, where i denote the *i*-th sample, and j denotes the *j*-th ensemble member. For each sample, we count the number of members whose predicted values are smaller than the true values, represented as n, which can take on values of only $0-N$. Then, we count the number of samples (for all samples of S, we run 10000 samples of the ensemble forecast) under each n, defined by S_n. The ideal frequency of S_n is $S/(N+1)$, for which we expect the true value to have equal probability in the $N+1$ intervals. We calculate the relative frequency $P_n = \dfrac{S_\mathrm{n}}{S/(N+1)}$. The distribution of P_n is plotted in Fig. 11. It shows that both the NLLV, ETKF and BV ensembles are under-dispersive. But the results for NLLVs show a flatter histogram, indicating the greater reliability of the NLLV ensemble system. The stability of the NLLV and the ETKF ensemble system is comparable. The results from BS and the Talagrand diagram are based on the ensemble experiment

in press

which adds perturbations to both fast and slow variables and we choose $X_3^{(f)}$ and $X_3^{(s)}$ to analyze. Similar results can be obtained from other error-addition schemes and variables (not shown). These results prove the better performance of NLLVs and ETKFs than BVs in the ensemble prediction.

4. Summary and discussion

It is still a huge challenge for ensemble prediction in multiscale systems (Vannitsem and Duan 2020). One important issue is how to generate appropriate perturbations for different time-scale variables. This issue has been addressed here by considering different time-scale initial perturbations in the prediction of different time-scale variables. Besides, the selection of ensemble generation schemes is very important. The NLLV method has been proved to have many advantages in ensemble forecasting (Feng et al. 2014; Feng et al. 2016; Hou et al. 2018). Therefore, we have explored how to add appropriate ensemble initial perturbations to a multiscale system, based on multiple initial perturbation methods. The results are as below.

Compared to the control forecast, the ensemble forecast can effectively reduce forecasting errors in the coupled model. Due to different error growth in different time-scale systems, the advantages of an ensemble forecast become apparent after a very short period of time in a fast subsystem, but after a relatively long period of time in a slow subsystem. After adding perturbations separately to a fast subsystem and a slow subsystem, we found that, as a result of coupled dynamic processes, whether adding perturbations to fast variables or to slow variables contributes to an

in press

improvement in the forecasting skill for fast variables and slow variables. In terms of initial perturbation methods, it is evident that independent perturbations (NLLVs and ETKFs) are relatively superior to the other kinds (BVs or RPs). The two of them had nearly equivalent prediction skill, and NLLVs won by a narrow margin. The ensemble forecasting system based on NLLVs or ETKFs is of higher quality than that based on BVs. In particular, when adding perturbations to slow variables, the highly independent perturbations (NLLVs and ETKFs) can capture initial uncertainty information quickly, giving them better prediction skill in the coupled system.

We may deduce that in a coupled ocean–atmosphere model, for the prediction of fast-scale variables (e.g. atmospheric variables), the ensemble forecast works on reducing the errors from the control forecast after a short period of time. However, for slow-scale variables (e.g. oceanic variables), the ensemble forecast may be effective in improving the medium and long-term forecasts. Considering air–sea coupling, adding perturbations to oceanic variables will contribute to an improvement in the forecasting skill for atmospheric variables and adding perturbations to atmospheric variables can also improve the forecasting skill for oceanic variables. For the prediction of atmospheric and oceanic variables, when adding perturbations to oceanic variables, independent perturbations may perform better in the ensemble forecast. These results may have important implications for the development of ensemble forecasts of the coupled model in the future.

In general, NLLVs and ETKFs have a better performance than BVs and RPs in a

in press

coupled model. The computations of ETKFs require much computing resources. Fortunately, ETKF perturbations can be a byproduct from the data assimilation. Compared to ETKFs, NLLVs are completed independent (orthogonal) and easy to calculate. Both NLLVs and ETKFs are expected to have a wide potential application in coupled models.

Nevertheless, since we obtained these results through a toy model, further research is however necessary to expand these results to realistic air–sea coupled models. Our research team have tried applying NLLVs in the Weather Research and Forecasting (WRF) model. It is expected that NLLVs will have a good performance in realistic air–sea coupled models. Besides, Vannitsem and Duan (2020) discovered that the fastest backward Lyapunov vectors are not the most suitable for initializing a multiscale ensemble forecasting system. So how to choose the appropriate NLLV modes in a multiscale ensemble forecasting system may be an important issue. Both will be addressed in the near future.

Acknowledgments. This work was jointly supported by the National Natural Science Foundation of China (Grant No. 41790474, 41975070).

in press

REFERENCES

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The Operational GFDL Coupled Hurricane–Ocean Prediction System and a Summary of Its Performance. *Monthly Weather Review*, **135,** 3965-3989.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, **129**, 420-436.

Boffetta, G., P. Giuliani, G. Paladin, and A. Vulpiani, 1998: An Extension of the Lyapunov Analysis for the Predictability Problem. *Journal of the Atmospheric Sciences*, **55,** 3409-3416.

Brier, G. W., 1950: Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, **78,** 1-3.

Buckingham, C., T. Marchok, I. Ginis, L. Rothstein, and D. Rowe, 2010: Short- and Medium-Range Prediction of Tropical and Transitioning Cyclone Tracks within the NCEP Global Ensemble Forecasting System. *Weather and Forecasting*, **25,** 1736-1754.

Candille, G., and O. Talagrand, 2010: Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society*, **131,** 2131-2150.

Demeritt, D., H. Cloke, F. Pappenberger, J. Thielen, J. Bartholmes, and M.-H. Ramos, 2007: Ensemble predictions and perceptions of risk, uncertainty, and error in

in press

flood forecasting. *Environmental Hazards*, **7,** 115-127.

Ding, R., and J. Li, 2007: Nonlinear finite-time Lyapunov exponent and predictability. *Physics Letters A*, **364,** 396-400.

——, 2012: Relationships between the limit of predictability and initial error in the uncoupled and coupled lorenz models. *Advances in Atmospheric Sciences*, **29,** 1078-1088.

Ding, R., J. Li, and B. Li, 2017: Determining the spectrum of the nonlinear local Lyapunov exponents in a multidimensional chaotic system. *Advances in Atmospheric Sciences*, **34,** 1027-1034.

Dong, B., R. T. Sutton, L. Shaffrey, and N. P. Klingaman, 2017: Attribution of forced decadal climate change in coupled and uncoupled ocean-atmosphere model experiments. *Journal of Climate*, **30,** 6203-6223.

Ehrendorfer, M., 1997: Predicting the uncertainty of numerical weather forecasts: A review. *Meteorologische Zeitschrift*, **6,** 147-183.

Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53,** 343-367.

——, 2004: Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. 54(6), 539-560. *Ocean Dynamics*, **54,** 539-560.

Feng, J., R. Ding, D. Liu, and J. Li, 2014: The Application of Nonlinear Local Lyapunov Vectors to Ensemble Predictions in Lorenz Systems. *Journal of the Atmospheric Sciences*, **71,** 3554-3567.

In Press

Feng, J., R. Ding, J. Li, and D. Liu, 2016: Comparison of Nonlinear Local Lyapunov Vectors with Bred Vectors, Random Perturbations and Ensemble Transform Kalman Filter Strategies in a Barotropic Model. *Advances in Atmospheric Sciences*, **33,** 1036-1046.

Feng, J., J. Li, R. Ding, and Z. Toth, 2018: Comparison of Nonlinear Local Lyapunov Vectors and Bred Vectors in Estimating the Spatial Distribution of Error Growth. *Journal of the Atmospheric Sciences*, **75,** 1073-1087.

Froude, L. S., L. Bengtsson, and K. I. Hodges, 2007: The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. *Monthly Weather Review*, **135,** 2545-2567.

Fu, X., and B. Wang, 2004: Differences of boreal summer intraseasonal oscillations simulated in an atmosphere–ocean coupled model and an atmosphere-only model. *Journal of Climate*, **17,** 1263-1271.

Hou, Z., J. Li, R. Ding, J. Feng, and W. Duan, 2018: The application of nonlinear local Lyapunov vectors to the Zebiak–Cane model and their performance in ensemble prediction. *Climate Dynamics*, **51,** 283-304.

Hunt, B., E. Kostelich, and I. Szunyogh, 2005: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, **230**, 112-126.

Larson, S. M., and B. P. Kirtman, 2017: Drivers of coupled model ENSO error dynamics and the spring predictability barrier. *Climate dynamics*, **48,** 3631-3644.

in press

Leith, C. E., 1974: Theoretical Skill of Monte Carlo Forecasts. *Monthly Weather Review*, **102,** 409-418.

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *Journal of computational physics*, **227,** 3515-3539.

Liu, Z., S. Wu, and S. Zhang, 2013: Ensemble data assimilation in a simple coupled climate model: The role of ocean-atmosphere interaction. *Advances in Atmospheric Sciences*, **30,** 1235-1248.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*,, **20**, 130-141.

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

——, 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34,** 505-513.

Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between Singular Vectors and Breeding Vectors as Initial Perturbations for the ECMWF Ensemble Prediction System. *Monthly Weather Review*, **136,** 4092-4104.

Mogensen, K. S., L. Magnusson, and J. R. Bidlot, 2017: Tropical cyclone sensitivity to ocean coupling in the ECMWF coupled model. *Journal of Geophysical Research: Oceans*, **122,** 4392-4412.

Mu M. and Z. Jiang, 2008: A new approach to the generation of initial perturbations for ensemble prediction: conditional nonlinear optimal

in press

perturbation. *Chinese Science Bulletin*, **53**(13), 2062-2068.

Ndione, D. M., S. Sambou, S. Kane, S. Diatta, M. L. Sane, and I. Leye, 2020: Ensemble forecasting system for the management of the Senegal River discharge: application upstream the Manantali dam. *Applied Water Science*, **10,** 126.

Palmer, T., R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, and L. Smith, 2006: Ensemble prediction: a pedagogical perspective. *ECMWF newsletter*, **106,** 10-17.

Palmer, T. N., 1992: Ensemble prediction. ECMWF Tech. Mem., **188,** 85pp.

Perlin, N., I. Kamenkovich, Y. Gao, and B. P. Kirtman, 2020: A study of mesoscale air–sea interaction in the Southern Ocean with a regional coupled model. *Ocean Modelling*, **153,** 101660.

Ratnam, J. V., F. Giorgi, A. Kaginalkar, and S. Cozzini, 2008: Simulation of the Indian monsoon using the RegCM3–ROMS regional coupled model. *Climate Dynamics*, **33,** 119-139.

Soloviev, A. V., R. Lukas, M. A. Donelan, B. K. Haus, and I. Ginis, 2014: The air-sea interface and surface stress under tropical cyclones. *Scientific Reports*, **4,** 5306.

Stephenson, D. B., C. A. S. Coelho, and I. T. Jolliffe, 2008: Two Extra Components in the Brier Score Decomposition. *Weather and Forecasting*, **23,** 752-757.

Talagrand, O., R. Vautard, and B. Strauss, 1997: Evaluation of probabilistic prediction systems. *Proceedings, ECMWF Workshop on Predictability*.

in press

ECMWF, 1-25.

Thompson, B., C. Sanchez, X. Sun, G. Song, J. Liu, X.-Y. Huang, and P. Tkalich, 2018: A high-resolution atmosphere–ocean coupled model for the western Maritime Continent: development and preliminary assessment. *Climate Dynamics*, **52,** 3951-3981.

Toth, Z., and E. Kalnay, 1993: Ensemble Forecasting at NMC: The Generation of Perturbations. *Bulletin of the American Meteorological Society*, **74,** 2317-2330.

——, 1997: Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, **125,** 3297-3319.

Vannitsem, S., 2017: Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics. *Chaos*, **27,** 032101.

Vannitsem, S., and W. Duan, 2020: On the use of near-neutral Backward Lyapunov Vectors to get reliable ensemble forecasts in coupled ocean–atmosphere systems. *Climate Dynamics*, **55,** 1125-1139.

Wang, B., Q. Ding, X. Fu, I. S. Kang, K. Jin, J. Shukla, and F. Doblas-Reyes, 2005: Fundamental challenge in simulation and prediction of summer monsoon rainfall. *Geophysical Research Letters*, **32,** L15711.

Wang, X., and C. H. Bishop, 2003: A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes. *Journal of the Atmospheric Sciences*, **60,** 1140-1158.

Wang, Z., D. Wu, D. Chen, H. Wu, X. Song, and Z. Zhang, 2002: Critical time span

in press

and nonlinear action structure of climatic atmosphere and ocean. *Advances in Atmospheric Sciences*, **19,** 741-756.

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 Series A-dynamic Meteorology & Oceanography*, **58**, 28-44.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus Series A-dynamic Meteorology & Oceanography*, **60,** 62-79.

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 Series A-dynamic Meteorology & Oceanography*, **58,** 28-44.

Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano, 1985: Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, **16,** 285-317.

Wu, C., and Coauthors, 2015: Intercomparison of Targeted Observation Guidance for Tropical Cyclones in the Northwestern Pacific. *Monthly Weather Review*, **137**, 2471-2492.

Zhang, S., M. J. Harrison, A. T. Wittenberg, A. Rosati, J. L. Anderson, and V. Balaji, 2005: Initialization of an ENSO Forecast System Using a Parallelized Ensemble Filter. *Monthly Weather Review*, **133,** 3176-3201.

in press

Zhang, Z., and T. N. Krishnamurti, 1999: A Perturbation Method for Hurricane Ensemble Predictions. *Monthly Weather Review*, **127,** 447-469.

Zou, L., T. Zhou, and D. Peng, 2016: Dynamical downscaling of historical climate over CORDEX East Asia domain: A comparison of regional ocean‐atmosphere coupled model to stand‐alone RCM simulations. *Journal of Geophysical Research: Atmospheres*, **121,** 1442-1458.

Zhou, T., L. Ding, J. Ji, L. Li, and W. Huang, 2019: Ensemble transform Kalman filter (ETKF) for large-scale wildland fire spread simulation using FARSITE tool and state estimation method. *Fire Safety Journal*, **105**, 95-106.

in press

Table 1. Physical parameters used in the coupled Lorenz model

Parameter	Description	Value
σ	Prandtl number	10
b	Physical dimensions of the layer	8/3
c	Relative time scale	10
r_s	Rayleigh number of the slow dynamics	28
r_f	Rayleigh number of the fast dynamics	45
ε_s	Coupling coefficient of the slow dynamics	10^{-2}
ε_f	Coupling coefficient of the fast dynamics	10

in press

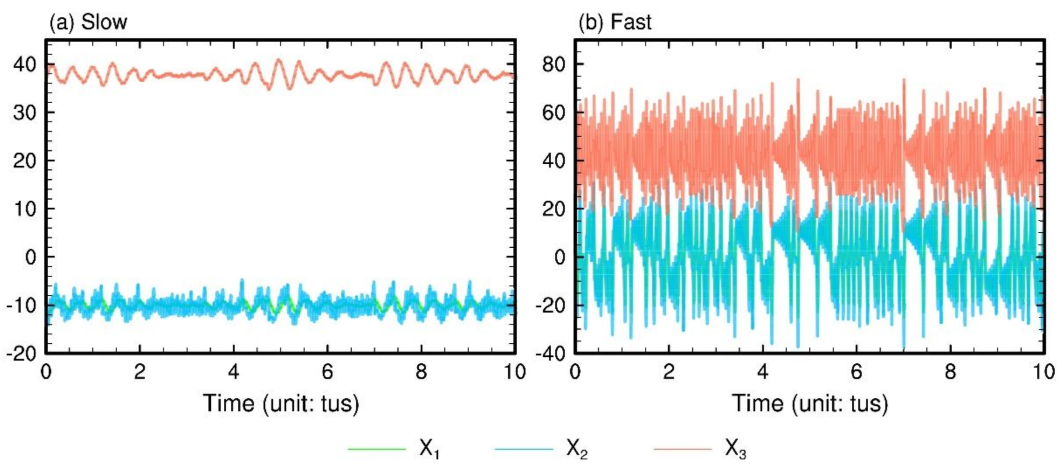

FIG. 1. Time evolution of variables for the coupled Lorenz model: (a) slow variables and (b) fast variables.

in press

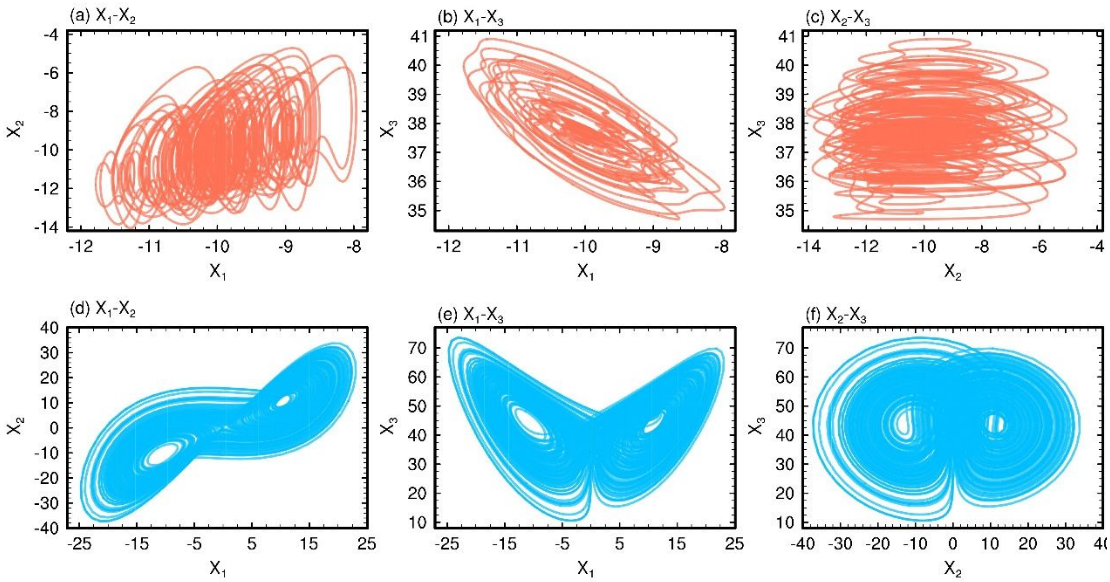

FIG. 2. Projections of the coupled Lorenz model on three two-dimensional planes: (a)–(c) for the slow variables, and (d)–(f) for the fast variables.

in press

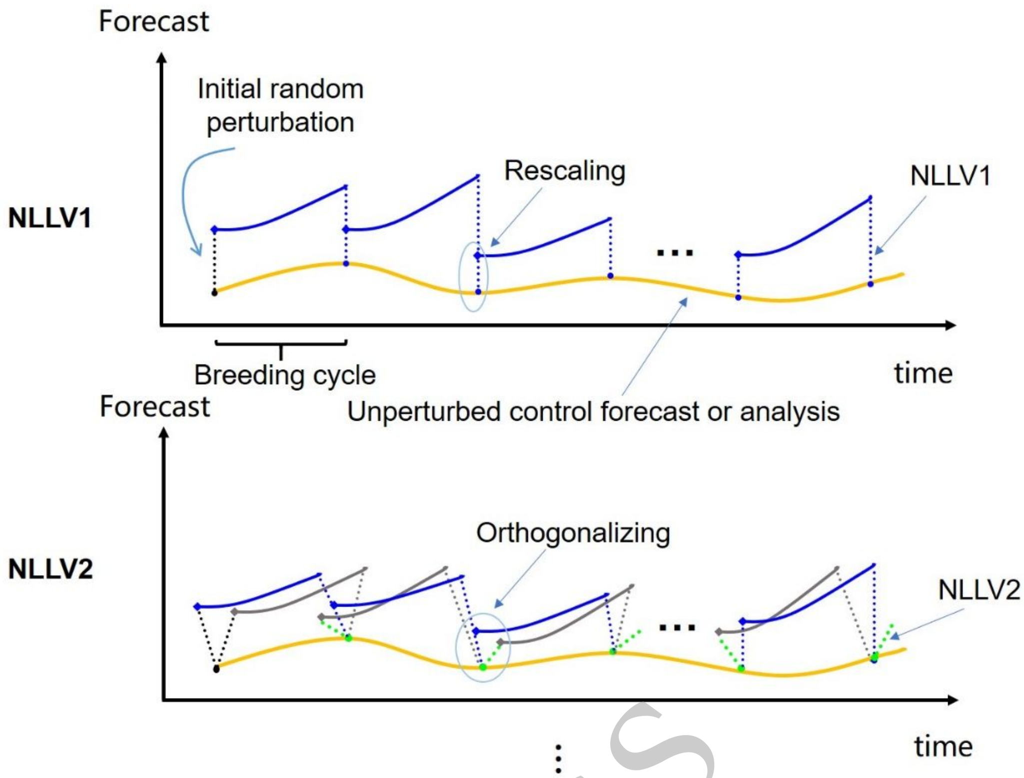

FIG. 3. Schematic diagram of the generation of NLLVs [adapted from Hou et al. (2018)]. The creation of NLLV1 is similar to the creation of BV. To acquire the NLLV2, a pair of RPs is initially added to the analysis state. The evolved perturbations (grey dashed line) are orthogonalized with the NLLV1 (blue dashed line) to produce the NLLV2 (green dashed line) using a Gram–Schmidt reorthonormalization (GSR) procedure. Similarly, NLLVn are orthogonalized with NLLV1, NLLV2, NLLV3, …, NLLVn-1.

in press

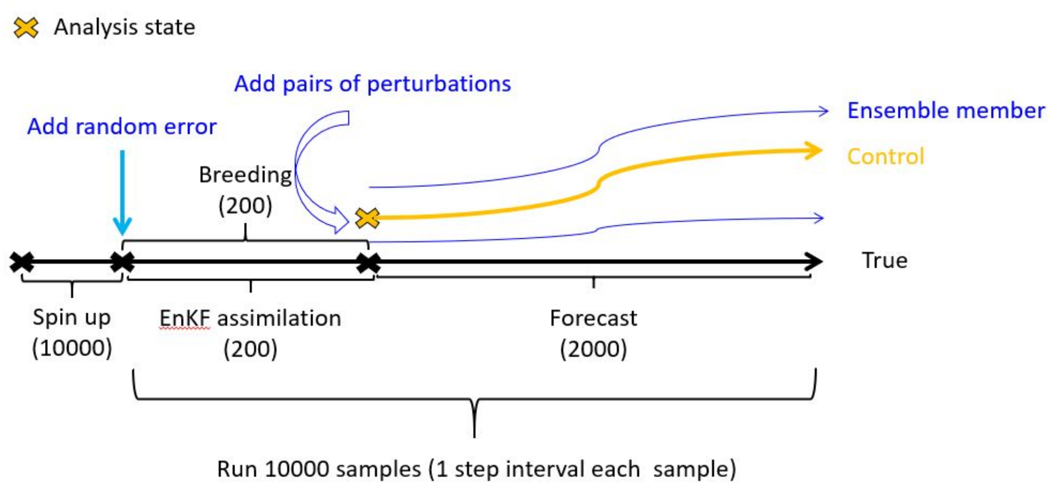

FIG. 4. Illustration of the initialization and forecasting procedure. Numbers represent the integration steps, and 1 step = 0.005tus.

in press

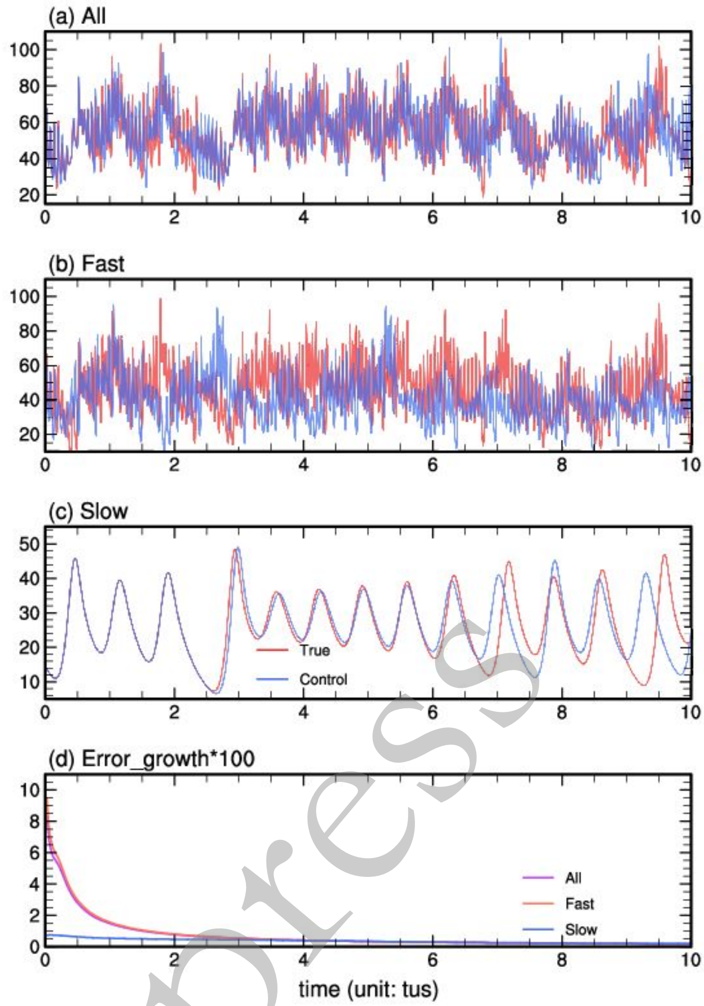

FIG. 5. (a)–(c) Evolution of control forecasts (light blue) against true state (light red) as a function of lead time, for (a) the whole system, (b) the fast subsystem, and (c) the slow subsystem (in the L_2 norm). (d) Mean growth rate in the form of Lyapunov exponent (value *100) of 10000 samples as a function of lead time from the coupled Lorenz model for the control run (the whole system (light purple), the fast subsystem (light orange), and the slow subsystem (light blue)).

in press

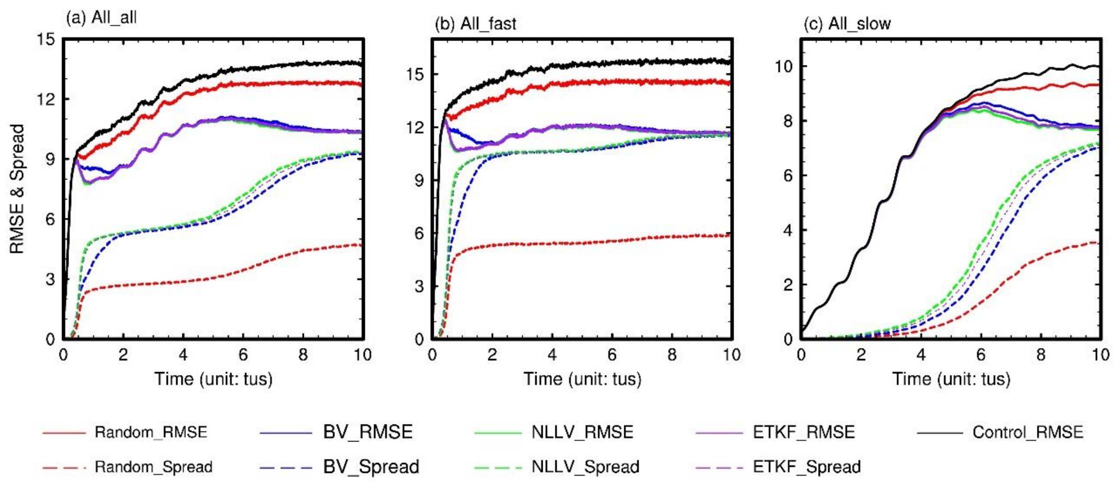

FIG. 6. Mean RMSE (solid lines) and ensemble spread (dashed lines) of 10000 samples as a function of lead time for the control run (black), RP method (red), BV method (blue), ETKF method (purple), and NLLV method (green) after adding perturbations to all variables. (a) the whole system, (b) the fast subsystem, and (c) the slow subsystem.

in press

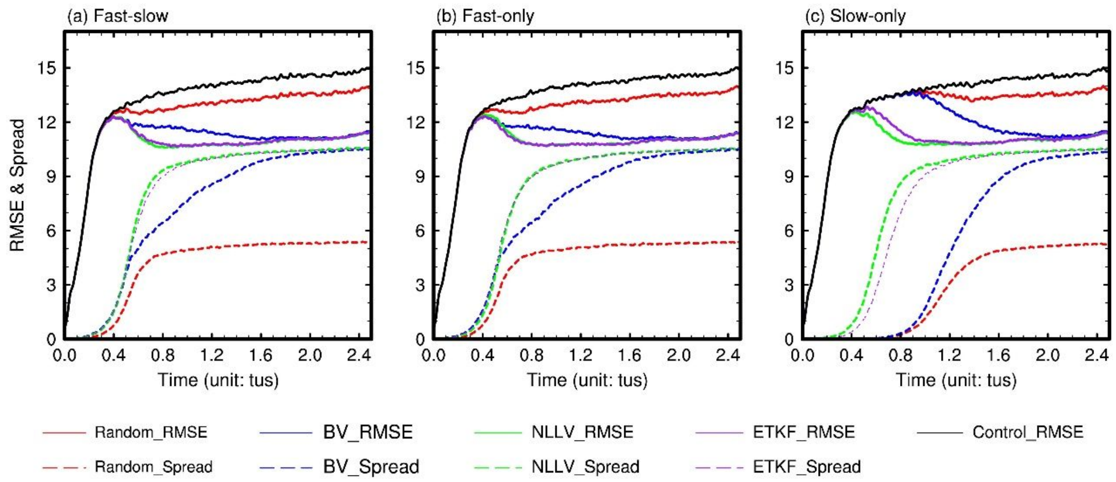

FIG. 7. Mean RMSE (solid lines) and ensemble spread (dashed lines) of 10000 samples in the fast subsystem as a function of lead time for the control run (black), random perturbation method (red), BV method (blue), ETKF method (purple), and NLLV method (green) after adding perturbations to different variables: (a) adding perturbations to both fast variables and slow variables, (b) adding perturbations only to fast variables, and (c) adding perturbations only to slow variables.

in press

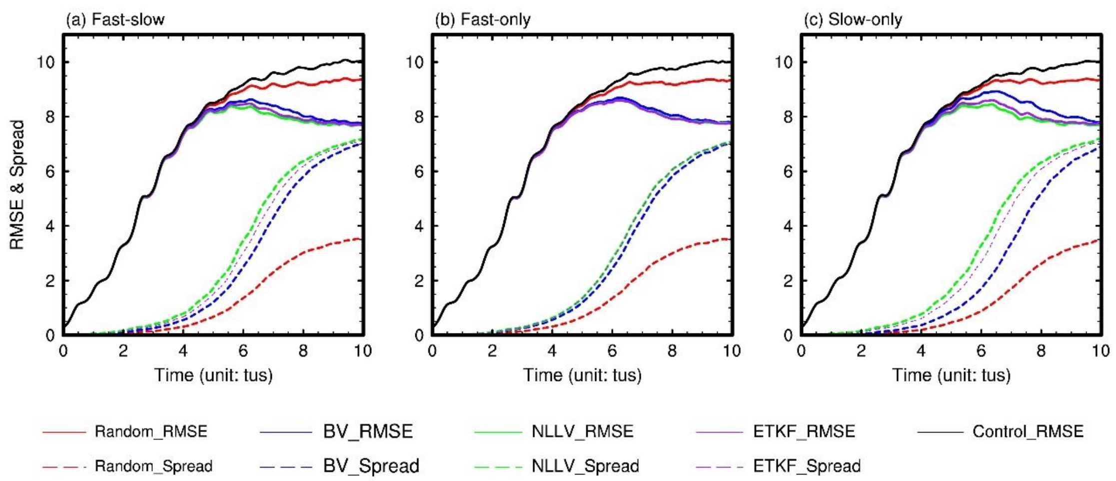

FIG. 8. Mean RMSE (solid lines) and ensemble spread (dashed lines) of 10000 samples in the slow subsystem as a function of lead time for the control run (black), random perturbation method (red), BV method (blue), ETKF method (purple), and NLLV method (green) after adding perturbations to different variables: (a) adding perturbations to both fast variables and slow variables, (b) adding perturbations only to fast variables, and (c) adding perturbations only to slow variables.

in press

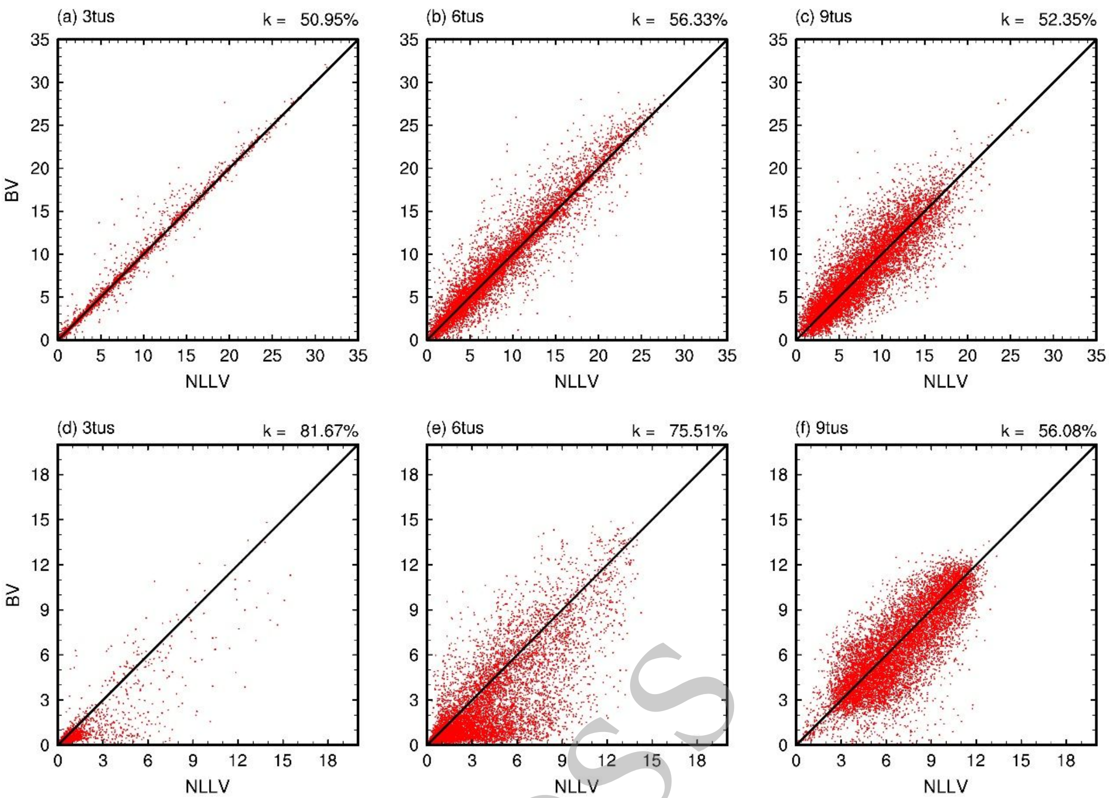

FIG. 9. (a)–(c) RMSE of 10000 samples based on NLLV and BV methods at (a) 3 tus, (b) 6 tus, (c) 9 tus in the slow subsystem. The upper right-hand corner indicates the ratio of samples where RMSE for the NLLV method is smaller than the RMSE for the BV method in (a)–(c). (d)–(f) The same as (a)–(c), but for an ensemble spread of 10000 samples. The upper right-hand corner indicates the ratio of samples where the ensemble spread for the NLLV method is larger than it is for the BV method in (d)–(f). (a)–(f) are based on the experiments which add perturbations to both fast and slow variables.

in press

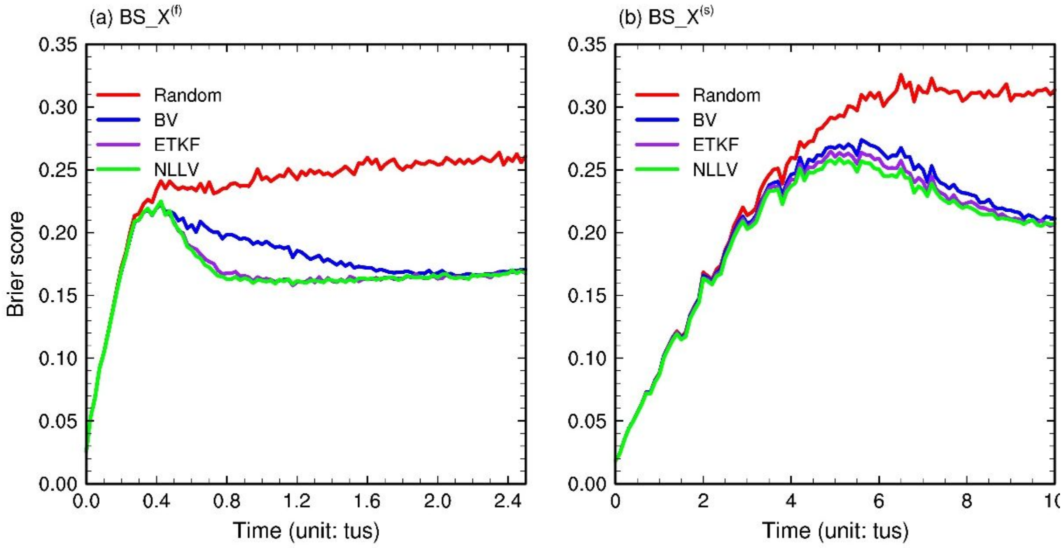

FIG. 10. (a) Basic Brier score (BS) for the event ϕ_1 (ϕ_1: where $X_3^{(f)}$ is the climatological mean to the distance of one standard deviation) of ensemble forecasts based on NLLVs (green line), ETKFs (purple line), BVs (blue line) and RPs (red line) as a function of lead time. (b) The same as (a), but for event ϕ_2 (ϕ_2: where $X_3^{(s)}$ is the climatological mean to the distance of one standard deviation)

in press

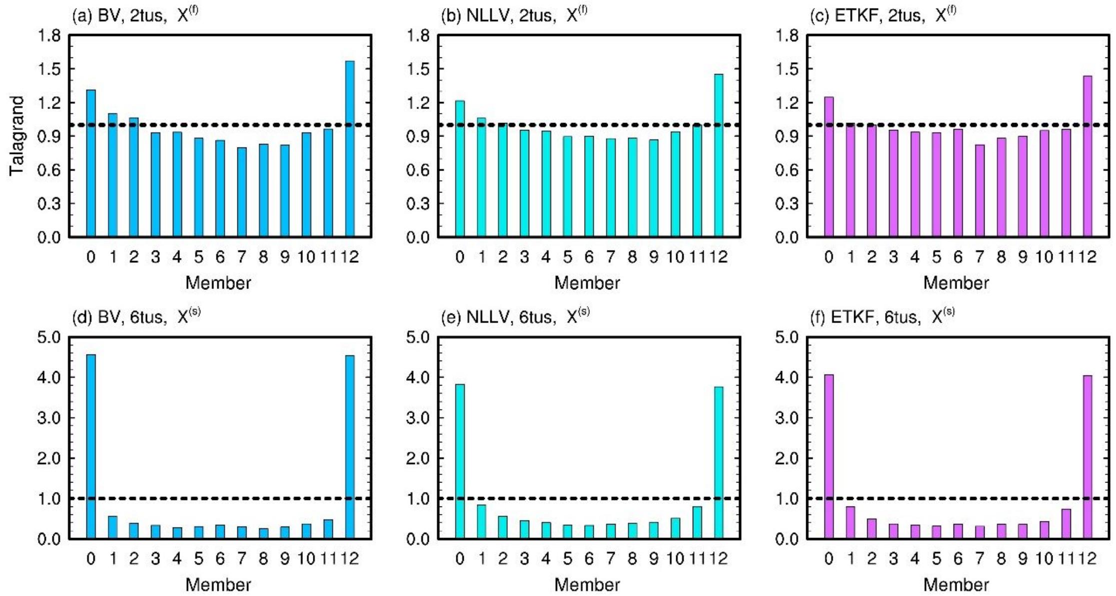

FIG. 11. The histogram of the Talagrand distribution for different member intervals. The horizontal dashed lines denote the expected probability. Ensemble forecasts based on (a) BVs, (b) NLLVs and ETKFs at 2 tus. (a)–(c) are based on the experiment which add perturbations to both fast and slow variables and predicts the variable $X_3^{(f)}$. (d)-(f) The same as (a)-(c), but at 6 tus and predicted variable is $X_3^{(s)}$.

in press